

Moving Multimodal Task Assessment Forward in Language Education via Learner Perspectives

Takuya Kojima

Moving Multimodal Task Assessment Forward in Language Education via Learner Perspectives

Nagasaki University Takuya Kojima

Abstract

This article considers assessment for multimodal tasks from learner perspectives and suggests more viable forms of assessment. In the current digitalised society, assembling multiple modes of expression as multimodality is prevalent in our daily communication. Scholars have recognised that this ability to use and interpret multimodal expressions, called multiliteracies, is increasingly important. There have been heated discussions on improvements in implementing multimodal tasks for learner multiliteracies development. However, discussions on assessment in this area have missed research into learner perspectives on such assessment. The current study takes an action research approach to explore learner perspectives on the multimodal composition of a university Japanese language course in Italy and its assessment practice qualitatively. Analysis of the participating student data revealed that nearly three-quarters showed subtle interest in their composition being multimodal. However, those students who focused on multimodality found that expressing feelings and involving readers—the aspects missing in scholarly discussions to date—as what counts as valuable in assessment for multimodal tasks. Drawing on these findings, this article argues that the following three actions can promote our multimodal assessment literacy to design and implement more viable assessment for multimodal tasks: 1) raising our awareness towards multimodality and multiliteracies, 2) integrating peer-assessment into our assessment practices, and 3) further exploring learner-generated criteria.

Key Words: multimodality, multiliteracies, assessment for multimodal tasks, learner perspective, peer-assessment, learner-generated criteria, rubric

1. Introduction¹

Our communication is undergoing a major change in the current digitalised society. We can observe the change in our communication with multiple modes of expression regularly appearing in an interwoven manner. These modes are understood as socially and culturally

¹ This article is based on the presentation given at Japanese Studies Association of Australia Conference 2021-Sustainability, Longevity and Mobility. The data were collected while the author was affiliated with Ca' Foscari University of Venice. However, he analysed the data after he moved to Nagasaki University.

shaped resources for meaning-making (Bezemer & Kress, 2008). Kress (2009) argues that each mode—written or spoken language, image, audio, gesture, and space—has its own unique affordances and limitations, so one mode cannot simply replace itself with another. In this article, “the use of several semiotic modes in the design of a semiotic product or event” is defined as *multimodality* (Kress & van Leeuwen, 2001, p. 20). A common example of multimodality is sending a text message with an emoji or two and a picture on a smartphone. The message is considered a multimodal product. Despite this change, scholars, teachers, and learners in the language education field have primarily paid attention to teaching and learning written and spoken language (Cope & Kalantzis, 2009; Grapin, 2019), although paralinguistic has occasionally been under consideration. In further developing the field, the New London Group (1996) has advocated for taking varieties of modes into consideration, proposing a new conceptualisation of literacy called *multiliteracies*.

Today, multiliteracies has gained momentum as a concept that questions the fundamental understanding of what counts as language and language education. Unlike mere literacy, multiliteracies is understood as our ability to use and interpret expressions constituted with multiple semiotic modes of which written and spoken language remains an integral part (Cope & Kalantzis, 2009). That is, we do not exclusively rely on written and spoken language as the only mode to express ourselves and understand ourselves, others, and society when we develop our multiliteracies. Assuring “the right to speak” (Peirce, 1995, p. 18), or more pointedly, the right to express ourselves in this way is vital for our wellbeing. The proliferation of multiliteracies may free us from circumstances in which we judge each other and determine hierarchical power relations based on our ability to use and understand written and spoken language alone (Grapin, 2019). Multiliteracies can also be particularly significant in relation to pursuing, actualising, and maintaining *Tabunkakyōsē* society, where people with diverse backgrounds and resources live convivially together (Illich, 1973; Jewitt, 2008). Thus, moving the field forward with multimodality and multiliteracies is considered a worthwhile endeavour.

To date, scholars and teachers have actively discussed what multiliteracies signifies and what kind of pedagogy best suits learners to develop their multiliteracies. In such discussions, engaging learners with multimodal tasks through which they interpret and produce multimodal products is considered meaningful practice. In contrast, discussions on assessment for multimodal tasks have just begun. In this article, assessment is signified as differ-

ent ways to understand learner performance, achievement, and change and the judgement made on them (Sato & Kumagai, 2011; Usami, 2016). The literature review identifies the need for increased attention on learner perspectives to engage in more fruitful discussions on assessment for multimodal tasks. Considering our changing communication landscape, as a Japanese language teacher, the author² made his first attempt to design and implement a course focusing on multimodal composition, one of the widely known multimodal tasks. Drawing primarily on learner perspectives in the attempt, this article aims to consider and suggest a possible direction for designing and implementing more viable assessment for multimodal tasks.

2. Literature Review

2. 1. Features of multimodal tasks

Discussions on assessment for multimodal tasks have caught scholars' and teachers' attention, having realised assessments informed by a print-based and language-focused view of literacy are no longer effective for assessing multimodal tasks. Two distinctive features in such tasks need to be noted to understand the reason behind this argument.

The first feature is marked by *semiotic harmony* created through multiple modes of expression. Kress (2009) explains that multimodal ensemble—the co-existence of more than one mode in a single unit of expression—and orchestration—the harmonisation of multiple modes in a single unit of expression—constitute the semiotic harmony. Semiotic cohesion is an often-used alternative term representing this feature (e.g., Levy & Kimber, 2009). Thus, assessment practice should capture this wholeness of multimodal tasks and end products.

The second feature emerges in the concept of *designing*. For instance, multimodal composition is no longer written but designed, and this term has two intertwined assumptions (Kress, 2009). One is that meaning-making occurs continuously when a multimodal product is designed. That is, we do not acquire existing knowledge, skills, and frameworks and reproduce them in designing. In turn, designing is practised with more than one answer.

² The study of this article took an action research approach. In action research, it is often recommended to refer to the researcher using the pronoun "I". Its purpose is to demonstrate teachers-as-researchers' awareness towards the importance of exercising their self-reflexivity (Burns, 2009). However, the author uses "the author" in the article to refer to himself to follow the journal's request.

These lead to another assumption—learning constantly occurs when designing. Assessment acknowledging and supporting such meaning-making and constant learning appears effective in assessing multimodal tasks. In this way, we can see those in language classrooms as not learners but designers creating social futures (Cope & Kalantzis, 2009; Grapin, 2019).

Notably, two versions of views on multimodality lie beneath multimodal tasks and their assessment. The first is the strong version of multimodality (Grapin, 2019), which underscores that we all strategically and constantly use multimodality and envisage all modes as equally valued based on their unique affordances and limitations. This version is what the originators envisage as multimodality (Kress, 2009; New London Group, 1996). The second view is the weak version of multimodality (Grapin, 2019), which signifies that language learners only rely on multimodality until they learn to express themselves solely using the target language. That is, non-language modes are considered as scaffolding. Non-language modes merely compensate a language mode; language is a privileged mode with which we can express everything. While these versions have no clear-cut boundary, Grapin (2019) argues that the weak version has spread into the mainstream of language education today. Understandably, the version, whichever it is, influences the design and implementation of the curriculum and assessment, as well as teaching and learning. Before shifting our attention to assessment for multimodal tasks, the following section overviews discussions on assessment in language education.

