Research Article

Cultivating Multimodal Literacy in a Japanese CLIL Context: An Action Research Study

ROOKS, Justin

Abstract

This study explores the design, implementation, and outcomes of a 14-week multimodal communication course conducted at a Japanese university, guided by an action research framework. Drawing on Burns's (2010) iterative cycle of planning, action, observation, and reflection, the course sought to deepen students' understanding of social semiotics, visual communication, and other theoretical perspectives relevant to multimodality. A mixed methods approach, including pre- and post-course image analyses, student surveys, teacher interviews, and assessment of final presentations, provided comprehensive data on learners' progress and perceptions. Findings indicate that students demonstrated marked improvements in their ability to deconstruct and interpret diverse texts, applying concepts related to multimodality and semiotics more effectively by the end of the term. Surveys further revealed high levels of satisfaction, with many students reporting heightened awareness of how social and cultural narratives are embedded in everyday media. While challenges arose in addressing abstract theory and varying English proficiency levels, the action research process facilitated timely pedagogical adjustments, such as simplified readings, interactive tasks, and iterative feedback loops, that supported learners' evolving needs. The results suggest that practitioner-oriented, reflective practice can successfully integrate multimodal analysis into a Content and Language Integrated Learning (CLIL) context, fostering advanced literacy skills and cross-cultural awareness. This study thus contributes to ongoing discussions on how action research may enhance the teaching of complex theoretical frameworks, pointing to broader applications in global higher education.

Keywords: *Multimodal communication, Social semiotics, Action research, Content and Language Integrated Learning (CLIL)*

Introduction

As communication technologies continue to evolve at an unprecedented pace, the ability to convey and interpret meaning across multiple modes—text, visuals, gestures, and audio—has

become increasingly important. Gunther Kress (2010), a pioneer in the field of multimodality, has aptly observed that communication is always and inevitably multimodal, and highlighted the necessity of developing multimodal literacy to engage with the complex interplay of modes in contemporary communication. In language education, this shift pushes learners beyond traditional reading and writing to engage with layered, media-rich texts that mirror real-world contexts. For Japanese students of English, incorporating multimodal elements can be especially beneficial: it not only develops language skills through authentic tasks but also introduces cultural and disciplinary content in a more engaging way.

Against this backdrop, Content and Language Integrated Learning (CLIL) has gained traction as a framework that allows educators to integrate subject matter and language learning simultaneously. In their pivotal book, *CLIL: Content and language integrated learning*, Coyle et al. (2010) define CLIL as "a dual-focused educational approach in which an additional language is used for the learning and teaching of both content and language" (p. 1). They develop this framework further by considering how this methodology enriches both subject knowledge and linguistic competence through their 4Cs Framework, which encompasses Content, Communication, Cognition, and Culture. In this approach, content refers to the disciplinary knowledge or subject matter forming the basis of instruction, and communication is the active use of language to articulate and engage with that content. Cognition involves developing students' critical thinking and problem-solving abilities so they can analyze and reflect upon new information, while culture highlights intercultural awareness and global citizenship. Together, these elements ensure balanced goals in CLIL, promote thorough scaffolding, and encourage active learning through collaborative, project-based methods. By integrating subject content with language development, CLIL provides an immersive environment where students engage with academic and professional materials while simultaneously honing their linguistic skills. This dual-focus approach encourages higher-order thinking, cultural awareness, and a more natural development of communication skills.

Within this CLIL framework, a course on multimodal communication aligns naturally with the 4Cs by integrating content mastery in tandem with English language development. By emphasizing *content*, the course focuses on core principles of multimodality and social semiotic theory. Through *communication*, students actively discuss and analyze these abstract ideas in English, thereby gaining familiarity with disciplinary terminology and functional language use. Higher-level *cognition* is cultivated as students critically examine how textual and visual components interact in various media, such as advertisements, news reports, or digital content, challenging them to synthesize ideas and explore deeper meanings. Finally, *culture* remains central, as the course examines how the production and interpretation of texts can differ across societies, preparing learners to engage in culturally diverse contexts. Weaving these elements into a cohesive curriculum highlights how CLIL fosters both subject mastery and language proficiency, demonstrating its relevance to multimodal literacy and broader educational contexts.

At Rikkyo University, and in Japan more broadly, the implementation of CLIL-based curricula in higher education is still evolving. Studies (Ikeda, 2013; Ikeda et al., 2022) have indicated that within the Japanese education system, CLIL shows significant promise for fostering both language proficiency and deeper engagement with subject content, but they also highlight the practical challenges of materials selection, language support, and teacher training. In line with this evolving focus, the curriculum at Rikkyo University has shifted toward a CLIL-based model with the intention of equipping students with the skills to engage deeply with content in English, while also introducing them to collaborative learning methods and various presentation formats (Ueno, 2024). Furthermore, it has been argued that a CLIL-based curriculum can help to cultivate multidimensional competencies that enable Rikkyo's students to navigate today's complex society with integrated knowledge, skills, and attitudes (Yamamoto & Niita, 2021). Although these changes represent a meaningful step in merging content and language skills, the integration of multimodality into such a curriculum remains largely underexplored. Limited research exists on how CLIL-based courses might be adapted to incorporate a multimodal focus that develops students' practical communication skills and multimodal competence. This gap underscores the need for a carefully designed course that addresses both linguistic competence and the ability to produce and interpret multimodal texts.

Literature Review

Multimodal Communication

Multimodal communication and social semiotics emphasize that meaning is constructed not only through language but also through various modes, including images, gestures, layout, and design. Halliday's (1978) concept of language as a social semiotic system emphasizes that language cannot be fully understood apart from the social and cultural contexts in which it is produced. Expanding on this idea, Kress and van Leeuwen (2006) argue that visual design follows structured "grammars" akin to verbal syntax, shaping relationships, emphasis, and hierarchy across multiple modes of communication. By extending Halliday's focus on context beyond language, they argue that the formal features of images, layouts, and other multimodal elements serve as resources that are designed to convey specific meanings. In addition, Jewitt (2009) highlights the growing prevalence of multimodal texts, from advertisements to educational resources, in everyday life and emphasizes the corresponding need for individuals to develop more sophisticated interpretive skills. Marcel Danesi (2019) also offers an insight into the nuances of multimodality and semiotics in a modern setting, by illustrating how metaphors and analogies underpin everyday cognition and communication, from the use of emojis to the elaboration of pop culture references. Multimodal texts often draw on well-known film scenes, celebrity personas, or iconic cultural symbols to evoke shared associations and emotional responses. This pragmatic orientation can serve as a helpful bridge in the classroom, showing students the real-world relevance of social semiotic and multimodal concepts. This is

particularly salient in second-language settings, where students must increasingly navigate not only the linguistic features of a text, but also the multimodal conventions of texts on a variety of diverse media platforms.

