

Article



Multilingual and Multimodal Pedagogy and Assessment to Support Young Learners in CLIL

Yuen Yi Lo ^{1,*}, Linfeng Wei ² and Kaly Po Tin Ching ¹

¹ Faculty of Education, The University of Hong Kong, Hong Kong SAR, China

² Shenzhen Middle School Meixiang (Primary Section), Shenzhen 518049, China

* Correspondence: yuenyilo@hku.hk

How To Cite: Lo, Y. Y., Wei, L., & Ching, K. P. T. (2026). Multilingual and Multimodal Pedagogy and Assessment to Support Young Learners in CLIL. *Studies of Applied Linguistics in Asia*, 1(1), 192–208. <https://doi.org/10.53941/sala.2026.100016>

Received: 20 August 2026

Revised: 15 September 2026

Accepted: 17 September 2026

Published: 8 October 2026

Abstract: In Content and Language Integrated Learning (CLIL), students learn content knowledge through their additional language (L2). While students may develop their L2 more effectively in CLIL, they need to overcome some challenges when mastering content knowledge through L2 academic language. Hence, teachers and students are encouraged to mobilise different languages and semiotic resources to communicate and construct knowledge in CLIL classrooms. Similarly, the mobilisation of multilingual and multimodal resources should be embraced in CLIL assessment, so that students can better demonstrate their learning outcomes. Yet, limited research has investigated multilingual and multimodal pedagogy and assessment in primary CLIL classrooms. This design-based research addressed this gap by planning and implementing multilingual and multimodal teaching across two cycles in a public primary school in China. Multilingual and multimodal assessment tasks, including one written test and one poster presentation task, were also experimented with the students. By comparing students' performance on the written pre- and post-tests, together with data collected from lesson observations and student interviews, this study shows the effectiveness of the teacher's multilingual and multimodal pedagogy in facilitating students' content and L2 learning. While students' performance on the multilingual and multimodal written post-test was not significantly different from that on the conventional written post-test, their content knowledge, spoken language and presentation skills, and creativity demonstrated through the poster design task were impressive. These findings not only help justify multilingual and multimodal pedagogy and assessment, but also illuminate CLIL pedagogy and assessment practices.

Keywords: Content and Language Integrated Learning (CLIL); multilingualism; translanguaging; multimodality; assessment

1. Introduction

With globalisation and international mobility, Content and Language Integrated Learning (CLIL), where students learn non-language content subjects through their additional language (L2), continues to spread to different educational contexts, particularly English-as-a-foreign-language (EFL) ones such as the European continent and Asia (Dalton-Puffer et al., 2022). The (assumed) integration of content and language learning is often regarded as an alternative approach to facilitating

foreign language education. However, there are also challenges that teachers and students need to overcome when learning abstract content knowledge and academic language in an L2 simultaneously (Fazzi & Menegale, 2023; Pun et al., 2024). Researchers have been advocating pedagogical translanguaging, where teachers and students mobilise different named languages and semiotic resources in their communicative repertoire to facilitate both content and language teaching in CLIL and other multilingual classrooms (Cenoz & Gorter, 2020; Li, 2018). Empirical studies showing how teachers and



Copyright: © 2026 by the authors. This is an open access article under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Publisher's Note: Scilight stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

students engage in multilingual and multimodal meaning-making process were mainly conducted in university and secondary school settings (e.g., He, 2025; Liu et al., 2026). It would be interesting to see how similar practices may also work in primary classrooms to support young learners (Liu & Lin, 2025).

While research has well documented pedagogical strategies and processes involving multilingual and multimodal resources, the same has not been extended to assessment, which remains largely monolingual and monomodal (Schissel et al., 2021). This represents a contradiction between instruction and assessment (Lin, 2024), and warrants the risk of underestimating CLIL students' learning outcomes (Lo et al., 2023). This design-based research (DBR) seeks to address such a misalignment by systematically incorporating multilingual and multimodal resources into teaching and assessment. The findings will illuminate the potential of multilingual and multimodal pedagogy and assessment, thereby yielding practical implications for CLIL.

2. Literature Review

2.1. CLIL for Young L2 Learners—Potential and Challenges

CLIL can be regarded as a variant of bilingual education programmes where students learn non-linguistic content knowledge through their L2. It is often discussed in relation to other programmes such as English Medium Instruction (EMI) and immersion, which are labels used to refer to programmes in other educational contexts, for different educational levels and students with different socio-linguistic profiles (Banegas & Zappa-Hollman, 2023; Cenoz et al., 2014). Following Dalton-Puffer et al. (2022), we conceptualise CLIL as a “context” for learning an L2. This language learning context can be positioned somewhere between instructed and naturalistic language learning environments. At the same time, CLIL is regarded as a pedagogical “approach” that integrates content and language teaching and learning. These two conceptualisations of CLIL nicely summarise the potential and challenges of CLIL.

Based on well-established Second Language Acquisition theories, it is believed that students can learn an L2 more effectively by being exposed to more L2 input, using more L2 to interact with others and express their meanings, and engaging in more authentic and meaningful tasks provided by content subjects (Lyster & de Zarobe, 2018). Hence, CLIL is often regarded as an alternative, or more innovative, approach to L2 teaching (Dalton-Puffer et al., 2022). However, as a pedagogical approach that integrates content and language learning, CLIL requires students to learn content knowledge and L2 simultaneously. Content knowledge involves concepts related to different topics in a discipline (e.g., Science, Geography) that can be cognitively challenging, while L2 includes both communicative language used in classroom

interaction and academic language associated with constructing content knowledge (e.g., subject-specific vocabulary, general academic vocabulary, sentence patterns expressing academic functions, genres) (Lin, 2015). Academic language is not intuitive and is challenging even for native speakers of the language (Moje, 2015). Hence, it is not easy for L2 learners to master academic language to express their content knowledge (Lo & Lin, 2019).

These potential benefits and challenges have been shown in empirical studies. Recent meta-analyses (Kaiypova et al., 2025; Lee et al., 2023, 2025) show that CLIL primary and secondary students outperformed their counterparts in mainstream education (i.e., learning content through their mother tongue [L1]) in their L2 performance, especially speaking for primary students and vocabulary for secondary students. However, CLIL students' academic performance or achievement was not on par with that of students learning through their L1, when their prior differences were controlled for. Hence, while CLIL can provide a conducive environment for L2 development, teachers need to better support their students in learning content knowledge through academic language in an L2. This is especially the case for young learners, who are still developing their oracy and literacy skills in both L1 and L2, not to mention L2 academic language (Fazzi & Menegale, 2023).

2.2. Multilingual and Multimodal Scaffolding in CLIL

To help CLIL learners overcome language barriers when learning content knowledge through an L2, researchers have urged teachers to mobilise all resources in learners' communicative repertoire, which include different named languages and other meaning-making modes such as visuals, gestures, models, experiments (Ruiz de Zarobe & Querol-Julián, 2025). Such advocacy aligns with the multilingual turn of foreign language education (May, 2013), which values learners' other language(s) and associated learning experience, and echoes the promotion of pedagogical translanguaging and trans-semiotising to facilitate knowledge co-construction in classrooms where students are learning through an L2 (Cenoz & Gorter, 2020). Lin's (2015) Multimodalities-Entextualisation-Cycle (MEC) presents a pedagogical framework for CLIL teachers to systematically incorporate different languages and semiotic resources to help students unpack difficult content knowledge, analyse the academic language involved, and repack (or entextualise) their understanding through appropriate academic language. Classroom-based studies have shown how teachers applied the MEC to support their students, shifting between everyday and academic language, L1 and L2, as well as using video clips, pictures, blackboard sketches, graphic organisers, writing and speaking templates (e.g., He & Lin, 2024). Similarly, other studies have demonstrated how teachers and

students mobilised different languages and semiotic resources (e.g., gestures, artefacts) to co-construct knowledge and language (e.g., Nikula et al., 2024). However, most classroom-based studies on multilingual and multimodal teaching in CLIL were contextualised in tertiary or secondary contexts, leaving primary classrooms with young L2 learners underexplored (Liu & Lin, 2025). Two recent studies provide some interesting insights. Amondarain-Garrido and Ruiz de Zarobe (2025) analysed the multimodal translanguaging episodes of three CLIL teachers in a Basque public primary school in Spain. They observed 159 multimodal translanguaging episodes from a corpus of 16 h of lesson videos (Grade 5–6, 10–12 years old). From these episodes, the three teachers combined languages (English, Basque and Spanish), gestures, movements around the classroom space, tools and resources to promote students' participation in the lessons. They also mobilised various communicative resources to express particular discourse functions such as asking a question, giving an explanation, indicating a procedure and giving positive feedback. These functions were important not only in knowledge construction, but also in establishing relationships among teachers and students.

Liu and Lin (2025) worked with teachers in two primary schools in Hong Kong to design CLIL units for Grade 1 students (6–7 years old), some with special educational needs. They analysed five videotaped lessons to identify how teachers collaborated spontaneously in class to orchestrate various semiotic resources including L1, L2, oral explanation, board writing, pictures and gestures to help the young learners understand some key thematic patterns of the target genre (e.g., the characteristics of an animal and how they can be described in a simple information report) and to manage classroom discipline. In one of the schools, the local and native English teachers collaborated to support students to present their written multimodal report in the form of a mini-drama, which was a new activity for the students. The teachers made good use of drawings, gestures and body movements to provide clear instructions and support students' performance.

