

Asynchronous Peer Teaching Using Student-Created Multimodal Materials

John Blake

Abstract—This paper reports on asynchronous peer teaching in which learners create multimodal explanations or tutorials for future cohorts. Japanese learners of English work individually or in teams to produce video and audio explanations. Multimodal explanations that meet the quality requirements are uploaded to the respective course websites housed on the university server. The aim is that student audio-visual developers learn during the creation process and student users learn from the multimodal resources developed. Each year a new cohort of students makes a new set of explanations. The mean quality of the multimodal explanations increases annually as the less useful or less popular video and audio files are replaced. This creates a continuous cycle of incremental improvement.

Index Terms—Digital artefacts, incremental improvement, materials development, multimodality, peer teaching, video production.

I. INTRODUCTION

This paper reports on asynchronous peer teaching in which Japanese learners of English create multimodal explanations or tutorials for future cohorts of learners. This approach was adopted in two elective courses that teach English using content and language integrated learning.

Online courses are an ideal medium for developing the receptive skills of listening and reading, since learners can progress through the course materials at their own pace, creating their own learning path through the course content. However, developing the productive skills of speaking and writing online is more challenging.

Typically, in language classes speaking practice involves learners working in pairs or small groups. Although presentations and monologues can be usefully practiced without an audience, conversation and discussion practice alone is rather artificial. This adds logistic and security dimensions to web-based courses in which learners in different locations work together virtually. Writing practice tends to be undertaken individually with teachers providing advice before, during and after the drafting and editing phases. Online written communication can be both synchronous and asynchronous.

Providing short audio or video explanations for learners to access on demand is one way to deliver asynchronous course content. Multimodal content creation is, however, time

intensive. The effectiveness of an approach can only be judged retrospectively and so it is possible for a teacher to dedicate hundreds of hours to developing content only to discover that the learners do not access it.

One potential solution is to get learners to create content for themselves. Rather than one teacher producing materials for many learners, the learners create materials for themselves, and so the workload can be more evenly distributed. This not only promotes a learner-centered approach but ensures that the content produced is viewed by other learners. Allowing learners to develop teaching materials provides them with a meaningful purpose, authentic audience and the opportunity to teach their peers.

This project aims to analyze the practice and pedagogy of learner-created material production. Ways to overcome practical challenges in the production of learner-created multimodal digital materials are also detailed.

This paper is organized as follows. Section II describes the educational context in which this research is situated. Section III provides an overview of the pertinent literature on student-created multimodal materials. Section IV describes the course framework, content specification and content creators. Section V discusses the three phases of creating multimodal materials, namely pre-production, production and post-production. Section VI introduces the online platform. Section VII evaluates the use of asynchronous peer teaching using student-created multimodal materials. A triangulated approach was adopted in which views of the learners, peers and tutor were collected. The final section (VIII) summarizes the lessons learnt and provides an actionable list of suggestions for teachers and materials developers contemplating adopting a similar approach.

II. BACKGROUND

This study was conducted at a public university dedicated to computer science and engineering, located in Fukushima prefecture, Japan. This is a relatively small, well-resourced university with a student population of around 1200. The faculty comprises approximately 40% non-Japanese. All undergraduates must write a graduation thesis and give a presentation in English to fulfill graduation requirements.

Class sizes are between 30 and 50 students. In most classes, there is a general reticence for most students to speak in English [1], [2]. The reasons for the lack of willingness to communicate vary among the students but include low levels of confidence and English language proficiency. It takes a high degree of confidence to speak in front of a class of peers. This is a barrier that many have not managed to overcome. The focus on grammatical accuracy rather than

Manuscript received September 19, 2020; revised January 12, 2021. This work was supported in part by a JSPS *Kakenhi* Grant-in-aid for Scientific Research (C) entitled "Feature visualizer and detector for scientific texts", Grant Number 19K00850.

