Check for updates

Multimodality and English for Special Purposes: Signification and Transduction in Architecture and Civil Engineering Models

Anne F. J. Hellwig*, Pauline T. Jones, Erika Matruglio and Helen Georgiou

School of Education, Faculty of the Arts, Social Sciences and Humanities, University of Wollongong, Wollongong, NSW, Australia

The applied disciplines of architecture and civil engineering require students to communicate multimodally, and to manipulate meaning across media and modes, such as image, writing or moving image. In their disciplinary studies for example, students must be able to transform the language of lectures and textbooks into models and diagrams. In their future workplaces, they will commonly be required to transform reports and legal documents into floor plans and digital & physical 3D models. Such multimodal literacy, however, is not typically reflected in their related subject-specific English language courses, especially in Germany, where a text-centric approach is favored. To better reflect the demands placed upon them, students in two courses of English for Architecture and Civil Engineering were tasked with creating digital, multimodal artifacts to explain a concept from either of these fields to a lay audience. The resultant artifacts used a wide variety of semiotic resources to make meaning, including a total of 26 separate architectural and civil engineering models. This is a quantity sufficiently large enough to invite closer examination, and also reflects the important role models play in the fields of architecture and civil engineering, both at university and in the workplace. This paper suggests that models of this kind exist within a system of signs, in which meaning is created in the relationships between the signs. The process of transforming one resource into another also invites the consideration of the artifacts in terms of the notion of "transduction", to discern how meaning changes between contexts, practices and modes and to contribute to existing literature on multimodal texts in tertiary education, particularly within a language-learning context.

Keywords: multimodality, English-second language, transduction, signification, architecture and civil engineering, models, education, systemic functional approach to multimodal discourse analysis

INTRODUCTION

All communication is inherently multimodal. Even a printed text communicates extra-linguistically through choices in font, spacing and layout (Kress, 2010). Improvements in the accessibility and affordability of digital technologies, however, have encouraged an increase in the creation of multimodal artifacts. Multimodal artifacts are texts that simultaneously employ a number of communicative modes. In this paper, we conceive of a mode as including, for example, speech,

writing, image, sound and gesture (Bezemer and Kress, 2008; Kress and van Leeuwen, 2021). These modes may occur in a print, screen or live media. Meanings are made in a variety of different modes, that is, meaning is always multimodal, and modes have differing "modal resources" or affordances (Gibson, 1977). While speech and writing, for example, are shaped by grammar and syntax, image is framed by considerations of size, color and shape (Bezemer and Kress, 2008, see below for further discussion). As multimodal communication becomes more common, students in the 21st century need to be able to communicate multimodally in order to be considered fully literate (Jewitt, 2009). This shift to increasing multimodal practice is not well reflected in educational settings in Germany, however, especially in the English language courses offered at university (Wilke, 2012). This is even the case for the English for Special Purposes courses which emphasize multimodal communication. Although students are expected to build physical and digital 3D models, create floorplans and use other forms of communication in their discipline-specific courses and future workplaces, their related English language courses largely ignore multimodal literacy in favor of focusing on language: the "four skills" of reading, writing, listening and speaking. The Council of Europe's Common European Framework of Reference for Languages (CEFR) – the preeminent framework for measuring language acquisition in Europe – has recently attempted to address this by deemphasising the four skills and incorporating "mediation" into the descriptors (Council of Europe, 2020). Mediation involves "…linguistic and semiotic reformulation… (of) terms, texts and discourse genres" (Council of Europe, 2015; p. 28), which bears a similarity to transduction, the movement of meaning between (groups of) modes (Bezemer and Kress, 2008)[1]. However, multimodal assessment tasks remain largely ignored in the German tertiary context.

In light of these shifts and challenges, this paper examines a multimodal assessment task assigned to students of two English for Architects and Civil Engineers courses at a university in Germany. Students were tasked with creating a 3–5-min video composition (VC) explaining a technical concept from the fields of architecture or civil engineering to a non-specialist audience using a variety of modes. Eighteen of these students subsequently agreed to participate in the research project reported on here. The data-set is comprised of the VCs and two collections of interviews, one with the student creator about the processes behind composing their artifact and another probing their interpretations of one artifact as an audience member.

The students employed text, images, graphological elements, video, animation, charts, formulae, diagrams, gesture, music, models, performance and a "prelude", or "upbeat" of silence in the creation of the 18 artifacts. The most intriguing of these were the models that are the products of architectural and civil engineering communities (Jewitt et al., 2016). These models were typically hand-constructed from a variety of 3D objects, images and text, and were employed in twelve of the eighteen VCs,

---

[1] The notion of mediation also has parallels to other concepts such as 'recontextualisation' (see Bezemer and Kress, 2008) and 'resemiotization' (see Iedema, 2003). These concepts, however, are beyond the scope of this paper.

making for a total of twenty-six models in the data-set. This frequency encourages closer investigation and raises the question of why so many students chose to make meaning in this way. This paper will examine how models mean within a system of signs, unpack some of the affordances of such models and explore what is gained and lost when transducing (Bezemer and Kress, 2008) from other modes to models of this kind.

BACKGROUND
Literature Review
Multimodal Assessment Tasks

Jewitt et al. (2001) argue that social semiotics allows us to see "…the process of learning as a *dynamic process of sign-making*" (p. 27, emphasis theirs). 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., 2015) and better prepare learners for future communication demands (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 emphasized 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 across a range of professional, educational and community contexts (Lotherington and Jenson, 2011). Multimodal projects using digital technologies can be found at all levels of education for at least the past twenty years, whether at primary (Burn and Parker, 2001 etc.), secondary (Nash, 2018 etc.) or tertiary levels. From the tertiary perspective, Georgiou (2020) suggests that multimodal, digital artifacts allow for "authentic and holistic assessment" (p. 13) and Nielsen et al. (2020) argue they "demonstrate different sorts of capabilities than are normally assessed" (p. 2391) (Turney and Jones, 2021). Guo (2004) p. 215, agrees, suggesting that ESP educators should emphasize multimodal, literacy practices "…to better prepare our students for their current and future academic and professional life".

It is no surprise then, that multimodal assessment tasks have been explicitly included in curricula all around the world, such as in the United States (Lapp and Fisher, 2011), Europe (Sindoni et al., 2019) and Australia (Australian, Curriculum, Assessment and Reporting Authority (ACARA), 2015). However, Early et al. (2015) suggest that multimodality is on the margins of the community, and Lotherington and Jenson (2011) claim that teachers of English as an Additional Language of Dialect (EAL/D) have been reluctant to embrace or even acknowledge multimodal literacy. Prior (2013) suggests that the "four-mode scheme" or four-skills approach - focused on reading, writing, listening and speaking - has dominated education at the expense of embodied, multimodal practice. Royce (2002) concurs, suggesting that multimodal communicative competence needs to be explicitly fostered in EAL/D classrooms. There is clearly an urgent need, then, for EAL/D educators, and ESP educators in particular, to embrace digital, multimodal assessment tasks.

Digital Multimodal Composing

In light of this, there has been an increasing emphasis placed upon the practice of digital multimodal composing (DMC) in tertiary, second language (L2) classrooms (Hafner, 2015). DMC is a process where student composers integrate multiple, digital tools and semiotic resources to produce a text (Jiang, 2017). Studies have shown that it has far-reaching pedagogical advantages, especially within an L2 context. Kim and Belcher (2020), working in a South Korean university writing course, compared learner outcomes when writing traditional essays to DMC in terms of three variables: syntactic complexity, syntactic accuracy and learners' perceptions of the tasks. They found that although the essays were generally more syntactically complex, they were not more accurate, and students found DMC to be more engaging and motivating. Hung (2019), working with digital storytelling at a university in Taiwan, found that DMC of this kind measurably improves a variety of L2 acquisition skills, including analyzing, evaluating and organizing information. Studies from around the world have corroborated these positive outcomes from DMC. Jiang and Luk (2016), writing from a Chinese tertiary context, found that multimodal assessment tasks increase students' "motivational capacity", while in Taiwan, Lee (2014) suggests that multimodal learning enhances students' motivation and self-confidence. At a university in the US, Dzekoe (2017) found that DMC improved students' L2 writing by helping them identify problem areas and increasing their fluency, while in the Dutch context, Vandommele et al. (2017) also found that DMC can promote the writing skills of L2 learners.

Within an ESP context, studies show similarly positive outcomes, especially in regard to the disciplinary implications of DMC. Hafner and Miller (2011, 2018), who tasked their students with creating digital, multimodal, scientific documentaries in an English for science course, suggest that DMC can engage learners with challenging and authentic opportunities for English-language use, particularly in their disciplinary studies. Kohnke et al. (2021), working with digital infographics in a discipline-specific English course at a Hong Kong University, concur. They suggest that DMC of this kind help students feel more confident demonstrating their discipline-specific language skills in professional contexts and prepare them for their future workplaces. Kohnke (2019), working this time with digital comic strips in an ESP context, found that this sort of DMC promoted collaborative learning and student engagement. There is a gap in the literature, however, around the effectiveness of DMC in the ESP courses specific to the disciplines of architecture and civil engineering. As Douglas (2013) argues, ESP language is precise and conforms closely to specific purposes, and more work needs to be done if educators are to better understand the impact of DMC in the context of English for Architects and Civil Engineers.

