

Article

Multimodal, Digital Artefacts as Learning Tools in a University Subject-Specific English Language Course

Anne FJ Hellwig

University of Wollongong, Australia

Received: 30 September 2021/Accepted: 15 February 2022/Published: 1 April 2022

Abstract

This paper explores the practice of using multimodal, digital assessment tasks assigned to students of an English for Architects and Civil Engineers course at a university in Germany. Students were tasked with creating multimodal video compositions and interviewed about the processes behind composing their artefacts. The goal was to interrogate to what extent multimodal assessment tasks such as these can promote the communication of technical concepts, facilitate nuanced opportunities for language development and develop the students as social agents. The artefacts were examined through the lens of Systemic Functional Semiotics, drawing particularly upon the Genre and Multimodality framework (Bateman et al., 2017; Bateman & Schmidt-Borcherding, 2018) and a recent approach to analysing multimodal artefacts developed by Turney & Jones (2021).

Keywords

Multimodality, TESOL, Systemic Functional Semiotics, Genre and Multimodality Framework

1 Introduction

The overt emphasis on digital, multimodal communication in the applied disciplines of architecture and civil engineering is not always reflected in related subject-specific English language courses. This is especially the case in Germany, where multimodal literacy has been neglected in favour of a text-centric approach to language education (Wilke, 2012). There is a pressing need to develop curriculum and assessment tasks in Teaching English to Speakers of other Languages (TESOL) education that better reflect the demands placed on English language students, both during their disciplinary studies and within their future workplaces. Additionally, the recent changes to the Council of Europe's Common European Framework of Reference for Languages (CEFR) have asked educators to substantially reimagine "... the user/learner as a social agent" (Council of Europe, 2020) and measure new competencies around mediation and plurilingualism. Further, educational policy around the world has increasingly emphasised the importance of digital literacy for 21st century social and civic participation (Lamb et al., 2017), especially with the recent move towards online education necessitated by the spread of Covid-19.

In light of these shifts and challenges, this paper focuses on one case in a multiple case study examining a multimodal, digital assessment task assigned to students of two English for Architects and Civil Engineers courses at a university in Germany. The case comprises a student-generated, multimodal video composition and two interviews, one with the student creator about the processes behind composing their artefact and another with a second student, probing their interpretations of the artefact as an audience member. The purpose of this study is to interrogate to what extent multimodal assessment tasks can promote the communication of technical concepts, facilitate more nuanced opportunities for meaning making in the target language and develop the students as social agents. The artefacts and interview data were examined through the lens of Systemic Functional Semiotics, drawing particularly upon the Genre and Multimodality framework (Bateman et al., 2017) and the Turney & Jones model for analysing multimodal artefacts (Turney & Jones, 2021). It is hoped that these analyses can help confirm the usefulness of multimodal assessments in the TESOL context, and a contribution can be made to the task of assessing language skills in an integrated way.

2 Background

2.1 Impetus

Educational policy around the world increasingly emphasises the importance of digital literacy for 21st century professional, social and civic participation, underscoring a pressing need to cultivate digital practices and skills. The World Economic Forum's "Future of Jobs" report, for example, suggests that "...by 2022, the skills required to perform most jobs will have shifted significantly..." (WEF, 2018, p. viii), which means that transdisciplinary skills with broad application are increasing in value. Digital, multimodal assessment tasks help improve students' digital literacy as well as their metacognition and communication skills, and these transferable skills are predicted to hold value in the job markets of the future (Lamb et al., 2017). Nevertheless, there appear to be clear gaps between policy standards and classroom practices relating to digital literacy (Gallagher & Rowsell, 2017), with studies tending to focus on the social uses of digital media rather than their disciplinary applications (Moreno, et al., 2015, *inter alia*).

The accessibility and affordability of digital technologies have also encouraged an increase in the creation of multimodal artefacts. Multimodal artefacts are texts that simultaneously employ a number of communicative modes, such as writing, image and film (Kress & van Leeuwen, 1996), and in the 21st century, students need to be able to communicate multimodally in order to be considered sufficiently literate (Jewitt, 2009). However, this shift to increasing digital, multimodal practice is not as well reflected in Germany, especially in the TESOL courses offered at university (Wilke, 2012). This is even the case for the English for Special Purposes courses which emphasise multimodal communication, such as the two English for Architects and Civil Engineers courses that form the basis of this paper. Although students are expected to build physical and digital 3D models, create floorplans and use other modes of communication in their discipline-specific courses and future workplaces, their related English language courses largely ignore multimodal literacy in favour of focusing on the four skills of reading, writing, listening and speaking.

However, the Council of Europe's CEFR – the preeminent framework for measuring language acquisition in Europe – has recently been changed to move away from the four skills, which they claim have "... increasingly proved inadequate to capture the complex reality of communication" (Council of Europe, 2020, p. 33). This has involved substantial changes to the way they assess language acquisition. Firstly, they have added the competence of "plurilingualism", which "...involves the ability to call flexibly upon an inter-related, uneven, plurilinguistic repertoire..." (Council of Europe, 2020, p. 30). They have also reimagined the learner as a "social agent" and redeveloped the assessment framework around "four

communicative modes”, of which the four skills form only a part. These modes of communication are: reception (listening, reading); production (speaking, writing); interaction (a social skill) and a new mode, mediation. It is this last mode that is the most complex, involving as it does all the other modes, and it is also of most relevance to this paper. According to the CEFR, mediation is an “...interpretation or reformulation of a source text” very similar to “resemiotization”, where meaning making shifts from context to context, from practice to practice, or from one stage of a practice to the next (Iedema, 2003, p. 41). Mediation, and its relationship to multimodal literacy, will be explored in more detail in 4.3 of the Results section. It will be suggested that multimodal assessment tasks are an excellent way to expand upon this area of language acquisition and provide educators with a way of satisfying the descriptors included within it. The multimodal assessment task explored in this paper involved the production of a 3-5-minute video composition (VC) explaining a concept from architecture or civil engineering to a non-specialist audience using a variety of modes. This task was worth 20% of the students’ overall grade.