Social Semiotics

In addition to the aforementioned theories, much of the groundwork of social semiotics can be traced to linguists and semioticians like Ferdinand de Saussure (1983), Roman Jakobson (1960), and Roland Barthes (1972, 1977) whose collective work introduced foundational concepts. Saussure's signifier-signified framework and his emphasis on "signs," laid the basis for understanding the symbolic nature of language, while Jakobson's communication model highlighted the interplay of factors such as context and contact in shaping messages. Additionally, Barthes's insights into myth and connotation revealed how cultural ideologies are naturalized in everyday signs and he highlighted how consumer goods, photographs, and advertisements reflect and perpetuate broader cultural narratives. Taken together, these theories offer a vital analytical lens for a course on multimodal communication by enabling students to interpret and produce texts that blend visual, linguistic, and cultural elements in meaningful ways. The Birmingham School of Cultural Studies provides an additional suite of useful concepts, particularly the work of Stuart Hall (1980) and John Fiske (1987). Hall's encoding-decoding model highlights how audiences bring their own experiences and social positions to interpreting media, which can lead to negotiated or oppositional readings of texts. Fiske, similarly, frames media consumption as an active process of meaning-making, reinforcing the concept that students are not just passive recipients of information but rather co-creators of meaning. Incorporating these ideas into a multimodal course helps students develop a critical awareness of how cultural context shapes both the production and reception of texts.

Digital Contexts

Extending the study of multimodality into modern, digital contexts adds a further layer of theory to those already considered. Because students now interact with an array of digital texts—ranging from short-form social media posts on platforms like X (formerly Twitter) and Instagram to fully multimodal environments like YouTube and TikTok—it is crucial for educators and students to consider how these texts differ in structure and function from traditional print-based forms. Central to understanding these forms is the concept of "affordances," which Hutchby (2001) describes as the ways technology both enables and constrains user actions: for instance, X's character limit promotes succinct communication, whereas longer, more nuanced discussion is less feasible. Papacharissi (2014) expands this perspective by illustrating how design features influence emotional connections and identity construction on platforms like Instagram or Facebook. At the same time, the content that circulates—be it memes, viral videos, or remix culture—also plays a defining role in shaping meaning. Shifman's (2014) analysis of memes, for instance, illustrates how humor, cultural

references, and visual remixing allow these digital artifacts to spread rapidly and influence public discourse. The synergy between technological affordances and the multimodal content itself fosters distinct online cultures, where emotional resonance, interactivity, and visual appeal are integral to how messages are produced, circulated, and interpreted. This underscores the powerful interplay between platform design and the evolving forms of digital communication.

Multimodal Communication and CLIL

Despite the robust theoretical landscape, relatively few published studies focus on teaching social semiotics and multimodal communication within a CLIL or second-language learning environment. Much of the existing literature on multimodality addresses native-speaker contexts, presupposing advanced proficiency in English and often neglecting the added linguistic challenges faced by international students for whom English is a second or foreign language. Where studies do exist, for example in *Multimodal teaching and learning: The rhetorics of the science classroom* (Kress et al., 2014), they tend to discuss broad instructional strategies for multimodal literacy, rather than detailed accounts of sequential, action-research-based interventions. This gap in the literature highlights the unique nature of the present project: it aims not only to introduce core semiotic concepts but also to do so within a Japanese university course taught in English, accounting for the dual challenge of language acquisition and conceptual understanding. As a result, this research has the potential to inform future curriculum design in similar multilingual and multicultural contexts, providing examples of how and when to scaffold complex theories through carefully selected and adapted materials and activities.

Methodology

Research Questions

This paper details the design, implementation, and evaluation of a 14-week multimodal communication course for Japanese undergraduates studying in English, grounded in CLIL principles and guided by an action research framework. The course was designed to enhance the students' ability to critically analyze a range of multimodal texts—such as advertisements, news reports, and public service announcements—by examining the interplay of text, visuals, and design. To investigate the effectiveness of this approach, the study addresses the following research questions:

1. How can a CLIL-based multimodal communication course foster the development of multimodal competence and literacy among Japanese university students studying in English?

2. Which teaching materials, scaffolding strategies, and classroom tasks are most

effective in supporting students' multimodal analysis skills within this context?

3. How do students perceive and experience the course content, activities, and multimodal frameworks introduced throughout the 14 weeks?

Employing a mixed-methods strategy that includes pre- and post-course tests, surveys, teacher interviews, and reflective journaling, the paper offers a comprehensive perspective on both course design and its impact on student learning. Ultimately, the findings provide practical insights for developing multimodal communication curricula that integrate social semiotics and multimodal analysis with language instruction, thereby equipping students with the critical skills and cultural awareness necessary to navigate diverse cultural and social environments.

Research Design

This study follows an action research methodology guided by Anne Burns's (2010) framework in *Doing action research in English language teaching*. According to Burns, action research entails "a self-reflective, systematic, and critical approach to investigating one's own teaching context" (2010, p. 2). Its cyclical nature—planning, action, observation, and reflection—encourages educators to continually refine their practice based on firsthand classroom data. As Burns explains, "teachers become active participants in the research process, identifying problems, implementing changes, and documenting the outcomes" (2010, p. 5). This practitioner-oriented model is especially apt for exploring innovative pedagogies, such as a course on multimodal communication, because it allows for real-time adjustments that respond to immediate classroom needs. By actively engaging in data collection and analysis, the teacher-researcher can assess the effectiveness of teaching strategies, materials, and theoretical frameworks in a more organic and iterative manner.

