Introduction: The nexus of multimodality, multimodal literacy, and English language teaching in research and practice in higher education settings

Belinda Crawford Camiciottoli (University of Pisa) (corresponding author)

Mari Carmen Campoy-Cubillo (Universitat Jaume I)

1. Introduction

As a field of scientific enquiry, multimodality came into its own in the 1990s. Pioneering studies by Roth (1994), Kress and van Leeuwen (1996), Lemke (1998), van Leeuwen (1999), and Martinec (2000) are examples of early research in this area that highlighted the key contribution of non-verbal communicative resources (e.g., gestures, visuals, music, bodily movement, facial expression) and their interplay to the construction of meaning during social interaction. Since then, there has been a growing interest in the study of multimodal phenomena among scholars with various disciplinary orientations and different perspectives, ranging from theory to application. Indeed, according to Jewitt (2009, p. 12), "multimodality can be understood as a theory, a perspective or field of enquiry or a methodological application".

In this Special Issue, we are concerned with pedagogical applications of multimodality in the context of English language teaching in higher education, with particular attention to both fostering and leveraging *multimodal literacy* among adult English language learners. Multimodal literacy is now a widely established concept that refers to the ability to successfully engage with texts that integrate different semiotic resources. It is defined by Walsh (2010) as the ability to

construct meanings through "reading, viewing, understanding, responding to and producing and interacting with multimedia and digital texts" (p. 213). This definition is clearly inspired by the notion of *multiliteracies*, which was originally formulated by New London Group in 1996. This was a group of scholars (including Norman Fairclough, James Paul Gee, and Gunther Kress) who met to propose ways to change teaching and learning paradigms in response to changing forms of communication. They stressed the need to go beyond the traditional interpretation of literacy in terms of the capacity to read and write (Gee, 1996), and urged educators to utilize new technologies to enhance multimodal literacy (New London Group, 1996). Thus, multiliteracies can be seen as drawing from (1) New Literacies Studies that interpret literacy as a sociocultural (vs. cognitive) phenomenon, driven by changing social practices involving digital communications and new media (Gee, 1996; Street, Pahl, & Rowsell, 2014) and (2) Multimodal Studies that apply new theoretical and methodological frameworks for analysing communication which integrates modes beyond verbal language (Kress & van Leeuwen, 1996; O'Halloran & Smith, 2011). The concept of multiliteracies is now widely applied in general education and the need for educators to develop multiliteracies among learners is beyond dispute (Royce, 2002; Jewitt & Kress, 2003).

Multimodal literacy has also become highly relevant for language teaching thanks to an increasing awareness that the multimodal approach can help students learn to exploit semiotic modes beyond verbal language (e.g., visual, gestural, spatial) to both understand and produce texts in the target language more effectively (O'Halloran, Tan, & Smith 2016). It may also heighten their awareness of the target culture, particularly in relation to different styles of non-verbal communication across cultures (Busà 2010, 2015). Of particular interest in the context of language learning is how non-verbal elements reinforce and/or add meaning to verbal expressions and thus facilitate comprehension, but also how multimodal approaches can serve to motivate learners (Shih, 2014). Over the years, there have been some studies pointing to the advantages of using techniques and resources that highlight non-verbal modes, such as images, gestures, and facial expressions,

during language teaching (cf. Sueyoshi & Hardison, 2005; Guichon & McLornan, 2008), thereby supporting the important role of multiple semiotic resources in the construction of meaning.

Linguists and practitioners working in the field of English language teaching are now called upon to find ways to adapt and integrate multimodal and multimedia resources for classroom materials and activities, representing an ongoing challenge in light of the ever-growing influence of multisemiotic digital input in the lives of learners, also outside the classroom (Street, Pahl, & Rowsell, 2011). This is now an imperative if we hope to keep pace with and effectively engage learners' growing interest and their own expertise in using multisemiotic digital resources. Indeed, rapid development and ongoing innovation in the field of digital technology over the last two decades have profoundly changed how we communicate and interact with others in all realms of life. This trend has been particularly evident in university-level education, as pointed out by O'Halloran, Tan, and Smith (2016):

> Changes in higher education, especially in the use of digital technology, have revolutionized traditional academic practices, with an increasing recognition of the need for students and teachers to develop multimodal competencies across a range of communicative platforms. (p. 256)

To accomplish this objective, it is necessary to increase awareness of the role of multimodality during the communication process by implementing new practices for teaching and learning in the university English language classroom. These may involve the effective exploitation of multimodal and multimedia resources during teacher-student interaction in the classroom, as well as explicit instruction on the role of multiple semiotic modes and their interaction in the construction of meaning (see for example Coccetta, this issue). Moreover, language teachers in general, as well as those working in online environments in particular, are further challenged to create modes and practices for language competency assessment that take multimodal communicative systems into consideration.