2.2 Theoretical background

Multimodality theory allows us to explore the ways in which the connections and combinations between modes can unsettle existing practices, forge new connections and animate new meanings. In this sense, multimodality is closely tied to semiotics, in particular, to social semiotics (Hodge & Kress, 1988). This approach draws on Halliday’s work on systemic functional linguistics (SFL), which sees language as a social action constituting culture (Halliday, 1994). Multimodal artefacts are those which communicate through a variety of modes simultaneously (Jewitt, 2005), and the lens of multimodality can help us to understand the complexities at play within them. Examples of modes include, but are not limited to, writing, images, music and architecture. At their essence, modes are shaped by culture to produce different kinds of knowledges, enact different social relationships and perform identities. However, the boundaries between what constitutes a mode can be somewhat diffuse, for example, the mode of “image” can usefully be broken down into image types, such as a photograph or a charcoal sketch. Indeed, even within a photograph there are different communicative elements, such as the choice of black & white or colour, and further still, within the realm of colour, communicative choices such as brightness, saturation and so on can be made.

For clarity, and in accordance with the work of Bateman et al. (2017), this paper conceptualises a mode as limited by the affordances of its materiality. They write, “The reach of a semiotic mode will usually be a refinement and extension of what its material carrier affords” (Bateman et al., 2017, p. 119); that is, the definitional boundaries of a mode are determined by the opportunities and constraints of its material context – such as which senses are engaged, whether and how time is involved, and so on - and also in terms of the discourse community within which it functions. It can be reasonably assumed that the student-creators of the video compositions (VCs) reported on here belong to one shared discourse community, narrowing the scope of available modes. Once the mode has been identified, it is then useful to think of modes as having different ‘modal resources’ (Bezemer & Kress, 2008), that is, within a mode like ‘image’, modal resources such as framing, composition and colour affect the meanings that are created and connoted (Kress and van Leeuwen, 1996).

2.3 Literature review

Jewitt et al. (2001, p. 27, emphasis original) argue that social semiotics allows us to see “...the process of learning as a *dynamic process of sign-making*”. Small wonder, then, that in the past twenty years in particular, multimodal assessment tasks have flourished in classrooms around the world. Scholars have found that multimodal assessment tasks can enhance student creativity and agency (McGinnis, 2007), provide opportunities for increased levels of student engagement (Pandya et al., 2018) and better prepare learners for the future (Hafner, 2015). Although the two concepts are distinct (Alvermann, 2017),

multimodal literacy has some overlap with the theory of “multiliteracies”, first proposed by the New London Group (1996) and further developed by Cope and Kalantzis (2009), who emphasised the socio-historical context of literacy. While the creation of multimodal texts may not be a new phenomenon, the increasing ubiquity of digital technologies has changed social practices (Lotherington & Jenson, 2011) and further democratised knowledge, with an increasing emphasis on collaboration and participation (Knobel & Lankshear, 2014).

Digital, multimodal projects have been popular at all levels of education for at least the past twenty years, whether at primary (Burn & Parker, 2001 etc), secondary (Nash, 2018 etc) or tertiary levels (Nielsen, *et al.*, 2016 etc). They have also been explicitly included in curricula all around the world, such as in Australia (ACARA, 2015), the United States (Lapp & Fisher, 2011) and Europe (EUMade4All, 2019). As such, it is no surprise that TESOL educators have also embraced multimodal literacy at all levels, including in primary (e.g. Grapin, 2019) and secondary education (e.g. Huang, 2019). From a tertiary TESOL perspective, Oldakowski (2014) argues that multimodal assessments deepen comprehension and promote engagement, while Zacchi (2016) asserts that multimodal meaning making is an effective way of traversing cultural differences in an increasingly globalized world. Jiang and Luk (2016), writing from a Chinese context, found that such multimodal assessment tasks increase students’ “motivational capacity”, with interviews indicating an increased sense of curiosity and cooperation, among other qualities. In Taiwan, Lee (2014) claims that multimodal learning practices enhance students’ motivation and self-confidence.

Nevertheless, digital, multimodal assessment tasks remain the exception rather than the rule within most TESOL curricula. Writing not only from the tertiary perspective but across their work with young children, adolescents, and adults, Early et al. (2015) suggest that multimodality is “on the margins... in the TESOL community” (p. 450-1) and that we may need to rethink our course design and rewrite our textbooks if we are to wholly incorporate multimodality into language education. Lotherington & Jenson (2011) further this claim, arguing that in all L2 teaching contexts, teachers have been reluctant to embrace or even acknowledge multimodal literacy, favouring instead the ‘flat literacies’ of paper-based assessment.

3 Data

3.1 Case context

The cases within the study were drawn from two courses of “English for Architects and Civil Engineers A”. This was a 4-credit point, 14-week course taught in the winter semester 2019-2020 at a university in Germany. It was designed for students with an English language level of intermediate to advanced (B2-C1 according to the CEFR scale), and the courses attracted both undergraduate and postgraduate students, (n = 38), 32 of whom were majoring in civil engineering, with only six students majoring in architecture. This bias was likely a consequence of the students needing a B2 English to graduate from Civil Engineering, while the Architecture students had no such requirement. For 20% of their grade, students were tasked with producing a 3-5 minute video composition (VC) explaining a concept from architecture or civil engineering to a non-specialised audience using a variety of modes. The number of modes could vary, resulting in dynamic, standalone, two-dimensional artefacts (the material parameters of these artefacts will be explored in more detail in 4.1 of the Results section). Students were asked to upload the artefacts to the sharing platform Moodle before they were screened in class, where they were encouraged to lead discussions around the VCs, offering feedback and asking questions of their fellow student creators.