Setting and Participants

The research took place at Rikkyo University's Ikebukuro campus, where a 14-week elective course on multimodal communication was offered to a mixed cohort of undergraduate students. The class met once a week, for 100 minutes, throughout the fall semester. The study included 25 undergraduate students from second, third, and fourth years, representing a range of English proficiency levels, with a minimum CEFR (Common European Framework of Reference for Languages) A2 standard recommended for enrollment. All students had previously completed mandatory English courses focusing on discussion, debate, presentation, and writing skills in their first year at the university. For this reason, they had some familiarity with university-level classes being conducted entirely in English. The students' majors spanned the humanities, providing a diverse range of academic and cultural perspectives. The classroom dynamic was positive, fostering thoughtful discussions around multimodal texts and how they function in various cultural contexts.

Research Instruments

Four primary methods of data collection were used to capture both qualitative and quantitative dimensions of student learning and course effectiveness.

Teacher Interviews

Semi-structured interviews were conducted with teachers who had previously taught the course. This step aimed to confirm or challenge my own approach to teaching the course and to gather further insight into effective strategies, and identify common pitfalls. Their feedback informed subsequent refinements to lesson plans and materials.

Classroom Observation

After each session, as a teacher-researcher, I maintained a reflective journal, documenting observations of student group work, engagement levels, and emergent questions or difficulties. This reflective process is central to Burns's (2010) action research framework, as it enables ongoing modification of teaching strategies based on immediate classroom realities.

Student Work Analysis

Throughout the semester, students completed assignments in which they analyzed multimodal texts—such as advertisements, news reports, and public service announcements—focusing on how text, image, layout, and other design features interact to create meaning. I devised rubrics anchored in the course content, emphasizing key concepts like Halliday's metafunctions, Kress's visual grammar, and other semiotics-related concepts. These rubrics measured the students' ability to apply theoretical frameworks to their analyses and offered a standardized measure of progress over time. Furthermore, I had students do pre- and post-course tasks that I also used to examine their understanding of multimodal concepts and texts.

Surveys

At the end of the course, students completed surveys regarding their engagement, comprehension of multimodal principles, perceived usefulness of course materials, and overall preferences. This survey provided valuable feedback on what aspects of the course most effectively facilitated learning and which areas required refinement.

By employing this multifaceted approach, the study aligns with Burns's emphasis on teacher-led inquiry, ensuring that each cycle of planning, action, observation, and reflection was firmly grounded in empirical classroom data. This methodology supports the iterative development of a multimodal communication curriculum tailored to the unique context of Japanese university students studying in English.

Ethical Considerations

Ethical approval for this study was obtained from the university's ethics review

committee prior to data collection. All participants were fully informed, both verbally and in writing, of the purpose of the research, the voluntary nature of their participation, and their right to withdraw at any time without academic penalty. Students signed a formal consent form allowing their anonymized work, survey responses, and classroom participation data to be used for research purposes. No personal information was disclosed in any publications or presentations arising from the study. These procedures ensured that the research was conducted in accordance with institutional guidelines and widely accepted ethical standards for educational research.

Implementation

Action research enables educators to investigate, evaluate, and improve their classroom practice in an iterative, evidence-based manner. Its emphasis on continual inquiry and adaptability makes it especially valuable for ensuring that lessons align with broader pedagogical goals while remaining responsive to student needs. In this project, action research underpinned the design and implementation of the 14-week course, guiding the iterative cycles of planning, action, observation, and reflection that shaped the syllabus, materials, and assessments. The following sections detail each phase of this action research process, illustrating how it provided a structured yet flexible framework for delivering a learner-centered, multimodal curriculum.

Planning

Identifying Goals and Choosing Framework and Concepts

The first step in planning this 14-week course involved clarifying the specific goals and outcomes students would be expected to achieve. In keeping with CLIL principles, I wanted them to refine their English language skills while developing a deeper understanding of multimodality, social semiotics, and cultural analysis. Establishing these clear objectives aligns with Nation and Macalister's (2010) emphasis on having well-defined learning goals that guide both content selection and classroom activities, thereby providing learners with a sense of direction and measurable milestones. Once these aims were set—such as recognizing key principles of multimodal communication, applying theoretical frameworks to real-world texts, and critically examining how culture shapes meaning—I evaluated numerous theoretical approaches to determine which would best serve these ends. To avoid overwhelming students with too many complex ideas at once, I then organized the material into a schedule that incrementally introduced more challenging concepts in a logical progression. According to Nation and Macalister (2010), one core principle of curriculum design is to structure the material in a systematic, incremental manner so that learners remain motivated and are not overwhelmed by overly difficult material. Ultimately, I decided to break the curriculum down into three sequential units, each building upon the previous one to maintain coherence and

student engagement and to balance increasing conceptual complexity with practical application.

In the first unit, students explored essential multimodality through key ideas drawn from Halliday's (1978) systemic functional linguistics (SFL) and Kress and van Leeuwen's (2006) grammar of visual design, supplemented by Barthes's (1977) insights on the interplay of image and text. In the second unit, we shifted focus to foundational semiotics, examining Barthes's (1972) further discussions of myth, alongside concepts inherited from Saussure—such as the signifier-signified distinction—as well as incorporating genre, intertextuality, and both synchronic and diachronic modes of analysis. Finally, in the third unit, students turned to digital texts, including narrative structures, and online phenomena like memes, drawing on Danesi (2019) and Shifman's (2014) explorations of popular culture to understand how these multimodal artifacts circulate in contemporary media. This three-tiered framework enabled learners to build gradually on their interpretive skills, transitioning from core understandings of multimodal principles to more advanced semiotic theories, and ultimately applying these insights to the dynamic and often participatory world of digital communication.