2. 2. Assessment in language education

Assessment has been increasingly integrated into discussions for quality language education in the field. It is now common to understand that assessment has a washback effect (Alderson & Wall, 1993; Davies, 2013). The washback effect signifies how the design and implementation of assessment influences teaching practice and learning activities (Alderson & Benerjee, 2001). Washback becomes positive when assessment, teaching, and learning coordinate and, conversely, functions negatively when they do not. Utmost caution is needed on washback in reality, as its effect cannot be determined simplistically; one assessment may differently influence teaching and learning or two individuals even in the same circumstance (Alderson & Wall, 1993). Considering this phenomenon, McNamara (2001) argues that the boundary between assessment, teaching, and learning has been blurred. Discussing assessment has become more vital than before in pursuing and actualising quality

language teaching and learning regardless of goals, underlying theoretical assumptions, and contextual factors.

We have observed a shift from assessment *of* learning to assessment *for* learning in language education. These assessment views set fundamentally differing purposes. Assessment *of* learning aims to provide accountability, ranking, and certifying competence (William, 2011). With this view, assessment is implemented when a series of the learning process is completed. This view marks features such as objectiveness, reliability, validity, fairness, and practicality (Davies, 2013; Shimada, 2020). In cases of assessment *of* learning, assessors are teachers or the equivalent conducting a one-way assessment. Tasks in such assessments include true-or-false, multiple-choice, and fill-in-the-blank questions that can set a definite answer for each question. The assessors know the answer. Scholars have begun pointing out that assessor subjectivities are hard to eliminate, one assessor may not give the same assignment the same opinion every time, and the unproductive power relationship between the assessing and the assessed emerge in one-way assessments (Davies, 2013; Usami, 2016). Consequently, individual subjectivity in assessments, intrapersonal assessor fluidity, and the mutually constitutive nature of assessment are being acknowledged to conduct more viable assessments.

Assessment *for* learning sets its purpose as promoting learning (William, 2011). In this assessment view, assessors are no longer only teachers but also learners themselves and their peers, which multiplies viewpoints in assessment and places learners closer to the centre of attention. This has led to introducing self-assessment and peer-assessment in the field, especially for performance-oriented tasks. Further, the implementation of portfolios and rubrics has begun in order to place more focus on learning processes. In particular, the value of rubrics is widely recognised in assessment *for* learning. They clarify selected aspects learners should attend to, alongside a scale from 'inadequate' or 'not achieved' to 'excellent'. This helps teachers assess learners more transparently, and learners self-assess and improve their work throughout the learning process (Reed, 2008). Observing this trend, Davies (2013) asserts that "the earlier questions for testers – 'how?' and 'what?' – [has] become 'why?' and 'should we?'" (p. 11). Taking these into consideration is of critical importance when discussing assessment for multimodal tasks.

2. 3. Assessment for multimodal tasks

To date, the perspectives of scholars and teachers have dominated discussions on assessment for multimodal tasks. For instance, among 47 studies on assessment for multimodal tasks reviewed by Anderson and Krachorsky (2019), 42 employ teacher-led assessment practice. Thus far, scholars and teachers have committed to developing and sophisticating assessment metalanguage to devise the ideal rubric. Drawing on Hafner and Ho's (2020) and Tan et al.'s (2020) studies, Table 1 summarises major assessment criteria for multimodal tasks. The theory-driven rubric informed by the New London Group (1996) focuses on linguistics, visual, gestural, auditory, and spatial design (Hung et al., 2013). The practice-based rubric developed by teaching practitioners takes artefact, rhetorical skills, substance, process management and technical skills, and habits of mind into consideration (Eidman-Aadahl et al., 2013). The flexible and adoptive rubric adjusted for differing multimodal tasks in reality includes rhetorical awareness, stance and support, organisation, convention, and design for medium (Burnett et al., 2014). Finally, the cohesion and tension rubric looks at visual-linguistic cohesion and tension, and audio-visual cohesion and tension (Fajardo, 2018).

| | | |
|---------------------|---|-----------------------------|
| Theory-driven | Linguistic, visual, gestural, auditory, and spatial design | Hung et al. (2013) |
| Practice-based | Artefact, rhetorical skills, substance, process management & technical skills, and habits of mind | Eidman-Aadahl et al. (2013) |
| Flexible & Adoptive | Rhetorical awareness, stance & support, organisation, convention, and design for medium | Burnett et al. (2014) |
| Cohesion & Tension | Visual-linguistic cohesion & tension, audio-visual cohesion & tension | Fajardo (2018) |

Table 1: Summary of example assessment criteria in rubrics for multimodal task assessment

The summary above illustrates an expansion of assessment focus from language to non-language modes, task management capability, and ways of thinking in assessment for multimodal tasks. However, literature frequently reports that assessments for multimodal tasks in various contexts tend to pay greater attention to the language mode (e.g., Qu, 2017; Tan et al., 2020). One of the major causes of this is accountability pressure on teachers teaching 'a language' (Siegel, 2012). According to Murray et al. (2009), the lack of teacher knowledge, skills, and experiences relating to multimodal tasks also contributes to this tendency. Further, as assessing multimodal tasks takes more time than assessing traditional print-based tasks, teachers are likely to focus on language-related criteria to hasten the assessment process (Hafner & Ho, 2020). Ross et al. (2020) suggest it is vital to first equip teachers with *multimodal assessment literacy* to facilitate effective improvement of learner

multiliteracies.

Scholars and teachers have not fully ignored learner perspectives in assessment practice and its discussions for multimodal tasks. For instance, some studies focus on multimodal task self-assessment by learners using teacher-generated criteria, although studies on peer-assessment are scarce (Anderson & Krachorsky, 2019). Such studies examining self-assessment report that even for multimodal tasks, self-assessment assists learners to analyse the quality of their work and improve it; that is, benefit autonomy in their multiliteracies development (Hung et al., 2013; Newfield et al., 2003). However, learners often find assessment criteria descriptors in rubrics unclear and puzzling (Cox et al., 2010). Moreover, Godhe (2013) reports that learners rarely refer to such descriptors while engaging in multimodal tasks. These studies imply that teacher-generated and pre-established assessment criteria are inadequate, at least to a certain extent. Besides, learners may have limited readiness to take on board when it comes to self-assessment for multimodal tasks.

A few studies have focused on learner-generated assessment criteria for multimodal tasks. Adsanatham (2012) finds that such criteria end up being quite similar to those originally envisaged by the class teacher when all varieties are summarised into a comprehensive rubric. It is due to the influence of the existing and dominant view on language assessment and regular teacher comments on the learner-generated criteria. Consequently, students tend not to favour such a process and result. Silseth and Gilje (2019) point to the potential gap between what to value for teachers and learners in multimodal tasks. In summary, what counts as the value in multimodal tasks needs further exploration from learner perspectives.

Further, it is notable that multiliteracies scholars attentive to learner perspectives are becoming increasingly sceptical towards using rubrics for assessing multimodal tasks (Newfield et al., 2003; Silseth & Gilje, 2019; Wyatt-Smith & Kimber, 2009). Reed (2008) examines non-native pre-service English teachers working on a multimodal task to create and present a physical semiotic representation as their future students likely do. When their presentations are assessed using a rubric, they report the rubric “flattened” what they achieved in their multimodal products, shifting their “magical” work to “mundane” work (Reed, 2008, p. 36). Reed, in line with the other scholars, asserts that rubrics diminish the linkage between multiple modes in multimodal products and performances. Considering these cri-

tiques, further scrutiny of rubrics appears worthwhile.