Students in both courses were also invited to participate in this research project, which had two phases of data collection: the artefacts themselves were collected, and after semester’s end participants

were interviewed about the processes of creating the artefacts and their perspectives as audience members. Of the 38 students in the two classes, 17 agreed to participate in the first phase of the study (artefact collection), with 7 of those also agreeing to the second phase (the interview). The semi-structured interviews lasted between 19:56 and 46:11 minutes and involved a pre-prepared interview protocol. There were seven general questions relating to the processes behind designing the artefacts as well as a number of questions developed in response to specific elements of the VCs. There were also four additional general questions relating to their experience as audience members. The data was then investigated inductively, using NVivo to develop themes and code the responses.

3.2 Case study: “Human comfort in relation to architectural spaces”

The case reported on in this paper is comprised of a video composition and the information provided in two post-task interviews, one with the student composer (Student C) and one with a fellow-student audience member (Student A). This artefact was selected as it was considered to exemplify the task, as well as including six modes with varying affordances: text (both written and spoken), image (hand-drawn sketches, cartoons and photographs), music, various typographical and layout elements, film and gesture. Student C is a C1 level (advanced) English language learner in an Architecture track program. Her artefact is 5:15 minutes and her interview lasted 36:02 mins. Student A is a B2 level (intermediate) language learner in a Civil Engineering program, and his interview lasted 46:11 minutes, of which 9:17 minutes were devoted to responses to Student C’s artefact. The remainder of Student A’s interview is not relevant to this paper and is not included in the data drawn upon here.

4 Results and Discussion

4.1 RQ1: To what extent can digital, multimodal assessment tasks promote the communication of technical concepts?


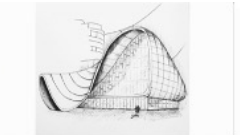

Measuring the impact of these video compositions on students’ understanding of technical concepts is far beyond the scope of this qualitative study. However, one way of approaching this question is by examining the artefact through a lens developed by Bateman and Schmidt-Borcherding (2018) in their quantitative study of the effectiveness of educational videos in terms of learner uptake and engagement. After analysing the results of a knowledge test and an engagement survey, they suggest that a successful educational video should establish clear expectations and avoid sensory overload. The more successful videos “...prepare their audiences for their messages audiovisually and then use this preparation for presenting new information” (Bateman & Schmidt-Borcherding, 2018, p. 4), a process they divide into the two ‘discourse units’ of ‘scaffolding’ and ‘development’. Units that scaffold information prepare audiences for what to expect later in the video, and units that develop information “elaborate or extend what has been introduced previously” (Bateman & Schmidt-Borcherding, 2018, p. 11).

In order to perform a fine-grained analysis of these very complex artefacts, they argue that the constraints and affordances of the media used must first be recognised in order to identify what is intentional and what is a result of a limitation of the medium (as this in turn restricts which semiotic modes can be employed). It is therefore important to identify the parameters of the physical situation in which VCs take place, what Bateman, et al. (2017, p. 96) term “...the ‘canvas’ that meaning is inscribed on...”. According to their definitions, a video composition changes over time and is therefore “dynamic”; it is also viewed rather than participated in, making it “observational”, and this is done through a computer screen, rendering it “two-dimensional”. Further, the artefact itself cannot change, making it “immutable”, and it can be rewatched for a limited time (until the course content is removed

from Moodle), making it “partially transient”. To analyse such a dynamic, 2D, immutable, observational, partially transient artefact, the artefact must be broken down into smaller, trackable units of meaning. The term “presentational micro-event” (PMEs) is useful here, as it facilitates looking at the artefacts as a collection of “...unit(s) of meaningful behaviour that may be distributed across several coordinated sensory channels...” (Bateman & Schmidt-Borcherding, 2018, p. 6).

Figure 1

A Visual Representation of Some Presentational Micro-events (PMEs) Sharing Meaning Intersemiotically

				
2:02 - 2:03	2:04 - 2:05	2:06 - 2:27	2:28 - 2:29	2:30 - 2:31
or a street space.	All in all,	the interior and exterior space must therefore have a balanced relationship in order to function. The choice of location also plays an important role here. The buildings have to establish a connection with the surroundings, to fit in, and further, to enter into a dialogue with other buildings. The architect then decides whether	dominance	or adaption into the frame is the right thing to do

This is perhaps best illustrated with an example. In Figure 1, a section of the artefact is presented. In this image, five still frames of the video are presented above a text of the accompanying narration. The VC has been broken down into five PMEs, and it is possible to see here how meaning is shared across modes, or sensory channels. Three noteworthy moments occur at 2:02 minutes, when a photograph of a highway appears on screen with the narration of the term “street space”; and at 2:28 minutes, when the architectural term “dominance” is accompanied by a sketch of Zaha Hadid’s Heydar Aliyev Centre in Azerbaijan; and again at 2:30 minutes, when a sketch of a terrace house co-occurs with the narrated term “adaptation”. The example at 2:28 minutes should perhaps be unpacked: this is a good example of how an image can co-instantiate the technical field of discourse as it draws upon a shared repertoire of architectural knowledge. For the initiated, the Heydar Aliyev Centre is one of the most recognisable of Hadid’s buildings. It carries with it some of the meaning of the term “dominance”, because Hadid’s designs are synonymous with the kind of architecture that ignores the context of its surroundings and has “...neither respect nor reference to its locality” (Bayley cited in Fairs, 2015). Student C follows this PME with a sketch of a terrace house at 2:30 minutes to co-instantiate the meaning of “adapting to the cityscape”, shown in synchrony with the narrated term, “adaption”. In this way she invites her audience to unpack the technical nominalisations of “dominance” and “adaption” by showing us a visual example of each. These can also be considered “grammatical metaphors”, as they package complex processes as single elements within the clause (Macnaught et al., 2013); that is, the activities of “dominating a cityscape” or “adapting to a cityscape” are bundled into technical terms of considerable complexity (‘dominance’, ‘adaption’), especially for English language learners. As such, Student C has chosen to depict buildings that exemplify the processes of dominating or adapting to the cityscape in order to help the audience “make sense” of these technical terms by sharing the “work” of making meaning intersemiotically.

