Exploring Verbal and Non-Verbal Expressions of ESP Undergraduates' own Voices and Identities

ARÁNZAZU GARCÍA-PINAR*
Centro Universitario de la Defensa (Spain)

ABSTRACT

Given the ubiquity of digital technologies in all sorts of academic contexts, it is generally assumed that many undergraduates' writing tasks will include verbal and visual modes these days. The interweaving of different modes allows students to express different multidisciplinary and individual identities while they become agents and designers of different L2 learning tasks. Using an interpretative qualitative approach, the present study explores the authorial voice and stance that four engineering undergraduates enacted in their presentation slides for an in-class oral presentation. Data sources included screen capture, classroom observation, and interview transcripts. Findings revealed that behind students' collaborative compositional processes there are complex multimodal decisions that help them express their identities and enhance their engagement in the L2. Students perceived their presentation slides as artefacts to accommodate their audience and as means through which they were able to represent themselves as agents and designers of the discipline of engineering. Based on the results, this study highlights different pedagogical implications and ideas for English for specific purposes (ESP) contexts.

KEYWORDS

Writing tasks; Multidisciplinary and individual identities; Multimodal decisions; Agents; Compositional processes.

1. INTRODUCTION

In the last two decades, literacy education for learners has been re-conceptualized as digital technologies permeate all sorts of academic contexts. Students, frequently referred to as

"digital natives" (Prensky, 2001), seem particularly adept at using all types of technology formats and show full competence over many of the tasks they produce. They become in many cases strategic designers by incorporating different modal ensembles in their learning products. In the context of language learning, the use of digital technologies is also increasingly common. Language proficiency skills for communication in educational contexts have been reassessed to incorporate multimodal ability. With image and speech forming a semantic unit, it is more and more difficult to understand one without the other (Zhao, Djonov & van Leeuwen, 2014). The meaning that arises across these semiotic resources, or "intersemiosis" (O'Halloran, 2005) is often more profound (Lim, 2004) than the meaning that a single semiotic resource can produce on its own. Also, these linkages produce semiotic cohesion, that is, meanings are made across different semiotic resources in multimodal discourse (Liu & O'Halloran, 2009).

In recent years there has been a growing interest in understanding how language learners can benefit from a multimodal approach (Burguess & Rowsell, 2020; Canals 2021; Grapin, 2019; Jewitt, 2013; Miller, 2009; Potts, 2013; Walsh, 2015, among others). Many of these studies focus on researching the different multimodal configurations that new technologies facilitate and how students construct meaning and communicate from these. Different theories have researched language-image relations (Bateman, 2014; O'Halloran & Lim, 2009). The understanding of emerging interactions between images and language seems increasingly relevant to reconceptualize literacy and literacy pedagogy from a multimodal standpoint (Unsworth & Cleirigh, 2009).

Although there has been increasing attention on the importance of multimodality in students' written texts and to research their composing processes and authorial agency (Archer, 2010; Cimasko & Shin, 2017; Jiang, 2018; Lim & Polio, 2020; Tardy, 2005), research into multimodal collaboration practices and how these have an influence on students' L2 engagement and identity exploration remains somewhat scarce (Freschi & Cavalari, 2020; Magnusson, & Godhe, 2019; Ponzio & Deroo, 2021). The current study was designed to further develop the understanding of undergraduates' complex multimodal composing processes and how this implied and projected identity expression.

2. MULTIMODALITY

One of the first calls that prompted the urgency to reconsider learning models and underscored the necessity to design new frameworks to change the direction of teaching and learning was addressed by the New London Group (1996). This group of prominent literacy educators and researchers intended to problematize the social and cultural reshaping of the communicational panorama that globalisation and new technologies had given rise to. A new approach to literacy and pedagogy, which they referred to as "multiliteracies", could widen

the limitations of more traditional approaches that did not encompass the shifting semiotic landscape that was transforming the ways in which people communicated. The authors advocated for the acceptance of new relationships among different modes, specifically geared to the articulation between visual image and the word, as an available resource for making meaning in the classroom. The New London Group's influential manifesto on new prospects for teaching and learning literacy triggered from 2000 onwards an interest in multimodality among researchers in different educational contexts.

The availability of digital technologies these days broadens the scope of multimodality. Communication has for a long time been multimodal. People have usually resorted to a wide range of modes to communicate, and texts have always been multimodal. However, the introduction of digital technologies takes a significant turn to theorise and analyse how, for instance, moving and still images are configured and combined with other modes, such as writing, speech, and layout. Screen-based texts enable new multimodal configurations where still and moving images, colour, layout, sound, writing and speech, body movement, and gestures play a part in communication, interpretation, and representation.

The new reshaping of practices enabled by the use of digital technologies makes their study especially relevant to multimodality. New technologies can change how knowledge and meaning are construed, how communication is produced, and how the world is presented. Inevitably, all this leads to research the different and meaningful ways students can best benefit so that they can distinguish how verbal and non-verbal modes articulate discourse in digital screen-based texts, where the boundaries between modes become less and less clear, and the dynamics of meaning-making come to the forefront (Jewitt, 2006). Since the screen is the prevailing medium globally (Johnson & Kress, 2003: 7), it seems reasonable to heightening pedagogical research interest in how meaning-making is realised through different forms of representation (i.e., linguistic, visual, audio, spatial and gestural). One important role of the second language teacher, thus, is to teach language learners the interpretation of the different ways that technology affords to represent meaning these days (Kern, 2006).