Overall, the literature indicates that learner perspectives have shed new light on the assessment for multimodal tasks. However, it also calls for further studies taking learner perspectives into account to consider more viable assessments for multimodal tasks. Thus, the present study aims to consider assessment for multimodal composition, one of the most widely known multimodal tasks, from learner perspectives. The learner perspectives this study draws on are those of the author's students at an Italian university. Thus, the author uses an action research approach to discuss both his educational practice, focusing on multimodality, and his students' experiences. This study's guiding research questions are as follows: 1) how do the students view multimodal tasks? and 2) what do the students value when assessing their own multimodal compositions? The first question aims to touch upon the students' underlying assumptions regarding multimodal tasks, which may influence how they engage in and view the assessment of multimodal tasks (Grapin, 2019). The second question directly examines that which the students want to be valued in their multimodal tasks. Drawing on its findings, this study suggests more viable forms of assessment for multimodal tasks in language education.

3. Context of the Study

3. 1. The program and course

The current study takes an action research approach to examine the author's educational practice for a third-year Japanese language course in a Japanese Studies program at an Italian university. Students in the three-year program undertake six Japanese language courses and courses related to Japanese Studies, such as Japanese art, Japanese literature, history, film studies, and sociology. The Japanese language goal for the students is to pass at least the N3 level of Japanese Language Proficiency Test (JLPT) and reach at least B1 of the Common European Framework of Reference (CEFR) by the time they graduate. Upon entering their third year, the students major in either humanities or economics in Japanese Studies. The third-year Japanese course under discussion was for students majoring in humanities in the Japanese Studies program.

The fifteen-week course of about 100 students consisted of lectures, tutorials focusing on grammar and conversation, and free conversation classes. The author coordinated the

course but taught the lectures only. The course aimed to not only improve students' Japanese language knowledge and proficiency but also nurture analytical and critical judgment skills and autonomy as a key aspect of life-long learning. To address such aims, with the program's request to focus on writing and his own academic interest in art and language education, the author set up three themes for the lectures—criticality, creativity, and artistry—and decided on a 1,000-character multimodal composition in Japanese as an assignment for the lecture part of the examination. The genre of the multimodal composition was discussion. Students were encouraged to be aware of the multifaceted nature of their chosen topic and to analyse it from multiple angles. The author also asked the students to raise their creative awareness by drawing on personal experiences and thoughts in the multimodal composition. Another requirement was for the students to incorporate artistic expressions such as photographs and drawings into their multimodal compositions and *design* them together with the text, aiming to improve their artistry or multiliteracies (New London Group, 1996). In the weekly lecture, the author sought to assist students in choosing a topic, learning strategies for writing compositions, building Japanese expressions by using the short discussion essays related to the topics shown in Table 2, and knowing software use for their multimodal composition.

| | | | |
|---|---|----|---------------------------------------|
| 1 | Introduction: Multiliteracies and Multimodality | 9 | Art: Art Thinking |
| 2 | Conveying Messages/Assessment for Art | 10 | Discussion: Students Deciding a Topic |
| 3 | Technology: AI & Human | 11 | Communication: Typing Words |
| 4 | Sports: Art Sports | 12 | Gender: Equality or Equity |
| 5 | Education: Adult Recurrent Education | 13 | Economy: Diverse Places to Live |
| 6 | Food: Slow Food Movement | 14 | Education: Child Education |
| 7 | Explaining Multimodal Composition | 15 | Summary and Reflection |
| 8 | Environmental Issue: Banning Plastic Bags | | |

Table 2: A summary of topics and contents of the lectures of the course

3. 2. Teaching approach

Throughout the lecture series, the author highlighted multiliteracies, multimodality, their relevance to the multimodal composition, and copyright issues. This aimed to allow students to deepen their understanding of such concepts and actively, and then self-reflectively, engage in their multimodal composition. Generally, one lecture sequenced through 1) understanding the multimodal composition and learning key expressions for the essay of the week, 2) dictation or three-sentence composition, 3) reading the essay of the week, and 4) exchanging opinions regarding the essay.

In Lectures one and two, the author introduced multimodality and multiliteracies. Varieties of sample multimodal compositions, including the ones he designed, were shared in Lecture seven. In Lecture eight, the students designed and submitted a 500-character multimodal composition on the beginning, current, and future of their Japanese learning, as a preparatory practice for the 1,000-character composition for the examination. Using the selected samples of the student compositions and the essay about art thinking, the students further thought of how they would comprehend artworks and their multimodal compositions in Lecture nine. In Lecture 13 and thereafter, the author gave feedback on the students' 1,000-character multimodal compositions as many times as they wished until they submitted their final version for the examination. Approximately half of the students revised drafts one or two times, while one-quarter of the students did so three or four times, and the remaining quarter did so five or six times.

Importantly, this course was taught in Italy from September to December 2020, at the time of the early stages of the global coronavirus pandemic. The course commenced as a hybrid mode, with a limited number of students attending in person and the rest attending online. However, soon after the semester started, all classes moved to a fully online mode. At the end of the semester, we went back to the hybrid mode. This unpredictable and ever-changing circumstance made the student learning and the author's teaching preparation challenging. The students faced internet issues and suffered from severe stress throughout the semester. Consequently, the author did not know who and how the students would attend the lectures. Overall, the lectures were relatively teacher-centred while aiming to get as much student input as possible.

3. 3. Assessment design and practice

The author used the rubric composed of teacher-generated criteria to assess students' multimodal compositions; he held greater authority in the assessment. He chose this due to limited expectations of student contribution to developing assessment practice under the severe condition of the pandemic. At that time, he was also unaware of critiques on rubric use (e.g., Reed, 2008). He set up six criteria (see Table 3) worth five points each, totalling 30, in line with the university grading system (see Appendix for the rubric in full). The rubric, prepared in English, Italian, and Japanese, was presented throughout the course period to support assessment *for* learning despite being used at the end of the learning process in effect (William, 2011). The multimodal composition score and the scores for the other types of

examinations—for tutorials such as Kanji, grammar, dictation, short translation, and conversation tests—were averaged out to determine each student’s grade for the course.

| | |
|--------------------------------------|---|
| Criticality (5 points) | Demonstrating the ability to question given information, others’ ideas, and your own ideas |
| Creativity (5 points) | Demonstrating the ability to express your own ideas as well as support the ideas with your own experiences, including what you did, saw, heard, read, tasted as such |
| Artistry (5 points) | Demonstrating the ability to use artistic (multimodal) expressions together with the text information |
| Delivery (5 points) | Demonstrating the ability to: 1) create a good overall impression (design, readability, & contents) 2) organise a structure & enhance a flow and, 3) check and increase accuracy (vocabulary, Kanji, & grammar) |
| Newly learned expressions (5 points) | Demonstrating the ability to integrate newly learned Japanese expressions during the period of Lingua Giapponese 3.1 in 2020 into the text |
| Self-assessment (5 points) | Demonstrating the ability to analyse, evaluate and tell your own improvement (i.e., process) and performance (i.e., outcome) regarding contents, Japanese language, design, and/or ways of thinking, as well as the ability to provide reasons behind the self-assessment |

Table 3: Assessment criteria and brief descriptions for each criterion for the multimodal composition

Here are additional explanations about the assessment practice. Regarding criticality, the author looked into not only contents—how the students considered a chosen topic from multiple angles—but also critical uses of the Japanese language (i.e., the uses of hedging). For artistry, he decided not to assess the images themselves used in each composition as he was incapable of doing so. Instead, he paid close attention to the images’ relevance to the core message in the text, which he repetitively emphasised in his lectures. Additionally, he included the description ‘design’ in assessing an overall impression within the delivery criterion. In this way, the assessment considered the use of colour, typeface, space, and any other non-language modes of expression. For newly learned expressions covering lecture and tutorial content, the students added an expression list containing words, Kanji, or grammar and their translation or explanation after the body of the text. This showed that they understood the expressions they used in their compositions and helped their peers read them. The author shared the compositions of students who permitted him to with their peers (names and student numbers were hidden) once the student grade was finalised. Thus, he encouraged the students to remain reader-oriented when designing their compositions.