Preparing Adapted Materials

I knew that many of the primary texts would pose challenges for Japanese students studying in English. Their complex terminology and culturally specific references would necessitate pedagogical grading or adaptation, ensuring that the students would be able to engage with the core ideas without being overwhelmed by difficult syntax or unfamiliar cultural allusions. Therefore, during the preparation phase, I created simplified primers—400 to 500-word summaries—of complex theories like Halliday's SFL or Barthes's mythologies. These primers were designed to maintain conceptual accuracy while reducing linguistic complexity, ensuring that learners at around the CEFR A2 level could follow the core ideas without excessive frustration. As Burns observes, "action research encourages teachers to be responsive to the unique linguistic and cultural environment of their learners" (2010, p. 15). In total, ten weekly readings were created, and I decided that it would be most beneficial to give them to students as homework and preparation for the following lesson. This meant they had a full week to read and consider the material and would be more prepared to discuss the material in the following lesson.

Additionally, I designed PowerPoint presentations that would build upon each reading and address each weekly topic. The presentations were designed to be as multimodal as possible, integrating diagrams, images, and concise bullet points. These presentations illustrated key theoretical concepts using real-world examples, including familiar images from advertising and social media. The rationale was to give students a visual point of reference, showing them how semiotic principles manifest in everyday texts. This approach also reflected an important element of action research: "making learning visible" (Burns, 2010, p. 36) by demonstrating theoretical principles in a format that resonated with students' own communicative experiences.

Teacher Interviews

In the first few weeks of the course, I interviewed two teachers who had taught the Multimodal Communication course in the previous semester at Rikkyo, including the professor who had created and designed the original program. The teacher interviews provided valuable insights into the challenges and successes of teaching multimodal communication in a second-language teaching context. Both teachers had structured their courses around similar theoretical frameworks: rhetorical analysis and social semiotics. One had organized the course through the analysis of podcasting and advertisements while the other had adopted a project-based approach, prioritizing collaborative learning and real-world multimodal materials. Both teachers suggested that their goal was to help students understand the complexity of modern, multimodal communication by providing a structured framework through which the students would be able to approach the material.

One of the biggest challenges both teachers faced was helping students move beyond their previous learning experiences to critical analysis. According to one teacher, this involved "getting students to break out of that culturally normative process" through which they typically processed and analyzed written and visual material, to thinking more analytically about the process of meaning creation. Both teachers also emphasized the need for careful scaffolding, which for one teacher sometimes involved some Japanese-language support for rhetorical concepts, and for the other teacher this meant simplifying assignments after realizing students struggled with theoretical complexity. He reported that rhetorical analysis, for example, needed a lot of scaffolding. It was suggested that introducing it too early in the course had left students confused. By emphasizing the practical elements over the theoretical supporting material, it was noted that "thinking about how to include elements of the theories we had studied, helped them to cement their knowledge." As these interviews were conducted either at or near the start of term, I was able to utilize these insights and corroborate them with my own early observations.

The most important insights gleaned from these interviews were the importance of gradually introducing theoretical concepts, incorporating visual aids, balancing theory and practice, and fostering collaboration. These points directly informed several adjustments to the course design. In Week 1, students understood "multi" but struggled with "mode," prompting the addition of a brief presentation using images and layout diagrams to clarify the concept. Advice on balancing abstraction with modelling guided a Week 4 demonstration in which I deconstructed a digital image into setting, participants, and symbolic elements, helping students translate concepts like modality and salience into concrete analytical steps. Further recommendations about guided practice led to a Week 5 genre-analysis session using annotated advertisements and news stories to connect theoretical constructs to everyday texts. Collectively, these targeted revisions transformed interview insights into practical, actionable strategies that supported students' interpretive development across proficiency levels.

Action

According to Burns (2010), the action stage of an action research cycle involves implementing planned strategies in the classroom to address the identified objectives and challenges, thereby moving beyond theoretical deliberation and into concrete practice. For this project, action entailed delivering lessons organized around the three previously established course units, while employing a consistent instructional structure that balanced reading review, targeted presentations, and practical application tasks. Each class followed a predictable sequence designed to foster both conceptual clarity and student engagement.

Homework Review

Students began by discussing the prior week's reading in small groups, guided by pre-set prompts. This arrangement allowed them to clarify any confusing points and support peers who struggled with the material—an approach particularly beneficial in mixed-level settings. Group interactions also helped students feel more at ease than they might in a whole-class plenary, enabling them to explore and compare interpretations in a collaborative atmosphere.

Presentation

Following the group discussion, I delivered a concise presentation—typically 10–15 minutes—that consolidated the homework material and illustrated how theoretical concepts operate in authentic contexts. For instance, I would examine a magazine advertisement using Kress and van Leeuwen's framework of modality and salience, thereby demonstrating how abstract ideas manifest in real texts. This segment aimed to deepen their understanding before transitioning to hands-on activities.

Applied Analysis

In pairs or groups of three, students selected and deconstructed multimodal texts—advertisements, social media posts, or other relevant materials—applying each lesson's core concepts to their analysis. To facilitate this process, I set up a discussion forum on the Canvas Learning Management System (LMS). This enabled me to monitor their progress and provide targeted feedback both in real time during class and online. The collaborative element of this task was designed not only to foster peer-to-peer learning but also to encourage students to relate theoretical principles directly to the digital and print media they encounter in daily life.

Group Reflection and Discussion

The lesson concluded with a brief reflection period or whole-class discussion, during which students offered insights, highlighted difficulties, or inquired about finer points of the material. Concurrently, I recorded my observations in a reflective journal, reported below, so that any difficulties or recurring themes could be addressed in subsequent lessons. This closing step further exemplified the cyclical nature of action research, ensuring that classroom

experiences and student feedback consistently informed the next phase of instruction.

At the end of each unit—lessons 5, 9, and 12—the standard lesson flow was slightly modified to provide time for introducing and discussing the upcoming assessments. For the first two units, students were required to submit written reports in which they analyzed images using the conceptual tools covered in that unit. In the third and final unit, however, they prepared group presentations. Additionally, the first unit served as a pilot of sorts, allowing for systematic observation and subsequent adjustments in keeping with the cyclical nature of the action research framework.