Digital communication technologies also give rise to new challenges and demands for teachers who often become designers of learning (Lim & Hung, 2016; Selander, 2008) and develop a semiotic awareness of how the choices they make when they use different multimodal resources to meaning making have a direct and unique influence on students' learning experiences (Lim, 2021). Teachers' language, gestures, positioning, movement, and the use of semiotic technologies contribute to students' literacy and to the improvement of students' engagement and performance in the classroom. Research on gestures (Kendon, 1980; McNeill, 1992, 2005), for instance, evidence that they may contribute to enhancing

listening comprehension, assisting in clarifying verbal meanings, and conveying additional information. Gestures that are accompanied by speech can increase listener's comprehension by developing in the listener a cognitive simulation or mental representation of the message. When it comes to assessing learners' progress in acquiring a second language, gestures also play a relevant role. Stam (2008: 3) highlights the importance of regarding verbal and nonverbal communication as part of students' communicative competence to obtain a valid sense of their L2 proficiency. In particular, the author refers to co-speech gestures as gestures that "occur only during speech, particularly with elements of high communicative dynamism (i.e., new, focused, or contrastive information) and are phonologically, pragmatically, and semantically synchronous with speech" (2014: 1). These gestures, the author notes, perform a key role in language teaching and interaction because they can provide information about students' proficiency. The spontaneous movements of arms and hands that take place during speech, for instance, show what students are thinking and how they are thinking (2008: 37).

Institutions that offer English for Specific Purposes (ESP) courses often include multimodal perspectives in their curricula. A fundamental role of ESP educators is to understand the different ways students will write and communicate as members of diverse workplaces and enable the acquisition and command of specialized discourses in the classroom (Hyland, 2017). The pervasiveness of digital communication entails today a reassessment of the different ways the writing and speaking skills have been traditionally approached to offer ESP learners optimum opportunities that allow them to engage with texts that are increasingly multimodal and frequently digital. Different digital teaching resources lend themselves to optimally prepare ESP students for "authentic scenarios of discipline-specific communication" (Kohnke, Jarvis & Ting, 2021: 2) and to deliver ESP content efficiently by adapting different modes (i.e., text, images, and sound, among others). Writing in ESP contexts moves away from the regulated conventions and puts language and aspects of design (e.g., layout, use of images, color, and font) on an equal footing. The design of students' slides and the writing of technical reports and essays ESP students elaborate, often position them as designers of new and dynamic *multimodal ensembles* or material products with different types of meaning-making combinations that form a coherent whole (Kress, 2010: 162). As Kress highlights (2005: 20), the notions of *design* and *the agency* of the individual become relevant in current teaching environments and in students' assignments:

> We need the notion of design, which says: In this social and cultural environment, with these demands for communication of these materials, for that audience, with these resources, and given this interest of mine, what is the design that best meets these requirements?

The composition of oral presentations students prepare for the ESP course these days move across multiple modes of text design that enhance and explore new creative

possibilities. Different organizational structures, vocabulary choice, visual representations and color in presentation slides allow students to express cultural, educational, linguistic, and disciplinary identities in ways the written texts alone cannot. Their voices therefore can be expressed through both linguistic choices and "through choices in prominent nonlinguistic modes" (Haffner, 2015: 492). And when students collaboratively reflect on the different modes that may best represent what they aim to transmit to their audiences, they engage in communities of practice (Lave & Wenge, 1991).

3. AGENCY AND ENGAGEMENT IN THE LANGUAGE CLASSROOM

Learners' active investment and involvement in the learning process has long been regarded as a central issue in all sorts of educational contexts. Different approaches to language learning and teaching put agency at the centre of the learning process. Approaches such as learner identity (e.g., Norton, 2000), action-based teaching (e.g., van Lier, 2007), and motivational self-regulation (Ushioda, 2008) posit that learners will make the most of their learning experiences when they feel they are in charge of their own learning, they are treated in their own right (van Lier, 2007: 47), and they have opportunities to make the learning material their own (Dörnyei & Kubanyiova, 2014: 36). Regarding Norton's approach of learner identity, learners will be more agentic when they can choose what they regard as beneficial to their imagined identities and when they invest or divest from "the language or literacy practices of particular classrooms and communities" (Darvin & Norton, 2017: 7). Agency is the defining construct in van Lier's (2007) action-based teaching approach, who highlights that learners must be "respected as speakers in their own right and as agents of their own educational destiny" (47). According to Ushioda (2008), it is very important that motivation comes from the learner and is not regulated by the teacher. Also, it is crucial that learners see themselves "as agents of the processes that shape their motivation" to "develop skills in regulation their own motivation on which good language depends" (Ushioda, 2008: 30). Whenever learning is conceived "to be about meaningful, active, and passionate knowing-being-doing by people as actors of history and agents of the world-in-the-making", agency is important to "meaningful and transformative" teaching-learning (Stetsenko, 2019: 9). According to the author, the teaching-learning process involves helping students develop their ability to adopt their own stands and roles in the world with imagination, creativity, and novelty. In this sense, the L2 teacher adopts quite a relevant role as they can facilitate reasonable freedom and autonomy among students regarding different aspects involved in the design and making of specific L2 tasks such as in-class presentations (i.e., design of slides, and way of working together in a collaborative support group).

One important factor in promoting students' sense of agency is to allow them to "have choice and to express their voice where possible" (Mercer & Dörnyei, 2020: 46). Practices of

production and dissemination that encourage students to take control over certain aspects of their learning will also promote their autonomy (Benson, 2001; van Lier, 1996) and their feeling of being in charge of their own learning. According to Ajzen's (1998) theory of planned behaviour and the concept of *perceived behavioural control*, engagement in the activity will be likely to occur if the student senses a certain degree of ownership and control over their learning. In this regard and given the sound mastery of digital technologies students have these days, agency and autonomy in the composing of different multimodal products such as presentations slides can considerably contribute to strengthening student engagement while discussing their design decisions and provide one another feedback.