This study aims to enrich this body of research by exploring how systematic incorporation of multilingual and multimodal resources during instruction may facilitate primary students' learning in CLIL classrooms in China, where bilingual education programmes have grown exponentially at different educational levels (Hu et al., 2023).

2.3. Underexplored Multilingual and Multimodal Assessment in CLIL

While multilingual and multimodal pedagogy has been widely researched (particularly in secondary and tertiary classrooms), a similar line of research has been lacking in assessment (Schissel et al., 2021). Scholars have repeatedly argued for allowing multilingual students to mobilise their whole communicative repertoire in content

assessment (e.g., through providing bilingual questions, allowing students to express their answers in any languages) (Gorter & Cenoz, 2017; Shohamy, 2022). The rationale behind this is to enable students to demonstrate their learning more effectively, without language barriers imposed by their L2. Similarly, multimodal resources such as visuals, tables and graphs should be incorporated in assessment so as to allow students to understand the questions and retrieve relevant information from their long-term memory more easily (Lindner et al., 2017). Students should also be allowed to use non-linguistic resources to express their content knowledge (Grapin & Llosa, 2022). The scant empirical evidence to date has shown that when students were allowed to use all linguistic and multimodal resources when tackling content assessment, their performance was better (Archila et al., 2021; Grapin & Llosa, 2022). In other words, multilingual and multimodal assessment can gauge students' actual learning outcomes more effectively. However, it is intriguing that teachers have been adopting multilingual and multimodal pedagogy, but assessment remains largely monolingual and monomodal (Lin, 2024; Mearns et al., 2025). Some obstacles to multilingual and multimodal assessment may include other stakeholders' beliefs in monolingualism or expectations about the "official" language of communication in CLIL (Oattes et al., 2018), concerns about fairness between CLIL and non-CLIL students, and wash-back of the high-stakes examination (Lo, 2026; Otto & Estrada, 2019). More empirical studies are needed to explore the potential of multilingual and multimodal assessment, and see how it can be incorporated into schools' or teachers' practices. This study therefore attempts to align multilingual and multimodal pedagogy and assessment, and illustrate how this alignment may facilitate primary school students' learning in CLIL. It addresses two research questions:

- (1) How can multilingual and multimodal pedagogy and assessment facilitate CLIL students' content learning?
- (2) How do primary school students perceive the affordances of multilingual and multimodal pedagogy and assessment?

3. Methodology

3.1. Overall Research Design

This study adopted a design-based research (DBR) methodology to investigate how CLIL lessons, integrated with multilingual and multimodal pedagogy and assessment, can facilitate content and language learning among young EFL learners. DBR is appropriate for this study because it seeks to improve educational practices through iterative cycles of design, implementation, and analysis (Philippakos et al., 2021). It brings researchers and practitioners together to co-design and iteratively

test pedagogical innovations in authentic classroom contexts (Amundrud & Rasmussen, 2026).

In this study, the research team and the teacher (second author) jointly designed two cycles of teaching. They selected the unit topics, developed the lesson plans, and constructed the pre- and post-tests based on students' levels. The teacher implemented the planned lessons to scaffold student learning, and gained insights for designing more CLIL lessons. The research team provided the theoretical grounding in CLIL and designed tests to evaluate the effectiveness of the pedagogy and assessment formats. This researcher–practitioner partnership enabled the DBR approach to investigate how the pedagogical innovation worked in this specific EFL classroom context.

Both teaching cycles focused on human body systems (more details will be presented later), which were challenging for primary learners and required the ability to describe scientific processes. Selecting topics from the same subject area minimised cross-disciplinary variability and made it easier to track student progress. After collaborating on the lesson plans, the teacher implemented four lessons in each cycle, while the researcher observed, provided feedback, and collected data. Reflections from Cycle 1 informed the design of Cycle 2, enabling iterative refinement of both pedagogy and assessment.

3.2. Research Context and Participants

This study was conducted at a public primary school in a district in Shenzhen, a region known for its great educational resources and faculty quality. The school has six grade levels, 42 classes, and approximately 42–46 students per class. The school admits students from the local area through the district's allocation system, and its students have performed slightly above the district average in written examinations. Family backgrounds vary, though most families are middle-class.

The school encourages teachers to explore new pedagogical approaches and supports professional development through extra-curricular programmes. These programmes include academic, arts, and sports electives, created by teachers and implemented on Thursday afternoons. While the school's English team had explored CLIL-related topics for two years, the lack of resources meant that their practice was largely confined to adapting content from existing English textbooks and remained language-focused. This prompted the teacher (the second author) to develop a genuine CLIL course that would integrate both content and language learning, with greater disciplinary rigor and scientific literacy in the public school. It is designed for students who are keen to develop their ability to express themselves across disciplines in English, both orally and in writing. The school is very supportive of this elective.

The teacher of this CLIL elective, Miss W, majored in English education for her undergraduate and holds a

master's degree in CLIL. When the study was conducted, she had been teaching English at the school for two years, and had tried incorporating CLIL-based topics into her regular English classes, covering Science and Arts. Through collaborating with the research team, Miss W developed a clearer understanding of how to integrate multilingual and multimodal strategies in CLIL contexts.

The participating 33 students (Grade 5–6, aged 10–12) enrolled in this year-long CLIL elective course, attending two 40-min lessons per week. They chose the course out of personal interest, rather than being selected by the school. Their parents learned about CLIL through online information, and were willing to assist with extra practice and homework. The students demonstrated above-average English proficiency, performing well in daily communication and English examinations, as well as strong abilities across other subjects. However, they had never studied academic content such as Science through English in their regular classes, making the CLIL elective a considerable challenge.

3.3. Design of Multilingual and Multimodal Pedagogy Across Two Cycles

Each teaching cycle consisted of four 40-min lessons across two weeks. The first teaching cycle focused on “The Heart”, expecting students to learn about the heart structure, heart rate and blood flow processes. The second cycle focused on “The Digestive System”, progressing from organ functions, digestive processes to healthy eating habits. The two cycles were implemented in two semesters respectively, with a five-month interval.

Across both cycles, multilingual and multimodal resources were employed to support student learning. The design was informed by the MEC (Lin, 2015). First, transanguaging was systematically planned to reduce EFL learners' anxiety and relate the new subject-specific terms to their daily life experience. For example, Miss W provided L1 translation equivalents for key terms (e.g., “Right Atrium—右心房” in Cycle 1; “esophagus—食管” in Cycle 2), and used L1 to explain complex processes such as blood flow direction or the function of gastric juice, allowing students to grasp scientific concepts before producing English description. Students were also permitted to use Chinese during discussion activities when they lacked English vocabulary or expressions.

Transanguaging also occurred spontaneously during classroom interaction. This usually happened when the teacher wanted to engage students or elicit their responses to some questions. After students expressed their ideas in L1, the teacher would paraphrase it in L2, thereby demonstrating the appropriate academic language (i.e., repackaging). Excerpt 1 serves as an example to illustrate such episodes (see the Appendix A for the transcription conventions). In the excerpt, which took place in Lesson 3 of Cycle 2, the teacher described a

symptom in English and used L1 to bridge to the target vocabulary.

Excerpt 1. The teacher was talking about problems related to the digestive system.

Miss W: *When you have this problem, it's hard to poop. The problem is called ____?*

[No response from students]

Miss W: *In Chinese, it's ____.*

S1: 便秘 (constipation).

Miss W: *Yes, constipation. Follow me, cons-ti-pa-tion.*

When it comes to multimodal resources, in Cycle 1, the teacher employed multimodal resources such as 3D heart models, YouTube animations, and pictures with arrows to indicate blood flow. Students also participated in activities such as measuring their pulse before and after exercise, and completed a line graph worksheet to record their heart rate changes (Figure 1). In Cycle 2, the teacher used different multimodal strategies: students sequenced physical organ cards to understand the food journey, completed a matching worksheet to pair organs with their functions, participated in a digital matching game for organ-function competition, and mimicked food movement on the blackboard with pictures (Figure 2).

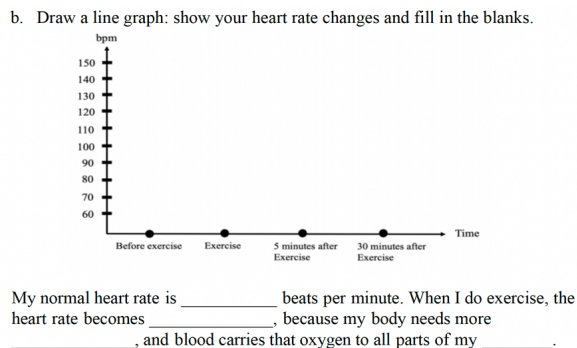


Figure 1. Line graph worksheet for recording heart rate changes before and after exercise.