Observation

While the observation and reflection sections appear here in consecutive format, it is important to note that in action research the process is recursive, meaning you begin to examine and interpret in a dynamic way from the outset (Burns, 2010). This study followed that principle by documenting observations and initial analyses simultaneously, yet each section is separated to maintain clarity of the process. In practice, data collection and interpretation evolved in a fluid, overlapping manner. My reflective journal formed the core of these observations, supplemented by student questions, small-group discussions, and interaction data from Canvas LMS. By focusing first on what was recorded, this section illustrates how systematic note-taking and thematic categorization helped organize key insights. Qualitative data in action research involves inductive coding, meaning that categories of study emerge from the data (Burns, 2010). Applying this approach, the central themes that emerged were language use and terminology, mixed-level engagement, and task structure and clarity.

Language Use and Terminology

Many students demonstrated intermediate to high English proficiency but encountered difficulties with certain theoretical phrases, including “multimodal” itself. Learners easily recognized “multi” but struggled with “modal.” Similar terminology issues arose as the course progressed. Moreover, toward the end of the first teaching unit, recurring errors emerged in describing the layout and placement of items in images using prepositional phrases— “*in* the left” instead of “*on* the left,” for example—as well as expressing appropriate degrees of certainty in terms of what meanings emerged during their analysis— “it means” instead of “it *could* mean.”

Mixed-Level Challenges

Data from Weeks 2 and 3 indicated that students spent between 5 and 20 minutes on the weekly readings, reflecting considerable variation in reading fluency within the same class. Some students understood the readings quite easily, while others needed to make extensive notes—often writing Japanese translations to fully grasp the text—and yet others found the material too dense or time-consuming to fully understand. Such discrepancies highlight the

complexity of catering to multiple proficiency levels. Several journal entries further underscored the tension between challenging more advanced students and supporting those who struggled with foundational concepts. These issues became particularly pronounced when tasks demanded higher-level thinking, such as when analyzing images with minimal guidance or applying theoretical constructs without concrete examples.

Task Structure and Clarity

By Week 4, the data showed that an unstructured image-search assignment overwhelmed certain students who lacked clear search terms or parameters, undermining both engagement and motivation. This underscored the importance of providing well-defined instructions, especially when students had to analyze multimodal elements like modality in unfamiliar contexts. In Week 5, a similar obstacle arose with genre distinctions: students had difficulty differentiating the features of advertisements, news stories, and public service announcements, suggesting that explicit instruction on genre features was needed. Both of these observations point to the value of scaffolding complex tasks, clarifying key definitions, and offering concrete examples. These are strategies that can help students navigate a wide range of multimodal texts more confidently.

Overall, the categorized data highlighted a range of challenges in teaching advanced multimodal and semiotic concepts in a linguistically diverse environment. While enthusiasm for the material was generally high, issues surfaced around language usage, reading comprehension, mixed-level engagement, task clarity, and genre recognition. These observations provided a critical foundation for the reflective and adaptive processes that would guide subsequent instructional decisions in the action research cycle.

Reflection

Based on these ongoing insights, I implemented several targeted interventions to address challenges around language use, varied reading proficiency, task clarity, and genre-related misunderstandings.

Language Use and Terminology

As mentioned, one recurring issue was students' difficulty in describing images accurately, especially in using prepositional phrases to describe features in a layout, and clarifying levels of speculation or certainty. Observations drawn from class discussions and online posts also showed that many learners struggled with the vocabulary needed to articulate meaning, layout, and symbolism. In response, I created a detailed booklet (Appendix A) that included illustrative sample images and model sentences demonstrating how to comment on visual elements and interpret their significance. After distributing this resource via Canvas and reviewing it in class, learners who required further practice and benefited from structured guidance were able to consult the guide. On the other hand, more advanced students could opt out if they felt sufficiently confident, thereby maintaining an appropriate degree of challenge

for everyone.

Mixed-Level Challenges

Reflective notes from Weeks 2 and 3 highlighted the tension between challenging more advanced learners and ensuring lower-level students were not overwhelmed by technical concepts. To address this, I differentiated tasks by incorporating multiple difficulty tiers in class activities, enabling students to complete the number of tasks most suited to their level of confidence and ability. As mentioned in the previous section, the optional supplementary materials I provided, such as simplified examples of language use, offered further support for those who needed it. In addition, I adapted the assigned weekly readings by substituting technical terms like *paradigm* and *syntagm* with *semiotic choice* and *semiotic combination* respectively. This preserved core conceptual meaning while lowering the linguistic barrier. Beyond vocabulary adjustments, I simplified sentence structures and general word choices, thereby reducing the cognitive load for lower-level learners. This multifaceted approach, informed by my early reflections, aimed to keep higher-level students engaged while ensuring those who struggled received adequate scaffolding.

Task Structure and Clarity

Observations also indicated that unstructured activities, like the Week 4 image-search exercise, posed significant barriers to student engagement. By making search parameters clearer—suggesting that they find images related to "university life," "favorite products," or "sports teams"—students found it easier to locate relevant images and apply course concepts. In a similar vein, the Week 5 confusion over identifying advertisements, news stories, and public service announcements led me to introduce explicit genre overviews. Devoting a portion of class to clarifying each genre's defining features, and providing concrete examples for comparison, helped students feel more confident in their analyses. These modifications reflect Burns's notion that "the reality of classroom life" (2010, p. 135) should guide ongoing curricular adjustments.

Taken together, these interventions illustrate how the cyclical nature of action research facilitated dynamic, context-aware decision-making. By systematically observing and categorizing student challenges, I could tailor strategies that balanced theoretical rigor with practical accessibility, ensuring the course remained responsive to learners' evolving needs. These insights subsequently led to a new cycle of planning, action, observation, and reflection reinforcing the iterative nature of the action research framework.

Findings from Student Work and Course Data

This section synthesizes multiple data sources to evaluate student performance during the 14-week course. These sources include: an analysis of the students' final presentations, a comparative pre- and post-course image analysis, and a student survey. Together, they provide

a comprehensive view of how students applied theoretical concepts, their evolving multimodal literacy skills, and their subjective assessments of the course's efficacy.

Student Assessment

In the final weeks of the semester, students collaborated on group-based final presentation projects. They were required to apply a suite of multimodal and social semiotic theories to a variety of texts connected by a topic of their choosing. I evaluated each project in relation to the following three criteria: the depth of their analysis and their ability to deploy concepts learned during the course, their ability to identify and evaluate the societal attitudes, ideas, and beliefs present in their chosen texts, and their skill in communicating these ideas.