Students' designs processes help them express different expressions of individuality that can enact and enhance their self-perception of individuals with identities and personalities but also with goals and intentions within the language learning context. This is precisely what posits Ushioda (2009) in her motivational approach, the "Person-in Context Relational View of Motivation", which proposes a dynamic interaction between a person's agency and their background, motives, goals, as well as the activities and social relations this person is embedded in. Learners, as Ushioda (2009: 215) contends, need to be viewed as real individuals within particular contexts, with identities and personalities, and with goals and intentions. In advocating for an Action-based teaching, van Lier (2007: 47) also underscores the need to locate human agency at the center of the learning process and to look at learners as agents of their educational destiny, as "people with their own lives, aspirations, worries, dreams and identities … in need of forging productive identities that link the personal self to the new worldly demands presented by the new language".

From a multimodal perspective, students' multimodal creations might be inherently motivating (Dörnyei, 2020). Multimodal composing leads to opportunities for self-directed learning, making learners feel agentic while they actively participate as sign makers who exploit combinations of different modes. As "designers of their own learning practices" (Kress & Selander, 2012: 265), they get involved in interaction design processes and meaning-making negotiations by selecting the most apt mode for their communicative purposes, while they project their wishes and proposals "forward into an imagined social future" (267). The design of presentations slides, as we shall see in future sections, gives students different options to communicate their ideas. The slide format, usually space-constrained, makes students carefully consider the affordances of each mode; that is, how writing and imaging can convey different kinds of meaning and to weigh up which information to include or exclude. This requires students to be active and engaged with the L2 environment that is "full of potential meanings" (van Lier, 2000: 246). These meanings will become available when students actively participate and make use of all the affordances that emerge from their interactions with the digital text.

Learning occurs because there is engagement, or "transformative engagement" (Bezemer & Kress, 2016: 38), through which students collectively negotiate the meaning and engage in communication, collaborating with the teacher and other students. Motivation is likely to stem from the will to achieve recognition, connection, and belonging. Though research on multimodal composition has mainly focused on the positive aspects of these practices for promoting student motivation (Henry, 2019, 2021), few studies have closely examined the processes of composing with verbal and nonverbal modes for fostering student agency, voice and engagement in the L2. In an attempt to advance understanding of how the discursive choices in undergraduates' multimodal composing processes can foster engagement and provide scope for agency and disciplinary expression, we aimed to address the following research questions:

1. How did students approach their presentation slides to incorporate verbal and non-verbal expressions of their own voices and identities?
2. How did the use of the visual mode contribute to making students feel agentic and engaged in the L2?

4. METHODS

4.1. Pedagogical context and Participants

The course described here is a one-semester course in English for technology at a University in Spain. The four and a half-credit course is available to third year students and is designed to improve students' communication skills and to enhance their knowledge of specialised language. Different tasks are contextualised in daily engineering contexts to practice oral and written skills.

Through purposeful sampling (Patton, 1990), four focal engineering undergraduates were chosen for in-depth analysis of their multimodal compositional processes. They were John (male, 20 years old), Paul (male, 20 years old), Ana (female, 20 years old,) and Daniel (male, 21 years old). Participants were recruited based on the extent to which they had efficiently and extensively resorted to the visual mode in their in-class presentations. The four participants had decided to team up to prepare and deliver the in-class presentation; all of them had prior experience giving group English presentations that involved using the presentation program Google Slides. Regarding their English proficiency in relation to the *Cambridge English Placement Test,* which assessed their reading, listening, and language knowledge skills, this oscillated between B2 and C1. To respect the participants' privacy, confidentiality and anonymity were ensured. Accordingly, the participants' original names have been replaced by pseudonyms.

4.2. Data Collection and Analysis

The research data was collected during a 4-week course period and through a combination of field observation, oral performance recordings, interviews, and students' Google slides. As an action study, the lecturer was the only researcher. This study takes a qualitative interpretive approach (Davis, 1995), which aims to use the variety of data obtained from multiple perspectives to construct an understanding of students' engagement in multimodal composition processes. Field observations took place during the sessions students started to work on the oral presentations in the classroom. During one session students had to form groups, choose the TED Talk they would use as a baseline for their oral presentations, and establish which role each of them would have in these presentations. Writing field notes provided useful first insights into students' design decisions and their authorial stances that shed light on their relationships to their audience, that is, their classmates and English lecturer. Also, classroom observation allowed the researcher to realize that participants' engagement with this multimodal task provided a means of understanding the different ways motivation played out.

The main aim in conducting the individual, semi-structured interviews with the 4 focal participants was to allow them to identify specific design decisions and to detail the rhetorical implications of their modal choices. Another objective was to obtain a closer look at how the verbal and visual expressions in the Google slides they had designed contributed to projecting disciplinarity and individuality. Regarding the analysis of the verbal mode, this focused on the lexical choices students made to describe their research in their slides. Analysis of the verbal mode also examined closely the responses to the questions of the interviews where they referred to the design of the Google slides. The study on non-verbal modes focused on visual modes, and more specifically, on the choices students made on color, typeface, use of images, and layout to design the presentation slides. The study of the type of gesture students made when they carried out their oral presentations is outside the scope of the present article. Beat, iconic gesture, and the head gestures students often used to call attention to some specific points in their visuals or to assist them in highlighting the significance of specific words and stretches of their discourses reinforced nonetheless the importance of these modes to construe "relationships of solidarity with the audience" (Hood & Forey, 2005: 291) and express attitude and engagement.