John Blake is with the University of Aizu, Aizuwakamatsu, Japan (e-mail: jblake@u-aizu.ac.jp).

communicative competence throughout junior high and high school has made students acutely aware of their shortcomings in terms of accuracy, which in turn affects fluency. Silent students make no mistakes, and so that is an option that many select. In the Japanese education system schoolteachers tend to teach lockstep with the teacher speaking at the front of the class and students sitting quietly at individual desks in rows [3]. The result is that students are not used to active participation.

University students, like many people, live multimodal lifestyles, accessing their smartphones multiple times each day (or hour) to check their preferred social networking sites. A substantial share of their online surfing is spent on video-sharing social networking service (SNS) platforms, such as YouTube, Facebook and Instagram.

With the high quality of the video output of modern smartphones and the ease of editing video online, video creation is no longer the onerous task that it was in the 1990s. At that time, students had to loan and carry heavy video cameras and video editing was conducted off-line using specialist software. Nowadays many students use the video functionality of their smartphones and many have experimented with basic video editing in their leisure time. In each class, there are some students with extensive experience of taking photographs and making videos. This is unsurprising as YouTuber or influencer rank highly in the target jobs that Japanese school students aspire to.

The covid-19 pandemic has forced teachers worldwide to deliver tuition online via learning management systems (e.g. Moodle), videoconferencing platforms (e.g. Zoom) or other systems (e.g. websites and SNS platforms). This transition made it necessary for many teachers to move from paper-based, face-to-face tuition to online digital delivery of multimedia content.

The term *multimedia* is more frequently used than *multimodal* by the public. However, according to Lauer [4], academics prefer the term *multimodal* because of its emphasis on design and process. Multimodal texts are ubiquitous online [5]. Apkon [6] describes the popularity of sharing videos online as the “global visual conversation” (p.24), and “the primary mode of communication...that transcends languages, cultures and borders” (p.24).

III. RELATED WORKS

Video has been harnessed in mainstream teaching and language teaching for many years. Willis [7] describes the four key roles for video in the classroom as 1) language focus, 2) skills practice, 3) stimulus, and 4) resource. In the 1990s, teachers began adopting an active viewing approach using various techniques to engage learners in activities, such as predicting, identifying and recreating [8]. Digital video has been shown to enhance the emotional investment of learner videographers, which in turn, positively affects the cognitive and affective processes [9]. An extensive literature search failed to uncover any studies in which learners create multimodal explanations to teach either their current or future cohorts.

However, video has been used extensively in interview practice, and there is a sizeable body of research on learning

by teaching. Video interviewing was used successfully to prepare students for studying abroad [10]. Although the purpose of this study differed, the logistic challenges and video creation processes are very similar. Engin [11] used student-created digital videos to help students learn how to write academically. Students in Engin’s project, however, stated a preference for explanations from their teacher rather than their peers due to doubts over the quality of peer explanations. Benefits to student video creators have been shown to include increased understanding of content and the development of generic employability skills [12].

Through the creation of videos, learners increase their understanding. *The best way to learn is to teach* is a well-known aphorism. The learning-by-teaching effect has been shown to exist in numerous studies [13]-[15]. Research has shown the *protégé effect* in which learners put more effort into acquiring knowledge to teach someone else than to learn for themselves [16]. According to cognitive psychologists, the retrieval benefit, that is the act of retrieving the memories of the content, may explain these effects [17].

IV. COURSE ORGANIZATION

Learners create multimodal content, such as a teaching video or explanatory audio files, either as practice activities or to fulfil assessment requirements. During the creation process, they are expected to learn the content matter. Although the content of each of the elective courses differed greatly (applied logic vs. natural language processing) both could be broadly classified under the umbrella of computer science. The syllabus from each course focused on developing and applying sets of knowledge, behaviors and skills. The multimodal explanations were aimed at knowledge enhancement.

The most important elements in the course organization are the course framework, content specification and content creators, each of which is described in turn below.