3.4. Design of Multilingual and Multimodal Assessment Across Two Cycles

To investigate whether multilingual and multimodal assessment could support students' content and language learning, students were permitted to use their L1 (Chinese) in the pre-test when they did not know how to express their knowledge in English. The post-test was designed in two versions in each cycle. Set A was a conventional English-only version, while Set B incorporated L1 for key terms and more visual representations (e.g., diagrams, illustrations). Both versions had the same content questions but different linguistic and visual support. For example, in Cycle 1,

students were asked to sequence the order of blood flow processes. Set A only included sentences, but Set B included both sentences and pictures. In the "case study" (data-based) section where students had to read one or two scenarios and answer the subsequent questions, Set A only presented the scenario with words, while Set B provided a picture representing the scenario. Figure 3 shows the differences in the post-test of Cycle 1.

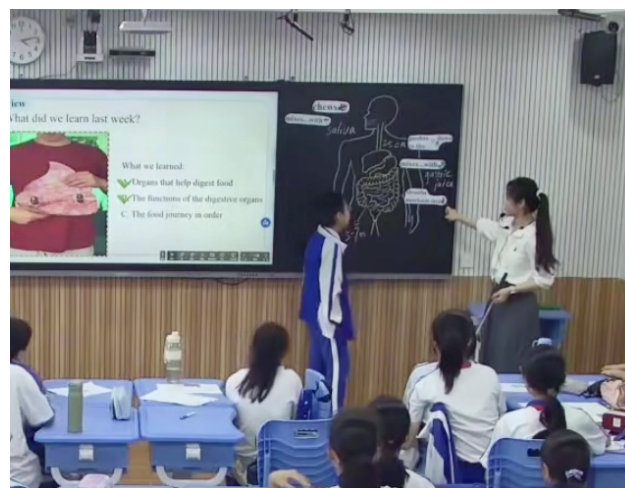


Figure 2. Student presenting the food journey using a blackboard sketch and labels.

L1 was also incorporated into Set B to provide more support. For example, in the Cycle 2 post-test, there was a cloze section asking students to fill in some blanks to describe the functions of different organs in the digestive system. In Set A, students completed the questions with no L1 support. In Set B, they were told they could use Chinese if they could not recall the English terms, with L1 cues provided for key terms (e.g., *squeezes* (挤压)) (see Figure 4).

In addition to the conventional written tests, a multimodal performance-based assessment task was implemented in Lesson 4 of both teaching cycles. In this assessment task, students worked in groups of 5–6 to create a poster summarising the unit content and delivered an oral presentation. The poster combined visuals (e.g., labelled diagrams, line graphs) with text (subject-specific words, explanatory sentences, key notes) to demonstrate content understanding (some examples will be presented in the Findings section). To support students to complete this assessment task, the teacher provided specific instructions for the poster content and a speaking template for oral presentation. This poster assessment task was formative in nature, as it was not marked by the teacher or the research team. However, the teacher invited the students to give comments on others' posters after each presentation, with the prompt "One thing I learned: ...; One thing I want to suggest/say: ...".

Case study (4 marks)

11. Before Amy starts rope jumping, her average heart rate is 70 bpm (beats per minute). After 1 minute of jumping, her heart rate has rose to 110 bpm. After 5 minutes of resting, her heart rate returns to 80 bpm. After 30 minutes of resting, her heart rate returns to 70 bpm.

Case study (4 marks)

11. Before Amy starts rope jumping, her average heart rate is 70 bpm (beats per minute). After 1 minute of jumping, her heart rate has rose to 110 bpm. After 5 minutes of resting, her heart rate returns to 80 bpm. After 30 minutes of resting, her heart rate returns to 70 bpm. The graph below shows Amy's heart rate at different time points.

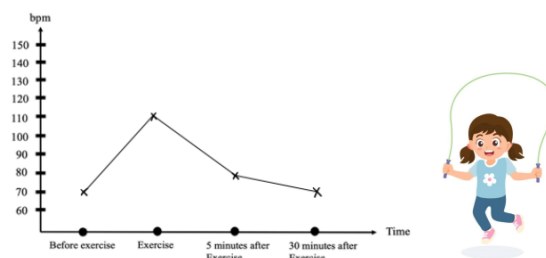


Figure 3. Comparison of the data-based question in Set A (left, text-only) and Set B (right, with visuals) of the Cycle 1 post-test.

Part D: Fill in the blanks

11. Complete the sentences below describing the functions of different organs in the digestive system. Please note that they do not follow the order.

- The esophagus squeezes and _____ food down to the stomach.
- The small intestine absorbs _____ into the blood.
- The mouth _____ food to make it into smaller pieces and mixes the food with saliva.
- The stomach further breaks down food with _____ juice.
- The large intestine absorbs water and makes _____.

Part D: Fill in the blanks

11. Complete the sentences below describing the functions of different organs in the digestive system. Please note that they do not follow the order. *If you do not remember the English words, you can use Chinese.*

- The esophagus squeezes (挤压) and _____ food down to the stomach.
- The small intestine absorbs (吸收) _____ into the blood.
- The mouth _____ food to make it into smaller pieces and mixes the food with saliva (唾液).
- The stomach further breaks down food with _____ juice.
- The large intestine absorbs water and makes _____.

Figure 4. Comparison of the cloze section in Set A (left, English only) and Set B (right, L1 incorporated) of the Cycle 2 post-test.

3.5. Data Collection and Analysis

To address the two questions concerning how multilingual and multimodal pedagogy and assessment could support student learning in CLIL, and their perceptions of the innovations, data were collected from multiple sources across both teaching cycles.

First, students' learning outcomes were captured with the written pre- and post-tests in each teaching cycle. Considering the limited CLIL experience of the students in this study, the questions in the tests did not impose very high productive linguistic demands. Instead, students were expected to demonstrate their understanding of the topic through multiple-choice questions, True/False statements, diagram labelling, order sequencing, matching and short data-based questions (i.e., they read short scenarios and answered questions with the content knowledge). The questions of the post-test were basically the same as those in the pre-test. However, they were reshuffled, or some questions appeared in a different section (e.g., some questions in the multiple-choice section in the pre-test were placed in the T/F statement section in the post-test). Such arrangement aimed to reduce the test-retest effect. The maximum score of the tests in the two cycles was 20 and 24 respectively.

At the beginning of each teaching cycle, all students completed the same written pre-test, which followed the conventional design (i.e., largely monolingual and monomodal). At the end of the teaching cycle (i.e., in Lesson 4), students completed a written post-test, which had two versions (conventional version vs multilingual and multimodal version) as described in the previous section. These two versions were randomly distributed to students. The tests assessed students' learning outcomes of the two cycles, while the two versions of tests revealed any impact of the test design. Repeated measures two-way ANOVA was performed on the pre- and post-test scores in each cycle, gauging the main effects of "time" (i.e., before and after the teaching) and "version of test" (i.e., the conventional vs multimodal version), and any interaction effects between the two variables.

Second, students' group posters were collected to demonstrate student learning at the end of each cycle through a multimodal assessment task. The combination of text and visuals in the posters, together with the transcribed oral presentations during the lesson, could provide additional evidence of student learning at the end of each unit, thereby complementing the written test results.

Third, all eight 40-min lessons (four in each cycle) were video-recorded and transcribed for subsequent

analysis. The first author was present during the last two lessons of each cycle and took field notes. These lesson recordings could provide a detailed picture of teacher-student interaction and whether teachers' pedagogical strategies worked. To address the specific research questions of this paper, episodes demonstrating the potential impact of teachers' multilingual and multimodal strategies were selected and analysed.

Fourth, a semi-structured group interview was conducted with six students per cycle (12 in total) to elicit their perceptions of the CLIL elective (e.g., whether they liked the elective, any challenges they encountered), the teacher's strategies in helping them learn, and their experience with multimodal assessment tasks (including the two versions of the post-test and the poster task). The students were selected purposively by the teacher, including students with different levels of L2 proficiency and attempted different versions of the post-test. The two interviews lasted for 21 min and 18 min respectively. Some students responded in Chinese (L1) while others in English, depending on their own preference. These interviews were audio-recorded and transcribed verbatim for subsequent thematic analysis. The research team read through the interview transcripts and coded quotes relevant to the research questions. Some codes were then grouped into larger themes to illustrate major patterns of students' perceptions of CLIL, multilingual and multimodal pedagogy and assessment.

Taken together, these multiple data sources enabled triangulation of findings across quantitative (test scores), qualitative (transcripts and interviews), and product-based (artefacts) perspectives (Cohen et al., 2017). Themes emerging from these different sources of data were constantly compared. When similar findings were observed, qualitative data such as lesson excerpts and interview quotes could elaborate on the quantitative data; when contrasting patterns were found, qualitative data could explain the potential reasons behind or provide other perspectives. For example, lesson episodes where the teacher used L1 and/or multimodal resources were carefully analysed in terms of teacher-student interaction (e.g., teacher's questions and scaffolding, students' responses), so as to infer their potential impact on student learning. The observations of these episodes were then compared with students' responses during the interviews, regarding their appreciation of teacher's use of multimodal resources. In addition, students' scores on pre- and post-tests were analysed with inferential statistical tests. The results were then elaborated and explained by students' perceptions of the two versions of tests during the interviews.

4. Findings

The two research questions are closely connected, in the sense that question 1 asks about what happened and

what worked, while question 2 seeks to understand students' perceptions and interpretation of their experience. Hence, the findings presented in this section will combine different sources of data to address the two questions together.