The students added their self-assessment score and justification written in Japanese, English or Italian after the expression list. The author’s primary intention in the self-assessment was to direct students’ focus on the designing process, which he could not observe. He also asked the students to interpret the self-assessment criterion freely and decide their focus, whether language, visual, design, ways of thinking, or a mixture of them, without much constraint. In this way, he aimed to make the self-assessment student-centred.

Overall, he attempted to balance the assessment for language aspects, non-language aspects, and ways of thinking aspects, although the language aspects were relatively underlined. Figure 1 shows one of the student multimodal compositions as an example.



Figure 1: A sample of student multimodal composition designed by Camilla (pseudonym)

4. Study Design

This study was conducted using an action research approach. In the field of education, teachers-as-researchers conduct action research to generate knowledge which directly impacts their educational practice, and impacts a wider audience through sharing research outcomes, often in the form of practical theories (Burns, 2009; Somekh, 2008). In this study, the author was a teacher-as-researcher. According to Somekh (2008), action research often involves a cyclic trial-and-error process, one with which this study did not comply; thus, this study is not action research. However, the author considered this study the beginning of his trial-and-error process, developing into action research when continued. Besides, this study has other characteristics of action research which are described as follows.

The value of action research lies in giving teachers voices (Burnes, 2009). However, some argue that the outcomes may be considered subjective, raising questions as to the trustworthiness of action research. Thus, scholars advocating action research often underline the importance of teachers-as-researchers' self-reflexivity and collaboration with others, including students, colleagues, and researchers (Somekh, 2008). The author collaborated with his students while exercising his self-reflexivity on his assumptions and influence over his student thinking and actions. In this way, action research becomes co-constructive. Using multiple emic views, teachers-as-researchers can attempt to provide a thick description regarding their teaching context, experiences, and voices (Burns, 2009), which the author aimed to achieve in this study. In so doing, action research often draws on types of data associated with qualitative research such as interview transcripts, journals, observation notes, open-ended questionnaires, and samples of student work. This study drew on student work, as is explained later. This does not mean that quantifying the data and presenting numbers as part of outcomes are excluded (Burns, 2008). Somekh (1993) calls it a "chameleon-like" practice (p. 29). This flexibility helped the author reach the goal of making sense of his teaching, his student learning, and the classroom and broader context, which in turn potentially guides his innovative action for positive change.

4. 1. Participants

This study involved 25 voluntarily participating students who had recently completed the third-year Japanese language course under discussion. These students were selected because this study drew on an action research approach, focusing on the author and his stu-

dents. They majored in Japanese Studies with a focus on humanities, meaning that they were mainly studying Japanese art, literature, history, and films in addition to their language courses. Seventeen were female and eight were male, aged between twenty and twenty-five. Although a few had mixed backgrounds, such as Slovenian-Italian or Turkish-Italian backgrounds, all were Italian. Their language backgrounds were not known in full; however, all used Italian in their studies and everyday lives as their first language. The students' Japanese language proficiencies varied widely, falling between N4 and N2 on JLPT or A2 and B1 on CEFR; that is, they could comprehend written and spoken Japanese regarding everyday topics and some familiar topics outside their daily routines. With some assistance, they could understand the news broadcasted on television and the contents of newspapers. Their writing and speaking levels were slightly surpassed by their reading and listening levels. This is typical for most foreign language learners whose target languages are not widely used outside their classrooms, resulting in limited output opportunities.

4. 2. Data collection and thematic analysis

The primary data of this study are the 25 multimodal compositions designed by the students, with a particular focus on the justifications provided by the students for their self-assessment scores. The compositions can inform how the students engaged in their multimodal compositions and show the outcomes of their engagements. The self-assessment descriptions can reveal how the students perceived their multimodal compositions and what they valued in the task and the end products. The author notes that these descriptions were unintentionally divided into two categories as a result of his unclear instruction to the students. The first category includes descriptions that follow the original instructions for the assessment. The descriptions in this category have their own structure and focus (see Figure 1 as an example). The second category includes descriptions that make references to the rubric criteria used by the author as a teacher, which differ from the author's original intention. All self-assessments in this category include scores for six criteria, averaged to determine the final self-assessment score, and justifications for the score in each criterion. Naturally, the descriptions in the second category mention artistry, although to what extent varies. This means that the student descriptions in the second category were led to mention multimodality. The author remained mindful about the distinction between these two types of descriptions.

The author employed a thematic analysis, as Burns (2009) recommends for studies taking an action research stance. Overall, the student multimodal compositions and self-assessment descriptions were constantly compared to evaluate the relevance of the descriptions to their compositions. However, the thematic analysis primarily focused on references to multimodality—“the use of several semiotic modes in the design of a semiotic product or event” (Kress & van Leeuwen, 2001, p. 20)—in the student self-assessment descriptions. The concepts presented in the literature review also guided the analysis process, such as assessment *of* and *for* learning, semiotic harmony, designing, and the strong and weak versions of multimodality. At the same time, the author remained attuned to the emergence of any other theme within the data set.

According to Nowell et al. (2017), thematic analysis is “a method for identifying, analysing, organising, describing, and reporting themes found within a data set” (p. 2). Drawing on this understanding, the author iteratively read the data in a printed form without using any software, made notes, and identified the data segment referring to multimodality and the themes expressing student perspectives. In so doing, the fact that many students did not mention the multimodal aspect of their composition became apparent. At this stage, the author decided to count the number of students referring to multimodality—quantifying the data. A close reading of the student descriptions mentioning the multimodal aspect showed that they viewed non-language modes as supplemental. Almost all brief descriptions referring to multimodality—especially the ones using the teacher’s marking criteria—fell into this category. These two were grouped as the first and most salient theme of this study: *subtle interest in the composition being multimodal*. Then, another iterative and close reading process revealed two themes in several student rich descriptions regarding their compositions being multimodal. These are relevant to their *readers* and *feelings*. This article presents three marked themes from this analysis. All participant names have been replaced with pseudonyms.

5. Findings

5. 1. Subtle interest in the composition being multimodal

The first and most striking theme is the subtlety of student references to the significance of multimodality in the self-assessment descriptions. Among 25 participating students, 15 mentioned multimodality; another six mentioned multimodality without using the teacher-

generated criteria. That is, nearly three-quarters of the student participants either placed little emphasis on multimodality (10 students) or was led to write about it (nine students). For instance, Antonella wrote nothing about multimodality in her self-assessment description. Motivated by a Japanese animation movie about deaf people, her learning experience of sign language, and her encounter with deaf people, her composition discusses the importance of considering deaf people’s hardship, learning sign language, and introducing sign language teaching and learning in Italian schools. Her balanced use of the image, colour, and space put together with the text resulted in a high-quality and, more importantly, personally satisfying and prideful multimodal composition (see Figure 2).