A. Course Framework

For each course a simple website was created using the same tailor-made template. Each course is divided into ten units, each of which was allocated one webpage. Each webpage was divided into discrete sections (more precisely Div Objects in the HTML DOM). The sections were labelled as objectives, reading activity, listening activity, watching activity, thinking activity, creation activity, assessments and summary. The heading in each section is marked with an emoticon representing the activity type so students can scroll down and see the type of activities they will be engaged in without having to read any text. In addition to the course website, there is a password-protected course platform housed on Moodle, an open-source learning management system (LMS). This provides two options for materials to be shared: open or closed. It is hoped that students give permission for their artefacts to be shared freely online on the course website, but the final decision remains with the student.

B. Content Specification

The course tutor created a wish list of student-generated materials. Each wish list was categorized and prioritized. For example, in the applied logic course, the wish list included ten formal fallacies. The creation of an explanation for a formal fallacy could be set as an individual assignment so that in a class of 50 students, five students produce an explanation for each formal fallacy. Alternatively, the class could be divided into 10 groups, each of which submits a collaborative video. However, because of covid-19 and the logistic difficulties, groupwork was encouraged but not made mandatory.

Requirements specifications were produced and phrased as assessment criteria, e.g. sound must be audible. This enables students to check the suitability of their created content prior to submission.

An additional quality control step is included to ameliorate any concerns over the accuracy and quality of content of published videos. This step was conducted by the class tutor when assessing submissions. The best quality submission for each concept was selected to be uploaded to the course website or LMS, subject to permission.

C. Content Creators

The content creators were undergraduates enrolled in the two credit-bearing elective courses. Students worked alone or in teams depending on the multimodal resource to be created. Simple artefacts, such as short audio recordings were set as individual tasks while for more challenging tasks, such as the creation of a one-minute video explanation, students were encouraged but not compelled to work in teams.

Individual tasks were assigned based on the final digit of the student identity number. This ensured that ten different types of artefacts could be created for each assignment. Group tasks were assigned and managed using Trello, a web-based *kanban*-style project management system that uses movable labels grouped into lists. Each label contains the details of one project. Three lists were used: audio, video and software. Once a group finishes the audio content creation project, they choose a project from the next list. This acts as an incentive for groups to complete their projects quickly so that they can select their preferred project. The software creation tasks are not described in this paper as they are not directly related to multimodal material development.

V. MULTIMODAL MATERIAL CREATION

The process of creating audio and video explanations differs. Audio production is more straightforward and far less time-consuming than video creation.

When creating audio materials, teams were given only one piece of advice “sound is king”. Audio explanations although technically simple, rely on clarity of speech. Most students recorded their explanations multiple times until they were satisfied with the quality of their recording.

When creating video, teams had to plan the pre-production, production and post-production stages in sufficient detail. Each of the plans were approved by the course tutor before students progressed to the next phase. Students were given basic instructions and examples of storyboarding, screencasting, videoing, editing and subtitling. Teams were

given complete creative freedom.

To help learners prepare quality videos effectively and efficiently, the three-phase process is used. Table I shows the deliverables, i.e. the artefacts that learners submit in each phase in the process. The more complex the video, the more important the phases and deliverables are. Live action videos which use multiple actors require substantial planning to ensure that videos are produced in a timely manner with sufficient quality. Videos created digitally such as narrative slideshows and animation are less complex and can be edited post-production, and so interim quality checks of these deliverables were less important.

TABLE I: THREE-PHASE VIDEO CREATION PROCESS

Phase	Deliverables
Pre-production	script, storyboard, shot list
Production	video footage
Post-production	edited footage, subtitle text

A. Pre-production

In this phase students prepare either the complete script or a skeleton plan of what to say. For live action videos a storyboard and a plan of the type of shots to be used were required. The creation of the storyboard and shot list helps student teams shorten the video production time.

B. Production

The recording or creation of the video is the production stage. Capturing digital images on computer screens is more straightforward than videoing actors. The simplest type of video using an actor is a talking-head in which only the head of the actor is in shot. However, this requires a student to volunteer to star in a video. Most videos created did not involve actors due to the lack of willingness of students to appear in videos. Other forms of video include narrated slideshows and screencasting, both of which alleviate the need for actors and the necessity to consider lighting or camera shots.