Theoretical Background
Multimodality

In this paper, we refer to multimodality as the combination of different modes for the purposes of constructing our experience, enacting social relations and organizing meaning (Jewitt et al., 2016). Many researchers draw on social semiotics in which a "mode" is seen as a socioculturally-shaped set of semiotic resources for meaning-making (Kress, 2010), composed to represent different kinds of knowledge and enact different social relationships and identities. Examples of modes include, but are not limited to, writing (O'Halloran, 2008 etc.), images (Painter et al., 2012 etc.), music (Barton and Unsworth, 2014), film (Bateman et al., 2017 etc.) and architecture (Ravelli and McMurtrie, 2015). A mode differs from a semiotic resource in that a mode has "…organizing principles that are recognized within a community" (Jewitt et al., 2016, 71). A semiotic resource is less socially organized, and includes such things as gaze, dress or scent. One way of conceptualizing this distinction is to consider gesture. In the context of language-driven communication, gesture typically functions as a semiotic resource, adding meaning to the conversation, with language as the dominant system ("paralinguistics") (Ariztimuño et al., 2022). However, in other communities, gesture communicates in an organized way and is therefore considered to be a mode. For example, the hearing-impaired have developed gesture into many complex systems, such as Auslan and other sign languages. Similarly, doubles-players of tennis have also advanced gesture from semiotic resource to semiotic mode by developing a principled system of hand signals to communicate serve types to their partners, such as a "body serve" or "wide serve".

The boundaries between modes, however, are notoriously slippery. Certain questions arise, such as whether "image" is a mode in itself, or if a distinction should instead be made between image types, such as a photograph or a charcoal sketch. Even within a photograph there are different communicative elements, such as the choice of black & white or color, and further still, within the realm of color, choices around brightness, saturation and so on can be made (Kress and van Leeuwen, 2021). Perhaps a "mode" will always resist exhaustive classification and depend "…on sign makers acting within the needs and understanding of a particular community" (Bezemer and Kress, 2008, p. 171). In the absence of universal guidelines, then, and in accordance with the work of Bateman et al. (2017), this paper conceptualizes 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" (p. 119). That is, the definitional boundaries of a mode are determined by the opportunities and constraints of its material context – what the medium affords – and also in terms of the discourse community within in which it functions. Once the mode has been identified, it is then useful to think of modes as having different "modal resources" (Bezemer and Kress, 2008). Much like semiotic resources, modal resources include communicative elements, such as framing, composition and color, which affect the meanings that are created and connoted (Kress and van Leeuwen, 2021).

Transduction

The frequency of models in this data-set invites the consideration of the artifacts in terms of the "epistemological commitment" necessitated by the changing of modes. An epistemological commitment is an "unavoidable affordance" imposed by the second mode and not present in the first mode from which meaning was transduced (Bezemer and Kress, 2008). Such

commitments have pedagogical and semiotic implications as they require an imaginative elaboration that both constrains and expands meaning potential. The process of transduction engenders different patterns of gains and losses with the movement of meaning between modes. This has partly to do with changes in the materiality as well as the media or "canvas" (Bateman et al., 2017) upon which the mode is realized, and also with the varying cultural histories that have led to the creation of a particular mode. Different resources are also made available to makers of different modes, and these have "…implications for generality and specificity …(and) the arrangement of constituents" (Bezemer and Kress, 2008, p. 182). For example, when meaning is transduced from still to moving image, decisions around temporality must now be made. "Elements can now appear and disappear, and through that, movement can be suggested" (Bezemer and Kress, 2008, p. 182). This will be explored in more detail in Section Transduction Between Text, Image and Diagram to Model below. An additional layer of theoretical complexity is added by the process of moving between a first language to the target language of English (Canagarajah, 2012; Smith et al., 2017), as well as by the interpersonal shifts in the social relationships of the student composers and their fellow classmates, the audience (Bernstein, 1996). However, a detailed exploration of these nuances is unfortunately beyond the scope of this paper. Instead, this paper will explore some of the practical and theoretical implications of model-making. Some of the ramifications of model-making in the context of EAL/D and ESP education will be elaborated in the discussion section. Model-making as a system of signs will then be explored, and the affordances of models examined in more detail, with particular attention paid to how meaning is transduced between models and other modes.

DATA

The cases within the study were drawn from two courses of "English for Architects and Civil Engineers A". This was a 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 C1-level English to graduate from a Master of Civil Engineering, while the architecture students had no such requirement. Although the course was an elective for all students, the civil engineering students needed to certify their English levels by successfully completing any C1 English course (or passing an internationally recognized C1 English exam) before they were able to receive their Master of Civil Engineering. As such, this course, with its content focused on the language of civil engineering and architecture, was a popular way of gaining this compulsory language accreditation. The architecture-track students were not obligated to certify their English levels before graduation, and therefore attended in lower numbers. Within

the course, students were required to successfully demonstrate listening, speaking, reading and writing skills, with each skill weighted at 20% of their grade. For the final 20%, students were tasked with producing a 3–5-min video composition (VC) explaining a concept from architecture or civil engineering to a non-specialized audience using a variety of modes. Students were asked to upload the artifacts to the sharing platform Moodle before they were screened in class, where they were encouraged to lead discussions around the VCs, offer feedback and ask questions of their fellow student creators. They were also invited to participate in this research project, which had two phases of data collection: the VCs themselves, and a two-part interview (exploring their perspectives as both composers and audience members). Seventeen of the 38 students agreed to participate in the first phase of the study (artifact collection), with 7 of those also agreeing to the second phase (the interview). An additional artifact was also included in the data-set (Student X's "Exemplar VC"), which was made by a student from a previous semester and was shown as an archetypal VC to all of the participants in this study. The semi-structured interviews lasted between 19:56 and 46:11 min and involved a pre-prepared interview protocol. The data were then investigated inductively, using NVivo to explore the extent to which the assessment task promoted the communication of technical concepts, facilitated nuanced opportunities for meaning making and developed the students as social agents (Hellwig, 2022).

In order to show the nuance and complexity of the 18 student-made artifacts, certain modes have been subcategorised (see **Table 1**). The degree track of each student is also noted as either architecture (A) or civil engineering (CE). The results reveal that all the students utilized the mode of text in spoken form as a voiceover ($n = 18/18$) and a considerable number also chose written text in the form of labels ($n = 16/18$), larger chunks of text ($n = 5/18$), or "live" writing ($n = 5/18$). Diagrams were the next most frequent modal selection ($n = 15/18$), followed by images (photos sourced from the internet, $n = 15/18$; photos taken by student, $n = 13/17$; drawings, $n = 5/18$) and architectural/civil engineering models ($n = 12/18$; see in **Appendix A** Table 1 for a detailed table of all the models). Charts ($n = 8/18$), graphological elements ($n = 8/18$) and gesture were also well represented, occurring in around half of the data-set. Formulae ($n = 5/18$), music ($n = 5/18$), moving image ($n = 2/18$), animation ($n = 1/18$), and performance ($n = 1/18$) were also present in smaller numbers, as was the unusual mode of "prelude" ($n = 1$), where an "upbeat" of silence was included for emphasis. In terms of textual features, most students included opening ($n = 17/18$) and closing phases (15/18) such as a title slide or some concluding text, e.g., "Thanks for watching", and around half employed transitional devices common to software editing suites, such as wipes and fades ($n = 8/18$).

DISCUSSION

The paper has both practical and theoretical implications. The first part of this discussion (Section From Text to Model) will explore the practical implications for model-making in the

TABLE 1 | Table of modes used.

			Mode 1: text				Mode 2: image			Mode 3: graphological element	Mode 4: video		Mode 5: animation	Mode 6: chart	Mode 7: formula	Mode 8: diagram	Mode 9: gesture	Mode 10: music	Mode 11: prelude	Mode 12: model		Mode 13: performance	Mode 14: textual feature		
	Modes used	A/CE	Text (spoken)	Text (written)	Text (labels etc)	Text (live writing)	Image (photo from internet)	Image (photo-student's own)	Image (drawing)	Graphological elements	Video (from internet)	Video (student's own)	Animation (from internet)	Chart	Formula	Diagram	Gesture	Music	Prelude	Model	Intervention with model	Performance	Textual features (wipes, fades etc)	Textual features (opener)	Textual features (closer)
X	6	CE	X		X		X	X	X	X		X				X	X	X		X	X		X	X	X
A	5	CE	X				X	X		X		X				X	X							X	
B	7	CE	X		X	X	X	X		X		X		X		X	X			X			X	X	X
C	6	A	X	X	X		X		X	X		X					X	X	X					X	
D	4	A	X			X						X				X									X
E	6	CE	X	X	X		X	X	X	X		X			X	X				X	X			X	X
F	3	CE	X	X	X		X							X										X	X
G	6	CE	X		X	X	X	X		X		X				X	X			X	X		X	X	X
H	6	CE	X		X							X		X	X	X				X	X			X	X
J	6	CE	X	X	X		X				X			X	X	X							X	X	X
K	4	CE	X	X	X		X									X				X			X	X	X
L	7	CE	X		X			X		X		X		X	X	X				X	X	X		X	X
M	4	CE	X		X		X					X				X				X	X		X	X	X
N	6	CE	X		X		X			X	X		X	X						X			X	X	X
O	6	A	X		X	X	X					X			X	X	X	X		X	X			X	X
P	4	CE	X		X		X							X		X							X	X	
Q	6	A	X		X		X		X			X				X	X	X		X	X			X	X
R	7	CE	X		X	X	X		X			X		X		X	X	X		X	X			X	X
	Total		18	5	16	5	15	6	5	8	2	13	1	8	5	15	7	5	1	12	9	1	8	17	15

context of EAL/D, particularly in terms of ESP. The second part of the discussion (Section Theoretical Implications: Model-Making as a System of Signs) will focus on the theoretical implications of model-making, and how the models can be conceived as a system of signs.