Student C helps her audience understand her technical concepts in other ways as well. She frequently uses parallelisms and redundancies across the narration and the images, or the “audio and visual sensory channels”, to support her audience’s understanding. As can be seen in Figure 1 at 2:06 minutes, the terms of the text of the narration in colour coincide with the moment the bullet point of written text appears on screen. Student C has taken the time to temporally coordinate her visual information in order to support her audience’s conceptual understanding: when she says “balanced” in the narration, the bullet point “balanced relationship” appears; when she says “connection”, the bullet point ‘connecting with the surrounding’ appears; and when she says “to enter into” the bullet point “dialogue with other buildings”

appears. This attention to detail shows not only that she has harnessed the affordances of the medium effectively and appropriately, but that she has given considerable thought to its pedagogic potential and her relationship to the audience.

Further, returning to Bateman and Schmidt-Borcherding's discourse functions of "scaffolding" and "development", Student C has placed scaffolding segments at regular and appropriate intervals throughout the artefact. They can be found at four intervals: at 0:35 minutes, when she introduces a taxonomy of the elements to consider when designing interior space; again at 1:34 when she taxonomises exterior space in the same manner; at 2:06 (see Figure 1), when she relates the building to its environment; and again at 2:39 minutes, when she categorises the ways in which individuals experience the built environment. This strengthens her participation in the genre of a 'system explanation', which will be elaborated upon below. It also helps the audience anticipate the meanings to come and focus their attention upon the most salient concepts.

She also makes meaning intersemiotically by repeating a sketch of a face throughout the VC. This sketch occurs at the beginning and end of the artefact, and, crucially, re-occurs when new information is being scaffolded as a "visual reminder" for the audience to focus their attention. The image also remains in the background of the artefact as a sort of visual "anchor", albeit at varying degrees of magnification. You can see it clearly at 2:04 minutes, for example (Figure 1), but it is also present in the background at 4:27 (Figure 2), although the magnification transforms it into an indistinguishable blur of pixels. It works as an almost subliminal cohesive device, guiding the audience into unfamiliar conceptual territory while retaining a "familiar face". The fact that this is a hand-drawn sketch by the participant herself of a singer she enjoys speaks to the rich vein of interpersonal meaning that is present in the artefact but beyond the scope of this paper. It is worth noting, however, that eye-tracking research suggests that faces on screen, however small, attract and fixate the gaze (Wang & Antonenko, 2017, cited in [Bateman & Schmidt-Borcherding, 2018](#)), and it could be said that the sketch was included in an attempt to hold the audience's attention.

Summary of RQ1: (RQ1: To what extent can digital, multimodal assessment tasks promote the communication of technical concepts?)

Bateman and Schmidt-Borcherding (2018) suggest the constraints and affordances of the media employed – the "canvas" – should be identified in order to separate intentional from inadvertent meaning making. After this, the artefact can be broken down into smaller units of meaning - termed "presentational micro-events" (PMEs) - to track how meaning is made. Such an approach illuminates how Student C consistently shares the semiotic labour across modes to help her audience understand the technical concepts she communicates. For example, certain technical terms employed in the narration are elaborated visually in the form of sketches, and she frequently uses parallelisms and redundancies to support her audience's understanding. A successful educational video should also establish clear expectations and avoid sensory overload. According to Bateman and Schmidt-Borcherding (2018), this can be achieved through two processes they term "scaffolding" and "development". Scaffolding, that is, preparing the audience for what to expect, is observable in Student C's VC when she categorises elements of space and the built environment. She then 'develops' these taxonomies by elaborating upon what has already been introduced. All of these elements suggest that digital, multimodal assessment tasks such as this one can very effectively promote the communication of technical concepts.

4.2 RQ2: To what extent can digital, multimodal assessment tasks facilitate more nuanced opportunities for meaning making?



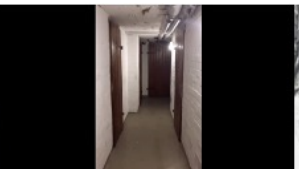

In order to understand meaning making, we need to situate the artefact in the context of the culture, and one of the best ways to do this is to identify which genre it is participating in. Genres can be seen as 'recurrent configurations of meaning' ([Rose & Martin, 2012](#), p. 53) which make sense to the discourse communities that comprise the culture. As such, it is important that the students correctly identify and

reproduce the genre required by the task. In order to better determine multimodal task fulfilment, Turney & Jones (2021) have developed a method of analysing tertiary-level, student-generated, educational videos. Drawing upon the Genre and Multimodality model (Bateman et al. 2017), they suggest first identifying the genre of the artefact before examining the media used and then exploring how the artefact unfolds intersemiotically, realizing configurations of register variables (patterning of field, tenor and mode). The medium and some of the intersemiotic meanings realised in the artefact have been briefly touched upon in the previous section, but it is worth unpacking the VC in terms of its genre. In the Martinian systemic functional perspective, genre is situated in the stratified context plane, departing from a strictly Hallidayan perspective which associates genre with mode (Martin, 2009). For an artefact to participate in a genre, it must have a characteristic structure and observable stages and phases contributing to the achievement of its social purpose (Derewianka & Jones, 2016). It is also realised in terms of three simultaneously occurring parameters: the field, tenor and mode of discourse. While field is concerned with the topic of the text and its representation of the world, tenor focuses on the relationships between the participants, and mode is concerned with the text type and its organisation (Martin, 2009). Identifying and understanding the boundaries of genre is essential to teaching and assessment, especially in a TESOL context, where the differences between genres may be more difficult for second language learners to identify and reproduce.