4.1. Multilingual and Multimodal Scaffolding for Content and Language Learning

As described in the Methodology section, although the students on the CLIL elective were self-selected and their English proficiency was above average in their school, they had limited experience of learning content knowledge through English. Hence, several students in the interviews admitted that the CLIL approach posed some challenges to them. One student highlighted the "new words" he encountered in the units, and he had to "repeatedly revise them, practise how to say them, remember how to write them" (Interview, Cycle 2). Another student added, "Sometimes it's not only the scientific terminology, but perhaps it's the first time we encountered the science knowledge" (Interview, Cycle 2). These two students summarised the challenges in both content and language aspects.

Despite these challenges, the majority of student interviewees shared that they found the CLIL units "interesting". One student reflected:

When I first experienced it [the CLIL elective], it's very innovative. Its coverage is broader than normal English lessons, and it involves more English than normal Science lessons. We can also learn more knowledge.

(Interview, Cycle 1)

This quote pinpoints the nature of CLIL and its position in this school—exposing students to more authentic English use through learning knowledge in other disciplines. A few student interviewees also mentioned a stronger sense of achievement and the benefits of CLIL for their language learning, particularly their communicative ability. This is probably due to the poster design and presentation task in the CLIL units. This point will be explicated later.

Miss W acknowledged the potential challenges that students may encounter on the CLIL elective. Hence, when planning the units and designing the teaching materials, she and the research team intentionally incorporated multilingual and multimodal resources as scaffolds. As observed from the lessons, these scaffolding resources were effective in facilitating content and language learning.

First, Miss W employed a wide variety of multimodal resources to help students better understand the scientific concepts involved. The choice of these multimodal resources depended on the concepts to be explained. For example, the two teaching cycles expected

students to learn the location of different organs in the circulatory and digestive system respectively. The teacher then used pictures to show these organs, their structure (e.g., the heart), and their location in the body. These can reinforce students' learning of these organs and their spatial awareness. Models (e.g., a big heart with four chambers) and other realia (e.g., a pump) were brought to the classroom to let students better understand the concepts. To enable students to understand the biological processes involved (e.g., how blood flows in the body; how food goes through the digestive system), Miss W and the research team carefully selected short and comprehensible video clips or animations. In Cycle 1, the teacher also used a line graph to show the changes in heart rate before and after doing exercise (so as to explain the function of our heart). She engaged students to plot a graph about their own heart rates.

The purposeful selection of multimodal resources, when combined with language (starting from subject-specific words, later to phrases and sentences), shown on the PPT slides and worksheets, was effective in helping students unpack and repack content knowledge. In particular, the two teaching cycles expected students to be able to describe some biological processes, which typically involved sequential explanation. This genre can be difficult, especially for young learners with limited CLIL experience. Here, the teacher made good use of video clips and animations to help students better visualise the processes. This was then followed by pictures, repeated oral and written description. Take the description of blood flow in Cycle 1 as an example. Miss W first showed a 2-min animation about how the blood flows through the four heart chambers and to the other parts of the body. She then provided four pictures for students to arrange in small groups. After students shared and agreed upon the correct sequence, the teacher showed and explained the sentence describing each picture (i.e., each step). Next, the teacher asked one student to go to the screen and draw some arrows showing the blood flow in each picture/step, while the rest of the class described the steps in sentences. The teacher further invited one student to describe the whole process in front of the whole class, and each student had a chance to practise describing the processes in their own groups. The sequential explanation was further consolidated by a written worksheet, where the students needed to arrange the pictures in order and then fill in the blanks. These instructional steps closely aligned with the MEC (Lin, 2015), with unpacking of content knowledge with multimodal resources, engaging students in reading and language analysis, and finally enabling students to express their content knowledge through L2 academic language.

Comparatively speaking, there were not many instances of teachers' or students' use of L1 in the two teaching cycles. As described in the Methodology section, most instances of L1 were planned to help students map

the subject-specific words to their existing L1 knowledge. Very often, Miss W introduced the target subject-specific terms and elicited the L1 equivalents from students (as Excerpt 1 above shows), so as to make sure students knew the terms. Beyond the word level, Miss W sometimes encouraged students to explore new concepts with L1. Excerpt 2 is an example of L1 use for this purpose. It took place in Lesson 1 of Cycle 1, when the teacher wanted students to compare arteries and veins. Figure 5 is the screenshot of the lesson video that accompanied Excerpt 2. Excerpt 2: Differences between arteries and veins [Lesson 1, Cycle 1; 25:40–28:42]

Miss W: *But what are some differences between arteries and veins? Let's discuss in groups first. Start.*

[Ss discussed in groups for about 2 min; some discussed in L1 and some in L2].

Miss W: *Who wants to share something? [pointing to a student to invite him to speak].*

S1: *動脈的是流動得快 (Arteries flow faster).*

Miss W: *Yes, a good point. [inviting another student to answer with gesture].*

S2: *Arteries is, arteries is, rich in, rich in oxygen, and, and.*

Miss W: *Veins?*

S2: *Veins are poor, are poor in oxygen.*

Miss W: *Wow, very good. Anything else? [waited for 4 s] Can we see veins or arteries?*

Ss: *Yes, yes.*

Miss W: *Please show [grabbing a student's arm]. So we can see some veins, but arteries.*

Ss: *We can't.*

Miss W: *Yes, it's hard to find them, but we can feel it's beating [touching her own wrist]. Okay? And there is one big difference between them. [pointing to the PPT slide]. Arteries carry blood away from the heart. Veins carry blood back to the heart. Let's see this picture [pointing to a simple heart picture on the blackboard]. The arteries, away or back?*

Ss: *Away.*

Miss W: *So this is artery [drawing an arrow away from the heart on the blackboard and writing the word "artery"]. Vein?*

Ss: *Carry back to.*

Miss W: *Back to. Very good. Vein [pause] back [drawing an arrow back to the heart on the blackboard and write the words "vein" and*

“back”]. Artery [pause] away [writing the word “away” near “artery”]. Right, let’s read these two sentences together. Artery, one, two, go.

*Ss: Arteries carry blood away from the heart.
Veins carry blood back to the heart.*

In Excerpt 2, Miss W asked students to brainstorm some differences between arteries and veins in their small groups first. This was one of the key concepts of Cycle 1. When students discussed in groups, some used L1 and others kept using English. Similarly, when the teacher invited volunteers to share their answers, the first student used L1, while the second used L2. Starting from there, the whole discussion or teacher-student interaction switched back to the target L2, supported by gestures and visuals (e.g., looking for veins or arteries in their body, annotating the picture on the blackboard, reading aloud complete sentences on the PPT slide). We could tell that this multilingual and multimodal teaching episode was effective with what happened in Lesson 3, which took place one week after Lesson 2. At the beginning of the lesson, Miss W recapped what students had learned in the previous week. When she asked for the differences between veins and arteries, one student volunteered and said, *“Veins carry blood to our heart, and arteries carry blood from the heart.”* When the teacher further prompted, *“from the heart?”*, the student elaborated, *“from the heart to all parts of the body”*.

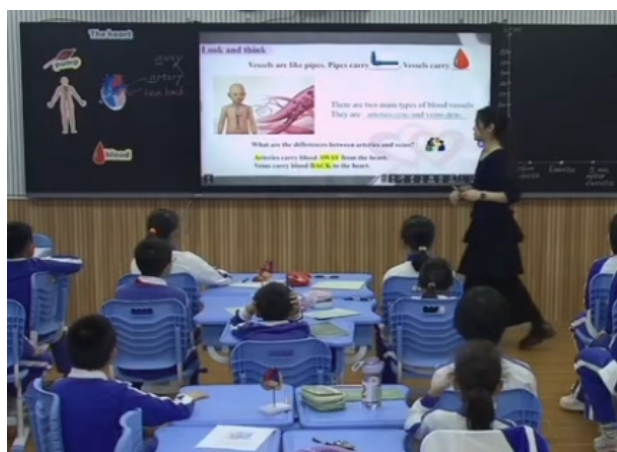


Figure 5. Screenshot of the lesson video showing what happened in Excerpt 2.

The effectiveness of the teacher’s scaffolding strategies was also recognised by the students. The majority of student interviewees could identify some strategies employed by Miss W, particularly the use of multimodal resources across the two cycles. They highlighted *“video clips”*, *“pictures”* and *“models”*, which helped them understand the concepts more easily. One student specified:

I think if it was all in English at the beginning, it would be difficult for me to understand. However,

if some pictures were added, I could roughly understand, for example, the heart, what its structure is. If there was only English description, it would not be that clear.

(Interview, Cycle 1)

Similarly, another student highlighted that the use of pictures to show the location of different organs (e.g., *“esophagus”*) in Cycle 2 made the concepts more concrete. These responses triangulated what was observed in the lessons.

4.2. Multilingual and Multimodal Assessment for Expressing Content and Language Learning

One major aim of this study was to experiment with multilingual and multimodal assessment design. Students’ performance on the written tests will be presented first, followed by the poster assessment task.