Practical Implications: Model-Making and English for Special Purposes

Designing and constructing models such as these involves a complex series of purposeful semiotic choices. The implications of these decision-making processes for language learning, particularly in the context of ESP, is broad. Four main implications are worth examining in more detail. Firstly, the data suggests that there is a clear link between the students' purposes, the choices they make in their models and the quality of their explanations. Their choices reflect what the student considers to be "criterial" (Kress, 2010, 70) and what aligns with their pedagogical purposes, and the quality of their explanations could be said to be improved when their decisions enhance the clarity of their explanations and deepen audience comprehension of the technical concepts involved. This involves the decision to choose a model over another way to communicate a concept, and also in using a model to share the semiotic labor of making meaning across modes. All models ($n = 18$) were shown in simultaneity with at least one other mode: usually text in the form of narration, but also other modes such as diagrams, as in the case of Student O (see **Figure 9**). These parallelisms support audience understanding and make the meanings of technical terminology clearer, especially for an EAL/D audience. For ESP learners in particular, this technical terminology is vital for successful participation in their disciplinary studies and future workplaces (Ehrenreich, 2010). By using models, and by sharing the work of making meaning across modes, the pedagogical purposes of these VCs are more effectively achieved.
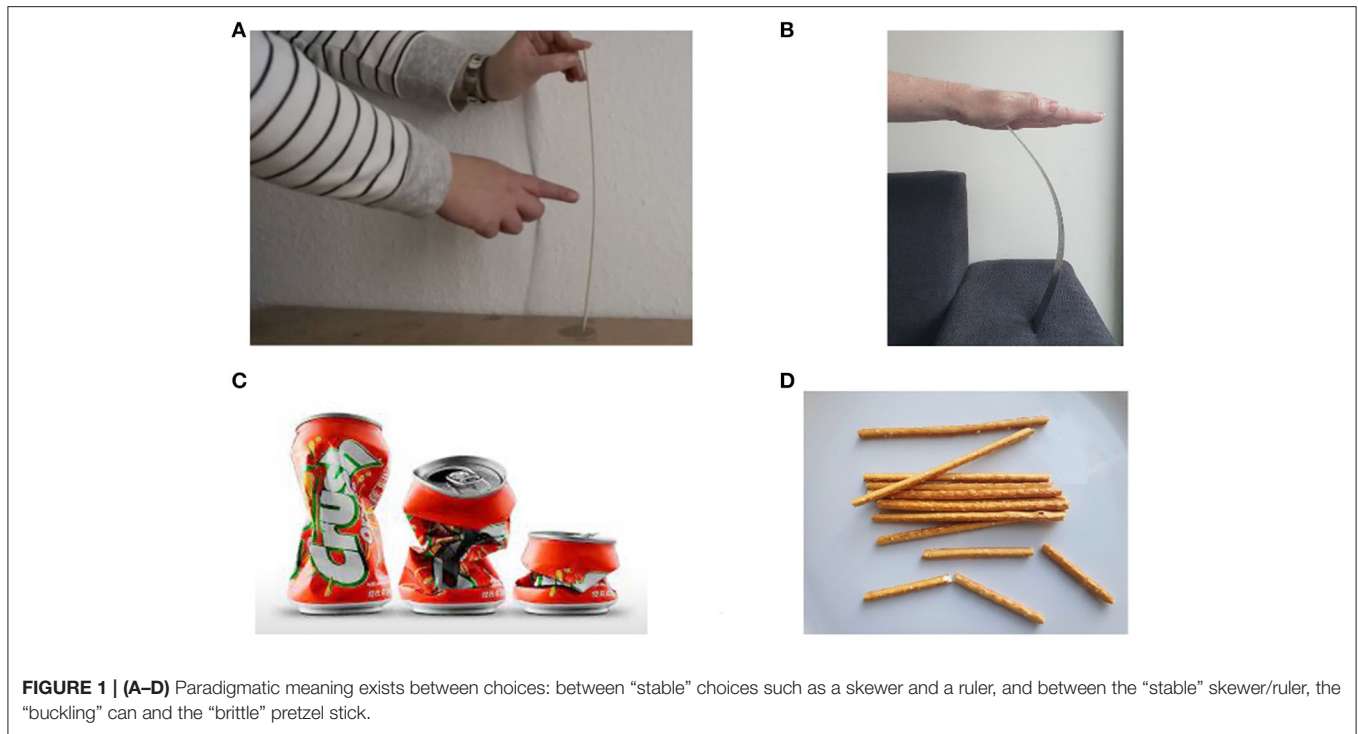
The link between purpose, semiotic choice and quality of explanation is also corroborated in the interview data collected after the VCs were submitted. As will be discussed in detail in Section Theoretical Implications: Model-Making as a System of Signs, Student E could have potentially used a plastic ruler instead of a bamboo skewer to illustrate the concept of "stability", as both objects share similar material properties suited to her explanatory purpose (see **Figure 1**). In her interview data, however, she explains that she decided against using a ruler for pedagogical reasons, commenting, "…I thought about using a ruler, but the ruler I have at home is, like, a clear plastic so… I thought it would be difficult to see". Similarly, Student G's choices were also motivated by the considerations of her audience. She claims that "I needed a model to put in there, because it makes it more understandable for others to follow" and because she herself had difficulty understanding this concept in her lectures. In the interview data she discusses how she wanted a model in addition to the diagrams and formulae she also used because the particular affordances of a model allowed her to show "…the forces, how they are carried out… How this works is the most important part of the video, I

think, and when I have the model you can imagine what it looks like".

Secondly, the contributions from this study corroborate the findings of many existing studies in the field. Multimodal assessment tasks allow students to demonstrate aptitudes not typically assessed (Nielsen et al., 2020), such as transducing complex concepts between modes. They also appear to increase students' motivation, as suggested by the findings of Jiang and Luk (2016) and Lee (2014). Student E claims that the task of making a VC motivated her and her classmates to produce a high standard of work despite this course being an elective ("you can really see how much thought everyone put into it… like, I do this course, like, for fun but I still didn't want to put something bad as a video and you can really see that everybody really thinks (the same way)"). Other students claimed to find the task enjoyable: Student A commented that "I'm enjoying (making the VC) right now because that it breaks the old rules of school", while Student C claimed, "I had a lot of fun doing this". Interestingly, the interview data also suggests that the students found value and perhaps an increased sense of self-worth in the task. Student A remarked that "It means a lot to me to make something like this", Student C reflected that "the video is a little piece of my heart", and Student E thought "it's something to be proud of". Their enthusiasm was not limited to the creative process of making VCs, but also to the screenings of the other VCs in class. Student A commented, "I love it too, to see the other videos. How far can our imagination go? That's what's really exciting!".

Thirdly, multimodal assessment tasks such as this one can facilitate more effective assessment of complex skills. As mentioned previously, 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 (Hellwig, 2022). For example, in order for students to become CEFR accredited, they are now expected to be able to "…interpret and describe reliably (in Language B) detailed information contained in complex diagrams, charts and other visually organized information (with text in Language A)". Diagrams, charts and other visual modes are employed frequently in the VCs (see **Table 1**), and the models in particular are complex examples of "visually organized information" that requires nuanced interpretation. Another descriptor asks students to "…explain technical terminology and difficult concepts when communicating with non-experts about matters within their own field of specialization" (Council of Europe, 2020, 119-122). Tasks aimed at explaining technical concepts to lay audiences using a variety of modes would satisfy both of these descriptors and allow teachers to grant CEFR accreditation in a principled and measurable way.

Finally, and in accordance with Georgiou (2020), this study suggests that the use of multimodal, digital artifacts – and the use of models in particular – increase the authenticity of assessment tasks. Digital and physical 3D models are frequently used in the workplaces of civil engineers and architects and including them in English language classrooms reflects existing professional practices. Although it should be noted that the models used in these workplaces are typically not pedagogical in

FIGURE 1 | (A–D) Paradigmatic meaning exists between choices: between "stable" choices such as a skewer and a ruler, and between the "stable" skewer/ruler, the "buckling" can and the "brittle" pretzel stick.

purpose, models are nevertheless used to communicate with all stakeholders involved in the construction process. For example, models are used by architects to persuade clients to invest; by suppliers to explain product information to architects and civil engineers; between architects and civil engineers to communicate ideas between themselves; and by architects and civil engineers to inform the public of upcoming works. In this sense, then, making meaning with and connoting meaning from models is a valuable skill worth developing in the ESP classroom, especially in courses aimed at architects and civil engineers.