C. Post-production

Most students used Apple iMovie to edit videos as it is installed on computers in the iLabs and CALL classrooms available on campus for student use.

Rather than subtitle manually, most students uploaded their content to YouTube to use its speech recognition technology to automatically generate subtitles. The accuracy of the subtitles was checked and alterations made as necessary. This step helped students notice possible pronunciation issues when the machine learning algorithm was unable to recognize their speech. The caption file can be downloaded as a SubRip text (SRT) file.

VI. ONLINE PLATFORM AND MATERIALS

The digital artefacts produced were uploaded to the two course websites to be played on demand by users.

Multimodal explanations that meet the quality requirements are uploaded to the course websites housed on the university server. The quality standards were:

- 1) Clear audio (with mean of approximately -12dB)
- 2) Video that is at least standard HD (1280x720 px)
- 3) Error-free text
- 4) No intrusive background noise
- 5) No offensive content

Some explanations were housed on the password-protected learning management system due to quality issues and/or lack of permission from students to share materials freely. Although some students created video materials using live action, namely actors on location, none of those materials met the quality standards to be shared online. Typical problems that impinged on the quality were lighting issues and camera wobble.

The six types of video explanations that were created by students and met quality expectations are shown in Table II.

TABLE II: SIX TYPES OF VIDEO FORMATS

Type	Description
Animation	Animations can be created using specialist software or platforms (e.g. Animaker and Powtoon). Slideshow software, such as PowerPoint, can also be used by incorporating transitions to create simple animations.
Slideshow	A series of slides are created using PowerPoint, Keynote, etc. A video recording of the slideshow can be created directly in the software.
Screencast	Software (e.g. Camtasia and Screencast-O-matic) is used to capture the whole or part of the computer screen. This enables the creator to record whatever applications are in use.
Pen and paper	A digital overhead projector visualizer is used to capture handwriting on a piece of paper held under the visualizer. Visualizers are available in most classrooms on campus.
Khan Academy style	A screencast can capture the text and drawings on a tablet (e.g. Wacom Bamboo Pen Tablet). The writing can be handwritten or typed. The drawings may be drawn freehand using an interactive pen or by using a mouse. Tablets are available in most classrooms on campus.
Puppet	A hand puppet is used as the central character. The creator controls the movement of the puppet.

VII. EVALUATION

Various types of videos were created, including Khan-style tablet capture and slideshows with voice over. The student-created content and the content-creation process were evaluated by learners, peers and course tutors.

A. Learner Feedback

Feedback from language learners on student-created multimodal explanations was sought using observation, focus group interviews and a small-scale questionnaire survey. Observations revealed that students tended to watch only the first part of many videos.

In focus groups, students noted that they enjoyed listening to the recordings made and watching the videos created by their classmates. On further questioning they explained that they usually stopped listening to or watching explanations

once they understood the content. However, for particularly interesting ones they would listen to or watch the whole recording. Table III shows the results of a survey on typological preferences.

TABLE III: POPULARITY OF SIX TYPES OF VIDEO

Type	Positive	Neutral	Negative	Overall rank
Animation	9	2	0	1
Slideshow	4	6	1	2
Screencast	3	7	1	3
Pen and paper	3	6	2	4
Khan Academy style	2	7	2	5
Puppet	2	4	5	6

Animation was overwhelmingly the most popular format and the use of a puppet was the least.

Based on feedback from content creators, the six types of videos were ranked according to the technical and logistic ease of creation and the time needed to create the video. The results of the rankings provided by student content creators in the applied logic course are shown in Table IV. The rankings were the results of a student-led class discussion/negotiation.

TABLE IV: EASE AND SPEED OF CREATION OF SIX TYPES OF VIDEO

Type	Ease of creation	Speed of creation	Overall rank
Pen and paper	easy	fast	1
Khan Academy	quite easy	average	2
Screencast	quite easy	quite fast	3
Puppet	average	fast	4
Slideshow	quite difficult	quite slow	5
Animation	difficult	slow	6

It can be seen that the most popular format, animation, required more time and effort to create. A notable advantage of the animated format was that the digital artefact, namely a slideshow presentation with embedded sound and animations could easily be edited. This enables simple mistakes, such as typos in the slideshow or mistakes in the audio recording, to be corrected.