Part of the challenge of creating and developing practices to promote multimodal language competence is the fact that multimodal environments frequently call for a mixed skills approach. This means that in multimodal task design the teacher needs to value the importance of the different interacting modes and analyse whether a particular task involves mainly an individual skill, or whether a second or third skill become relevant to successfully complete the task. It also means paying attention to the number of modes in the event and the importance they have in that particular situation.

Thus, when designing multimodal tasks, we need to consider *modal density*, a concept which combines the notions of *modal intensity* and *modal complexity*, as defined by Norris (2004). The former refers to the "weight" that a mode may carry, while the latter indicates the degree of intricacy or complex connections created among modes. Thus, modal density does not necessarily imply complexity, but when it does, the teacher needs to detect that complexity in order to either simplify the input or design an adequate and fair response option or assessment method. Thus, input and output modes and their complexity necessarily become part of the multimodal equation. Therefore, multimodal task complexity involves considering: (1) mode interaction in task design and mode choice as a task variable, (2) assessment methods related to the task (diagnostic, formative, summative) that take into account the interplay between input and output modes, and (3) the cognitive demands of the task while accessing, processing and understanding information from different verbal and non-verbal sources. In sum, on the part of the teacher, multimodal task complexity implies being aware of the multimodal architecture complexity.

Given the ongoing evolution in technologies that can be harnessed for educational purposes, it is important to broaden our knowledge of the multimodal dimension of classroom resources, also by means of new analytical methods and instruments for multimodal discourse analysis, for example software for the multimodal annotation of texts (see Bonsignori, this issue). Such tools can shed new light on *intersemiotic complementarity*, i.e., how different modes "complement each other in the ways that they project meaning" (Royce, 2007a, p. 63). This knowledge can then be

implemented to design and develop up-to-date multimodal approaches for the English language classroom in higher education settings that capitalize on new affordances characterized by multiple modes of communication and different types of media.

The aims of this Special Issue are to (1) advance the current state of research-based knowledge about how multimodal and multimedia resources can be leveraged to enhance multimodal communication practices in English language teaching in higher education, and (2) to provide a platform for original research-based practical applications that incorporate innovative multisemiotic resources and techniques, thereby offering new perspectives on the benefits of the multimodal approach when teaching English for both general and specific purposes at the university level. In the following section, we discuss the role of multimodal literacy in the context of enhancing language proficiency as the underlying objective of English language practitioners.

2. Multimodality and language proficiency

The notion of proficiency in a second language has typically focused on the linguistic dimension. However, thanks to ongoing advances in digital technology which offer new communicative contexts and genres, there is a growing awareness of the importance of fostering multimodal literacy in English language teaching contexts (Royce, 2007b). As Hafner (2014, p. 655) points out, English language teaching should "be expanded beyond the traditional focus on speech and writing to the production of multimodal ensembles, drawing on a range of other semiotic modes".

An example of the heightened interest in including multimodal literacy in language teaching is also reflected in the updated version of the *Common European Framework of Reference* (2001-2018). In the 2001 *Common European Framework of Reference for languages: Language, Teaching, Assessment* (CEFR) document, there was no mention of the concepts of "multimodality", "(communicative) mode" or even "online" reception, production or interaction. In contrast, the new 2018 CEFR project specifically addresses online communication by providing two new scales for

"online conversation and discussion" and "goal-oriented online transactions and collaboration" (p. 80 ff.). Both categories recognize the importance of the typical multimodal activity that is intrinsic to the use of the Web, with different degrees of synchronous or asynchronous computer-mediated interaction (Abrams, 2003).

Thus, online conversation and discussion is regarded as a multimodal phenomenon where specific weight is given to how speakers deal with social exchanges in open-ended environments, where no definite timing, answer, or rule is necessarily provided. As stated in the framework (2018, p. 98), there are a number of relevant new concepts to bear in mind:

- instances of simultaneous (real time) and consecutive interaction, the latter giving time to prepare a draft and/or consult aids;

- participation in sustained interaction with one or more interlocutors;

- composing posts and contributions for others to respond to;

- comments (e.g., evaluative) on posts, comments and contributions of others;

- reactions to embedded media;

- the ability to include symbols, images, and other codes for making the message convey tone, stress and prosody, but also the affective/emotional side, irony. etc.