Interviews were conducted after in-class oral presentations had finished. Each interview lasted about 30 minutes, was voice-recorded, and fully transcribed. Following thematic content analysis (Dörnyei, 2007), the interview transcripts were thoroughly read through three times. The analysis went through open and selective coding. During the first phase, the data was coded manually line-by-line to obtain closer familiarity with the overall organisation. Different categories were assigned. Special attention was given to nouns,

adjectives, and adverbs that conveyed students' motivational intensity and engagement that derived from working collaboratively and with real material. During the second phase, expressions and utterances that expressed disciplinary-specific terminology which suggested students identified themselves as members of their engineering community were highlighted. Also, the second phase involved singling out the expressions that conveyed participants' modal preferences. Participants' responses that were closely related were grouped (Loewen et al., 2009). To ensure coding quality, an independent coder was invited to examine and code the data. Five major themes emerged as significant for understanding ESP undergraduates' own voices and identities: 1) effectivity of images as a means of conveying better students' individuality; 2) engagement and interest of the audience; 3) role of the verbal mode in students' presentation; 4) affordance of different non-verbal modes; 5) agency and engagement in students' online media creations.

4.3. Findings and discussion

This section aims to describe the dynamic processes in students' multimodal composing. Taking a social semiotic approach to communication, participants' visual designs were broken down into basic components to gain useful knowledge about how these worked together (Ledin & Machin, 2020). The following findings are in accordance with the guiding research questions: (1) Verbal and non-verbal expressions of students' own voices and identities, and (2) Agency and engagement in students' online media creations. Each theme is discussed using extracts from the interviews with the four students.

4.3.1. Verbal and non-verbal expressions of students' own voices and identities

The focal task of the course is an in-class oral presentation in which students conducted a technological study on the main topic developed in the TED Talk they choose from a pre-prepared list of different technology-oriented talks. Particularly, the four participants selected the TED Talk *"How we are using drones to deliver blood and save lives"*, a demonstration talk where Keller Rinaudo, co-founder of Zipline, explains to the audience how he and his team have created a delivery system to transport blood and plasma to different countries in Africa. The oral presentation had to be visually complemented with presentation slides and participants decided to use Google Slides, because, as detailed in the interviews, this presentation program let them create and edit files online while allowing for real-time collaboration. Despite the inherent style of Google Slides, they had to make different multimodal choices to construct meaning on account of decisions that had to do with the concept of "modal aptness" (i.e., modal best-fit), modal affordance (i.e., what is possible to convey easily with a mode), and visual salience (i.e., features of a composition are given

ranking in relation to others through size, colour, or foregrounding to capture the viewer's attention) (Kress & van Leeuwen, 1996). Thus, the use of different layouts and visual representations in students' slides allowed them to express different field and personal identities of expression (Tardy, 2005) in ways the verbal mode alone could not.

In the presentation slides designed for the in-class oral presentation, the four participants portrayed themselves as members of a disciplinary community through the use of verbal and visual elements. Though students' accounts implied certain individual variations, their choices lay within a variety of community practices. Regarding the visual mode, participants' choices involved color, typeface, and layout. These semiotic resources symbolized the connection and the logic of their multimodal design. In terms of the verbal mode, the slide space constraints demanded students to transform key ideas in their presentation into concise language, this mainly consisting of headings and short phrases. Language, thus, played a significant role in participants' multimodal wholes and students regarded this mode to be as important as the other non-verbal resources. All of them gave due attention to writing good scripts and presenting their narration effectively to their audience. This fact made them resort to the L2 strategically. Students' slides, however, contrasted in the extent that they used verbal text. Some slides were completely reliant on visuals, others were heavy in visuals but included verbal text. This oscillation related to the fact that each participant had a specific task in the presentation, and some of these tasks lent themselves to being more verbal or visual than others. Daniel expressed the need to use only keywords and short phrases in some slides and portrayed himself as analytic (see Figure 1):

All of us agreed on using only words or short phrases but only when we estimated visuals could not substitute these. This slide is a good example, because I had to analyze the speaker's speech in relation to the strong and weak points throughout his talk. I had to resort to adjectives and some key phrases, I could not express these aspects through visuals. (Daniel, 20/04/21)

Figure 1. Daniel's slide on the analysis of TED speaker's performance, April 2021.

Though Paul explained some concern that the verbal text in most of his slides was a little bit lengthy, he was inclined to use more prose-like text (see Figure 2).

Figure 2. Paul's slide on origins of drones, April 2021.

Another important factor had to do with the confidence in the language skills each participant had. Though proficient in the L2, all of them regarded the use of effective images as a means of conveying better their individuality. Getting the attention of their audience was also paramount and they considered the display of effective visuals helpful for that purpose. John admitted that he searched for some type of peer engagement and recognition through the selection of well though-out images and by avoiding redundant verbal text in the slides he designed. The images he used invoked his disciplinary identity in assuming the audience's acquaintance with the topic he talked (see Figure 3):

Definitely, a picture is worth a thousand words, right? When I selected the images for this slide, I was thinking of my audience, my classmates. I wanted to draw their attention and I knew that some of the images, in particular the last one with the design of that tiny drone could reach the audience much more clearly than if I would have tried to explain it verbally. (John, 21/04/2021)
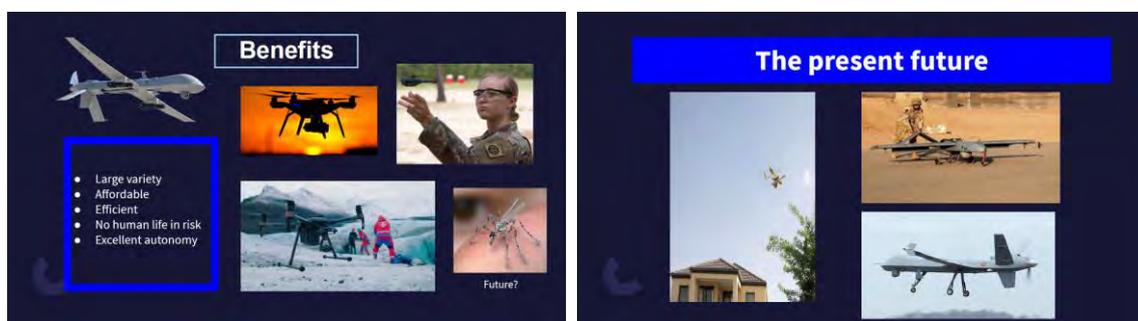
Figure 3. John's slides that exemplify specific uses of drones these days, April 2021.