Figure 2: A multimodal composition designed by Antonella

The close reading of her self-assessment description reveals that her focus was exclusively on her language and how her personal experience with sign language and deaf people led to choosing the topic and making the discussions. Besides, she concluded her description by pointing out the insufficiency of her Japanese proficiency to describe the movie—“*such an incredible piece of art*”³—and desire to improve her Japanese to be capable of doing so in the future. An underlying assumption in this description is the idea that words, if proficient enough, can describe the movie fully—the weak version of multimodality (Grapin, 2019):

³ The quotes of the student writing are original. However, the author modified them to enhance the readers’ understanding of the text information to the extent that the modification does not change the original meaning. When some information is added, the author inserted it using square brackets.

I am very proud of how this report turned out. I feel like I'm getting a better grasp of the language, and this report had a huge role in understanding how to express certain thoughts and how to use specific grammar forms. The topic I chose to talk about, sign language, is a complicated matter. After experiencing first-hand what some of the most frequent hardships that deaf and mute people go through [in] daily life, I decided to write about why it would be a good idea to teach sign language in schools, and why some people oppose it. I hope my delivery was clear enough [for the readers] to understand why the topic is so important and cannot be ignored anymore by society and the government. I also did not focus too much on the movie, 'A Silent Voice', mostly because in my opinion, it is such an incredible piece of art that it cannot properly be described in words, especially not with my level of Japanese. I can only hope I will further improve and be able to express my thoughts and feelings better in the future. (Antonella)

Further, even the students who mentioned multimodality mostly showed their subtle interest in it, whether they were led to mention it or not. For instance, they wrote descriptions such as *"the fact that I chose to talk about an existing anime, with its own character design books and episode stills, gave me material[s] to visually support my writing"* (Aisha), *"the artistic expressions contributed to build[ing] the overall image giving the report a good final touch"* (Francesca), and *"the pictures that I chose from the movie 'Spirited Away' can help to understand better the scene I mentioned [in the text]"* (Elisa). The terms such as *"support"*, *"a good final touch"*, and *"help"* describing the role of non-language modes hint that such modes play a secondary role, and their compositions being multimodal does not mean considerable significance. However, all student participants mentioning multimodality did not think in a similar way. Federica below was one of the very few exceptions.

5. 2. Reader involvement

Federica, who used the teacher-generated criteria in her self-assessment, wrote a lengthy description on artistry, highlighting a notable theme: *involving readers*. She designed a multimodal composition discussing the relationship between human beings, animals, and foods (see Figure 3). She pointed out how human beings see animals differently depending on contexts, whether the same animals are pets, wildlife, or livestock. However, her stance was to protect animal lives and dignity from beginning to end no matter where we are. To do so, she discussed the importance of human beings' critical engagement with media, which often controls their thoughts and actions without being noticed.

5. 3. Expressing feelings multimodally

When turning our attention to the students writing their self-assessment descriptions relatively freely and mentioning multimodality without being led, one noteworthy theme emerged; *the multimodal expression of feelings*. Camilla, who designed the multimodal composition as shown in Figure 1, wrote her self-assessment description further elucidating the theme. Her composition discusses how city pop, a Japanese music genre, helped her maintain her mental health under the severe living and study conditions brought about by the coronavirus pandemic. As shown in the excerpt below, she developed her self-assessment description while centring her attention on her feelings; for example, “*I used [Japanese language] as a mean[s] to express what I feel*” and “*I tried to transmit what I feel*”. Her reference to multimodality appears at the end when her focus shifts to non-language modes of her composition. She wrote that she wanted readers to perceive the 80s’ city pop soft and bright vibes that she perceives using the colour, neon light style, urban-style picture, and album sleeves as parts of the design:

Writing this report made me feel more confident with the Japanese language since I used it as a mean[s] to express what I feel... Using correctly new expressions [was] necessary to allow a better comprehension of the report by the reader, to whom I tried to transmit what I feel when I listen to city pop. With this aim in [my] mind, I tried my best to use words and expressions in Japanese carefully ... Finally, through this pink and white design, the neon light style, the urban-style picture and the albums’ cover’s images, I wanted the reader to perceive what I perceive when it comes to city pop: 80s’ soft and bright vibes. (Camilla)

Despite their appearance at the end of the description, it is difficult to justify that Camilla viewed non-language modes as less privileged than a language mode in her composition. Her description indicates that she values how her feelings are expressed through Japanese language, colours, style, and pictures without separating them or prioritising any. That is, she achieved semiotic harmony (Kress, 2009). Next, what the themes presented so far mean to our consideration of and engagement in more viable assessment for multimodal tasks will be explored in the discussions and implications below.

6. Discussions and Implications

6. 1. Raising awareness towards multimodality in language courses

First and foremost, the findings have suggested the need to raise teacher and learner awareness towards recognising multimodality as an object for assessment in a language course to make the assessment more viable. In so doing, the reasons why the student participants in this study had limited awareness towards multimodality call for possible explanations. The influence of the marking criteria that the author set up and explained is hardly disputable. Matter-of-factly, a larger proportion of focus was placed on language and elements related to language in the rubric. This aligns with previous studies. The way teachers conceptualise multimodality and explain it to their students shape how their students conceptualise and engage in multimodality (Adsanatham, 2012). This indicates an urgent need to raise teacher awareness towards the strong version of multimodality (Grapin, 2019) and enhance their multimodal assessment literacy (Ross et al., 2020) for assessment for multimodal tasks to take on the meaning.

Another possible reason may relate to contextuality; that is, a language course in the Japanese Studies program. The students were likely to understand the distinction between the language courses and content courses soon after commencing the program, considering the set structure. This idea should have become more concrete as they progressed to their second and third years. Indeed, they studied artistry focusing on non-language aspects of paintings, calligraphy, poetry, and movies as such in the Japanese art and film studies courses. This might mean that the students in this study had greater readiness to engage in multimodality than students focusing on language learning only. Nonetheless, thinking of, utilising, and developing an eye for non-language modes of expressions were quickly forgotten when they came to one of the language classes. It is understandable that learning *a language* is the priority in language classes (Siegel, 2012). However, language teaching needs to address that what counts as language is fast-changing in the current society and support the learners to become aware of it.

These altogether led three-quarters of the student participants to address their multimodal compositions with the weak version of multimodality in mind (Grapin, 2019). It is no surprise that Antonella, in this study, claimed that she can describe a movie and express her *“thoughts and feelings”* only in words if her proficiency improves. The author had believed

that his emphasis on the value and importance of understanding and using multimodality, as the New London Group (1996) argues, was more than sufficient. However, it turned out to be insufficient. Shedding light on multimodality and its assessment in language courses may require far more overt and emphasised teacher actions upon students than currently imagined.