These CEFR requirements for successful online communication bring to the fore the relevance that non-verbal modes have in combination with verbal modes. The last two concepts specifically demand some degree of multimodal competence and the development of multimodal strategies (see multimodal sociocultural strategies below, section 2.1).

The multiple verbal and non-verbal resources described in the CEFR play an important role in real life communication, particularly in service and business-related areas. In many professional situations nowadays, negotiations are carried out through a number of online devices that usually imply multimodal competences. This is reflected in the CEFR (2018, p. 100) scale named "goal-

oriented online transactions and collaboration", a scale where "multimodality is increasingly a key feature and resource, and the descriptors therefore assume the exploitation of different online media and tools according to context".

However, although the use of different modes has been acknowledged, there are still many questions to answer, regarding both the description of what different communicative modes contribute to communication and which multimodal elements need to be observed in receptive, productive, and interactive tasks, so that they may be valued and assessed. Generally speaking, there seems to be more awareness of the use of different means and media than of the use of modes within those media. In fact, in the *CEFR Companion Volume with New Descriptors Provisional Edition* (2017, p. 225) the scale for "*Multimodal* conversation and discussion" (our emphasis) is introduced, but is later changed in the 2018 version into "*Online* conversation and discussion" (our emphasis) (p. 232). In the changed new version, key ideas are lost, such as the importance of understanding communicative modes and their affordances, as well as mode ensembles as essential units including verbal and non-verbal communicative modes. It is the understanding of how multimodality operates in communicative environments that can provide us with the necessary tools to identify, analyze, interpret, and apply multimodal communication effectively. Only this understanding of multimodality may provide us with the tools to evaluate multimodal proficiency levels. Beltrán-Palanques and Querol-Julián (this issue) address this need by situating modal density and mode ensembles in relation to a particular event with students at two different language proficiency levels. The implications of this approach are that multimodality in pedagogical contexts should also be related to a communicative situation or event, and that researchers should ask themselves to which extent a particular event calls for specific verbal and non-verbal support, and how appropriate performance would differ from one proficiency level to another.

The topics of the articles in this Special Issue respond to the need for more detailed descriptors in the field of multimodality and language teaching. They deal with various aspects of the four core abilities of language proficiency (i.e., reading, writing, listening, and speaking) from a

multimodal perspective. In the following sections, we discuss relevant work relating to each of the four skills that has drawn on multimodality as an underlying conceptual framework. In this way, we hope to provide readers with a useful backdrop for the innovative contributions that follow which address the four skills to different degrees and from a variety of perspectives.

2.1 Reading

The majority of rich multimodal reading experiences take place in computer-mediated environments. The reading paths that these environments facilitate are structured in syntagmatic and paradigmatic configurations (Francesconi, 2016). The configuration of semiotic resources co-occurring in a page make up the syntagmatic multimodal patterning, while mode interaction across pages enhanced by the combination of hypertextual and multimodal pathways, or what Lemke (2002) calls *hypermodality*, constitute the paradigmatic multimodal structures. These paradigmatic and syntagmatic multimodal configurations make it possible for the reader to take a more active part in the e-reading process, since choices need to be made to continue reading and interpreting information through the mode selection by choosing one reading path or another.

In order to be able to actively select and interpret information in multimodal tasks, readers also need to develop specific reading strategies that tend to be more complex than those followed in reading texts in which the verbal content predominates and which have low interaction levels. In this regard, Liu's (2013) framework distinguishes four interpretative strategies for visual images: meta-interpretative strategies, perceptual strategies, analytical strategies, and sociocultural strategies. Although this author focuses on images as a mode, his framework can be used to understand mode co-deployment in multimodal texts. In fact, Liu's framework could be extended for the interpretation of multimodal comprehension strategies. Each of Liu's strategies comes into action in the observation, interpretation, and participation in multimodal texts. Multimodal comprehension strategies may also be needed for one or more modes, and different strategies may be needed in the different modes at the same time while processing a multimodal ensemble (i.e.,

different modes may also call for different strategies in the interpretation and participation in a particular event). Thus, the extended framework would first include multimodal meta-interpretative strategies that are related to the awareness of multimodal meaning construction. This means that modes in the multimodal ensemble may create new meanings that do not necessarily coincide with the separated meanings brought to mind by each of these modes individually, or in other mode ensembles. Once these meta-interpretative strategies are operational, we must then pay attention to multimodal perceptual strategies, the abilities to notice different elements in a multimodal text, or what the reader is able to perceive. Perceptual strategies are essential in the sense that what is not perceived cannot be interpreted and part of the information may be lost or misunderstood. Multimodal analytical strategies are part of the comprehension process and imply making sense (beyond literal meanings) of the interrelationship within and among modes. In this issue, for instance, students in the paper presented by Bellés-Fortuño need to combine multimodal perceptual and analytical strategies. They first use their multimodal perceptual strategies and then later develop multimodal analytical strategies in interpreting the medical leaflet genre and being able to transfer multimodal information to their own medical leaflets. Finally, multimodal sociocultural strategies should enable readers to perceive, analyse and interpret cultural, social, scientific, ecological and political ideologies within a particular culture. In Wawra's paper, multimodal meta-interpretative and sociocultural strategies are necessary in the analysis of political cartoons where ideology predominates as a component of the multimodal texts.