The task question explicitly asked students to “explain a concept from the fields of architecture or civil engineering...”, and unsurprisingly, a great number of language features present in the narration of Student C’s VC were typical of the “explanation” genre. Looking firstly at how she constructs her field of discourse, there are a marked number of generalised participants (e.g., “interior and exterior spaces”) and nominalized abstract concepts (e.g., dominance, adaption) as well as causal relationships (e.g., “These experiences are triggered by the senses...”) and a considerable amount of technical and specialized vocabulary (e.g., “form dimensioning”). It is harder to ascertain which of the explanation genres the artefact is participating in, as it does not unproblematically conform to any one. It can, however, be read as a system explanation, the explanation genre concerned with the relationships and interactions between different parts of a system. This genre typically begins by identifying the Phenomenon, describing the System, explaining the interaction between the Components of the System and concluding with a Generalisation (Derewianka & Jones, 2016, p. 205-6). These elements are broadly observable in Student C’s VC, with “the interior”, “the public space”, “the architect” and “the human comfort” functioning as the Components. In this way, the interaction between the Components reflects the student’s own sequential process through the narration: Interaction One is between the interior/exterior and human senses, or as she terms it, the “first step” of cognitive appraisal; Interaction Two occurs between the interior/exterior and human emotions (“the affective reaction”), and Interaction Three is between the interior/exterior and human aesthetics, or what she calls the third process of “aesthetic reaction”. The interaction between the levels is less explicitly realised, but can be seen in her conclusion (4:38 – 5:13 mins), where she narrates, “the relationship between space, human and content is what gives the place its expression”.

Figure 2

Intersemiotic Meaning Making with Film, Image and Text

			
4:11 - 4:18	4:19 - 4:20	4:21 - 4:26	4:27 - 4:37
or a church or mosque where you feel delightful, peaceful. Or a space that triggers	discomfort, and therefore	radiates a feeling of coldness such as a basement. The third process would be	the aesthetic reaction, which is taste dependent and different from each person. Does the room correspond to my style? Is it aesthetically appealing? Is the design modern, timeless or minimalistic?

Having identified the genre, the question remains as to what extent multimodal assessment tasks such as this one can facilitate more nuanced opportunities for meaning making. Perhaps the most effective method of exploring this is to identify where meaning is made with notable depth or subtlety. Figure 2 depicts one such section of the artefact. Much like in Figure 1, the still frames in Figure 2 were taken from the VC between 4:11 and 4:37 minutes, the text of the narration is printed underneath, and the section is divided into four PMEs. What makes this section different to the one shown in Figure 1 is the use of video, shot by a friend of the participant at her request. Two separate videos were recorded, one at 4:11 – 4:18 minutes and another at 4:21 – 4:26 minutes, both of which share the meaning of the narration in subtle and interesting ways. In the first video, a hand-held camera pans upwards, simulating the craning of a neck as Student C narrates “... a church or mosque where you feel delightful, peaceful”. Neck craning both connotes and is a physical response to awe, and this gesture corroborates the meaning made in her narration. Similarly, a second film, also suggestive of a first-person experience with the use of a hand-held camera at head height, carries some of the meaning of the narration at 4:21 minutes (“... radiates a feeling of coldness, such as a basement”), with the jolting movement of someone walking into a dark, narrow space. Student C is sharing the work of making meaning across modes and is also elaborating upon the experiences she describes, inviting her audience to share in her experiences and affiliate around her values (Knight, 2010): feeling “peaceful” and “delightful” as they “gaze” upwards at a light-filled dome in a mosque, and feeling “cold” in the dark basement. Exploring the interpersonal patternings instantiated in the tenor of this artefact is beyond the scope of this paper, but such fine-grained attention to detail also guides the audience to engage with and comprehend the key concepts explained in her artefact and emphasised in her title “Human comfort in architectural spaces”. Although some additional information is provided in the narration that is not manifested visually, her most salient ideational meanings are almost unfailingly supported across more than one mode simultaneously. Whenever she presents key ideas, for example, she both narrates and visually displays the key points on screen (see Figure 2 at 4:27 minutes).

Similarly, the information presented visually supports the meanings being made in the audio text, even if this is not always performed flawlessly. When asked about her inclusion of what appears to be a cartoonish drawing of David Bowie at 4:27 minutes (see Figure 2), alongside drawings of a man in a suit and a woman with a flower crown, Student C commented in her interview data that “...they’re all different, so I just wanted to include that or stress that”. For the student creator, these animations expand upon the meanings made in her narration (“...the aesthetic reaction, which is taste dependent and different from each person. Does the room correspond to my style?”), and so what at first appears unnecessary or distracting, upon closer examination is yet another example of her semiotic decision making. Such “representation” produces a sign that is focused on Student C’s interest, rather than “...the assumed interest of the recipient of the sign” (Kress, 2010, p. 71). Although the cartoons she chose may not be as transparent in their meanings as other elements in the artefact, they are nevertheless examples of motivated representation rather than meanings made inadvertently.

Summary of RQ2: (To what extent do digital, multimodal assessment tasks facilitate more nuanced opportunities for meaning making?)

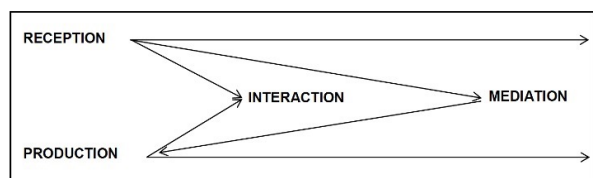
For meaning to be made effectively, the artefact should be appropriately situated within the context of the culture. TESOL students in particular often struggle with identifying and participating in genres and reproducing the stages and phases that contribute to their structures. This multimodal assessment task facilitates opportunities for effective meaning making in a broad sense by necessitating that the artefact participates in one of the explanation genres, as Student C’s does. The task also facilitates more nuanced opportunities for meaning making, as evidenced by the rich and subtle meanings made by, for example, certain camera movements in her filmed segments to variously communicate awe and claustrophobia.

4.3 RQ3: To what extent do digital, multimodal assessment tasks develop the students as social agents?

As mentioned in section 2.1 of the Background, the Council of Europe has redesigned their CEFR assessment criteria in “...a move away from the matrix of four skills...” and towards “... real-life language use (Council of Europe, 2020, p. 33)”. They have not only added a new competence, plurilingualism, but have also expanded on the four skills to add two new “communicative modes” (see Figure 3), of which one, mediation, is of particular relevance here. Mediation emphasises “...the constant movement between the individual and social level in language learning, mainly through its vision of the user/learner as a social agent” (Council of Europe, 2020, p. 36). It also bears a striking resemblance to Iedema’s “resemiotization”, the process of transforming meaning across contexts and practices (Iedema, 2003). The theoretical background for this conceptual shift is based, however, on the work of Vygotsky (1978) and sociocultural theory, as well as the ecological model (van Lier, 2000) and complexity theories (Piccardo, 2015, all cited in Council of Europe, 2016, p. 16).