6. 2. Peer-assessment as indispensable

The reader involvement, which Federica valued, calls for our willingness to integrate peer-assessment to make assessment for multimodal tasks more viable. She suggests that readers are indispensable for her work to be complete. The most easily reached readers in the process and at the end of designing are peers in most educational contexts. Learners focusing on their own tasks may not reach out to one another voluntarily. The author senses that the effectiveness of his attempt to prompt students' awareness of peers as potential readers might have been limited. That is, peer-assessment could have been an encouraging device for paying attention to peers' work. To date, peer-assessment has been almost absent in the relevant studies (Anderson & Krachorsky, 2019). Changing this landscape can help learners remain reader-oriented, give feedback to each other, and develop end-products involving audiences more. This aligns with assessment *for* learning in the sense that the assessment integrates more varieties of perspectives and becomes mutually constitutive (Usami, 2016; William, 2011). Indeed, such peer-assessments—analysing, evaluating, and commenting on various multimodal designs—should contribute to promoting learner multimodal assessment literacy (Ross et al., 2020). This, in turn, likely enhances their multiliteracies development (Alderson & Wall, 1992; McNamara, 2001). Considering these, the author argues that peer-assessment can enrich assessment for multimodal tasks overall and, further, create an optimal learning cycle for multiliteracies development.

6. 3. Integrating learner-generated criteria

The expression of feelings raised by Camilla can be understood as a signal to invite learner-generated criteria in assessment for multimodal tasks more. Kress (2009) notes that non-language modes closely relate to the feelings or emotions of multimodality designers. Those advocating multiliteracies often underscore the significance of designer *voices* enhanced in multimodality (e.g., Nelson, 2008). The voices may not be enhanced as envisaged only with a language mode that is good at, for instance, rational and logical explanations of causes and effects. Feelings and emotions are concepts that pictures, drawings, sounds,

movies, colours, and, spaces are good at expressing (Nelson, 2008). Considering these features of multimodality, Camilla directing her attention to expressing her feelings multimodally seems natural. Nonetheless, discussions led by scholars and teachers on assessment for multimodal tasks have seldom touched upon considering how feelings are expressed multimodally, other than Anderson and Kachorsky (2019) referring to “affect” in their review (p. 328). This point illustrates that referring to learner perspectives and welcoming learner-generated criteria are critically important to implement quality assessment for multimodal tasks.

As Davies (2013) suggests, an important question is now posed: should we assess such items as ‘feelings’? The answer can be yes for now, although further discussions are indeed needed. The expression of feelings is what learners, like Camilla, may find valuable, whether in the designing process or in the designed product. Acknowledging and valuing it in assessment practice are likely to encourage learners to engage in multimodality actively and develop multiliteracies, unlike the case of Adsanatham (2012) and Godhe (2013). In so doing, setting up one definite model drawing on the assumption that a teacher can objectively assess the criterion under discussion, or any other learner-generated criteria, may not be the best approach, as hinted by Cox et al. (2010), Godhe (2013), and Reed (2008). That is, learner self-assessment in which fluid assessor subjectivities are not considered dismissible appears more workable for learner-generated criteria in assessment for multimodal tasks (William, 2011). There might be many other learner-generated criteria for assessing multimodal tasks. Beginning to use them is likely to create a gap between the criteria to which a teacher and a learner refer, as Silseth and Gilje (2019) expect. This gap can be seen as rather positive, given that diversifying perspectives in an assessment is the way to go considering the assessment *for* learning trend in the field.

6. 4. Diversity in what counts and moving beyond rubrics

The importance of ensuring diversity in ‘what’ counts comes under the spotlight when teacher-assessment, self-assessment, and peer-assessment are all involved in assessment for multimodal tasks. Setting up shared criteria for all in the same cohort has not faded away, even after the emergence and proliferation of assessment *for* learning (Adsanatham, 2012; Hafner & Ho, 2020; Tan et al., 2020). This way of thinking might be necessary considering the significance of fairness in the current education system. However, such practice appears built on the premise that all interpret the descriptors of the same rubric and then

apply the interpretation to assess performance, achievement, or change in the exact same way. The author questions the premise as suggested in previous studies (Cox et al., 2010; Godhe, 2013). Thus, the time has come to move on.

We can begin with setting up unique criteria for teacher-assessment, self-assessment, and peer-assessment each to have diversity in what counts in a class or a course. The author suggests this way as, realistically, it is debatable whether allowing each individual to have unique criteria is manageable, especially in a course with a teacher and 100 students or more. Considering Adsanatham's case (2012), it appears ideal that learner-generated criteria constitute self-assessment and peer-assessment. Such assessment criteria need to be discussed and established as learning progresses. This is nothing negative as it resonates with the processual nature embedded in multimodality (Kress, 2009). Yet, past studies indicate that such attempts might face challenges when it comes to assessment for multimodal tasks (Adsanatham, 2012), compared to educational practice not focusing on multimodality (e.g., Thomson, 2008). This work becomes further challenging in the current education system, which often requires teachers to complete assessment design well before the learning process begins. However, the author argues that these do not mean our efforts result in failure. In his attempt, Camilla and Federica demonstrated that they could end up writing thick descriptions of themselves as designers creating their social futures (Cope & Kalantzis, 2009) and their multimodal compositions. It makes us consider that taking a little more time on discussing assessment for multimodal tasks *within practice* in classrooms can guide us in the desired direction.

Before closing this article, the author reflects on the use of rubrics in assessment for multimodal tasks. The essential value of multimodality lies in multimodal ensemble, orchestration, and resultant semiotic harmony (Kress, 2009). Keeping such wholeness of multimodality in mind is indispensable to understand its real value and design effective and captivating multimodal products. Reflecting upon his grading, the author recalls assessing each composition by referring to each criterion in sequence as shown in the rubric. He determined the score for each criterion and calculated the overall grade by adding the scores. The sub-criterion of 'overall impression' does exist in the rubric. However, considering the overall impression almost always came at the end of the grading of each composition as "*a good final touch*", as with Francesca. In the process of assigning a grade, the wholeness of the student multimodal compositions of which he is not hesitant to feel genuinely proud

was broken into pieces. The students constantly referring to the rubric might have been led to see their ‘magical’ work as ‘mundane’ work, as in Reed’s study (2008), although it was not clearly evident in the data set. Using rubrics can be beneficial in many ways, as Reed (2008) also admits. The author asserts that they are workable for assessing the development and achievement regarding task management skills, technical skills, and similar skills with these or anything quantifiable such as the newly learned expressions in this study. However, whether a rubric is the best tool for assessing multimodal designs is called into question again. Such scepticism may not disappear in our best efforts to capture the wholeness, as with Fajardo’s (2018) primary attention on cohesiveness, if rubrics remain the default. The author suggests embarking on an ambitious project to design an entirely new assessment technique and tool for multimodal products. Such assessment needs to take the wholeness into consideration at its best while being assessor-friendly to be feasible in the current educational system.

7. Conclusion

This article has considered assessment for multimodal tasks from learner perspectives taking multimodality and multiliteracies as a starting point. One conclusion the author has drawn is that assessments for multimodal tasks still face challenges and need further discussions. Both teachers and learners in language education need more awareness, skills, knowledge, and experience regarding multimodality and multiliteracies, as Murray et al. (2010) note. The author underlines that we must work more on raising awareness. In the current digitalised society, our communication is increasingly multimodal, although this is not often noticed. With our *multimodal mindset*, we can actively and reflectively engage in conversations about assessment to nurture our multimodal assessment literacy (Ross et al., 2020). Considering the shift towards assessment *for* learning, which values learner-centred assessment, learners cannot be left behind. In this article, the author has illustrated that students have a lot to offer for such conversations. They already have an eye for multimodality or wake it up when prompted. We can draw on learner perspectives more to further the purposeful discussions on assessment for multimodal tasks. The current study is small-scale, touching upon an educational practice at an Italian university. Its strength lies in carefully unpacking learner perspectives in detail, but it is one of the cases. More studies casting light on learner perspectives will be fruitful to design and implement viable assessment for multimodal tasks.