A key strength observed across many submissions was the ability to contextualize the chosen topic—whether it was Japanese train culture, the role of idols in Japanese pop culture, or evolving gender norms in toy advertisements—within broader socio-cultural frameworks. In several high-scoring presentations, students were able to articulate how textual and visual elements interact to convey ideological or cultural messages. For example, one high-scoring group compared Japanese and American New Year's advertisements. They explained how the handwritten font and warm color palette in the Japanese ads evoked family intimacy, while the sharp gold-and-black contrasts in the American ads signaled luxury and celebration. They linked these design choices to broader cultural values, demonstrating an ability to connect multimodal features to societal meanings.

Despite these positive outcomes, some presentations demonstrated only minimal integration of course concepts. While these students occasionally referenced relevant theoretical terminology, they often did not connect it systematically to the visual or textual evidence. For example, one lower-scoring group analyzed a public service announcement about distracted driving but relied largely on literal description. The students stated that the image used "high modality" and "strong salience," but did not explain how: they described "a man looking at his phone" without discussing the significance of the dim background, or the contrast between the bright screen and the darkened environment. Nevertheless, most students displayed greater textual awareness and cultural sensitivity than they had at the start of the course, though some still needed explicit guidance in applying multimodal theories more thoroughly.

Pre- and Post-Course Image Analysis

To gauge progression in the students' analytical abilities, I administered a pre-course image analysis in Week 1—a Dior advertisement featuring Johnny Depp—and a post-course image analysis in the final week—a Chanel advertisement featuring the actor Gaspard Ulliel (Appendix B). A comparative review of these analyses revealed concrete improvements in theoretical application and level of detail.

Early responses in the pre-course image analysis revealed broad, class-wide patterns: most students focused primarily on surface-level attributes such as celebrity presence, product

placement, or general mood, offering descriptors like “masculine,” “cool,” or “strong.” While a few students referenced elements such as the desert setting or Johnny Depp’s persona, the majority lacked a coherent framework for linking these observations to broader semiotic processes. Students also rarely attended to textual features, including the Dior logo, font, or tagline, indicating limited awareness of how different modes combine to produce meaning.

In contrast, the post-course analysis showed a clear overall shift across the class. Students more frequently identified specific visual cues—color, lighting, composition—and explained how these elements contributed to meanings such as “luxury,” “exclusivity,” or “sophistication.” Many also began integrating references to textual modes, including logo placement and background selection, to argue how Chanel constructs brand identity through multimodal choices. Although a small number of students did not consistently use the full range of terminology taught in the course, the class as a whole demonstrated a more structured and theoretically informed approach to analysis.

An illustrative example of one student’s progress demonstrates the shift many students made from basic description to informed multimodal analysis. In the pre-course task, this student’s comments on the Dior advertisement focused largely on naming visible components—“There is a man in desert…. Next to him, there is a perfume”—and offering general impressions such as “braveness are present,” with little explanation of how any element produced these effects. By contrast, the same student’s post-course analysis of the Chanel advertisement showed more attention to multimodal features and textual analysis. She observed that the “tone of color… is very dark and not bright,” allowing her to “catch the feeling of braveness,” and she linked the model’s pose and suit to ideas of “confidence and strength.” She also identified the brand’s messaging, explaining that Chanel “wants to tell how men will be able to get confidence by putting perfume.” This illustrates how instruction in concepts related to multimodality enabled her to develop a more coherent interpretation grounded in theory.

Student Surveys

To supplement these performance-based measures, a mixed-methods survey (Appendix C) was conducted in the final week of the course. A total of 19 students participated, providing both quantitative Likert-scale ratings and qualitative feedback through open-ended questions. The instrument aimed to assess overall satisfaction, engagement, perceived comprehension of multimodal theories, and recommendations for course enhancement.

Overall Satisfaction and Engagement

On a five-point Likert scale—ranging from 1 = “Poor” to 5 = “Excellent”—a majority of respondents (N = 19) rated the course experience as either “Very Good” (4) or “Excellent” (5), representing 68.4% of the class. A large majority of students—84.2% of respondents—rated the course content as either “Somewhat Engaging” (42.1%) or “Very Engaging” (42.1%). This suggests that most learners found the material stimulating and relevant. However, a small

minority (10.5%) selected "Neutral," and one respondent rated the content as "Not Engaging," indicating some variability in individual interest and learning preferences. Additional feedback implied that group discussions and applied activities contributed to sustained engagement, whereas exclusively lecture-based segments were perceived as less stimulating.

Confidence in Understanding Multimodal Concepts

Students' self-assessments of their mastery of core concepts showed a clear upward shift over the semester (Figures 1 and 2). Before instruction, 53% of students rated their understanding of multimodality as "Poor," 47% rated visual communication as "Poor," and 58% rated semiotics as "Poor," with very few selecting "Very Good" or "Excellent." After the 14-week course, these distributions changed substantially: for multimodality, 53% rated their skills as "Very Good" and 11% as "Excellent"; for visual communication, 63% selected "Very Good" and 16% "Excellent"; and for semiotics, 47% chose "Very Good" and 5% "Excellent," with no "Poor" ratings remaining in any category. These gains suggest that the combination of theoretical instruction and applied analytical tasks significantly strengthened students' perceived mastery. A small minority, however, still reported difficulty with abstract concepts such as myth and metonymy. This indicates the continued importance of scaffolding and concrete examples when teaching complex theories in a second-language environment.

Figure 1

Students' Self-Assessed Understanding of Content and Concepts Before Instruction.

Figure 2

Students' Self-Assessed Understanding of Content and Concepts After Instruction.

Effectiveness of Course Structure and Activities

When asked to evaluate the course structure—which combined lectures, discussions, and practical analysis—83% of respondents rated it "Effective" (4) or "Very Effective" (5). Open-ended comments reinforced the value of group-based tasks in promoting critical engagement and language practice. For example, one student said, "I think the group discussion at the beginning of the class was very useful because I was able to understand the contents which I could not understand by myself by talking with classmates." Several respondents advocated for even more collaborative opportunities, suggesting that repeated interactions with diverse peers facilitated deeper exploration of ideas.