Theoretical Implications: Model-Making as a System of Signs

Although Ravelli and McMurtrie (2015) drawing upon the work of O'Toole (1994), have written extensively about architectural space, minimal attention has been paid to architectural and civil engineering models and the semiotic choices involved in their design. "Model" refers here to hand-constructed physical models, although one student included a rendering, that is, a still image of a 3D digital model (for a table outlining the models in detail, see **Appendix A**). The models were constructed from a variety of 3D objects, images, graphological elements (e.g., plus signs) and text. As such, they could be seen as "modes comprised of modes" or "…patterns of patterns" (Martin, 2011, 246), arranged in layers of increasing abstraction, with the strata related by metaredundancy. This suggests that one stratum realizes the next, with the patterns at one level redounding with (coming back upon) patterns at the next level (Martin and Rose, 2007). Following Martin (2011), this paper conceives of the models as composed of three ranks: "element", or model

part; "ensemble", or the model itself; and "intervention", where the student interacts performatively with the model to create a series of meaning-making moments syntagmatically, analogous to stringing clauses into sentences ($n = 9$, see **Figure 2**). In order to understand how the models mean, this paper follows Martin (2011), who uses Saussure (1983) conception of signification. Such an approach suggests that in addition to meaning being made syntagmatically, meaning is also made paradigmatically (through difference, or "*valeur*"), and models mean through expressions of choices that have been made from within a system of signs. In other words, the models themselves do not mean anything; rather their meanings are made in the relationships between the choices that were made and the choices that could have been made but were not.

In order to demonstrate *valeur* more clearly, this paper will touch briefly upon an example from the data. This example, a model drawn from Student E's VC, is used to demonstrate the technical concept of "stability". Stability is the ability of a structural component to resist compressive (glossed here as "downward") forces, either by not "buckling"[2] at all, or by deforming under pressure but returning to its original form after the forces are removed. Student E chooses to demonstrate this concept by exerting downward force on a bamboo cooking skewer (see **Figure 1A**). The meaning of "stability" is not contained within the skewer, however. Instead, the meaning is made in relation to other objects with similar properties that were not chosen but could have been. For example, she could have

---

[2]Buckling is a sudden, longitudinal deformation caused by high compressive stress on a structural component (e.g. when a column bends sideways under a heavy snowfall).

FIGURE 2 | The system of models with realization statements.
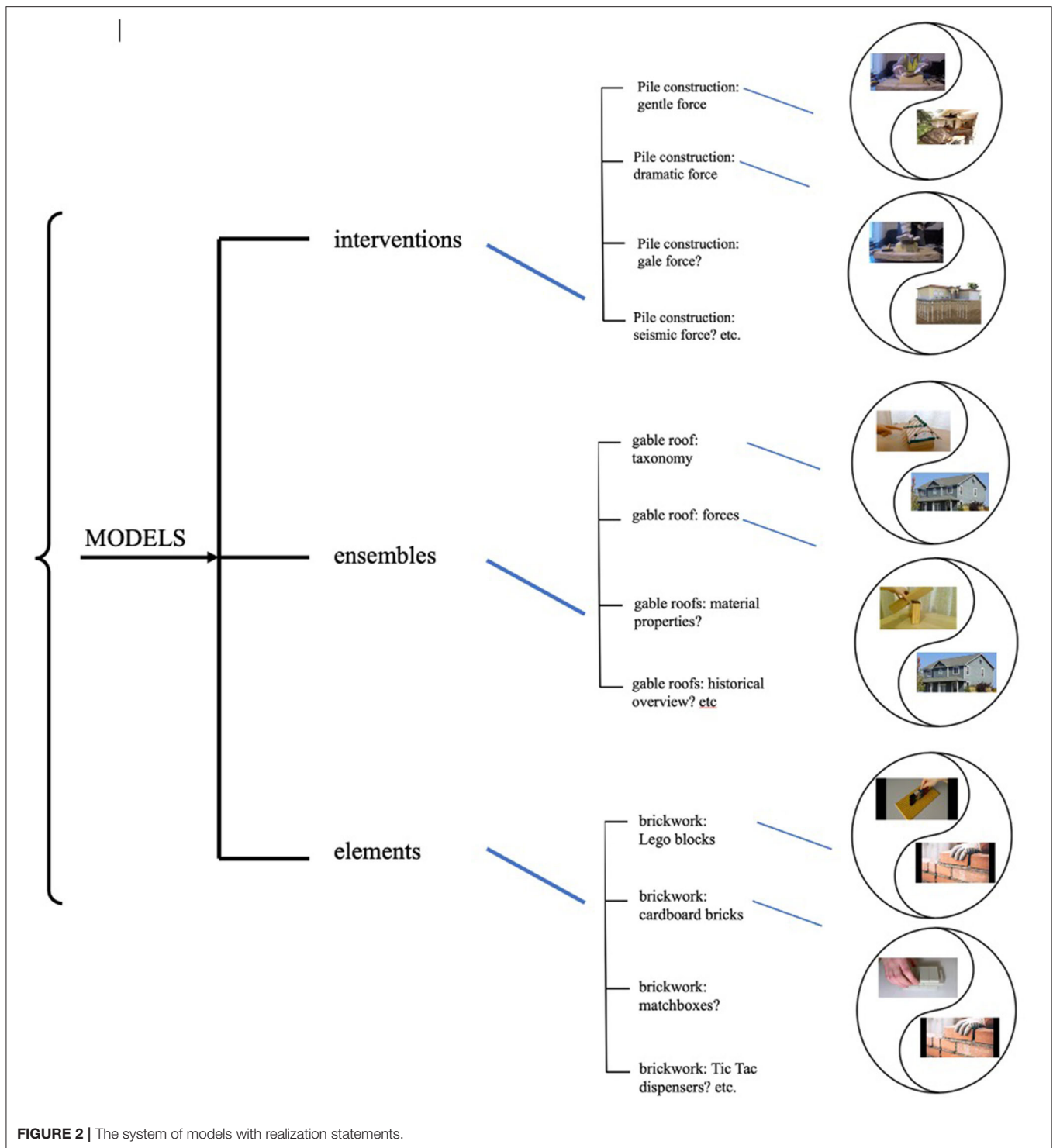
chosen to demonstrate "stability" with a metal ruler because it also returns to its original shape after the force is removed (see **Figure 1B**). However, an object that would remain bent or snap under pressure would not have been suitable to her purposes. To visualize this more clearly, imagine three outcomes for an object under downward forces. An object could buckle and return

to form ("stability"), buckle and remain deformed ("buckling"), or experience sudden collapse ("brittleness"). Now imagine that Student E had also chosen to demonstrate these three outcomes by finding suitable household objects to illustrate each outcome and applying pressure to each one in turn. She might have chosen, for example, an empty drink can for "buckling" or a pretzel

stick for "brittleness" (see **Figures 1C,D**). Even though Student E only used one model in her VC, the meaning of her model could be said to exist between this model and the ones not used, between "skewer/stability", "drink can/buckling" and "pretzel stick/brittleness". As such, her meaning is made paradigmatically and exists between choices in a system, even if the other choices are not immediately visible.

This paper will further elaborate upon this by showing how *valeur* is observable at all three ranks. Firstly, three models drawn from three different VCs will be used to demonstrate *valeur* at each rank, before focusing in more detail upon models used in one VC, Student O's "The typical Palladian villa design". For reference, **Appendix A** is a table of all the models in the data-set and what is harnessed at the ranks of element, ensemble and intervention. At the rank of "element", *valeur* exists between model parts employed in different models for similar purposes, and those that were not chosen but could have been. Both Student M and Student Q, for example, constructed models to explain brickwork (see **Figure 3**), and paradigmatic meaning exists between Student M's Lego blocks, Student Q's hand-built cardboard bricks and any number of other possible semiotic choices with similar material properties, such as matchboxes or Tic Tac dispensers.

Similarly, *valeur* is observable at the rank of ensemble, where models communicating similar concepts are analyzed in contrast to each other and to what could have been constructed but was not. Student G, for example, made two different models of gable roofs: the first model consists of fourteen wooden skewers, two set squares, an amount of plasticine and two books while the second model is constructed of a piece of foam and a travel chess board (see **Figure 4**). This means that the elements of skewer and foam exist in paradigmatic relation to each other as alternative representations of rafters, as do the chessboard and books, which connote the purlins (horizontal roof beams) of the structure. However, the models also make meaning at the rank of ensemble, with each model serving a different pedagogical purpose. The first model offers a classifying taxonomy of a gable roof, where each skewer represents a different technical feature, such as a "ridge board", a "top purlin", a "middle purlin" and so on. Student G interacts with the model through a series of deictic gestures as she enumerates these features. The second model, however, was designed for the purpose of demonstrating how forces are transferred through the rafters and into the purlins, and as such, what is criterial about the model shifts. In the second model, the absence of plasticine is semiotic as it shows that it is gravitational force rather than adhesion that connects the rafter to the purlin. The intervention of Student G placing the rafter on the purlin also helps communicate this and is part of the *valeur* between the models and their differing purposes. In interview data, Student G commented, "I just wanted to show, like, what I used, the Play-Doh, it's just, like, a connection thing but it's not like this in reality and you don't use glue… And I thought you can't see this on the other model, so I just wanted to make sure that it's different, that it's really clear that it's the gravity holding it in place". The two models also demonstrate one of the ways in which the disciplines of architecture and civil engineering differ from each other. While the first model shows the elements of a structure and the space created by them, pertinent to the discipline of architecture, the second model is concerned with how forces are carried by structural elements, a primary concern of civil engineering.