Visual images were used to describe students' research and results through pictures and contributed to better express aspects of the world. Drawing on Halliday's (1978), Kress and van Leeuwen (1996) theorized that language always fulfills three communicative metafunctions: (1) ideational function by means of which language is able to express ideas about the world; (2) interpersonal function by means of which language communicates the personal stance of how people relate to these ideas about the world; and (3) textual function by means of which language organizes the meaning into coherence parts. This organization of meanings has been used in multimodal studies to research how different resources fulfill these functions. In the present study, and supporting Kress' (2003) argument, students expressed how the images they used in the presentation slides helped them better than words to represent aspects and phenomena of the world, construe their experiences, and communicate ideas about things and situations. As the students explained, the images they used contributed to further expressing ideational content and to aligning their audience with some truths. As participants pointed out, the images of different drones contributed to explaining better the topic of their presentation and helped them show the audience what was out there in the world, particularly, the varied ways drones are used these days. In space-constrained presentations slides, the efficiency and self-explanatory affordance of the visuals they chose to show the versatility of drones was crucial. In multimodal terms, these images expressed high modality; that is, represented closely "what we would expect to find in the real world" (Ledin & Machin, 2020: 63). Most images in students' slides claimed visual truth with high naturalistic modality resembling the way "we would see that something if we saw it in reality" (80) (see Figure 4):

> I selected these images because they can demonstrate how the use of drones is improving. They are currently being used for cartography purposes, in search and rescue missions and in catastrophic situations of different types. I think the fire image is quite representative. I think the images I used engaged my audience. (Ana, 21/04/2021)

Figure 4. Ana's slides showing some current uses of drones, April 2021.

The images in the slides designed by Daniel aimed at shaping his audience's experiences of the world so that they could engage with the reality they depicted. Figure 5 shows Rwanda's blood distribution; the red dots represent the hospitals and the blue ones the distribution centers. Thus, the image of the slide helped him to bring to light Rwanda's current blood demand problem while efficiently avoiding a lengthy verbal description. As he explained:

> Though in my presentation I did not explicitly refer to this slide verbally, I think the images could speak for themselves. The map showed my audience Africa's current health dilemma in terms of time and blood demand and storage. (Daniel, 20/04/2021)

Figure 5. Daniel's slide that shows Zipline blood distribution in Rwanda, April 2021.

As can be interpreted from participants' accounts, the images they selected contributed to better explaining aspects and phenomena of the world better. Presumably, these images also helped the audience understand the current situation in the use of drones. Additionally, the use of images that were real aroused some motivation among the participants, who acknowledged that working with actual images encouraged them to investigate more thoroughly the topic they had to present (see Figure 6). Paul's account shows this aspect:

> These two images certainly made me want to know more about Zipline. I wanted to know how this company provided many countries with instant access to very important medical supplies, such as blood. I was especially interested in their logistics and delivery system and the type of people that worked for the company. (Paul, 21/04/2021)

Figure 6. Paul's slides that show Zipline workers and drones, April 2021.

4.3.2. Agency and engagement in students' online media creations

Ensuring that learners remain engaged in the face of the distractions that currently surround them is one of the big challenges teachers constantly face these days. Teachers become in many cases designers of experiences for students (Schlechty, 2011) aiming at creating motivating tasks with clear objectives that may encourage language learning and use. It is common enough to find L2 lecturers who design tasks that are personally, academically, and professionally meaningful to students (author, 2020). Yet, to captivate them it is also necessary that students feel competent and capable enough to accomplish these tasks and experience the whole process as enjoyable. Regarding these last aspects, Mercer and Dörnyei (2020: 106) contend that for a task to be sufficiently interesting, engaging, and enjoyable, the teacher must consider four important issues: (1) mode of working, (2) the medium of working, (3) the social structure, and (4) the form of task output students create. The mode of working is related to how much time, freedom and resources students have on the task. The medium of working has to do with the different language skill areas that are used in carrying out the task. The social structure involves the place, the ways and with whom students work. Finally, the form of task output concerns the degree of attractiveness involved in the creation of the task output as perceived by students so that they can establish different connections to what they do. The present study reflected on the four issues detailed above. Regarding the mode of working, participants enjoyed reasonable freedom. Among others, they could choose the presentation program, the TED Talk they would use as a baseline for their oral presentation and establish the specific role each of them would take on in the task. In relation to the social structure, data obtained through class observation and interviews suggest that the preparation of the group presentation fully engaged the participants because they were able to realise that this task was academically and personally fulfilling and had a clearly defined and salient outcome. The four participants were fully aware of the implicit facilitative role of the visual mode in any presentation, and the quest for the 'good images' engaged the group.