4.2.1. Students’ Performance in the Pre-Test and in the Two Versions of Post-Tests

Students completed the written pre-test and post-test for each teaching cycle. The latter had two different versions (conventional vs multilingual and multimodal), which were randomly distributed to students. Table 1 summarises the descriptive statistics of the test scores across the two cycles, and Tables 2 and 3 present the results of the two-way repeated measures ANOVA. Before conducting inferential statistical tests, the assumptions of normal distribution of data and homogeneity of covariance matrices were checked. Skewness and kurtosis values fell within acceptable thresholds for normal distribution (± 2), except for the post-test scores in Cycle 1 (skewness = -1.71 ; kurtosis = 3.16). Box’s test of equality of covariance matrices was not significant (Cycle 1: Box’s $M = 1.19$, $F(3, 397445.48) = 0.37$, $p = 0.78$; Cycle 2: Box’s $M = 3.97$, $F(3, 140813.21) = 1.21$, $p = 0.31$), indicating that the assumption of homogeneity of covariance matrices was met.

The tables show that all students’ scores improved significantly ($p < 0.001$) in the post-test in both teaching cycles, and the effect size was high (0.6 or above). This echoes the findings presented in the previous section—the teacher’s teaching was effective in supporting students to master the content knowledge through L2. When examining the answer scripts, students’ performance improved especially in the section of data-based questions. In this section, students were asked to read one or two scenarios related to the topic and answered some questions. This section therefore imposed higher cognitive and linguistic demands on students, as they needed to apply their content knowledge to new situations and express their knowledge in sentences. Hence, in the pre-test, many students got very low marks in this section, as they mainly repeated what was described in the scenario, or they simply did not attempt to answer. When it came to

the post-test, most students could provide at least some accurate responses. For example, in Cycle 2, one student got 8 marks (out of 24) in the pre-test. She did not match the correct organs with their corresponding functions, and she wrote “I don’t know” to one data-based question (with two sub-questions). Then, in the post-test, this student got 21 out of 24 marks. Such a noticeable improvement could be attributed to her correct answers when describing the functions of the organs in the digestive system. In the data-based questions, she could point out the problems (e.g., “constipation”) and explain the reasons behind (e.g., “eat too much chips”, “not eat food rich in dietary fibre”). These serve as an example to show that the two teaching cycles could achieve their content and language objectives.

Table 1. Descriptive statistics of students’ performance on the tests across two cycles.

Version of Post-Test	Cycle 1 (N = 29)		Cycle 2 (N = 27)	
	Pre-Test M (SD)	Post-Test M (SD)	Pre-Test M (SD)	Post-Test M (SD)
Conventional	9.94 (2.29)	19.16 (2.07)	17.92 (3.43)	20.85 (3.34)
Multimodal	9.07 (2.62)	18.21 (2.68)	15.93 (4.32)	20.93 (2.46)

Table 2. Results of the two-way repeated measures ANOVA in Cycle 1.

Effect	DF1	DF2	F	p	Partial Eta-Squared
Time	1	28	272.78	<0.001 **	0.68
Version	1	28	1.75	0.197	0.02
Version × Time	1	28	0.00	0.946	0.00

** $p < 0.01$.

Table 3. Results of the two-way repeated measures ANOVA in Cycle 2.

Effect	DF1	DF2	F	p	Partial Eta-Squared
Time	1	25	37.12	<0.001 **	0.6
Version	1	25	0.68	0.416	0.03
Version × Time	1	25	2.5	0.127	0.09

** $p < 0.01$.

However, there were no statistically significant differences between students’ attempting the two versions of post-tests, as shown by the main effect of test version and the interaction effect between test version and time (i.e., pre- vs post-test). In other words, the multilingual and multimodal design of the post-test was not more effective in helping students express their content knowledge. However, in Cycle 2, the increase in scores between pre- and post-tests among students attempting the multilingual and multimodal post-test was more obvious than those taking the conventional test. The lack of statistical significance may be due to the small sample size (with fewer than 15 in each group) and the topics involved, which will be explicated in the discussion section.

Despite the lack of statistically significant results, some students did observe the differences between the two versions of the post-test and highlight some potential benefits of the multilingual and multimodal version of test. A few students mentioned that the visuals incorporated into the test could help them understand the content of the test more easily. As they had seen some pictures in the lessons (e.g., the four stages of blood flow in the heart), they could even “not read the text” and answer the questions immediately. In the data-based questions, one student mentioned that the pictures helped them understand the scenario better. For example, in the written test of Cycle 2, one data-based question was about a boy’s unhealthy eating habits. Instead of reading the text, the student “could easily see and understand the deep-fried food shown in the picture, and could quickly write the answer” (Interview, Cycle 2). A couple of students also highlighted the emotional or affective aspect of the multimodal test, as the test would be “more interesting” if there were pictures, and they would “have headache if they could only see text” in the test paper.

In terms of the use of L1 in the tests, not many students mentioned the incorporation of L1 in the tests. When the interviewer prompted whether L1 could help, one student explained, “Perhaps it could help before we learned the topic, but we already knew most of the words after learning”, and hence the use of L1 may not be particularly useful in the post-test. In a similar vein, when examining the test scripts, only a few students used L1 to answer the questions, even though the teacher allowed them to do so in the pre-test and in the multilingual and multimodal version of the post-tests. Interestingly, L1 could help on some occasions, but not on others. The latter reflects that if students did not master the concepts per se, it did not matter which language they used. However, for some students who did understand the concepts, but lacked the appropriate language, L1 could better reflect their actual learning. For instance, the student who improved from 8 to 21 marks in Cycle 2 reported above employed L1 in her post-test and managed to accurately describe all the functions of the organs in the digestive system. Figure 6 shows her answers in that particular section.

It can be seen that the student wrote L1 in questions (a) and (d). The correct answer to question (d) is gastric juice, which was quite difficult for young learners. Some other students could not spell the words correctly and did not get the mark. Hence, this student’s use of L1 could help her get the point. For question (b), the student actually wanted to write the Chinese of “nutrients” (i.e., 营养), but she seemed to forget the second character and wrote the phonetic spelling instead (“yǎng”). This is a very interesting example of translanguaging, in the sense that the student mobilised all her meaning-making resources to convey the intended meaning.

Part D: Fill in the blanks

11. Complete the sentences below describing the functions of different organs in the digestive system. Please note that they do not follow the order. *If you do not remember the English words, you can use Chinese.*

- a) The esophagus squeezes (挤压) and transport 输送 food down to the stomach.
- b) The small intestine absorbs (吸收) nutrients 营养 into the blood.
- c) The mouth chews food to make it into smaller pieces and mixes the food with saliva (唾液).
- d) The stomach further breaks down food with gastric juice 胃液 juice.
- e) The large intestine absorbs water and makes waste.

Figure 6. A student's answers to the cloze questions in Cycle 2 post-test (English translation inserted).

4.2.2. Students' Performance on the Poster Assessment Task

In addition to the written tests, students were asked to create a poster at the end of each teaching cycle, so as to demonstrate their learning. While the teacher would give specific instructions and a presentation template, students were allowed some flexibility in deciding on the content and design of their posters. From the posters collected from both teaching cycles, we could see that this productive multimodal assessment task not only allowed students to demonstrate their understanding of the content, but also catered for learner diversity and provided a venue for students to show their creativity.

In Cycle 1, the poster was expected to include:

- A labelled diagram of the heart's structure
- A line graph comparing resting and post-exercise heart rates (collected in Lesson 2)
- A simple explanation of why exercise increases heart rate (can be written or oral)
- A description of the blood flow processes
- Tips for protecting the heart

Figures 7 and 8 show the posters created by two groups of students. Both posters included the required information listed above. The students also used the colour red a lot, probably because it is usually associated with the heart or blood. They also included symbols representing the heart and heart rate in the posters. All these show that students in general know how to choose visuals related to the content. Comparing the two posters, it seems that the one in Figure 8 includes more written elaboration of the blood flow processes to go with the given pictures. This may reflect different levels of students' English proficiency.

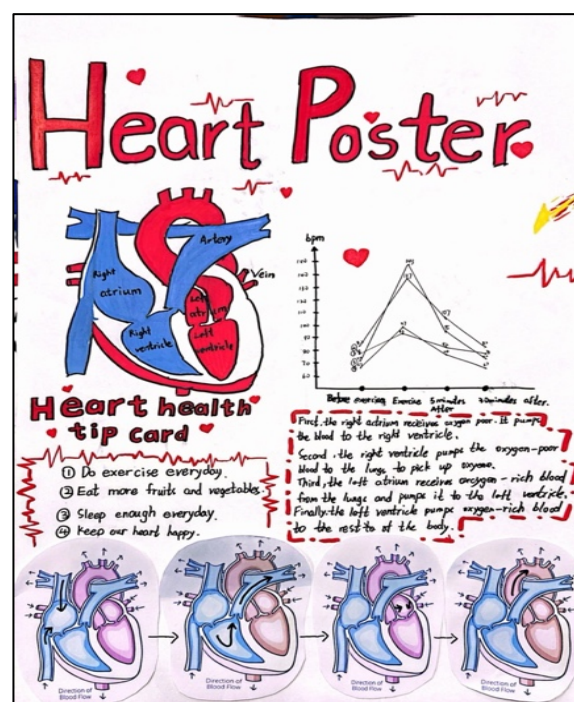


Figure 7. "The Heart" poster from one group.