Finally, *valeur* exists at the rank scale of intervention. In the Exemplar VC shown to students as an example of a VC, Student X creates a model to explain how piles (foundational, underground posts) resist forces and maintain structural integrity (see **Figure 5**). With the sponge representing soft ground, the saucer a one-story building and the nails, the piles, it is the intervention of Student X that serves his most pertinent semiotic purpose. His strength performs the role of gravity, wind loads and snow forces on the model. In the first intervention, Student X compresses the model with his fingers and the indentation produced is enough to show the instability of the saucer/building. After reinforcement with nails/piles, he repeatedly leans on the model with his full body weight, creating such pressure as to cause the camera to judder, yet without causing the saucer to tilt. It is clear, then, that what is most relevant to his communicative purpose is his intervention into the model, and *valeur* exists between his interventions rather than between the models themselves. If nails were added and no force applied, Student X's meaning would not have been as clearly made.

Now that *valeur* has been demonstrated at all three rank scales, this paper will focus on just one VC, Student O's "The typical Palladian villa design". In this artifact, Student O constructed four models, and these will be used to show how her models make meaning at all three ranks of element, ensemble and intervention. After this, her models will be analyzed in terms of what is lost and gained as she transduces the texts, images and diagrams drawn from textbooks and university lectures into the models constructed for her VC.

Student O: "The Typical Palladian Villa Design"

Student O created four models for her VC (see **Figure 6**). Three of these were composed of pieces of colored paper, which some could consider to problematise the boundary between diagram and model, however, they are treated here as models for a number of reasons. Paper is inherently three-dimensional, and the way the student-composer intervenes with the paper pieces transforms them into models. In models 2, 3 and 4 she layers model elements on top of each other, effectively "building" her model to show the way the structural members relate to each other spatially. The diagram, or "floorplan" she includes in the corner of the screen emphasizes this departure from diagram to model: what she has created is something distinct from a floor plan, it has a depth, however flat, and can be manipulated to demonstrate her concepts. It stands in contrast to image 5 (see **Figure 6**), where the same elements are used differently to construct something more like an architectural diagram. Here, Student O films the elements in static relation to each other. She does not intervene with the model or offer the "mediated focalisation" visible in her other models, where her hands "…can be seen emerging from the edge in the foreground of the picture", thereby encouraging us to adopt her vantage point (Painter et al., 2012, p. 23).
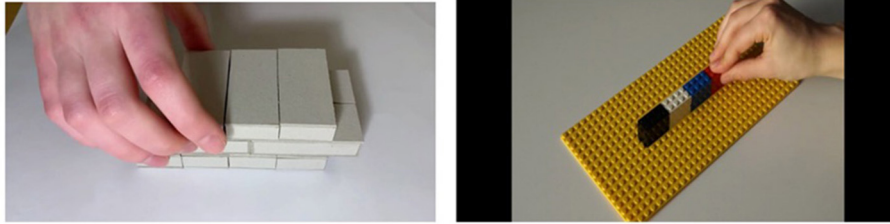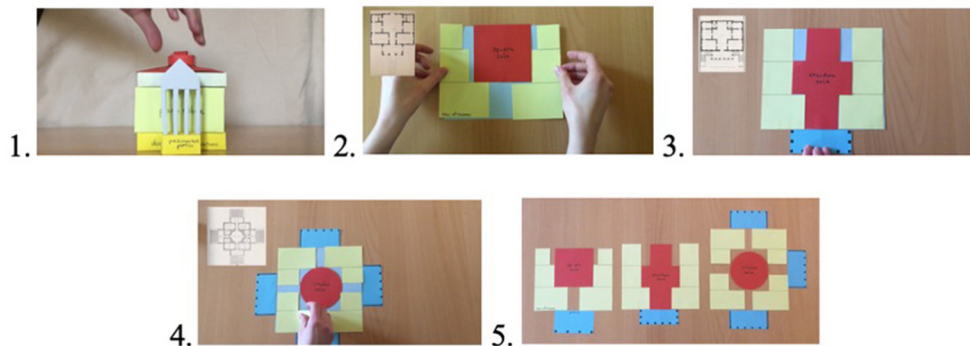
FIGURE 3 | Student Q (left) and Student M (right) chose different elements to explain similar concepts.
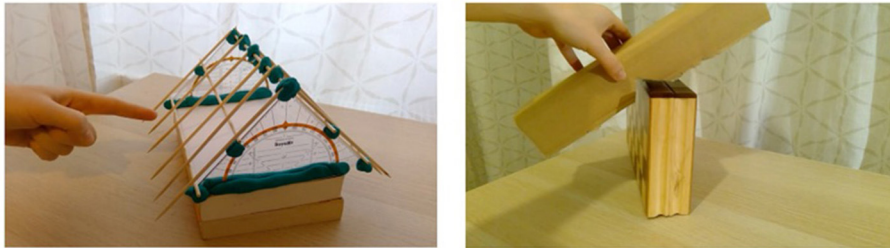
FIGURE 4 | Two models of gable roofs (Student G).

FIGURE 5 | Two interventions into the same model (Student G).

FIGURE 6 | Student O's four models and the architectural diagram (at 5.).

Finally, these "flat" models share a precedent in the multimodal discourse community of architecture, of which Student O is an apprentice. When positing ideas for new developments, paper modeling is commonly used to share concepts quickly and inexpensively (see **Figure 7**). In this type of modeling, tracing paper is used to layer potential structures onto a space without permanently altering the floor plan or the digital 3D model, which can become time consuming and expensive. This allows alternative visions to materialize in space, while those not selected can be easily discarded. As such, this paper considers

such "flat" models a different semiotic resource to diagrams, and their affordances shall be explored in more detail below.

As mentioned previously, models make meaning at all three ranks, and Student O's models can be used to reiterate this. At the rank of "element", for example, choices have been made between differently shaped *salas* (a type of room), e.g. square, cruciform and circular, as well as the *sala* shapes not selected. At the rank of ensemble, the three models she exemplifies – the Villa Chiericati, the Villa Foscari and the Villa Rotonda – are contrasted with each other and with other Palladian villas absent from her artifact (see **Figure 8**). Lastly, she constructs different reading paths for her audience by intervening into her models in different ways. These choices – and the ones not chosen – make meaning at the rank scale of intervention. These decisions around "arrangement" (Bezemer and Kress, 2008) will be explored in more detail in Section From Diagram to Model.

Transduction Between Text, Image and Diagram to Model

In order to better conceptualize the affordances of models, this section explores what is lost and gained as meaning is moved from three different modes to models: from text to model, from image to model and from diagram to model. These are three of the most commonly employed modes in the discipline of architecture, and it is likely that Student O would have drawn on these to construct her VC, and in particular, her models. The intention here is also to show how these transductions fundamentally alter the meanings made and have the potential to transform perceptions of the concepts explained.

From Text to Model

It is firstly important to note that the text shown in **Figure 9** is unlikely to have been the text from which the model is transduced. It is the verbal text that accompanies the intervention into the model, narrated by the student herself, and was most plausibly written in collaboration with her model planning. As she did not opt to participate in the interview stage of data collection, no particular text can be isolated as a source text. Instead, it is likely that the model was transduced from a multiplicity of texts, sourced from any number of German- or English-language lectures, journal articles and the like. However, the spoken text forms a useful starting point for considering some of the elements that are lost or gained when moving meaning from text to model. Firstly, because the sample text is drawn from the narration, the model co-creates the meaning in the verbal text, sharing the "work" of making meaning intersemiotically. Student O has also taken the time to temporally coordinate her text with her interventions into the model, so that, for example, the segment of text "…the porticos surrounding the cubic building on all four sides" is timed with the placement of porticos on the four sides of her ground plan. In this way, meaning is spread across semiotic resources and "…distributed across several coordinated sensory channels…" (Bateman and Schmidt-Borcherding, 2018; p. 6).

Another important consideration is what happens to the model in terms of specificity vs. generality (Bezemer and Kress, 2008). It could initially be said that there is a loss of specificity as meaning moves from text to model. While the verbal text states that "the floors are joined by small spiral staircases", staircases and other architectural details are absent from the model. Other pertinent information, such as that it was built in 1,565 or was inspired by the Pantheon in Rome, are also not communicated. However, in the spoken text that accompanies her interventions into the models (i.e., where she effectively "builds" them from the ground up), she states, "Palladio's villa design followed a recurring system which evolved over time". Student O's purpose, then, appears to be to show how much of Palladian design is generated by a "grid process" of aggregation and subdivision (Benrós et al., 2012; 523), rather than to exhibit decoration or communicate the historical impact of the building.