Arguably, this search triggered group flow and cooperation. In terms of the form of task output or how attractive the task was for the participants, qualitative data allowed the researcher to state that participants experienced some type of enjoyment while designing their presentation. The feeling of pleasure might have derived from knowing they were doing something personally rewarding and from having to research and talk about a technological topic they all enjoyed: the current use of drones. This fact might have had some type of influence on heightening participants' sense of mastery to design appealing multimodal slides to complement their oral presentation. Thus, it can be contended that the multimodal task captivated students on two levels: *activity appeal* and *content appeal* (Mercer & Dörnyei, 2020). The following excerpt illustrates participants' positive emotionality:

> I really liked preparing this presentation and researching curious data about drones to be able to involve my audience, well… ultimately my classmates. So, I learnt and enjoyed. (Ana, 21/04/2021)

The freedom students enjoyed regarding the different aspects described above (i.e., choice of topic to research, design of slides, and way of working together in a collaborative support group), provided them with a sense of autonomy. The degree of responsibility participants had for the outcomes of their presentation contributed to increasing their sense of self-efficacy and their need for recognition and belonging with significant others, that is, their classmates and the lecturer assessing their oral presentation. Participants acknowledged having had an audience in mind during the preparation and design of their presentation. Paul, for instance, aimed at finding attractive and shocking images to captivate his classmates and teacher:

> We had home field advantage because our audience had the same tastes and like the latest designs of drones. Also, I thought of the lecturers who were evaluating the presentation and I wanted to surprise them by showing tiny drones they might not have seen before. We wanted our audience to learn. (Paul, 21/04/2021)

Apart from the verbal mode and the images, participants made a consistent use of colour, layout and typeface to realize their interests as makers of the communicative task and to achieve a particular communicative purpose. Color was a semiotic choice students prioritized. Interviews' accounts showed how the participants were quite aware of the meaning potentials of the colors they chose for their designs. Thus, the dominant use of blue in all the slides served participants to communicate ideas, attitudes, and to create coherence. Paul expressed how he and the other members of the group had decided to mostly use blue

due to its association with knowledge, science, and truth, to subsequently invoke an engineering orientation:

> The choice of blue for the background in all the slides was adopted unanimously. We have seen many other engineering presentations that typically use a blue background. And we wanted to avoid a colorful design overall. I think that all shades of blue can inject designs with seriousness and rigor. So, yes, we used blue deliberately. (Paul, 21/04/2021)

The use of blue for background and white as the font color in all the Google slides also helped the participants to create coherence and classifications. All the headings with the same status used a saturated blue. The background color in all the slides was dark blue. As Ana explained in the interview, coding the slides with the same color contributed to creating some link across their presentation, coordinating the different parts each student had prepared, and having an integrated coherent whole. Color, therefore, was carefully thought through:

> Though each of us designed their own slides, we had previously planned to use the same colors for the background, for raising the titles of all the slides and for the verbal part we could insert. So, though each of us enjoyed reasonable freedom to choose the visuals for the slides, color and typeface was the same for all. Some of us have had the opportunity to see different presentations that use all types of colors and typeface. Though the content might be superb, you get the idea that there's some disorganization and incoherence. (Ana, 21/04/2021)

Color is also used deliberately to create salience. This mode was used to show relations between elements, to "create a more level feeling where parts are of equal salience" (Ledin & Machin: 173). A darker and more saturated blue, for example, was used in all the slides to raise all the title, all of them at the same level across the document. Participants also resorted to typeface to create coherence. Particularly, they used different font sizes to create a hierarchy of salience. Thus, the slides used two different font sizes: one larger in size with increased weight at the top of all the slides that introduced a new section and a smaller version of the same size to write keywords or phrases (see Figures 2 and 5). This coherent use of typeface helped participants bring different kinds of relationships across their design. The interrelatedness between color, typeface and visuals in the participants' design made meaning. Different qualities and features of these semiotic resources were used for their overall design.

The way they organized the different semiotic resources contributed to creating an overall wholeness. Images, color, and layout were therefore intentionally orchestrated and used as rhetorical hooks to attract their audience. Students made strategic choices regarding the multimodal choices available to them in a specific communicational environment. In

some slides, for example, the images of the different types of drones stood out and were the most salient elements in terms of size and foregrounding (see Figures 2 and 3). There were other slides where elements were given salience through foregrounding. This is the case of Figure 1, where Daniel had to talk about the speaker and the way he delivered his TED Talk. The speaker is deliberately placed at the front of the composition and takes up more space on the design of the slide. Yet, though the speaker is clearly given salience in the slide, participants succeeded at connecting the different parts through the use of color, words and typeface.

The interactive design of the multimodal slides positioned the four participants agentively as they resorted to the affordances offered by different modes for making meaning and to the aptness of fit between mode and audience. Just as they aimed at designing visually appealing slides, so too they had a sense of their viewers' expectations. Engagement in the multimodal task was to a large extent motivated by a desire for recognition and connection.

5. CONCLUSION

This interpretive qualitative case study explored ESP students' verbal and non-verbal expressions in the process of designing the presentation slides for an in-class group presentation. Students' compositional processes followed complex multimodal decisions that contributed to expressing their authorial voices and provided them with a sense of mastery and autonomy. Different organizational structures and visuals choices in participants' presentation slides allowed them to express linguistic and disciplinary identities in ways the written texts alone could not. The combinations participants made of image and text clearly allowed for different possibilities not afforded by the verbal mode alone. Additionally, the slides students designed had an audience in mind as they had developed a sense of their classmates and lecturer expectations. Audience, goal, and medium were key factors that gave shape to their multimodal designs.

As language teachers, we should take note of the potentials of multimodality to promote students' agency when expressing discipline and improving academic language. Rather than taking for granted students' digital composing skills, the L2 lecturer is encouraged to develop the multimodal competence in the classroom along with other skills so that students can discern the most appropriate way to communicate and represent a particular topic (Kress, 2005). Students' proactive investment in the multimodal competence will certainly promote their L2 motivation. Non-verbal modes can provide them with an extra layer of confidence to express their individuality and linguistic identity when they struggle with their language skills. Also, the specific design decisions and the rhetorical implications of their modal choices can contribute to a sense of enjoyment and to a more meaningful learning experience as students make the learning material their own.