All the groups in Cycle 1 simply described the poster in their oral presentation (i.e., describing the structure of the heart, the different heart rates, the blood flow processes, etc.). When it came to Cycle 2, students became more creative in the poster design and oral presentation. For example, Figure 9 shows the poster designed by one group, which closely aligned with the teacher's instructions. The students described the key stages of digestion during their oral presentation. Figure 10 shows the poster designed by another group of students. The text in the poster was similar to the one in Figure 9. However,

when this group did the presentation, they designed a role play, pretending to be a strawberry going through the different organs in the digestive system. Excerpt 3 shows part of their oral presentation:

Excerpt 3. Part of the oral presentation of one group (with the poster in Figure 10) [Lesson 4, Cycle 2; 30:25–31:25]

S1 (narrator): *Let's follow a strawberry on its journey.*

S2: *Hello, I am a strawberry. Today, I am eaten by a human. First, I go into a black hole. Look, there are some white stones. Importantly, what's happening here? (...) [turning to S3] Who are you?*

S3: *Hello, I am Miss (...). You can call me (...). I can help you to take a trip to, in a human body. Let's go. [S2 and S3 jumped]*

S2: *Really? But where are we?*

S3: *Now we are in the mouth. Be careful! Those, those white stones will chew you. They are called teeth.*

S2: *Ah, it hurts! [covering her arm] Let's go. [holding S3's hand and jumping] Oh, I am going down! Why am I falling down? Where are we?*

S3: *Oh, we are in the esophagus. It will push you down to the stomach ...*

holding any script, and they just performed naturally. Some other groups also performed drama, pretending to be doctors and patients, or news reporters giving health advice. The teacher and research team were very impressed by students' presentations.

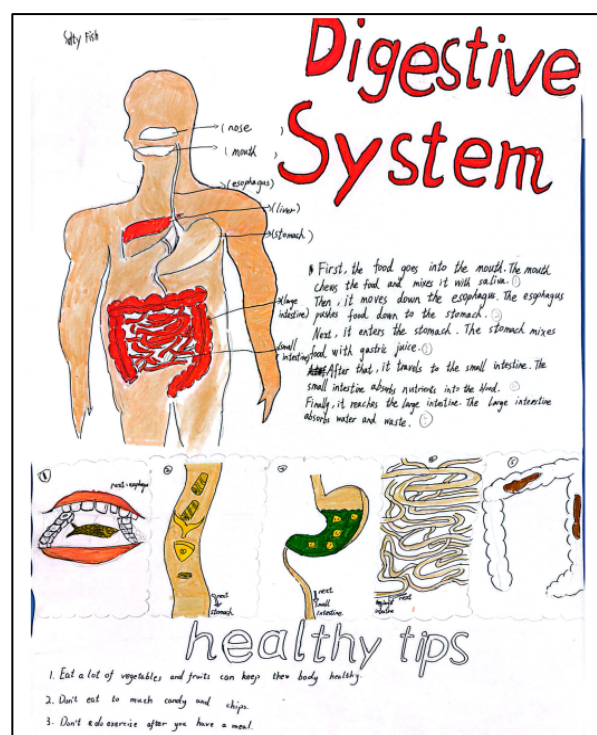


Figure 9. "The Digestive system" poster from one group.

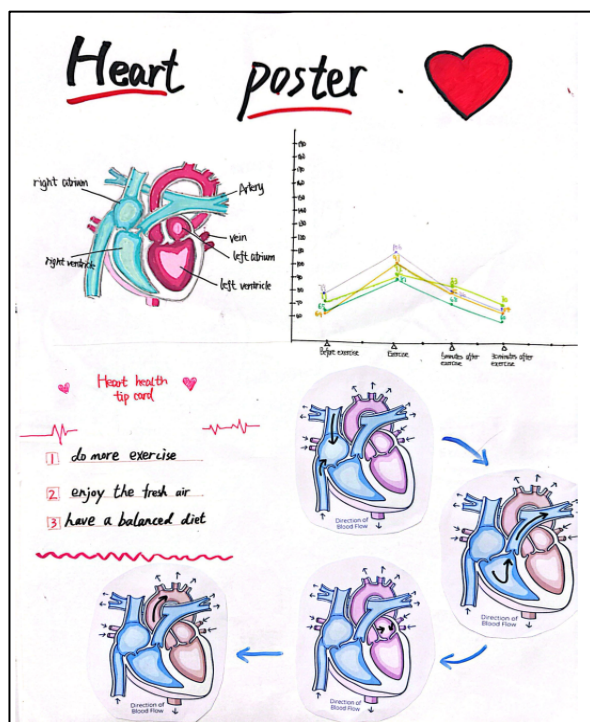


Figure 8. "The Heart" poster from another group.

The journey of the strawberry through the digestive system went on, with S2 and S3 role-playing and describing their "experience". The students were not

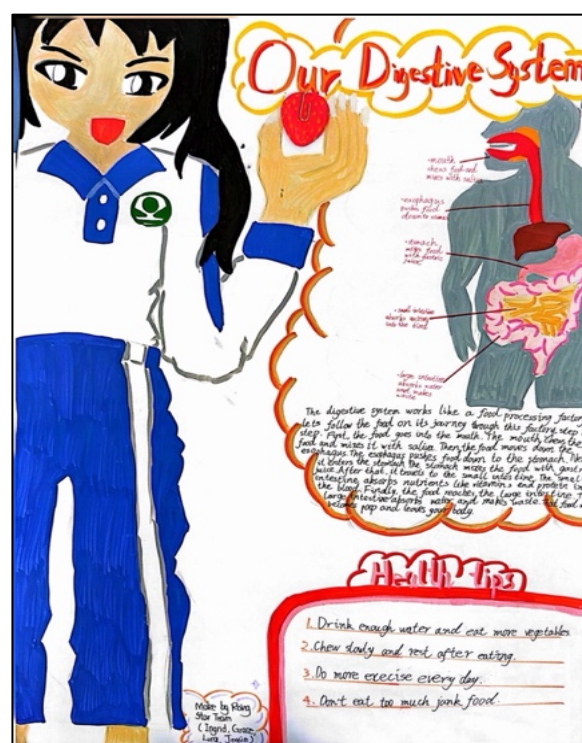


Figure 10. "The Digestive system" poster from another group.

During the interviews, the students also appreciated the poster design and presentation task. First, some students reflected that this assessment task could reinforce their learning, as they needed to “*present orally*” or “*demonstrate on the stage*”. More students felt that the poster design task was “*more interesting*” than written tests, as they could draw and share their ideas with others through speaking. Echoing the researchers’ observation, a few students mentioned that the task could “*stimulate their creativity*”. One student elaborated in this way:

I think this poster can stimulate our creativity. All the groups were not restricted by the traditional homework. We could be creative. And we were not restricted by this poster. We could also perform a short story or something.

(Interview, Cycle 2)

As the poster design task was allocated as group work, some students had learned how to work with others as a team. Every member in the team could write and draw, and they would discuss which colours to use and complete the task together. If one student did not know something, he/she could ask others and learn from them. This was unlike doing a test, which is individual work and whether one could finish it totally depends on his/her own knowledge.

Hence, while students’ performance did not vary much in the two versions of written tests, they were found to actively engage in the poster design and oral presentation task, which was multimodal in nature. They were eager to demonstrate their learning outcomes in creative ways by combining various semiotic resources, including written and oral language, colours, symbols, pictures, graphs, gestures and acting.

5. Discussion

This study set out to explore whether multilingual and multimodal pedagogy could facilitate content and language learning among young EFL learners. Further, it sought to experiment with multilingual and multimodal assessment to see if it may enable students to better demonstrate their content learning through their L2.

With different sources of data, our findings seem to reconfirm the benefits of multilingual and multimodal pedagogy in primary CLIL classrooms. Through the DBR approach, the research team and teacher co-designed two teaching units, by systematically incorporating students’ L1 and a wide variety of multimodal resources to facilitate both content and language teaching. In particular, L1 equivalents, pictures (together with annotation or labelling) and models are effective in introducing subject-specific key words; short video clips or animations, hands-on activities (e.g., rearranging pictures to show the correct sequence, measuring and recording heart rates) and small group discussion stimulate students to explore

and unpack difficult content knowledge. It is important to consolidate students’ understanding with guided oral and written repacking, or what Lin (2015) called “*entextualisation*”, so that students can express their content knowledge with appropriate academic language. These findings corroborate previous studies (e.g., Nikula et al., 2024) and provide more empirical evidence in primary CLIL classrooms, where young EFL learners are acquiring both content and L2 simultaneously (e.g., Liu & Lin, 2025). There has been some discussion about when students should start intensive exposure to L2 or experience CLIL or other bilingual education programmes. Johnson and Swain (1994) hypothesised that young learners may master content and L2 more easily, as the concepts involved may be less abstract and the associated language may be less complicated. Previous meta-analyses (Kaiypova et al., 2025; Lee et al., 2023, 2025) on students’ learning outcomes in CLIL also reveal different impacts on students’ L2 development, with primary students benefitting more in speaking. By documenting what happened in the CLIL lessons, this study can provide some explanation for the different impacts—in primary CLIL classrooms, the teacher tends to engage students in exploring content knowledge through discussion and other activities; the teacher may also put more emphasis on students’ sharing, reading aloud and oral presentation, instead of written practice, since young learners are still developing their literacy skills. These may have contributed more to students’ speaking skills. Of course, our study only involved one teacher, but her contributions to the design of the two CLIL units reflected her experience and understanding of primary EFL learners, which may be similar to those in other EFL contexts.