Acknowledgements

The author would like to express his gratitude to Professor Chihiro Kinoshita Thomson for her valuable guidance throughout the writing process of this article as well as Dr Yun-jing Zhen, Ms Emi Yoshida, Ms Bo Dong, and Associate Professor Todd James Allen for providing constructive comments on the draft of this article. He also appreciates the students participating in this study.

References

- Adsanatham, C. (2012). Integrating assessment and instruction: Using student-generated grading criteria to evaluate multimodal digital projects. *Computers and Composition*, 29(2), 152-174.
- Alderson, J. C., & Banerjee, J. (2001). State of the art review: Language testing and assessment (Part I). *Language Teaching*, 34(4), 213-236. doi:10.1017/S0261444800014464
- Alderson, J. C., & Wall, D. (1993). Does washback exist? *Applied Linguistics*, 14(2), 115-129. <https://doi.org/10.1093/applin/14.2.115>
- Anderson, K. T., & Kachorsky, D. (2019). Assessing students' multimodal compositions: An analysis of the literature. *English Teaching: Practice & Critique*, 18(3), 312-334. <https://doi.org/10.1108/ETPC-11-2018-0092>
- Bezemer, J., & Kress, G. (2008). Writing in multimodal texts: A social-semiotic account of designs for learning. *Written Communication*, 25(2), 166-195. <https://doi.org/10.1177/0741088307313177>
- Burnett, R. E., Frazee, A., Hanggi, K., & Madden, A. (2014). A programmatic ecology of assessment: Using a common rubric to evaluate multimodal processes and artifacts. *Computers and Composition*, 31, 53-66. <https://doi.org/10.1016/j.compcom.2013.12.005>
- Burns, A. (2009). Action research. In J. Heigham & R. A. Croker (Eds.), *Qualitative Research in Applied Linguistics* (pp. 112-134). Palgrave Macmillan.
- Cope, B., & Kalantzis, M. (2009). "Multiliteracies": New literacies, new learning. *Pedagogies: An International Journal*, 4(3), 164-195. <https://doi.org/10.1080/15544800903076044>
- Cox, A. M., Vasconcelos, A. C., & Holdridge, P. (2010). Diversifying assessment through multimedia creation in a non-technical module: Reflections on the MAIK Project. *Assessment & Evaluation in Higher Education*, 35(7), 831-846. <https://doi.org/10.1080/02602930903125249>
- Davies, A. (2013). Fifty years of language assessment. In A. J. Kunnan (Ed.), *The Companion to Language Assessment* (pp. 1-19). Wiley-Blackwell. <https://doi.org/10.1002/9781118411360.wbcla127>
- Eidman-Aadahl, E., Blair, K., DeVoss, D. N., Hochman, D., Jimerson, L., Jurich, C., Murphy, S., Rupert, B., Whithaus, C., & Wood, J. (2013). Developing domains for multimodal writing assessment: The language of evaluation, the language of instruction. In H. A. McKee & D. N. DeVoss (Eds.), *Digital Writing Assessment and Evaluation (Chapter 7)*. Computers and Composition Digital Press/Utah State University Press. Retrieved October 18, 2021, from https://ccdigitalpress.org/book/dwae/07_nwp.html
- Fajardo, M. F. (2018). Cohesion and tension in tertiary students' digital compositions: Implications for teaching and assessment of multimodal compositions. In H. De Silva Joyce & S. Feez (Eds.), *Multimodality across Classrooms: Learning about and through Different Modalities* (pp. 178-193). Routledge.
- Godhe, A. (2013). Negotiating assessment criteria for multimodal texts. *The International Journal of Assessment and Evaluation*, 19(3), 31-43. doi:10.18848/2327-7920/CGP/v19i03/48329.
- Grapin, S. (2019). Multimodality in the new content standards era: Implications for English learners. *TESOL Quarterly*, 53(1), 30-55. <https://doi.org/10.1002/tesq.443>

- Hafner, C. A., & Ho, W. Y. J. (2020). Assessing digital multimodal composing in second language writing: Towards a process-based model. *Journal of Second Language Writing*, 47, 100710. <https://doi.org/10.1016/j.jslw.2020.100710>
- Hung, H.-T., Chiu, Y.-C. J., & Yeh, H.-C. (2013). Multimodal assessment of and for learning: A theory-driven design rubric. *British Journal of Educational Technology*, 44(3), 400-409. doi:10.1111/j.1467-8535.2012.01337.
- Illich, I. (1973). *Tools for Conviviality*. Harper and Row Publisher.
- Jewitt, C. (2008). Multimodality and literacy in school classrooms. *Review of Research in Education*, 32(1), 241-267. <https://doi.org/10.3102/0091732X07310586>
- Kress, G. (2009). *Multimodality: A Social Semiotic Approach to Contemporary Communication*. Routledge.
- Kress, G., & van Leeuwen, T. (2001). *Multimodal Discourse: The Modes and Media of Contemporary Communication*. Edward Arnold.
- Levy, M., & Kimber, K. (2009). Developing an approach for comparing students' multimodal text creations: A case study. *Australasian Journal of Educational Technology*, 25(4), 489-508. <https://doi.org/10.14742/ajet.1125>
- McNamara, T. (2001). Language assessment as social practice: Challenges for research. *Language Testing*, 18(4), 333-349. <https://doi.org/10.1177/026553220101800402>
- Murray, E. A., Sheets, H. A., & Williams, N. A. (2009). *The New Work of Assessment: Evaluating Multimodal Compositions*. Retrieved October 18, 2021, from http://conlinejournal.org/murray_etal/index.html
- Nelson, M. E. (2008). Multimodal synthesis and the voice of the multimedia author in a Japanese EFL context. *Innovation in Language Learning and Teaching*, 2(1), 65-82. DOI: 10.1080/17501220802158826
- New London Group. (1996). A pedagogy of multiliteracies: Designing social futures. *Harvard Educational Review*, 66(1), 60-93. <https://doi.org/10.17763/haer.66.1.17370n67v22j160u>
- Newfield, D., Andrew, D., Stein, P., & Maungedzo, R. (2003). No number can describe how good it was: Assessment issues in the multimodal classroom. *Assessment in Education: Principles, Policy and Practice*, 10(1), 61-81. <https://doi.org/10.1080/09695940301695>
- Nowell, L. S., Norris, J. M., White, D. E., & Moules, N. J. (2017). Thematic analysis: Striving to meet the trustworthiness criteria. *International Journal of Qualitative Methods*, 16(1), 1-13. <https://doi.org/10.1177/1609406917733847>
- Peirce, B. N. (1995). Social identity, investment, and language learning. *TESOL Quarterly*, 29(1), 9-31. <https://doi.org/10.2307/3587803>
- Qu, W. (2017). For L2 writers, it is always the problem of the language. *Journal of Second Language Writing*, 38, 92-93. <http://dx.doi.org/10.1016/j.jslw.2017.10.007>
- Reed, Y. (2008). No rubric can describe the magic: Multimodal designs and assessment challenges in a postgraduate course for English teachers. *English Teaching: Practice and Critique*, 7(3), 26-41.
- Ross, J., Curwood, J. S., & Bell, A. (2020). A multimodal assessment framework for higher education. *E-Learning and Digital Media*, 17(4), 290-306. <https://doi.org/10.1177/2042753020927201>
- Sato, S., & Kumagai, Y. (Eds.) (2011). *Asesumento to Nihongokyōiku: Atarashi Hyōka no Riron to Jissen [Assessment and Japanese Language Education: Theories and Practices of New Assessment]*. Kuroshio Publication.
- Shimada, M. (2020). Hyōka [Assessment]. In S. Endo (Ed.), *Shin Nihongokyōiku o Manabu – Naze, Nanio, Dō Oshieruka [New Learning of Japanese Language Education – Why, What, and How to Teach]* (pp. 163-185). Sanshūsha.
- Siegel, M. (2012). New times for multimodality? Confronting the accountability culture. *Journal of Adolescent and Adult Literacy*, 55, 671-81. <https://doi.org/10.1002/JAAL.00082>
- Silseth, K., & Gilje, Ø. (2019). Multimodal composition and assessment: A sociocultural perspective. As-

essment in Education: Principles, Policy & Practice, 26(1), 26-42. DOI: 10.1080/0969594 X.2017.1297292