Impact on Media Literacy and Critical Thinking

The survey further revealed that most participants experienced a notable shift in how they interpret daily media like advertisements and content on social media. One respondent said, "I became consciously aware of things I normally don't look at," while another said, "Up until now, when I saw an advertisement, I would just think, "It's attractive and catchy," but now I've started to look for the reason behind it, thinking, "It's catchy because the composition of this part is well thought out." This qualitative finding aligns with the overarching course objective of fostering a more nuanced and critical perspective on everyday multimodal texts. Such self-reported transformations suggest that, despite occasional difficulties, students did internalize key analytical frameworks to a degree sufficient for practical application beyond academic tasks.

Suggestions for Future Course Enhancements

Open-ended survey responses spotlighted potential areas for refinement. While most students found the course material engaging and beneficial—71% rated these as "Effective" (4) or "Very Effective" (5)—a subset expressed an interest in having more time to focus on elements that were only covered in passing. They requested more time to look at videos, for example, and to think more about the construction of narratives. Likewise, some suggested that not enough time was devoted to affordances, especially as they apply to social media platforms. This reflected their interest and engagement in communication in more contemporary, digital contexts. These recommendations are consistent with my observation that concrete exemplars related to the students' real-world experiences and interests improve comprehension and maintain learner motivation.

By combining the final presentation assessment, comparative image analyses, and survey data, a comprehensive picture emerges of students' development in multimodal literacy. The evidence indicates that learners made substantial progress in dissecting and critiquing texts through a semiotic lens, transitioning from initial, largely descriptive observations to more rigorous, evidence-based observations. The majority of respondents reported high overall satisfaction and described increased media literacy. This suggested that the course successfully promoted critical thinking skills relevant to a variety of real-world communicative contexts.

Discussion

The findings from this study indicate that an action research approach, following Burns's (2010) cyclical methodology, can be highly effective in designing and implementing a course on multimodal communication for Japanese university students studying in English. Over the 14-week term, participants demonstrated growth in analyzing, interpreting, and critiquing a variety of cultural texts through a multimodal lens, and the systematic process of planning, teaching, observing, and reflecting proved central to this development. Regular use of reflective journals, coursework analysis, and classroom observations allowed for timely pedagogical shifts, from simplifying theoretical readings to providing structured search prompts for image analysis. Such interventions successfully addressed specific challenges—mixed proficiency levels, difficulties with abstract concepts, and gaps in task clarity—and contributed to demonstrable improvements in students' analytical depth and cultural insight.

From a theoretical perspective, these findings validate the view that social semiotics and multimodal discourse analysis provide an effective blueprint for helping learners interpret visual, textual, and design features in a more systematic manner. By progressing through three distinct units—beginning with fundamental multimodality, moving into foundational semiotics, and concluding with digital texts—students advanced from mostly surface-level observations to increasingly nuanced critiques. Explicit instruction in these theoretical frameworks sharpened their interpretive skills, revealing how cultural and ideological factors shape meaning across a range of communication modes. Moreover, by engaging with a diverse array of multimodal texts, students reevaluated their everyday interactions with media and emerged with a more analytical, critical approach to understanding how meaning is constructed and conveyed.

The study further suggests that multimodal approaches enrich the language-learning experience, particularly within a CLIL framework. While many existing studies focus primarily on making teaching materials more visually engaging, fewer detail explicit strategies for teaching social semiotics and multimodality to second-language learners. The data presented here underline the potential of tasks that merge theory with iterative, hands-on practice—such as image analyses, reflective writing, and guided discussions—enabling learners to decode and produce meaning across multiple genres of texts.

Several pedagogical implications emerge from these findings. First, targeted scaffolding is crucial when introducing abstract theories, particularly for mixed-level classes. Adapting high-level texts or developing concise primers can offer a more manageable entry point, ensuring that students grasp the conceptual core without becoming discouraged by overly technical language. Second, collaborative tasks—like group discussions and peer-review exercises—are vital for fostering critical engagement and confidence, enabling learners to share insights and support one another's interpretive efforts. Third, the effectiveness of frequent feedback loops—including immediate commentary, smaller checkpoints, and reflective journaling—underlines the value of cycling information back into subsequent lessons. This

helps students internalize theoretical concepts as part of an ongoing process rather than merely memorizing them for discrete assessments. Finally, diverse text selection can keep motivation high and appeal to a range of cultural interests, demonstrating how multimodal principles function across different contexts and genres.

Collectively, these observations highlight how an action research model enables educators to remain adaptive, systematically evaluating and refining their methods in response to real-time classroom data. In doing so, teachers can align theoretical rigor with practical accessibility, guiding students toward a fuller understanding of multimodal communication and ensuring that their evolving needs remain at the forefront of pedagogical decision-making.

Conclusion

This study set out to develop and implement a multimodal communication course at a Japanese university, drawing on Burns's (2010) action research model to equip students with analytical tools for dissecting multimodal texts while simultaneously refining their English-language skills. By merging theories about multimodality, social semiotics, and visual communication, the course facilitated a more holistic understanding of how meaning is constructed across multiple modes. While implementation challenges included mixed proficiency levels and the abstract nature of key theoretical concepts, empirical evidence, spanning final presentations, image analyses, and survey data, demonstrated that students evolved into critical, reflective interpreters of complex cultural texts. These findings hold promise for broader implementations of multimodal pedagogy within CLIL contexts, particularly where teachers seek to cultivate advanced literacy skills and cross-cultural awareness through iterative, practitioner-oriented approaches.

The data revealed notable improvements in students' abilities to identify cultural and ideological meanings, articulate nuanced interpretations, and apply specialized terminology related to this field of study. Survey responses corroborated these gains, showing high levels of satisfaction with the course design and emphasizing students' heightened awareness of how social and cultural narratives embed themselves in everyday media. Crucially, the action research process proved instrumental in balancing the theoretical complexity of multimodal frameworks with the practicalities of classroom engagement: by iterating instruction based on data-driven insights, the course remained responsive to student needs while meeting its academic objectives. Ultimately, the resulting gains in analytical sophistication suggest that students can thrive in settings where cultural, linguistic, and semiotic dimensions intersect—a conclusion that highlights the transformative potential of reflective, theory-informed teaching in global higher education.