In this sense, the model affords a much *higher* level of specificity than writing. For example, if we compare her model to an excerpt about this grid process (sometimes termed "the Palladian grammar"), we can more clearly see the gains of transducing from text to model (see **Figure 10**). This text has three complex nominal groups: "a system of notation disclosing the variable syntax of each villa", "the spatial tripartitions of the interiors" and "the color-coded models and axonometric drawings", requiring the reader to make abstractions based upon their technical knowledge. Student O, in contrast, is able to grant the non-specialist access to these concepts by the affordances of the model. Now that temporality can play a role in meaning-making, she can "unpack" the nominalisation "spatial tripartions" by physically dividing the space into three parts. Further, by using actual squares, circles and rectangles she can avoid the desiccated nominalisation "axonometric drawings" while still making the same meaning[3]. Similarly, instead of the term "color-coded models", the affordances of the model impel her to make specific choices around color (e.g., red, yellow and blue). These epistemological commitments of red circles for the *salas*, yellow L-shapes for the room groups and blue rectangles for the porticos help the audience more fully comprehend how "the interiors are mapped out". Finally, because these colors and shapes remain the same across several models, the audience is encouraged to recognize patterns, and something of what Sherer (2012) terms "the variable syntax of each villa" can potentially be better understood.

From Image to Model

We now turn to focus on the shift from image to model, and what is lost or gained in terms of the meanings made possible by both. The image in **Figure 11** was created by Student O for her VC, and is used multimodally in her artifact, with meaning shared between image and text (both a label and her narration are layered onto the image). The model, although still constructed of colored paper, is more traditionally three-dimensional, and constitutes six elements: the domestic and farm functions; the *piano nobile* (the first floor consisting of reception areas and bedrooms); the attic; the tent roof; and the dome. Firstly, it is clear that the model affords more generality than the image. The

---

[3]An axonometric drawing is a two-dimensional object that has been rotated on its axes to give it a three-dimensional appearance, like a drawing of a cube.
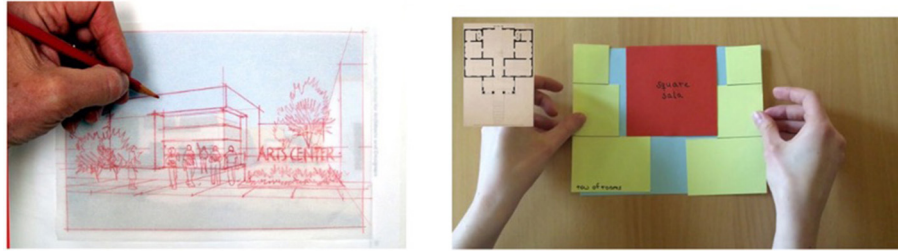
FIGURE 7 | Tracing paper modeling (left) compared with one of Student O's 'flat' models (right).
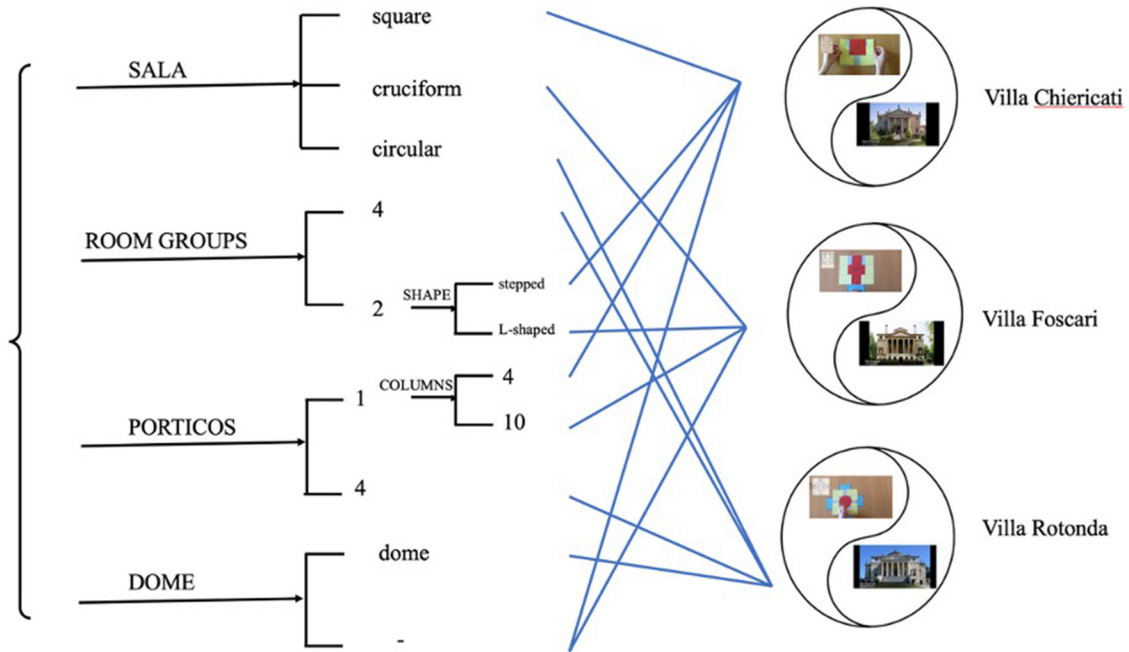

FIGURE 8 | Student O's closed system with three realizations at the rank of ensemble.

The Villa Rotonda was built in Vicenza and is Palladio's most famous villa design. The ground plan is a square. The outer shape is determined by the porticos surrounding the cubic building on all four sides. The symmetrical plan has a circular *sala* in the centre. Four groups of rooms are placed around the *sala* instead of the regular two. The passages between the *sala* and the four halls divide the plan into four equal parts. Each consists of a larger and a smaller space. The floors are joined by small spiral staircases. A flat, slightly-stepped dome protrudes from the pyramid roof.
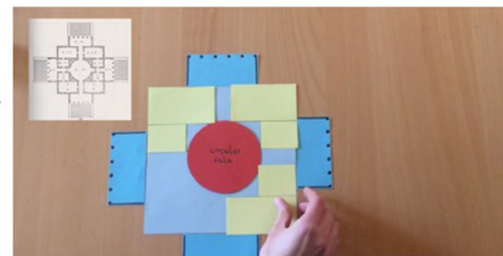
FIGURE 9 | Transducing a segment of narration to Model 4.

material and its properties are entirely absent from the model, represented instead by dramatically simplified paper shapes. Here, what is criterial for Student O, is not the color or texture of the masonry – a white stone, textured by the patina of age – nor the creeping vines and life-sized sculptures on the roof, nor the "pleasant hills" which surround it (Berzal de Dios, 2014, p.

178). Instead, the specificity of photographic image is transduced into the generality of colored paper and pen, a mode more apt to her pedagogical purpose. As mentioned previously, she is interested in communicating the elements of Palladian design rather than the particularities of the individual villas, and with the visual simplicity of her model she makes salient the structural

elements, their shapes, and their locational relationships with each other.

The shift from (still) image to (moving) interaction into model grants another gain. Although the villa is clearly large, something of its scale and grandeur is lost even in a photograph. This is because there is no "human scale": part of the reason why architectural renderings (two-dimensional slices of three-dimensional digital models) almost always include small, anonymous figures is to communicate the scale of the building and the space it occupies in relation to other buildings and the people who inhabit them. Although Student O doesn't provide human figures, her model nevertheless offers more of this sense of proportion. Her first element, a paper rectangular prism labeled "domestic/farm functions" helps the audience realize that the space is large enough not only for at least a kitchen and a barn, but also for domestic staff, agricultural staff and animals. When the *piano nobile* – an element nearly three times the size – is added on top, something of the magnitude of the villa is restored. The model, and the process of "building" it, more explicitly demonstrate the wealth and social standing of the villa owners than could a static photograph. The final gesture of "crowning" the building with a tent roof and "cordoning" it off from the wider public by a towering pedimented portico is yet more strongly evocative of the owners' status.

From Diagram to Model

Turning to the transduction from diagram to model, a number of additional observations can be made. Firstly, it is clear that a diagram was insufficient for Student O's purpose, as she provided a diagram in the corner of her VC while she intervened into the model (see **Figure 12**). As mentioned previously, by reusing elements across three of her models, she was able to convey the addition or subtraction of parts in a way that a still diagram could not. Furthermore, her interventions also hint at the role of architecture in organizing space and make the role of the architect more visible to her audience. This move from diagram to intervention into a model restores, in a sense, Palladio the architect to the concept of Palladian design. While a diagram of a Palladian villa depicts the Goal of the material process (the floorplan produced), an intervention into a model demonstrates how a Palladian villa was constructed and reminds the viewer of the active material processes involved (i.e., Palladio designed a villa). It is a move away from abstraction toward the real, and an apt choice for Student O as an apprentice of architecture.