While the potential of multilingual and multimodal teaching has been documented in previous studies, multilingual and multimodal assessment has been under-explored (Schissel et al., 2021). This indeed represents a misalignment between pedagogy and assessment (Lin, 2024). Hence, another aim of this study was to explore whether multilingual and multimodal assessment tasks may enable CLIL students to better demonstrate their learning outcomes. Together with the teacher, the research team designed two versions of the written post-test, one following the more conventional design and the other incorporating L1 support (e.g., translation of key words) and visuals. The rationale behind the latter was to help students understand the test questions more easily, so that they could retrieve the relevant concepts and address the questions. Students were also allowed to use L1 when they responded. These design principles corresponded to those advocated by other researchers (e.g., Gorter & Cenoz, 2017; Shohamy, 2022). The statistical analysis of the test scores did not show any significant differences in students’ performance between the two versions of test. This may be explained by two

reasons. First, as aforementioned, the sample size was relatively small, with fewer than 15 students taking each version of the test. Hence, even though some differences were observed in the increase in the mean scores of the two groups, especially in Cycle 2, they did not reach statistical significance. Second, the concepts involved in the two teaching cycles, although new to students, could still be regarded as concrete and contextualised. They were about the human body (the heart and the digestive system), something that students could feel and experience (e.g., the heart rate, constipation, indigestion). After they had mastered the concepts during teaching, the incorporation of multimodal resources in the written tests may not result in significant differences between the two groups. Some students did talk about the assistance of pictures in helping them understand the test questions. This may have helped them attempt the questions more quickly, but the accuracy of their answers may remain similar with the text-only version. These potential differences in the affordances of multimodal resources during teaching and testing, or in affecting the testing processes and outcomes will be worth further exploration.

Another attempt with multilingual and multimodal assessment was the poster design and oral presentation task at the end of each teaching cycle. The written post-test is the conventional assessment format, whereas this poster task could be regarded as an alternative or more innovative assessment strategy. It requires students to combine different language skills (written and spoken) and semiotic resources (including colours, symbols, pictures, gestures and body language) to express their understanding of the topics. The posters and oral presentations observed were impressive, where students demonstrated their content and language learning outcomes, as well as their creativity. These benefits of multimodal assessment are in line with previous research (e.g., Jiang et al., 2024). Multimodal or alternative assessment strategies in CLIL classrooms can provide students with another venue to demonstrate their learning outcomes. When used together with more conventional written tests serving the purpose of assessment of learning (i.e., to evaluate students' learning at a particular point of time), these alternative assessment strategies can provide additional evidence or data of student learning, which can inform further instruction and learning.

We have been discussing multilingual and multimodal pedagogy and assessment. However, it is interesting to note that both the teacher and students in this study did not use L1 frequently in both teaching and assessment. As presented in the findings section, the teacher mainly provided L1 equivalents for some subject-specific words and some students used L1 to express their thoughts, especially during the brainstorming or exploration stage. These instances of L1 use were often short, immediately followed by similar discussion or

reinforcement in L2. Similarly, we did not find much L1 use in the written tests or oral presentations. One may expect that young EFL learners, with their developing L2 proficiency, may use more L1 to communicate. However, this did not happen in our study, with two possible explanations. First, the teaching context was a CLIL elective, where students opted to join to improve their English. With this goal in mind, students may wish to maximise their use of English in the CLIL lessons. Second, when the teacher and students made good use of other semiotic resources such as visuals, video clips, gestures and symbols, the need for L1 use may not be very strong. Of course, the use of L1 and other semiotic resources are not mutually exclusive, and numerous studies have shown the pedagogical benefits of mobilising all the resources in one's communicative repertoire (e.g., Liu & Lin, 2025). Yet, the findings of this study may provide some insights into the affordances of using multimodal resources, together with carefully planned L2 scaffolding, in facilitating students' mastery of content and language, especially in contexts where L2 development is one explicit and strong motivation for the CLIL approach.

6. Conclusions

This study represents one of the few attempts to explore the alignment between multilingual and multimodal pedagogy and assessment, and whether this is effective in facilitating content and language learning among young EFL learners. From the two iterative cycles of this design-based research, some principles for multilingual and multimodal pedagogy and assessment could be drawn. First, students' L1 and other semiotic resources can be systematically planned and deployed during different instructional stages to support their content and language learning. Aligning with the Multimodalities-Entextualisation-Cycle (Lin, 2015), L1, visuals, realia and hands-on experiments are useful in unpacking new academic concepts for students at the beginning of the teaching unit. After students have mastered the key concepts, systematic oral and written scaffolding for L2 academic language can be provided, so that students can express their content knowledge in the target L2 by the end of the unit. Second, when designing conventional written tests to evaluate students' content knowledge, teachers can incorporate some L1 and representational visuals to make the questions more comprehensible for students. Whenever possible, teachers can also allow students to use L1 to demonstrate their content knowledge when necessary. These would help reduce the potential language barriers students encounter and enable teachers to better diagnose students' learning progress and difficulties. In particular, when teachers observe that some students can express accurate content knowledge with L1, they can follow up by reinforcing the appropriate use of L2 academic

language. Third, alternative assessment strategies or tasks (e.g., the poster design and presentation task in this study) seem to be well received by students, who could demonstrate their content and learning outcomes through other means. Hence, when the teaching schedule allows, teachers could incorporate alternative assessment tasks, which not only provide more data on students' learning progress, but also foster other skills such as multimodal literacies, communicative and collaborative skills.

We acknowledge some potential obstacles of implementing alternative assessment strategies, especially involving the use of other languages and multimodal resources. These include unclear assessment criteria, inconsistency in marking different outputs, and the dominance of monolingual and monomodal examinations (Liu & Lo, 2025; Yi & Choi, 2015). However, considering the fact that multilingual and multimodal assessment may provide more accurate data about students' learning progress, it is worth exploring how to better integrate such assessment formats into existing assessment practices or mechanisms. For example, alternative assessment strategies could be encouraged as on-going and formative assessment, the results of which would account for a proportion of the final results, together with more conventional assessment format such as the year-end examination. This combination may help teachers better diagnose students' learning progress and difficulties in both content and language aspects, and still prepare students for the high-stakes examination. The ultimate goal for exploring different assessment practices in CLIL and other bilingual education programmes is to assess all students in a more valid and fairer way (Shohamy, 2022), and any changes or breakaway from the conventional assessment practices would require the concerted efforts of researchers, teachers and policy makers.

Author Contributions

Y.Y.L.: conceptualisation, methodology, data collection, data analysis, writing, funding; L.W.: methodology, data collection, writing; K.P.T.C.: literature review, methodology, data analysis. All authors have read and agreed to the published version of the manuscript.

Funding

This work was supported by General Research Funds, the Hong Kong Research Grants Council [grant number 17615323].

Institutional Review Board Statement

This study has been approved by the Human Research Ethics Committee of the University of Hong Kong (reference number: EA220529).

Informed Consent Statement

Informed consent was obtained from all participants (the teacher, students and their parents) involved in the study.

Data Availability Statement

The data of this study are not available for sharing due to privacy considerations.

Acknowledgments

The authors would like to thank the school and students who participated in the study.

Conflicts of Interest

The authors declare no conflict of interest.

Use of AI and AI-Assisted Technologies

No AI tools were utilised for this paper.

Appendix A. Transcription Conventions

- S1, S2, S3, etc. = single student
- Ss = more than one student
- (...) = inaudible utterances
- (*italics*) = translation of Chinese utterances
- [] = nonverbal actions or author's comments
- __ (at the end of questions) = short pauses indicating blank filling questions
- ... = some lines deleted

References

- Amondarain-Garrido, M., & Ruiz de Zarobe, Y. (2025). Looking into multimodal translanguaging in CLIL: A case study in the Basque Country. *Journal of Multilingual and Multicultural Development*, 46(5), 1440–1458. <https://doi.org/10.1080/01434632.2024.2429419>.
- Amundrud, A., & Rasmussen, I. (2026). Interpreting tensions in teacher–researcher collaboration. *Teachers and Teaching*, 1–21. <https://doi.org/10.1080/13540602.2026.2713119>.
- Archila, P. A., Molina, J., & Truscott de Mejía, A. M. (2021). Fostering bilingual scientific writing through a systematic and purposeful code-switching pedagogical strategy. *International Journal of Bilingual Education and Bilingualism*, 24(6), 785–803. <https://doi.org/10.1080/13670050.2018.1516189>.
- Banegas, D. L., & Zappa-Hollman, S. (2023). Introduction. In D. L. Banegas & S. Zappa-Hollman (Eds.), *The Routledge handbook of Content and Language Integrated Learning* (pp. 1–7). Routledge.
- Cenoz, J., Genesee, F., & Gorter, D. (2014). Critical analysis of CLIL: Taking stock and looking forward. *Applied Linguistics*, 35(3), 243–262. <https://doi.org/10.1093/applin/amt011>.