Somekh, B. (1993). Quality in educational research: The contribution of classroom teachers. In J. Edge & K. Richards (Eds.), *Teachers Develop Teachers' Research* (pp. 26-38). Heinemann.

Somekh, B. (2008). Action research. In L. M. Given (Ed.), *The SAGE Encyclopedia of Qualitative Research Methods* (pp. 4-7). SAGE Publications.

Tan, L., Zammit, K., D'warte, J., & Gearside, A. (2020). Assessing multimodal literacies in practice: A critical review of its implementations in educational settings. *Language and Education, 34*(2), 97-114. <https://doi.org/10.1080/09500782.2019.1708926>

Thomson, C. K. (2008). Kaigai no Nihongokyōiku no Genba ni okeru Hyōka – Jikohyōka no Katsuyō to Gakushūsha Shudōgata Hyōka no Tēan [Assessment in the context of Japanese language education outside Japan – Proposing utilising self-assessment and learner-led assessment]. *Nihongokyōiku, 136*, 27-37.

Usami, Y. (Ed.) (2016). *Hyōka o motte Machi ni deyo – Oshieta Koto/Manada Koto no Hyōka toiu Hassō o koete [Bringing Assessment in a City: Moving beyond the Idea of Assessing What is Taught and Learnt]*. Kuroshio Publication.

Wiliam, D. (2011). What is assessment for learning? *Studies in Educational Evaluation, 37*(1), 3-14. <https://doi.org/10.1016/j.stueduc.2011.03.001>

Wyatt-Smith, C., & Kimber, K. (2005). Valuing and evaluating student-generated online multimodal texts: Rethinking what counts. *English in Education, 39*(2), 22-43. <https://doi.org/10.1111/j.1754-8845.2005.tb00615.x>

Appendix

Appendix: Rubric Used in the Course under the Study

| English Version (Official and Used for Marking) | | | | | | | | | | | | | | | | | | |
|---|--|--|--|--|--|--|---|--|---|--|---|--|---|--|------|--|-----|--|
| Lingua Giapponese 3.1 Classe 1 Final Report Marking Criteria | | | | | | | | | | | | | | | | | | |
| | 0.5 | | 1 | | Name | | 2.5 | | 3 | | 3.5 | | 4 | | Mark | | /30 | |
| | | | | | | | | | | | | | | | | | | |
| Criticality (5) Demonstrating the ability to question given information, others' ideas, and your own ideas | ○ Almost no expression showing author's criticality | | ○ Inappropriate uses of critical expressions, even when used | | ○ A limited number and range of critical expressions (1 to 2) ○ Not necessarily appropriate uses of critical expressions | | ○ A limited number and range of critical expressions (2 to 3) ○ Appropriate uses of critical expressions | | ○ Some number and range of critical expressions (3 to 4) ○ Sufficient degree of appropriate uses of critical expressions | | ○ A very good number and range of critical expressions (3 to 4 or more) ○ Appropriate uses of critical expressions | | | | | | | |
| Creativity (5) Demonstrating the ability to express your own ideas as well as support the ideas with your own experiences (including what you did, saw, heard, read (e.g. novels or mangas), tasted as such) | ○ Almost no description about author's ideas | | ○ Almost no description about author's experiences to support the ideas | | ○ Limited descriptions about author's ideas ○ Almost no description about author's experiences supporting the ideas | | ○ Some descriptions about author's ideas ○ Limited descriptions about author's experiences supporting the ideas | | ○ A good number of descriptions about author's ideas ○ Some descriptions about author's experiences supporting the ideas | | ○ Descriptions about author's ideas are central to and spread across the text ○ Ample and appropriate descriptions about author's experiences supporting the ideas | | | | | | | |
| Artistry (5) Demonstrating the ability to use artistic (multimodal) expressions together with the text information | ○ Almost no artistic expression | | ○ The uses of artistic expressions irrelevant to the text, even when used | | ○ Limited uses of artistic expressions ○ The uses of artistic expressions not necessarily relevant to the text | | ○ Moderate uses of artistic expressions ○ The uses of artistic expressions partially relevant to the text | | ○ Active uses of artistic expressions ○ The uses of artistic expressions closely relevant to the text | | ○ Exceptional uses of artistic expressions ○ The uses of artistic expressions closely relevant to the key points of the text | | | | | | | |
| Delivery (5) Demonstrating the ability to: 1) Create a good overall impression (design, readability, & contents) 2) organise a structure & enhance a flow and 3) check and increase accuracy (vocabulary, Kanji, & grammar) | ○ Poor overall impression ○ Poor structure and flow ○ Too many mistakes making the text incomprehensible | | ○ Acceptable overall impression ○ Acceptable yet confusing structure and flow ○ Many mistakes and ones hindering the reader's understanding are the majority | | ○ Good overall impression ○ Acceptable and understandable structure and flow ○ Some mistakes but the ones hindering the reader's understanding are limited | | ○ Great overall impression ○ Well-organised structure and smooth flow ○ A few mistakes and most of them hinder the reader's understanding | | ○ Exceptional overall impression ○ Exceptionally clear structure and seamless flow ○ A few mistakes but they do not hinder the reader's understanding | | | | | | | | | |
| Newly Learnt Expressions (5) Demonstrating the ability to integrate newly learnt Japanese expressions during the period of Lingua Giapponese 3.1 in 2020 into the text | ○ Almost no use of newly learnt expressions (1 to 2) | | ○ Limited uses of newly learnt expressions (3 to 4) | | ○ Sufficient uses of newly learnt expressions (5 to 6) | | ○ Good uses of newly learnt expressions (7 to 8) | | ○ Exceptional uses of newly learnt expressions (9 to 10) | | | | | | | | | |
| Self-Evaluation (5) Demonstrating the ability to analyse, evaluate and tell your own improvement (i.e. process) and performance (i.e. outcome) regarding contents, Japanese language, design, and/or ways of thinking, as well as the ability to provide reasons behind the self-evaluation | ○ Almost no effort ○ Almost no improvement ○ Dissatisfactory outcome | | ○ A little effort ○ A little improvement ○ A little satisfactory outcome | | ○ Sufficient effort ○ Sufficient improvement ○ Acceptable outcome | | ○ Good effort ○ Good improvement ○ Satisfactory outcome | | ○ Great effort ○ Great improvement ○ Great outcome | | | | | | | | | |