The multimodal strategy framework can be an important element in developing educational models to promote multimodal literacy and to be able to assess students' performance in this respect. In a language learning environment, the analytical and sociocultural strategies are bound to be related with more complex multimodal contexts and therefore with higher language proficiency levels. When teaching multimodal reading skills, it is important to bear in mind how modes interact and how they are processed by the reader. For example, visual images need to be interpreted in relation to written verbal text and vice-versa. Then, the reader may choose one or more of the

modes presented to construct meaning and thus is given more opportunities to interact or choose his or her own reading path. Understanding how learners develop strategies to cope with information in multimodal texts is therefore of paramount importance, as is discerning whether the strategies followed by learners are adequate or successful.

2.2 Writing

In the context of L2 writing competence, the multimodal approach may involve teaching students how to produce written texts on the basis of input from other semiotic resources. For example, Royce (2007b) proposes activities for the L2 writing classroom in which students practice creative narrative writing by extracting visuals from reading assignments, or even draw their own visuals, and then utilize them to write an original story. Similarly, Stein (2000) provides an example of a project in which students take photographs and create a poster that depicts literacy practices in various environments, which they then analyse in an academic essay. She refers to this process as "re-sourcing resources", (Stein, 2000, p. 336), where learners transform representational resources into a new context with new meanings. According to this author, incorporation of the visual mode was useful to help English language learners produce better structured and more logically flowing written texts. Liang (2010) describes a study involving an EFL writing class in which students made use of various multimodal prompts (audiovisuals, films, advertisements, art work) in a process-oriented activity to brainstorm ideas and then draft, revise, and edit expository texts by means of synchronous online peer response sessions. Thus, these studies highlight the potential of harnessing multiple modes and media to help L2 learners write more effectively.

The multimodal approach to L2 writing may also entail the creation of texts that combine multiple modes, i.e., *multimodal composition* (Selfe, 2007), whereby learners produce texts that integrate words, image, and sound by exploiting the affordances of digital technology (see DePalma & Poe Alexander, Bellés-Fortuño, and Chen, this issue). Bowen and Whithouse (2013, p. 7) refer to this process as *multimodal composing*, defined as "the conscious manipulation of the interaction

among various sensory experiences—visual, textual, verbal, tactile, and aural—used in the processes of producing and reading texts". Yang (2012) analysed multimodal digital stories (including voice narration, written text, images, music, animation) that were crafted by English language learners. The study not only shows how multimodal composition can be implemented in the English language classroom, but it also sheds light on how learners approach designing multimodal texts and assigning meanings to semiotic resources. Nelson (2006) illustrates a multimedia writing project in which L2 students created a digital essay that incorporated text, image, speech, and music, by means of multimedia software. Follow-up analysis based on the digital essays as well as supporting interviews and student journals suggested that multimodal authorship can benefit L2 learners by enhancing their ability to more effectively convey their authorial voices.

Multimodal composition can also be an effective way to motivate English language learners, also because it allows them to engage with the multisemiotic digital resources that the vast majority already embrace them in the context of their daily lives (Street, Pahl, & Rowsell 2011). Indeed, several studies have also highlighted the positive effects of multimodal composing in the classroom, serving to stimulate interest and provide English language learners with a sense of autonomy and self-efficacy (e.g., Jiang & Luk, 2016; Guichon & McLornan, 2008; Chen, this issue).

2.3 Listening

Traditionally, in most language learning situations, one-way audio listening has been the main practice (that is, audio-recorded texts with no interaction on the part of the listener and usually one correct answer for any comprehension question). The communicative approach to language learning has slowly brought in changes in relation to how the listening construct is understood and which task types and comprehension questions are more relevant for developing listening competence. We now discuss two-way listening (interacting with the interlocutor, see Beltrán-Palanques and Querol-

Julián in this issue) and may also extend the audio-only listening mode to the possibility of using other non-verbal modes and different media.