The inclusion of a multimodal perspective in the ESP curriculum leads the teacher to reflect on the suitability of one mode over another and carry out some type of gain and loss analysis: what may be gained and what may be lost if the design of an ESP curriculum steadily moves from representation through writing to representation mainly through image? Rather than underestimating the relevant role that language always plays, it seems reasonable to ask what is best for what purpose, why is an image used instead of a linguistic description or whether visuals expand or enhance the related message. Continued research into students' multimodal composing is likely to offer responses to these aspects and provide very interesting insights into how emerging technologies effectively engage and motivate students to achieve their learning outcomes.

REFERENCES

Ajzen. I. (1988). *Attitudes, Personality, and Behavior*. Dorsey Press.

Archer, A. (2010). Multimodal texts in Higher Education and the implications for writing pedagogy. *English in Education*, *44*, 201–213.

Bateman, J. A. (2014). *Text and Image: A Critical Introduction to the Visual/Verbal Divide*. Routledge.

Bezemer, J. & Kress, G. (2016). *Multimodality, learning and communication: A social semiotic frame*. Routledge.

Burguess, J. & Rowsell, J. (2020). Transcultural-affective flows and multimodal engagements: Reimagining pedagogy and assessment with adult language learners. *Language and Education*, *34*(2), 173–191.

Canals, L. (2021). Multimodality and translanguaging in negotiation of meaning. *Foreign Language Annals*, 1– 24.

Cimasko, T. & Shin, D-S. (2017). Multimodal resemiotization and authorial agency in an L2 writing classroom. *Written Communication, 34*(4), 387–413.

Darvin, R. & Norton, B. (2017). Language, identity, and investment in the 21st century. In T. McCarty & S. May (Eds.), *Language policy and political issues in education. Encyclopedia of language and education* (pp. 227–240). Springer.

Davis, K. A. (1995). Qualitative theory and methods in applied linguistics research. *TESOL Quarterly, 29*, 427–453. doi: 10.2307/3588070

Dörnyei, Z. (2007). *Research methods in applied linguistics*. Oxford University Press.

Dörnyei, Z. & Kubanyiova, M. (2014*). Motivating learners, motivating teachers: Building vision in the language classroom*. Cambridge University Press.

Dörnyei, Z. (2020). *Innovations and challenges in Language Learning Motivation*. Routledge.

Freschi, A. & Cavalari, S. (2020). Corrective feedback and multimodality: Rethinking categories in telecollaborative learning. *TESL Canada Journal*, *37*(2), 154–180.

Grapin, S. E. (2019). Multimodality in the new content standards era: Implications for English learners. *TESOL Quarterly, 53*(1), 30–55.

Hafner, C. A. (2015). Remix Culture and English Language Teaching: The Expression of Learner Voice in Digital Multimodal Compositions. *TESOL Quarterly, 49*, 486–509.

Halliday, M. A. K. (1978). *Language as a Social Semiotic System*. Edward Arnold.

Henry, A. (2019). Online media creation and L2 motivation: A socially situated perspective. *TESOL Quarterly*, *53*(2),1–33. doi: 10.1002/tesq.485

Henry, A. (2021). Student engagement with digital video production. *ELT Journal*, *76*(1), 109–118,

Hood, S. & Forey, G. (2005). Introducing a conference paper: Getting interpersonal with your audience. *Journal of English for Academic Purposes*, *4*, 291–306.

Hyland, K. (2017). English in the discipline: Arguments for specificity. *ESP Today*, *5*(1), 5–23.

Jewitt, C. (2006). *Technology, Literacy, Learning: A Multimodal Approach*. Routledge.

Jewitt, C. (2013). Multimodal teaching and learning. In C. A. Chapelle (Ed.), *The Encyclopaedia of Applied Linguistics* (pp. 4109–4114). Blackwell Publishing Ltd.

Jiang, L. (2018). Digital multimodal composing and investment change in learners' writing in English as a foreign language. *Journal of Second Language Writing, 40*, 60–72.

Johnson, D. & Kress, G. (2003). Globalisation, literacy and society: Redesigning pedagogy and assessment. *Assessment in Education*, *10*(1), 5–14. doi: [10.1080/09695940301697](10.1080/09695940301697).

Kendon, A. (1980). Gesticulation and Speech: Two Aspects of the Process of Utterance. In Mary. R. Key (Ed.), *The relationship of verbal and non-verbal communication* (pp. 207–227). Mouton.

Kern, R. (2006). Perspectives on technology in learning and teaching languages. *TESOL Quarterly, 40*(1), 183–210.

Kohnke, L. M. A., Jarvis, A. T., & Ting, W. C. A. (2021). Digital multimodal composing as authentic assessment in discipline-specific English courses: Insights from ESP learners. *TESOL Journal*. http://hdl.handle.net/10397/90297

Kress, G. (2003). *Literacy in the new media age*. Routledge.

Kress, G. (2005). Gains and Losses: New Forms of Texts, Knowledge, and Learning. *Computers and Composition*, *22*(1), 5–22. doi: 10.1016/j.compcom.2004.12.004.

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

Kress, G. & Selander, S. (2012). Multimodal design, learning and cultures of recognition. *The Internet and Higher Education*, *15*(4), 265–268.

Kress, G. & van Leeuwen, T. (1996). *Reading images: The grammar of visual design.* Routledge.

Lave, J. & Wenger, E. (1991). *Situated Learning: Legitimate Peripheral Participation.* Cambridge: Cambridge University Press.