By shifting from a diagram to a model, Student O also constructs a reading path for her audience that is not present in the diagram. This decision regarding the "arrangement" (Bezemer and Kress, 2008) of her model fundamentally alters the meanings she is making. It also changes how the audience relates to the concept she is explaining and perhaps affects their interpretation of it. While a diagram can be read in any number of ways – working from the outside in, for example, or moving through the central doorway into the reception area – the building of a model unfolds temporally, and in a linear sequence. Rather than a multiplicity of possible readings, this intervention constrains the audience's reading path to the one intended by Student O. This is perhaps a loss of sorts, but it also provides

an insight into the pedagogical purpose of the student creator, and what she considers to be her "…own interest, the needs of the matter to be communicated, and the characteristics of the audience" (Bezemer and Kress, 2008, p. 171).

CONCLUSION

This paper has attempted to unpack the complex semiotic resources of architecture and civil engineering models as constructed for an English for Architects and Civil Engineers classroom in Germany. In so doing, both practical and theoretical implications of model-making have been explored. In terms of the practical ramifications, there are four major considerations regarding model-making in the context of EAL/D and ESP education.

There is firstly a link between the students' pedagogical purposes, their semiotic choices and the quality of their explanations. Their explanations could be said to be improved when technical concepts are communicated clearly to deepen audience comprehension, and purposeful pedagogical choices are visible throughout the data-set and corroborated in the post-task interviews. Secondly, this study confirms the findings of a number of existing studies. Digital, multimodal assessments tasks – and model-making in particular – allow students to exhibit skills not usually included in standard EAL/D rubrics (Nielsen et al., 2020) and increase students' motivation (Lee, 2014; Jiang and Luk, 2016). Students also found the task to be enjoyable, valuable and to impact positively upon their self-perception. Thirdly, model-making – and multimodal tasks more generally - can facilitate more effective assessment of complex skills. The CEFR, for example, now requires students to be able to interpret visually organized information and explain technical terminology to non-experts, among other skills (Council of Europe, 2020). Digital, multimodal tasks offer teachers a purposeful and considered way of assessing these skills. Finally, model-making increases the authenticity of these assessment tasks, as models are used extensively in the construction process. In the ESP classroom, and in courses aimed at architects and civil engineers in particular, using models to make meaning is a valuable skill reflective of contemporary professional practices.

The data also suggests some intriguing theoretical implications. Model-making can be seen as a system of signs making meaning paradigmatically (Saussure, 1983; Martin, 2011). This paper conceives of models as composed of three ranks: "element", or model part; "ensemble", or the model itself; and "intervention", a performative interaction with the model. Models were also found to offer nuanced opportunities for meaning-making. In approaching the models as modes comprised of modes, we highlight the usefulness of a rich and dynamic conceptualization of mode, one that operates in relation to the materiality of the communicative artifact. In order to better illustrate some of the affordances of models, and the advantages they offer the meaning-maker, the models in Student O's VC were analyzed. Particular attention was given to what is lost and gained as she transduces meanings between models and three other modes (text, image and diagram).
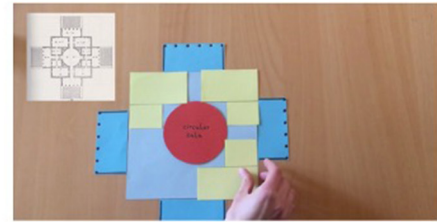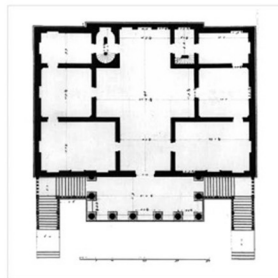
FIGURE 10 | Transducing an excerpt from an academic paper to Model 4.
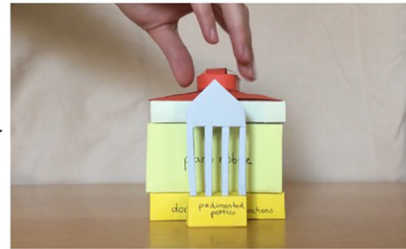
FIGURE 11 | Transducing a photographic image to Model 1.

FIGURE 12 | Transducing a diagram to Model 3.

Firstly, in terms of moving meaning from text to model, shifts occur in specificity and generality. Although some detail is lost, there was a considerable increase in the specificity of the meanings made in terms of Student O's purpose of explaining the patterns of "Palladian Grammar". Model-making improved the clarity of her explanation and illuminated the meanings obscured by the technical terms in the text. Secondly, in the shift from image to model, there was a move from the specific to the general in accordance with her pedagogical purpose. Model-making also offered the opportunity to make meanings not overtly apparent in the photograph, such as the magnitude of the villa in relation to its environment and inhabitants. Finally, in moving meaning from diagram to model, there was a narrowing of possible interpretations. Student O was also able to impact audience interpretations of the Palladian design process by "building" the models, which permitted temporal meanings in addition to the spatial meanings afforded by the diagram. This emphasized the role of the architect in organizing space, constrained the

audience's reading path of the models and influenced how they interpreted the models and the concepts involved. This paper suggests that architectural and civil engineering models are complex semiotic resources rich in theoretical potential and practical ramifications for the ESP classroom. It is hoped that with the theoretical tools such as these explicated here, emergent sign systems such as this one can be analyzed in more detail, and their pedagogical and epistemological impact better understood.

DATA AVAILABILITY STATEMENT

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

ETHICS STATEMENT

The studies involving human participants were reviewed and approved by IRMA Human Ethics, University of Wollongong

and the Ethikkommission der Fakultaet 2 der TU Braunschweig (Ethics Commission of Faculty 2, University of Technology Braunschweig). The patients/participants provided their written informed consent to participate in this study.

AUTHOR CONTRIBUTIONS

AH, PJ, EM, and HG contributed to conception and design of the study. AH acquired, analyzed and interpreted the data, and wrote the first draft of the manuscript. PJ, EM, and HG edited and added to the manuscript. All authors contributed to manuscript revision, read, and approved the submitted version.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fcomm. 2022.901719/full#supplementary-material

REFERENCES

Alvermann, D. (2017). The M word: dare we use it? *J. Adolesc. Adult Liter.* 61, 99–102. doi: 10.1002/jaal.665

Ariztimuño, L. I., Dreyfus, S., and Moore, A. R. (2022). *Emotion in speech: A systemic functional semiotic approach to the vocalisation of affect.*

Australian, Curriculum, Assessment and Reporting Authority (ACARA). (2015). Australian Curriculum - "Multimedia: English" Available online at: https://www.australiancurriculum.edu.au/resources/curriculum-connections/dimensions/?Id=67512andYearLevels=42681andsearchTerm=ACELA1433 (accessed June 27, 2022).

Barton, G., and Unsworth, L. (2014). Music, multiliteracies and multimodality: exploring the book and movie versions of Shaun Tan's' The lost thing'. *Austr. J. Langu. Liter.* 37, 3–20. doi: 10.3316/aeipt.201731

Bateman, J., and Schmidt-Borcherding, F. (2018). The communicative effectiveness of education videos: towards an empirically-motivated multimodal account. *Multimodal Technol. Inter.* 2, 59. doi: 10.3390/mti2030059

Bateman, J., Wildfeuer, J., and Hiippala, T. (2017). *Multimodality: Foundations, Research and Analysis – A Problem-Oriented Introduction.* Berlin: Walter de Gruyter GmbH doi: 10.1515/9783110479898

Benrós, D., Duarte, J. P., and Hanna, S. (2012). A new Palladian shape grammar: A subdivision grammar as alternative to the Palladian grammar. *Int. J. Architect. Comput.* 4, 521–540. doi: 10.1260/1478-0771.10.4.521

Bernstein, B. (1996). *Pedagogy, Symbolic Control and Identity: Theory, research, critique* (revised edn). Lanham, Maryland: Rowman and Littlefield Publishers.

Berzal de Dios, J. (2014). *The Spectacle of Space: Visual Experiences in the Early Modern Scenography of Italy (PhD Thesis).* Ohio State University, Columbus, OH, United States. Available online at: https://www-proquest-com.ezproxy.uow.edu.au/docview/2317707978?pq-origsite=primo (accessed June 28, 2022).

Bezemer, J., and Kress, G. (2008). Writing in multimodal texts: A social semiotic account of designs for learning. *Written Commun.* 25, 166–195. doi: 10.1177/0741088307313177

Burn, A., and Parker, D. (2001). Making your mark: digital inscription, animation, and a new visual semiotic. *Educ. Commun. Inf.* 1, 155–179. doi: 10.1080/14636310120091913

Canagarajah, S. (2012). *Translingual Practice: Global Englishes and Cosmopolitan Relations.* London: Routledge. doi: 10.4324/9780203073889

Cope, B., and Kalantzis, M. (2009), A grammar of multimodality. *Int. J. Learn.* 16, 361–425. doi: 10.18848/1447-9494/CGP/v16i02/46137

Council of Europe (2015). *"Education, Mobility, Otherness: The Mediation Functions of Schools".* Language Policy Unit. Strasbourg: Council of Europe. Available online at: www.coe.int/lang (accessed June 27, 2022).