References

Barthes, R. (1972). *Mythologies* (A. Lavers, Trans.). Hill and Wang. (Original work published 1957)

Barthes, R. (1977). *Image, music, text* (S. Heath, Trans.). Hill and Wang. (Original works published 1961–1973)

Burns, A. (2010). *Doing action research in English language teaching*. Routledge.

Coyle, D., Hood, P., & Marsh, D. (2010). *CLIL: Content and language integrated learning*. Cambridge University Press.

Danesi, M. (2019). *Understanding media semiotics*. Bloomsbury.

Fiske, J. (1987). *Television culture*. Methuen.

Hall, S. (1980). Encoding/decoding. In S. Hall, D. Hobson, A. Lowe, & P. Willis (Eds.), *Culture, Media, Language: Working Papers in Cultural Studies, 1972-79* (pp. 128–138). Routledge.

Halliday, M. (1978). *Language as social semiotic: The social interpretation of language and meaning*. Edward Arnold.

Hutchby, I. (2001). Technologies, texts and affordances. *Sociology*, *35*(2), 441–456. https://doi.org/10.1177/s0038038501000219

Ikeda, M. (2013). Does CLIL work for Japanese secondary school students? Potential for the 'weak' version of CLIL. *International CLIL Research Journal*, *2*(1), 31–43. https://urn.fi/URN:NBN:fi:jyu-202504083138

Ikeda, M., Izumi, S., Watanabe, Y. P., & Davis, M. (2022). *Soft CLIL and English language teaching: Understanding Japanese policy, practice, and implications*. Routledge.

Jakobson, R. (1960). Closing statement: Linguistics and poetics. In T. Sebeok (Ed.), *Style in Language* (pp. 350–377). MIT Press.

Jewitt, C. (2009). *The Routledge handbook of multimodal analysis*. Routledge.

Kress, G. (2010). *Multimodality: A social semiotic approach to contemporary communication*. Routledge.

Kress, G., Jewitt, C., Ogborn, J., & Tsatsarelis, C. (2014). *Multimodal teaching and learning: The rhetorics of the science classroom*. Bloomsbury.

Kress, G., & van Leeuwen, T. (2006). *Reading images: The grammar of visual design*. Routledge.

Nation, I., & Macalister, J. (2010). *Language curriculum design*. Routledge.

Papacharissi, Z. (2014). *Affective publics: Sentiment, technology, and politics*. Oxford University Press.

Saussure, F. d. (1983). *Course in general linguistics* (R. Harris, Trans. & Annot.). Duckworth. (Original work published 1916)

Shifman, L. (2014). *Memes in digital culture*. MIT Press.

Ueno, I. (2024). Bridging to English-medium instruction through CLIL: Student feedback based

on university pilot courses. *Journal of Foreign Language Education and Research*, *5*, 1–20.

Yamamoto, Y., & Niita, R. (2021). Action-oriented approach to curriculum development in CLIL courses: A theoretical and methodological framework. *Journal of Foreign Language Education and Research*, *2*, 122–135.

Appendices

Appendix A: Sample Page from Booklet for Describing Images

1. Image Layout: Prepositions of Place

Looking at the Image

On the left / right: There is a tree on the left side of the picture.

In the middle/center: A large building is in the center of the image.

At the top / bottom: At the top of the picture, we see the sky.

In the foreground / background: There is a bottle in the background.

People and objects in relation to each other

Next to: The man is standing next to a car.

Behind: Behind the building, there is a mountain.

In front of: There is a woman standing in front of the house.

Between: The lamp is between two chairs.

Above: Above the mountains, we can see clouds.

Below: Below the table, there is a cat.

Untitled (Chicago) by Vivian Maier (1975)

In the center of this urban street scene, two men and one woman, all wearing yellow clothing, stand out against the grey sidewalk and background buildings. **On the left**, one man leans forward toward a board; **on the right**, the woman walks towards the camera, holding a tan handbag. It seems this image captures a candid, everyday moment, yet the colors feel coordinated or staged. The repetition of yellow might symbolize conformity, coincidence, or even absurdity. This photo suggests how random patterns in public life can appear humorous, strange, or strangely beautiful.

Appendix B: Images for Pre- and Post-Course Analysis

The images were selected for their formal similarities. Both are advertisements for men's fragrance and have four predominant features: model, setting, product image, brand.

Pre-Course Image: "Sauvage" advertisement by Dior (Dior, 2015). Reproduced under fair use for educational purposes.

Post-Course Image: "Bleu" advertisement by Chanel (Chanel, 2010). Reproduced under fair use for educational purposes.

Appendix C: Student Survey

Section 1: Demographic Information

1. Student Year
2. How would you describe your English level?

Section 2: Course Experience

3. Overall, how would you rate your experience on this course?
4. Which concepts, if any, did you find difficult to understand? (Select all that apply)
5. Please explain why you found the above concepts difficult to understand. [Open-ended response]
6. Which materials were most useful for developing an understanding of the concepts? (Select all that apply)
7. How confident are you that your understanding of multimodal communication improved during the course?

Section 3: Understanding of Key Concepts

8. Rate your understanding of the following concepts before the course.
9. Rate your understanding of the following concepts after the course.
10. How helpful were the following materials in enhancing your understanding of multimodal communication?
11. What additional resources or materials would have helped you better understand the course concepts? [Open-ended response]

Section 4: Course Feedback

12. How engaging did you find the course content?
13. How effective was the course structure (lectures, activities, discussions) in helping you learn?
14. How effective was the class assessment (assignment, presentation, portfolio) in helping you learn?
15. Were there any parts of the course you think could be removed or shortened? [Open-ended response]
16. Were there any parts of the course you wish had more time or focus? [Open-ended response]
17. Did this course change the way you think about media, images, or advertisements? If so, how? [Open-ended response]
18. Please provide any additional comments or suggestions for improving the course. [Open-ended response]