Ledin, P. & Machin, D. (2020). *Introduction to Multimodal Analysis, 2nd ed*. Bloomsbury Academic.

Lim, F. V. & Hung. D. (2016). Teachers as learning designers. What technology has to do with learning. *Educational Technology, 56*(4), 26–29.

Lim, J. & Polio, C. (2020). Multimodal assignments in higher education: Implications for multimodal writing tasks for L2 writers. *Journal of Second Language Writing*, *47*(1).

Liu, Y. & O'Halloran, K. L. (2009) Intersemiotic texture: Analysing cohesive devices between language and images. *Social Semiotics, 19*(4), 367–388.

Loewen, S., Shaofeng, L., Thompson, A., Nakatsukasa, K., Seongmee, A., & Xiaoqing, C. (2009). Second language learners' beliefs about grammar instruction and error correction. *The Modern Language Journal, 93*, 91–104. doi:10.1111/j.1540-4781.2009.00830.x

Magnusson, P. & Godhe, A.-L. (2019). Multimodality in Language Education – Implications for Teaching. Designs for Learning, *11*(1), 127–137. DOI: https://doi.org/10.16993/dfl.12

McNeill, D. (1992). *Hand and mind: What gestures reveal about thought*. University of Chicago Press.

McNeill, D. (2005). *Gesture and Thought*. University of Chicago Press.

Mercer, S. & Dörnyei, Z. (2020). *Engaging language learners in contemporary classrooms*. Cambridge University Press.

Miller, L. (2009). Engineering lectures in a second language: What factors facilitate students' listening comprehension? *Asian EFL Journal*, *11*(2), 8–30.

New London Group. (2000). A pedagogy of multiliteracies. In B. Cope and M. Kalantzis (Eds.), *Multiliteracies: Literacy Learning and the Design of Social Futures* (pp. 9–38). Macmillan.

Norton, B. (2000). *Identity and Language Learning: Gender, Ethnicity and Educational Change.* Pearson Education.

O'Halloran, K. L. (2005). *Mathematical Discourse: Language, Symbolism and Visual Images*. Continuum.

O'Halloran, K. L. & Lim, F. V. (2009). Sequential Visual Discourse Frames. In E. Ventola & J. Moya (Eds.), *The World Told and the World Shown: Multisemiotic Issues* (139–156). Palgrave Macmillan.

Patton, M. Q. (1990). *Qualitative research and evaluation methods* (2nd ed., pp. 169–186). Sage.

Ponzio, C. & Deroo, M. (2021) Harnessing multimodality in language teacher education: expanding English-dominant teachers' translanguaging capacities through a Multimodalities Entextualization Cycle, *International Journal of Bilingual Education and Bilingualism*, 1–17. doi: 10.1080/13670050.2021.1933893

Potts, D. (2013). Plurilingualism as a Multimodal Practice. *TESOL Quarterly*, *47*(3), 600–614.

Prensky, M. (2001). Digital natives, digital immigrants, part 1. *On the Horizon*, *9*(5), 1–6.

Rowsell, J. (2013). *Working with Multimodality*. Routledge.

Schlechty, P. C. (2011). *Engaging Students: The Next Level of Working on the Work*. Jossey-Bass.

Selander, S. (2008). Designs for learning – a theoretical perspective. *Designs for Learning*, *1*(1), 4–22.

Stam, G. (2008). What gestures reveal about second language acquisition. In S. McCafferty, & G. Stam (Eds.), *Gesture: Second Language Acquisition and Classroom Research*. Lawrence Erlbaum Publishers.

Stam, G. (2014). Why gesture! *Al-Forum-The Newsletter of the Applied Linguistic Section*, TESOL, September 2014.

Stetsenko, A. (2019). Radical-transformative agency: continuities and contrasts with relational agency and implications for education. *Frontiers in Education*, 4. DOI:10.3389/feduc.2019.00148

Tardy, C. M. (2005). Expressions of disciplinarity and individuality in a multimodal genre. *Computers and Composition*, *22*(3), 319–336. doi: 10.1016/j.compcom.2005.05.004.

Unsworth, L. & Cleirigh, C. (2009). Multimodality and reading: The construction of meaning through image-text interaction. In C. Jewitt (Ed.), *The Routledge Handbook of Multimodal Analysis* (pp. 151–163). Routledge.

Ushioda, E. (2008). *Motivation and good language learners.* In C. Griffiths (Ed.), *Lessons from Good Language Learners* (pp. 19–34). Cambridge University Press.

Ushioda, E. (2009). A person-in-context relational view of emergent motivation, self and identity. In Z. Dörnyei, & E. Ushioda (Eds.), *Motivation, Language Identity, and the L2 Self* (pp. 215–28). Multilingual Matters.

van Lier, L. (1996) *Interaction in the Language Curriculum: Awareness, Autonomy and Authenticity*. Longman.

van Lier, L. (2000). From input to affordance: Social-interactive learning from an ecological perspective. In J. P. Lantolf (Ed.), *Sociocultural theory and second language learning* (pp. 245–259). Oxford University Press.

van Lier, L. (2007). Action-based teaching, autonomy, and identity. *International Journal of Innovation in Language Learning and Teaching*, *1*(1), 46–65. doi: 10.2167/illt42.0.

Walsh, M. (2015). Literacy education. In J. D. Wright (Ed.), *International Encyclopaedia of the Social & Behavioural Sciences, 2nd edn.,* (pp. 246–250). Elsevier.

Zhao, S., Djonov, E. & van Leeuwen, T. (2014). Semiotic technology and practice: A multimodal social semiotic approach to PowerPoint. *Text & Talk*, *34*(3), 349–375.