B. Peer Feedback

Peer feedback was sought from four professors in the school of computer science and engineering. Table V provides an overview of their comments and suggestions.

TABLE V: SUMMARY OF PEER FEEDBACK

Comments	Number of teachers
Time efficient	3
Approval of teach-to-learn approach	4
Win-win scenario	1
Concerns regarding quality of materials	2

Overall, the feedback was positive with most teachers noting the time-saving nature of using students to create content. Given the immense time pressure that many teachers have experienced when transitioning to online learning, this criterion alone could be sufficient for teachers to adopt this approach.

All peers approved of the adoption of a teach-to-learn approach. One teacher described the solution as a win-win scenario since students learn by doing the work normally associated with teachers. As students worked in the target

language and produced deliverables in the target language, this activity provided learners with the opportunity to learn English by using English.

Some teachers, however, expressed reservations about the quality of student-created materials. Most of the comments regarding materials related to incomplete or incorrect information contained in the student-created artefacts.

The tutor acknowledges that some explanations were inaccurate or incomplete. However, with appropriate rubrics, these could also make useful learning materials. (e.g. by using questions like: Which of the four aspects is not mentioned? Which of the three definitions is inaccurate?)

C. Tutor Reflection

Without the help of learners to create content, it would have been difficult if not impossible to produce tailor-made multimodal content in a timely manner. Knowing learners benefit from the teaching-to-learn effect means that the producers of digital artefacts learn and that future cohorts can also learn from them.

The main problem was not actually the content of the explanations, but the production quality. Despite emphasizing the importance of quality, a number of audio files were not publishable due to low sound levels or background noise. Production problems with video files included confusing jump cuts, continuity errors and camera shake when the camera operator did not use a tripod.

Both the students and teachers agreed that the best quality videos were not live action, but animated slideshows with narration. This format had multiple advantages including the lack of necessity for actors, and the creation of multiple file formats (e.g. .pptx, .wav, .mp4), making the digital artefacts easy to edit post-production.

VIII. LESSONS LEARNT

Creating high-quality content takes planning and preparation, but with a well-designed quality control process, student-created multimodal materials can help teachers rapidly develop a bank of publishable quality materials and help students learn the content. The main lessons learned are listed below as imperatives:

- 1) Ensure students know who the intended audience is.
- 2) Explain to students that they will learn by teaching.
- 3) Provide clear rubrics, e.g. make a one-minute video explaining concept X.
- 4) Provide clear assessment criteria, e.g. Video must be between 50 and 70 seconds. Sound must be audible. Subtitles must be accurate.
- 5) Provide examples and/or templates.
- 6) Show how lighting and sound affect video quality using good and bad examples.
- 7) Monitor video production progress and quality by requiring submission of storyboards and scripts.
- 8) Be aware that live action is the most difficult, most time-consuming type of video format.

IX. CONCLUSION

Once teachers relinquish the production of (some)

materials to learners, the teachers can focus on helping learners learn.

A rudimentary understanding of digital filmmaking makes it easier to advise learners on the problems and pitfalls. The more technical knowledge the teacher possesses, the more likely that better quality multimodal materials will be created. For example, knowing when to use a directional or bidirectional microphone, the rule of thirds to create visually appealing frames and how to use depth of field to control the focus of the audience will result in higher quality videos.

Each year a new cohort of students makes a new set of explanations. The mean quality of the multimodal explanations increases annually as the worst video and audio files are replaced. This potentially creates a continuous cycle of incremental improvement.

CONFLICT OF INTEREST

The author declares no conflict of interest.

ACKNOWLEDGMENT

I would like to thank Professor John Brine for his invaluable advice and suggestions on the practicalities of student-generated video production. My sincere thanks also go to all the students who actively participated in both elective courses and created the multimodal materials.