Council of Europe (2020). *Common European Framework of Reference for Languages: Learning, teaching, assessment – Companion volume.* Strasbourg: Council of Europe Publishing. Available online at: www.coe.int/lang-cefr (accessed June 27, 2022).

Douglas, D. (2013). "ESP and assessment," in B. Paltridge and S. Starfield (Eds.), *The handbook of English for specific purposes.* Wiley-Blackwell. p. 367–383. doi: 10.1002/9781118339855.ch19

Dzekoe, R. (2017). Computer-based multimodal composing activities, self-revision, and L2 acquisition through writing. *Langu. Learn. Technol.* 21, 73–95. doi: 10.31274/etd-180810-2304

Early, M., Kendrick, M., and Potts, D. (2015). Multimodality: Out from the margins of English language teaching. *TESOL Quarterly* 49, 447–460 doi: 10.1002/tesq.246

Ehrenreich, S. (2010). English as a business lingua franca in a german multinational corporation: meeting the challenge. *J. Bus. Commun.* 47, 408–431. doi: 10.1177/0021943610377303

Georgiou, H. (2020). Characterising communication of scientific concepts in student-generated digital products. *Educ. Sci.* 10, 18. doi: 10.3390/educsci10010018

Gibson, W. (1977). "The theory of affordances," in R. Shaw and J. Brandsford (Eds.), *Perceiving, acting and knowing: Toward an ecological psychology.* Mahwah, New Jersey: Lawrence Erlbaum Associates. p. 62–82.

Guo, L. (2004). "Multimodality in a biology textbook," in K. O' Halloran (Ed.), *Multimodal Discourse Analysis: Systemic Functional Perspectives.* London: Continuum. p. 196–219.

Hafner, C. (2015). Remix culture and English language teaching: the expression of learner voice in digital multimodal compositions. *TESOL Quart.* 49, 486–509. doi: 10.1002/tesq.238

Hafner, C. A., and Miller, L. (2018). *English in the Disciplines: A Multidimensional Model for ESP Course Design.* London: Routledge doi: 10.4324/9780429452437

Hafner, C. A., and Miller, M. (2011). Fostering learner autonomy in English for science: A collaborative digital video project in a technological learning environment. *Langu. Learn. Technol.* 15, 68–86. doi: 10.25134/erjee.v7i1.1495

Hellwig, A. F. J. (2022). Multimodal, digital artefacts as learning tools in a university subject-specific english language course. *Int. J. TESOL Studies* 4, 24–38. doi: 10.46451/ijts.2022.02.03

Hung, S.-T. A. (2019). Creating digital stories: EFL learners' engagement, cognitive and metacognitive skills. *Educ. Technol. Soc.* 22, 26–37.

Iedema, R. (2003). Multimodality, resemiotization: Extending the analysis of discourse as multi-semiotic practice. *Visual Commun.* 2, 29–57. doi: 10.1177/1470357203002001751

Jewitt, C. (2009). *The Routledge Handbook of Multimodal Analysis.* London: Routledge.

Jewitt, C., Bezemer, J., and O'Halloran, K. (2016). *Introducing Multimodality.* London: Routledge

Jewitt, C., Kress, G., Ogborn, J., and Tsatsarelis, C. (2001). Exploring learning through visual, actional and linguistic communication: the multimodal environment of a science classroom. *Educ. Rev.* 53, 5–18. doi: 10.1080/00131910123753

Jiang, L. (2017). The affordances of digital multimodal composing for EFL learning. *ELT J.* 71, 413–422. doi: 10.1093/elt/ccw098

Jiang, L., and Luk, J. (2016). Multimodal composing as a learning activity in English classrooms: inquiring into the sources of its motivational capacity. *System* 59, 1–11. doi: 10.1016/j.system.2016.04.001

Kim, Y., and Belcher, D. (2020). Multimodal Composing and Traditional Essays: Linguistic Performance and Learner Perceptions. *RELC J.* 51 86–100 doi: 10.1177/0033688220906943

Kohnke, L. (2019). Using comic strips to stimulate student creativity in language learning. *TESOL J.* 10, 1–5. doi: 10.1002/tesj.419

Kohnke, L., Jarvis, A., and Ting, A. (2021). Digital multimodal composing as authentic assessment in discipline-specific English courses: insights from ESP learners. *TESOL J.* 12, 115, doi: 10.1002/tesj.600

Kress, G. (2010). *Multimodality: A Social Semiotic Approach to Contemporary Communication.* Routledge

Kress, G., and van Leeuwen, T. (2021). *Reading Images: The Grammar of Visual Design.* (4th edn). London: Routledge. doi: 10.4324/9781003099857

Lapp, D., and Fisher, D. (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*. London: 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 Teach. Pract. Critique* 13, 55–67.

Lotherington, H., and Jenson, J. (2011). Teaching multimodal and digital literacy in L2 settings: New literacies, new basics, new pedagogies. *Ann. Rev. Appl. Linguist.* 31, 226–246. doi: 10.1017/S0267190511000110

Martin, J. R. (2011). "Multimodal semiotics: Theoretical challenges," in S. Dreyfus, S. Hood and M. Stenglin (Eds.), *Semiotic margins: Meaning in multimodalities*. Continuum. p. 243–270.

Martin, J. R., and Rose, D. (2007). *Working With Discourse: Meaning Beyond the Clause*. London: Continuum.

McGinnis, T. A. (2007). Khmer rap boys, X-Men, Asia's fruits, and Dragonball Z: Creating multilingual and multimodal classroom contexts. *J. Adolesc. Adult Liter.* 50, 570–579. doi: 10.1598/JAAL.50.7.6

Nash, B. (2018). Exploring multimodal writing in secondary English classrooms: a literature review. *English Teaching.* 17, 342–356 doi: 10.1108/ETPC-01-2018-0012

New London Group (1996). A pedagogy of multiliteracies: designing social futures. *Harvard Educ. Rev.* 66, 60–92. doi: 10.17763/haer.66.1.17370n67v22j160u

Nielsen, W., Turney, A., Georgiou, H., and Jones, P. (2020). Working with multiple representations: preservice teachers' decision-making to produce a digital explanation. *Learning: Res. Pract.* 6, 51–69. doi: 10.1080/23735082.2020.1750673

O'Halloran, K. L. (2008). Systemic functional-multimodal discourse analysis (SF-MDA): Constructing ideational meaning using language and visual imagery. *Vis. Commun.* 7, 443–475. doi: 10.1177/1470357208096210

O'Toole, M. (1994). *The Language of Displayed Art*. Madison, New Jersey: Fairleigh Dickinson University Press.

Painter, C., Martin, J. R., and Unsworth, L. (2012). *Reading Visual Narratives: Image Analysis in Children's Picture Books*. Sheffield: Equinox.

Pandya, J. Z., Pagdilao, K. C., Kim, A. E., and Marquez, E. M. (2015). Transnational children orchestrating competing voices in multimodal, digital autobiographies. *Teach. College Rec.* 117, 1–32. doi: 10.1177/016146811511700707

Prior, P. (2013). "Multimodality and ESP research," in B. Paltridge and S. Starfield (Eds.), *The Handbook of English for Specific Purposes*. John Wiley and Sons. p. 519–534. doi: 10.1002/9781118339855.ch27

Ravelli, L. J., and McMurtrie, R. J. (2015). *Multimodality in the Built Environment: Spatial Discourse Analysis*. London: Routledge. doi: 10.4324/9781315880037

Royce, T. (2002). Multimodality in the TESOL classroom: exploring visual-verbal synergy. *TESOL Quart.* 36, 191–205. doi: 10.2307/3588330

Saussure, F. D. (1983). *Course in General Linguistics*. (R. Harris, trans.). London: Bloomsbury

Sherer, D. (2012). Critical and palladian. *Log.* 26, 135–143.

Sindoni, M. G., Adami, E., Karats, S., Marenzi, I., Moschini, I., Petroni, S., et al. (2019). *The Common Framework of Reference for Intercultural Digital Literacies: A Comprehensive Set of Guidelines of Proficiency and Intercultural Awareness in Multimodal Digital Literacies*. Erasmus; Brussels: European Commission.

Smith, B. E., Pacheco, M. B., and de Almeida, C. R. (2017). Multimodal codemeshing: Bilingual adolescents' processes composing across modes and languages. *J. Second Langu. Writ.* 36, 6–22. doi: 10.1016/j.jslw.2017.04.001

Turney, A., and Jones, P. (2021). Approaching complex multimodal phenomena in educational settings-insights from theory. *Austr. J. Langu. Liter.* 44, 29–45. doi: 10.3316/informit.742971114110258

Vandommele, G., Van den Branden, K., Van Gorp, K., and De Maeyer, S. (2017). In-school and out-of-school multimodal writing as an L2 writing resource for beginner learners of Dutch. *J. Second Langu. Writ.* 36, 23–36. doi: 10.1016/j.jslw.2017.05.010

Wilke, B. (2012). Transformationen im Zusammenspiel von Bild und Text in Online-Reiseführern: Schlussfolgerungen für die Didaktik. *Testi e Linguaggi.* 6, 293–310.

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.