Figure 3

The Relationship Between Reception, Production, Interaction and Mediation (Council of Europe, 2020, p. 34)



Multimodal artefacts such as the one explored here position the learner as a social agent in screenings and uploadings of their work. The skills demonstrated throughout this task also bear a striking resemblance to some of the updated descriptors included in the appendix of the latest companion volume of the CEFR. This could be of benefit to TESOL educators wishing to accredit their students with CEFR certification. Examples of how multimodal assessment tasks could meet CEFR descriptors is provided in Figure 4. Multimodal assessment tasks such as this one could also be adapted to groupwork, which would incorporate some of the other new mediation competences, such as “Managing interaction” and “Collaborating to construct meaning”.

The Council of Europe have added eighteen new competences, with detailed descriptors provided for learners from beginner (A1) to advanced (C2) levels (Council of Europe, 2020). Of these eighteen, only the eight competences most relevant to multimodal literacy are shown here (Figure 4), along with one corresponding descriptor for both B2 and C1 levels. In order to demonstrate how closely aligned some of the new descriptors are with many of the learning outcomes of this assessment task, it is worth looking at one competence in closer detail, along with two of its related descriptors at both C1 and B2 levels. One example of demonstrating mediation competence at C1 level is that students should be able to “...explain (in Language B) the relevance of specific information found in a particular section of a long, complex text (in Language A)”, while at B2 level, they should be able to “...interpret and describe reliably (in Language B) detailed information contained in complex diagrams, charts and other visually organised information (with text in Language A)” (see Figure 4). This dovetails beautifully with the task requirements of the assessment explored here and is richly demonstrated in the artefacts. When students rephrase, or “mediate” the academic, German language of their lectures and textbooks into “everyday” English, they also fulfil the descriptors for both spoken and written mediation competence. This is visible not only in the labels and text boxes of their VCs, but also in the scripts they compose to prepare for their narration: six of the seven students interviewed reported that they wrote a text to read aloud in advance of their audio narration.

Further, within the category, “mediation strategies”, there are a further two sub-categories of particular relevance here: “strategies to explain a new concept and strategies to simplify a text” (see Figure 5). The descriptors here are remarkably relevant to this project: for example, students at C1 level are expected to “...explain technical terminology and difficult concepts when communicating with non-experts about matters within their own field of specialisation”. Similarly, at B2 level, learners should be able to “...explain technical topics within their field, using suitably non-technical language for a recipient who does not have specialist knowledge” (see Figure 5). The simplifying strategies are also highly pertinent: at C1 level, students are expected to “...make complex, challenging content more accessible by explaining difficult aspects more explicitly...”, while at B2 level, students should “...make concepts on subjects in their fields of interest more accessible by giving concrete examples...” (see Figure 5).

Figure 4

Some of the Mediation Descriptors Satisfied by this Assessment Task (Council of Europe, 2020, p. 198-241).

	Relaying specific information in speech or sign		Relaying specific information in writing
C1	Can explain (in Language B) the relevance of specific information found in a particular section of a long, complex text (in Language A).	C1	No descriptors available: see B2+
B2	Can relay (in Language B) the main point(s) contained in formal correspondence and/or reports (in Language A) on general subjects and on subjects related to their fields of interest.	B2+	Can relay in writing (in Language B) the relevant point(s) contained in propositionally complex but well-structured texts (in Language A) within their fields of professional, academic and personal interest.
	Explaining data (in graphs, diagrams, etc.) in speech or sign		Explaining data (in graphs, diagrams, etc.) in writing
C1	Can interpret and describe clearly and reliably (in Language B) the salient points and details contained in complex diagrams and other visually organised information (with text in Language A) on complex academic or professional topics.	C1	Can interpret and present clearly and reliably in writing (in Language B) the salient, relevant points contained in complex diagrams and other visually organised data (with text in Language A) on complex academic or professional topics.
B2	Can interpret and describe reliably (in Language B) detailed information contained in complex diagrams, charts and other visually organised information (with text in Language A) on topics in their fields of interest.	B2	Can interpret and present reliably in writing (in Language B) detailed information from diagrams and visually organised data in their fields of interest (with text in Language A).
	Processing text in speech or sign		Processing text in writing
C1	Can summarise (in Language B) discussion (in Language A) on matters within their academic or professional competence, elaborating and weighing up different points of view and identifying the most significant points.	C1	Can summarise in writing (in Language B) long, complex texts (in Language A), interpreting the content appropriately, provided they can occasionally check the precise meaning of unusual, technical terms.
B2	Can summarise (in Language B) the important points made in longer, complex texts (in Language A) on subjects of current interest, including their fields of special interest.	B2	Can summarise in writing (in Language B) the main content of complex texts (in Language A) on subjects related to their fields of interest and specialisation.
	Translating a written text in speech or sign		Translating written text in writing
C1	Can provide fluent oral translation (into Language B) of complex written texts (written in Language A) on a wide range of general and specialised topics, capturing most nuances.	C1	Can translate (into Language B) abstract texts on social, academic and professional subjects in their field (written in Language A), successfully conveying evaluative aspects and arguments, including many of the implications associated with them, though some expression may be over-influenced by the original.
B2+	Can provide oral translation (into Language B) of complex texts (written in Language A) containing information and arguments on subjects within their fields of professional, academic and personal interest.	B2	Can produce translations (into Language B) that closely follow the sentence and paragraph structure of the original text (in Language A), conveying the main points of the source text accurately, though the translation may read awkwardly.