REFERENCES

- [1] J. King, "Silence in the second language classrooms of Japanese universities," *Applied Linguistics*, vol. 34, no. 3, pp. 325-343.
- [2] S. Harumi, "Classroom silence: Voices from Japanese EFL learners," *ELT Journal*, vol. 65, no. 3, pp. 260-269, 2011.
- [3] S. Yoneyama, *The Japanese High School: Silence and Resistance*, Routledge, 2012.
- [4] C. Lauer, "Contending with terms: Multimodal and multimedia in the academic and public spheres," *Computers and Composition*, vol. 29, pp. 225-239, 2009.
- [5] B. Goldstein and P. Driver, *Language Learning with Digital Video*, Cambridge: Cambridge University, 2015.
- [6] S. Apkon, *The Age of the Image: Redefining Literacy in a World of Screens*, New York, NY: Farrar, Straus & Giroux, 2013.
- [7] J. Willis, "Implications for the exploitation of video in the EFL classroom," in *Video Application in English Language Teaching, ELT Documents*, J. McGovern, Ed. London: Pergamon Press, 1983, pp. 29-42.
- [8] B. Goldstein and P. Driver, *Language Learning with Digital Video*, Cambridge: Cambridge University, 2015.
- [9] P. Hakkarainen, "Designing and implementing a PBL course on educational digital video production: Lessons learned from a design-based research," *Educational Technology Research & Development*, vol. 57, no. 2, pp. 211-228.
- [10] J. Brine, E. Kaneko, Y. Heo, A. Vazhenin and G. Bateson, "Language learning beyond Japanese university classrooms: Video interviewing for study abroad," in F. Helm, L. Bradley, M. Guarda, and S. Thousney, Eds. *Critical CALL: Proceedings of the 2015 EuroCALL Conference*, pp. 91-96, 2015.
- [11] M. Engin, "Extending the flipped classroom model: Developing second language writing skills through student-created digital videos," *Journal of the Scholarship of Teaching and Learning*, vol. 14, no. 5, pp. 12-26, 2014.
- [12] P. Hager and S. Holland, "Introduction," in *Graduate Attributes, Learning and Employability*, P. Hager, and S. Holland, Eds., Dordrecht, The Netherlands: Springer, pp. 17-47, 2006.
- [13] L. Fiorella and R. E. Mayer, "The relative benefits of learning by teaching and teaching expectancy," *Contemporary Educational Psychology*, vol. 38, no. 4, pp. 281-288, 2013.
- [14] R. Kozma and J. Russell, "Students become chemists: Developing representational competence," in *Visualisation in Science Education: Models and Modelling in Science Education*, J.K. Gilbert, Ed. Dordrecht, The Netherlands: Springer, pp. 121-146, 2005.

- [15] H.-K. Wu and S. Puntambekar, "Pedagogic affordances of multiple external representations in scientific processes," *Journal of Science Education and Technology*, vol. 21, pp. 754-767, 2012.
- [16] C. C. Chase, D. B. Chin, M. A. Oppezzo, and D. L. Schwartz, "Teachable agents and the protégé effect: Increasing the effort towards learning," *Journal of Science Education and Technology*, vol. 18, no. 4, pp. 334-352, 2009.
- [17] A. W. L. Koh, S. C. Lee, and S. W. H. Lim, "The learning benefits of teaching: A retrieval practice hypothesis." *Applied Cognitive Psychology*, vol. 32, no. 3, pp. 401-410, 2018.

Copyright © 2021 by the authors. This is an open access article distributed under the Creative Commons Attribution License which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited ([CC BY 4.0](https://creativecommons.org/licenses/by/4.0/)).



John Blake started his career teaching English as a foreign language in schools, colleges and universities in East Asia and the United Kingdom. He holds postgraduate degrees in education, applied linguistics, business administration and computer science. He is a Senior Associate Professor in the Center for Language Research at the University of Aizu where he creates practical online tools that help people learn English. His research draws on corpus linguistics to analyze texts and computational linguistics to create rule-based and probabilistic-based pattern-searching tools or pipelines. Prof. Blake teaches courses on applied logic, natural language processing and thesis writing.