- Cenoz, J., & Gorter, D. (2020). Pedagogical translanguaging: An introduction. *System*, 92, 102269. <https://doi.org/10.1016/j.system.2020.102269>.
- Cohen, L., Manion, L., & Morrison, K. (2017). *Research methods in education* (8th ed.). Routledge.
- Dalton-Puffer, C., Hüttner, J., & Llinares, A. (2022). CLIL in the 21st century: Retrospective and prospective challenges and opportunities. *Journal of Immersion and Content-Based Language Education*, 10(2), 182–206. <https://doi.org/10.1075/jicb.21021.dal>.
- Fazzi, F., & Menegale, M. (2023). Doing CLIL with primary learners: From principles to practice. In D. L. Banegas & S. Zappa-Hollman (Eds.), *The Routledge handbook of Content and Language Integrated Learning* (pp. 225–237). Routledge. <https://doi.org/10.4324/9781003173151-19>.
- Gorter, D., & Cenoz, J. (2017). Language education policy and multilingual assessment. *Language and Education*, 31(3), 231–248. <https://doi.org/10.1080/09500782.2016.1261892>.
- Grapin, S. E., & Llosa, L. (2022). Multimodal tasks to assess English learners and their peers in science. *Educational Assessment*, 27(1), 46–70. <https://doi.org/10.1080/10627197.2022.2028139>.
- He, P. (2025). *Thematic-pattern-based “Concept + Language Mapping” (CLM): Content and language integration in a transdisciplinary perspective*. Routledge. <https://doi.org/10.4324/9781003391548>.
- He, P., & Lin, A. M. Y. (2024). Dynamic flows of translanguaging/trans-semiotizing in CLIL eco-social systems. *Language Awareness*, 33(2), 384–406. <https://doi.org/10.1080/09658416.2023.2234824>.
- Hu, H., Mohd Said, N. E., & Hashim, H. (2023). Sustaining Content and Language Integrated Learning in China: A Systematic Review. *Sustainability*, 15(5), 3894. <https://doi.org/10.3390/su15053894>.
- Jiang, L., Li, Z., & Leung, J. S. C. (2024). Digital multimodal composing as translanguaging assessment in CLIL classrooms. *Learning and Instruction*, 92, 101900. <https://doi.org/10.1016/j.learninstruc.2024.101900>.
- Johnson, R. K., & Swain, M. (1994). From core to content: Bridging the L2 proficiency gap in late immersion. *Language and Education*, 8(4), 211–229. <https://doi.org/10.1080/09500789409541392>.
- Kaiypova, F., Lee, H., Lo, Y. Y., & Lee, J. H. (2025). Effects of Content and Language Integrated Learning on secondary-level students’ content learning: A meta-analysis. *System*, 129, 103580. <https://doi.org/10.1016/j.system.2024.103580>.
- Lee, J. H., Lee, H., & Lo, Y. Y. (2023). Effects of EMI-CLIL on Secondary-Level Students’ English Learning: A Meta-Analysis. *Studies in Second Language Learning and Teaching*, 13(2), 317–345. <https://doi.org/10.14746/sslts.38277>.
- Lee, J. H., Lee, H., & Lo, Y. Y. (2025). The effects of EMI-CLIL on language and content learning in primary education: A multi-level meta-analysis. *Educational Research Review*, 47, 100666. <https://doi.org/10.1016/j.edurev.2025.100666>.
- Li, W. (2018). Translanguaging as a practical theory of language. *Applied Linguistics*, 39(1), 9–30. <https://doi.org/10.1093/applin/amx039>.
- Lin, A. M. Y. (2015). Conceptualising the potential role of L1 in CLIL. *Language, Culture and Curriculum*, 28(1), 74–89. <https://doi.org/10.1080/07908318.2014.1000926>.
- Lin, A. M. Y. (2024). Can the monkey king break through the ‘Jin-Gang-Quan’ (金剛圈)? Overcoming the multiple contradictions in EMI education. *Language and Education*, 38(1), 139–147. <https://doi.org/10.1080/09500782.2023.2284802>.
- Lindner, M. A., Eitel, A., Strobel, B., & Köller, O. (2017). Identifying processes underlying the multimedia effect in testing: An eye-movement analysis. *Learning and Instruction*, 47, 91–102. <https://doi.org/10.1016/j.learninstruc.2016.10.007>.
- Liu, J. E., & Lo, Y. Y. (2025). Multimodality in CLIL assessment: Implications for teacher assessment literacy. *Journal of Multilingual and Multicultural Development*, 46(5), 1459–1477. <https://doi.org/10.1080/01434632.2024.2357140>.
- Liu, J. E., Yang, M., Lo, Y. Y., & Lin, A. M. Y. (2026). Investigating the role of trans-semiotizing in developing a literature review: Meaning entanglement across multiple timescales. *Applied Linguistics Review*, 17(3), 1129–1159. <https://doi.org/10.1515/applirev-2024-0333>.
- Liu, Y., & Lin, A. M. Y. (2025). Spontaneous teacher collaboration in early primary CLIL classrooms with pupils of special educational needs (SEN): A dynamic translanguaging and trans-semiotizing perspective. *Journal of Multilingual and Multicultural Development*, 46(5), 1416–1439. <https://doi.org/10.1080/01434632.2024.2441408>.
- Lo, Y. Y. (2026). *Assessment in Content and Language Integrated Learning: Complexities, principles and practices*. Routledge.
- Lo, Y. Y., Fung, D., & Qiu, X. (2023). Assessing content knowledge through L2: Mediating role of language of testing on students’ performance. *Journal of Multilingual and Multicultural Development*, 44(10), 1013–1028. <https://doi.org/10.1080/01434632.2020.1854274>.
- Lo, Y. Y., & Lin, A. M. Y. (2019). Teaching, learning and scaffolding in CLIL science classrooms. *Journal of Immersion and Content-Based Language Education*, 7(2), 151–165. <https://doi.org/10.1075/jicb.00006.lo>.
- Lyster, R., & de Zarobe, Y. R. (2018). Introduction: Instructional practices and teacher development in CLIL and immersion school settings. *International Journal of Bilingual Education and Bilingualism*, 21(3), 273–274. <https://doi.org/10.1080/13670050.2017.1383353>.
- May, S. (Ed.). (2013). *The multilingual turn: Implications for SLA, TESOL, and bilingual education*. Routledge. <https://doi.org/10.4324/9780203113493>.
- Mearns, T., van Beuningen, C., Nederlof, N., & de Jong, N. (2025). Integration, demands and accommodations: An exploration of summative assessment practices in Dutch bilingual secondary education: An exploration of current practices. *Dutch Journal of Applied Linguistics*, 14. <https://doi.org/10.51751/dujal15756>.

- Moje, E. B. (2015). Doing and teaching disciplinary literacy with adolescent learners: A social and cultural enterprise. *Harvard Educational Review*, 85(2), 254–278. <https://doi.org/10.17763/0017-8055.85.2.254>.
- Nikula, T., Jakonen, T., & Kääntä, L. (2024). Multimodal practices of unpacking and repacking subject-specific knowledge in CLIL physics and chemistry lessons. *Learning and Instruction*, 92, 101932. <https://doi.org/10.1016/j.learninstruc.2024.101932>.
- Oattes, H., Oostdam, R., de Graaff, R., & Wilschut, A. (2018). The challenge of balancing content and language: Perceptions of Dutch bilingual education history teachers. *Teaching and Teacher Education*, 70, 165–174. <https://doi.org/10.1016/j.tate.2017.11.022>.
- Otto, A., & Estrada, J. L. (2019). Towards an understanding of CLIL assessment practices in a European context: Main assessment tools and the role of language in content subjects. *CLIL. Journal of Innovation and Research in Plurilingual and Pluricultural Education*, 2(1), 31–42. <https://doi.org/10.5565/rev/clil.11>.
- Philippakos, Z. A., Howell, E., & Pellegrino, A. (2021). *Design-based research in education: Theory and applications*. Guilford Press.
- Pun, J. K. H., Fu, X., & Cheung, K. K. C. (2024). Language challenges and coping strategies in English Medium Instruction (EMI) science classrooms: A critical review of literature. *Studies in Science Education*, 60(1), 121–152. <https://doi.org/10.1080/03057267.2023.2188704>.
- Ruiz de Zarobe, Y., & Querol-Julián, M. (2025). Multilingualism and multimodality in the CLIL/EMI classroom. *Journal of Multilingual and Multicultural Development*, 46(5), 1397–1415. <https://doi.org/10.1080/01434632.2025.2482734>.
- Schissel, J. L., De Korne, H., & López-Gopar, M. (2021). Grappling with translanguaging for teaching and assessment in culturally and linguistically diverse contexts: Teacher perspectives from Oaxaca, Mexico. *International Journal of Bilingual Education and Bilingualism*, 24(3), 340–356. <https://doi.org/10.1080/13670050.2018.1463965>.
- Shohamy, E. (2022). Critical language testing, multilingualism, and social justice. *TESOL Quarterly*, 56(4), 1445–1457. <https://doi.org/10.1002/tesq.3185>.
- Yi, Y., & Choi, J. (2015). Teachers' views of multimodal practices in K-12 classrooms: Voices from teachers in the United States. *TESOL Quarterly*, 49(4), 838–847. <https://doi.org/10.1002/tesq.219>.