Figure 5

Some Relevant Descriptors of Mediation Strategies (Council of Europe, 2020, p. 119-122)

		Strategies to explain a new concept		
		Linking to previous knowledge	Adapting language	Breaking down complicated information
C1		Can spontaneously pose a series of questions to encourage people to think about their prior knowledge of an abstract issue and to help them establish a link to what is going to be explained.	Can explain technical terminology and difficult concepts when communicating with non-experts about matters within their own field of specialisation. Can adapt their language (e.g. syntax, idiomaticity, jargon) in order to make a complex specialist topic accessible to recipients who are not familiar with it. Can paraphrase and interpret complex, technical texts, using suitably non-technical language for a recipient who does not have specialist knowledge.	Can facilitate understanding of a complex issue by highlighting and categorising the main points, presenting them in a logically connected pattern, and reinforcing the message by repeating the key aspects in different ways.
	B2	Can clearly explain the connections between the goals of the session and the personal or professional interests and experiences of the participant(s).	Can explain technical topics within their field, using suitably non-technical language for a recipient who does not have specialist knowledge. Can make a specific, complex piece of information in their field clearer and more explicit for others by paraphrasing it in simpler language.	Can make a complicated issue easier to understand by presenting the components of the argument separately.
		Strategies to simplify a text		
		Amplifying a dense text	Streamlining a text	
C1		Can make complex, challenging content more accessible by explaining difficult aspects more explicitly and adding helpful detail. Can make the main points contained in a complex text more accessible to the target audience by adding redundancy, explaining, and modifying style and register.	Can reorganise a complex source text in order to focus on the points of most relevance to the target audience.	
	B2	Can make the content of a text on a subject in their fields of interest more accessible to a target audience by adding examples, reasoning and explanatory comments. Can make concepts on subjects in their fields of interest more accessible by giving concrete examples, recapitulating step by step and repeating the main points. Can make new information more accessible by using repetition and adding illustrations.	Can simplify a source text by excluding non-relevant or repetitive information and taking into consideration the intended audience. Can edit a source text by deleting the parts that do not add new information that is relevant for a given audience in order to make the significant content more accessible for them. Can identify related or repeated information in different parts of a text and merge it in order to make the essential message clearer.	

Summary of RQ3: (To what extent do digital, multimodal assessment tasks develop the students as social agents?)

The Council of Europe has redesigned their CEFR assessment criteria to include plurilingualism and mediation, and the skills demonstrated throughout this multimodal assessment task are very closely aligned with a great number of the related descriptors. For example, students should be able to “... interpret and describe reliably (in Language B) detailed information contained in complex diagrams, charts and other visually organised information (with text in Language A)” as well as “...explain technical terminology and difficult concepts when communicating with non-experts about matters within their own field of specialisation” (Council of Europe, 2020, p. 119-122). In this sense, digital, multimodal assessment tasks such as this one contribute considerably to developing the students as social agents through “...the constant movement between the individual and social level in language learning” (Council of Europe, 2020, p. 36), as well as in the screenings and uploadings of their work.

5 Conclusion

Despite the ubiquity of multimodal communication, the skills involved are largely neglected in tertiary TESOL classrooms in Germany. The Council of Europe has attempted to address this with new competences, claiming that “...tasks in the language classroom should involve communicative language activities... that also occur in the real world” (Council of Europe, 2020, p. 32). However, a reluctance

to move away from the ‘four skills’ persists. This case study is an attempt to demonstrate the usefulness of multimodal assessment tasks by examining the results of one in close detail. Returning again to the research questions posed by this paper, it has been demonstrated that this task can very effectively promote the communication of technical concepts (RQ1) and provide opportunities for nuanced meaning making in English (RQ2) while simultaneously developing the students as social agents (RQ3). More broadly, tasks such as these can prepare students for their disciplinary studies and the job markets of the future, while also helping them become CEFR accredited, especially in terms of their mediation skills.

References

- Alvermann, D. (2017). The M word: dare we use it? *Journal of Adolescent and Adult Literacy*, 61(1), 99-102.
- Australian, Curriculum, Assessment and Reporting Authority (ACARA). (2015). Australian Curriculum - “Multimedia: English” Retrieved from <https://www.australiancurriculum.edu.au/resources/curriculum-connections/dimensions/?Id=67512&YearLevels=42681&searchTerm=ACELA1433>
- Bateman, J., Wildfeuer, J., & Hiippala, T. (2017). *Multimodality: Foundations, research and analysis – A problem-oriented introduction*. Walter de Gruyter GmbH.
- Bateman, J., & Schmidt-Borcherding, F. (2018). The communicative effectiveness of education videos: Towards an empirically-motivated multimodal account. *Multimodal Technologies and Interaction*, 2(3), 59.
- Bezemer, J., & Kress, G. (2008). Writing in multimodal texts: A social semiotic account of designs for learning. *Written Communication*, 25(2), 166-195.
- Burn, A., & Parker, D. (2001). Making your mark: Digital inscription, animation, and a new visual semiotic. *Education, Communication & Information*, 1(2), 155-179.
- Cope, B., & Kalantzis, M. (2009). A grammar of multimodality. *International Journal of Learning*, 16(2), 361-425.
- Council of Europe (2020). Common European Framework of Reference for Languages: Learning, teaching, assessment – Companion volume. Council of Europe Publishing, retrieved from www.coe.int/lang-cefr.
- Council of Europe (2016). Common European Framework of Reference for Languages: Learning, teaching, assessment – Developing illustrative descriptors of aspects of mediation for the CEFR. Council of Europe Publishing, retrieved from www.coe.int/lang-cefr.
- Derewianka, B., & Jones, P. (2016). *Teaching language in context* (2nd edition). Oxford University Press.
- Early, M., Kendrick, M., & Potts, D. (2015). Multimodality: Out from the margins of English language teaching. *TESOL Quarterly*, 49(3), 447–60.
- EUMade4All (2017). *Integrating multimodal digital literacy and English for international communication*. Retrieved from <https://www.eumade4ll.eu/>.
- Fairs, M. (2015). “Zaha Hadid is resentful and wronged” says Stephen Bayley. *Dezeen*. Retrieved at <https://www.dezeen.com/2015/10/05/zaha-hadid-resentful-and-wronged-says-stephen-bayley/>
- Gallagher, T., & Rowsell, J. (2017). Untangling binaries: Where Canada sits in the ‘21st Century Debate’. *McGill Journal of Education/Revue des sciences de l'éducation de McGill*, 52(2), 383-407.
- Grapin, S. (2019). Multimodality in the new content standards era: Implications for English learners. *TESOL Quarterly*, 53(1), 30-55.
- Halliday, M. A. K. (1994). *An introduction to functional grammar*. Hodder Arnold.
- Hafner, C. (2015). Remix culture and English language teaching: The expression of learner voice in

- digital multimodal compositions. *TESOL Quarterly*, 49(3), 486–509.
- Hodge, R., & Kress, G. (1988). *Social semiotics*. Cornell University Press.
- Huang, S. Y. (2019). A critical multimodal framework for reading and analyzing pedagogical materials. *English Teaching: Practice & Critique*, 18(1), 52-69.
- Iedema, R. (2003). Multimodality, resemiotization: Extending the analysis of discourse as multi-semiotic practice. *Visual Communication*, 2(1), 29-57.
- Jewitt, C., Kress, G., Ogborn, J., & Tsatsarelis, C. (2001). Exploring learning through visual, actional and linguistic communication: The multimodal environment of a science classroom. *Educational Review*, 53(1), 5-18.
- Jewitt, C. (2005). Multimodality, “reading”, and “writing” for the 21st century. *Discourse: Studies in the Cultural Politics of Education*, 26(3), 315-331.
- Jewitt, C. (Ed.). (2009). *The Routledge handbook of multimodal analysis*. Routledge.
- Jiang, L., & Luk, J. (2016). Multimodal composing as a learning activity in English classrooms: inquiring into the sources of its motivational capacity. *System*, 59, 1–11.
- Knight, N. (2010). Wrinkling complexity: Concepts of identity and affiliation in humour. In M. Bednarek & J. R. Martin (Eds.), *New discourse on language: Functional perspectives on multimodality, identity and affiliation* (pp. 35–58). Continuum.
- Knobel, M., & Lankshear, C. (2014). Studying new literacies. *Journal of Adolescent and Adult Literacy*, 58(2), 97-101.
- Kress, G., (2010). *Multimodality: A social semiotic approach to contemporary communication*. Routledge
- Kress, G., & van Leeuwen, T. (1996). *Reading images: The grammar of visual design*. Routledge.
- Lamb, S., Maire, Q., & Doecke, E. (2017). *Key skills for the 21st Century: An evidence-based review*. Project Report. NSW Department of Education, Sydney. Retrieved at <http://vuir.vu.edu.au/id/eprint/35865>
- Lapp, D., & Fisher, D. (Eds.). (2011). *Handbook of research on teaching the English language arts: Co-sponsored by the International Reading Association and the National Council of Teachers of English*. Routledge.
- Lee, H-C. (2014). Using an arts-integrated multimodal approach to promote English learning: A case study of two Taiwanese junior college students. *English Teaching: Practice and Critique*, 13(2), 55–67.
- Lotherington, H. and Jenson, J. (2011). Teaching multimodal and digital literacy in L2 settings: New literacies, new basics, new pedagogies. *Annual Review of Applied Linguistics*, 31, 226-246.
- Macnaught, L., Maton, K., Martin, J. R., & Matruglio, E. (2013). Jointly constructing semantic waves: Implications for teacher training. *Linguistics and Education*, 24(1), 50-63.
- Martin, J. R. (2009). Genre and language learning: A social semiotic perspective. *Linguistics and Education*, 20(1), 10-21.
- McGinnis, T. A. (2007). Khmer rap boys, X-Men, Asia’s fruits, and Dragonball Z: Creating multilingual and multimodal classroom contexts. *Journal of Adolescent & Adult Literacy*, 50(7), 570-579.
- Moreno, A., Navarro, C., Tench, R., & Zeffass, A. (2015). Does social media usage matter? An analysis of online practices and digital media perceptions of communication practitioners in Europe. *Public Relations Review*, 41(2), 242-253.
- Nash, B. (2018). Exploring multimodal writing in secondary English classrooms: A literature review. *English Teaching*, 17(4), 342-356.
- Nielsen, W., Georgiou, H., Turney, A., & Jones, P. (2016). Analysing student-generated digital explanations. In L. E. Dyson, W. Ng & J. Fergusson, (Eds.). *Mobile learning futures: Sustaining*

- quality research and practice in mobile learning*. University of Technology, Sydney.
- New London Group (1996). A pedagogy of multiliteracies: Designing social futures. *Harvard Educational Review*, 66, 60-92.
- Oldakowski, T. (2014). A multimodal assignment that enriches literacy learning: The problem. *InSight: A Journal of Scholarly Teaching*, (9), 70-77.
- Pandya, J.Z., Pagdilao, K.C., Kim, A.E., & Marquez, E.M. (2015). Transnational children orchestrating competing voices in multimodal, digital autobiographies. *Teachers College Record*, 117(7), 1-32.
- Rose, D., & Martin, J. R. (2012). *Learning to write, reading to learn: Genre, knowledge & pedagogy in the Sydney school*. CT Equinox.
- Turney, A., & Jones, P. (2021). Approaching complex multimodal phenomena in educational settings-insights from theory. *The Australian Journal of Language and Literacy*, 44(1), 29-45.
- Wilke, B. (2012). Transformationen im Zusammenspiel von Bild und Text in Online-Reiseführern: Schlussfolgerungen für die Didaktik. *Testi e Linguaggi*, (6), 293-310.
- World Economic Forum (WEF). (2018). *The future of jobs report*. World Economic Forum Publishing.
- Zacchi, V. J. (2016). Multimodality, mass migration and English language teaching. *Revista Brasileira de Linguística Aplicada*, 16(4), 595-622.

Anne FJ Hellwig is a PhD candidate in the School of Education at the University of Wollongong, Australia. ORCID: 0000-0002-5996-4088