The 2001-2018 CEFR has introduced specific listening events that are part of real communication (understanding communication between other speakers and listening as a member of a live audience), and basic media have been incorporated as part of the framework (announcements and instructions, radio and audio recordings). A separate scale has been developed (2018) for audiovisual skills where watching TV, film, and video are considered important receptive activities. Some non-verbal issues and cues are included in this scale, such as understanding contextual clues or exploiting visual information and (multimodal) typography (Serafini & Clausen, 2012). However, particular modes are associated with specific proficiency levels in the CEFR. For instance, exploiting visual information is related to the lower levels while interpreting contextual cues is related to the higher ones. This mode/proficiency relationship is not always the case: non-verbal resources form part of simple and complex listening communicative situations and modal density and complexity will determine the usefulness of these resources in specific events and their relevance for a particular group of learners.

There are more key issues in multimodal listening when learning a foreign language. As suggested in Campoy-Cubillo and Querol-Julián (2015), assessment criteria for multimodal listening tasks should bear in mind three important aspects that will determine task or test format. One of them is the relationship between question type and modes employed in the listening task/test. Secondly, we need to analyze the complexity of the listening task in terms of verbal and non-verbal information processing. Finally, teachers and test designers need to consider the way test administration procedures are aligned with the question type and with the complexity of the test (e-testing vs. paper).

From a professional point of view, listening effectively is key in all kinds of personal relationships, customer relationships (e.g., business), and service encounters (e.g., public services, health care). Language learners should be trained for these future professional needs by attending to

specific professional purposes and situations. The complexity of the listening purpose and type of information needed in real multimodal environments is evident if we consider the different listening types defined for professional purposes: attentive, active and empathetic (Itani & Inyang, 2015; Van Slyke, 1999). In these listening types, non-verbal information is an essential part of the listening experience (see Franceschi for a classroom experience with medical students, this issue)

In sum, the teaching and assessment of multimodal listening skills are especially challenging for teachers. The main difficulty lies in the creation of multimodal listening tasks that are viable in terms of mode combinations and test format.

2.4 Speaking

When developing speaking skills in the English language classroom, the multimodal approach is reflected in the notion that effective oral communication entails not only language proficiency, but also knowledge of how to use language appropriately in various social settings, including the paralinguistic (e.g., intonation, stress, voice quality) and extra-linguistic features (e.g., gestures, facial expression, body posture/positioning) that typically accompany speech during face-to-face interactions. (Shumin, 2002). Thus, it is important to enhance learners' awareness of the contribution of non-verbal elements during communication. For example, Rance-Roney (2010) advocates for explicitly teaching appropriate body language in the context of learning functional language and conversational strategies, i.e., appropriate ways to disagree or to interrupt. Audiovisual resources can also be used for awareness-raising purposes. Learners could view video clips and then discuss the non-verbal features that they noticed and offer interpretations of their meanings. Busà (2010) describes a classroom application in which students were exposed to various multimedia resources (e.g., still images, digital slides, audio files, YouTube videos) and then engaged in metalinguistic reflections on how both verbal and non-verbal communication were used in different settings. Thus, in addition to raising learners' awareness of the role of non-verbal communication, they also had the opportunity to practice speaking skills.

However, perhaps the most effective way to help learners become more proficient speakers is to implement activities (e.g. role-playing, pair work) that allow them to actually practice the conscious incorporation of non-verbal elements while speaking (see Franceschi, this issue). Training people to use non-verbal cues successfully is a well-established practice in professional contexts (Goman, 2008; Munter, 2009). Yet when working with L2 learners, "non-verbal elements such as intonation, gaze, facial expressions, body movements and posture play an absolutely crucial role […] and cannot be neglected if the aim of instruction is to achieve successful communication" (Busà, 2010, p. 53). However, in the language classroom, an approach that views speech as a *multimodal ensemble* (Kress, 2011) in which all modes contribute to meaning during social interaction is rather infrequent (Celce-Murcia, 2008). Fortunately, various types of performance or (re)enactment activities that highlight the role non-verbal communication have been recently been proposed for L2 settings. Busà (2015) illustrates a course that aimed to help EFL students learn to integrate verbal and non-verbal communication skills. She engaged the students in a series of oral activities, including both live classroom performances and self-produced videos. The students were then given feedback from both the teacher and their classmates in relation to communication style, with particular reference to gestures and body language. In the context of a business communication course with L2 learners, Palmer-Silveira (2015) explains how learners used physical space, head movements, hand/arm gestures, and eye contact during an activity aiming to help them deliver effective company presentations. The learners were then able to evaluate the effectiveness of their own non-verbal styles by observing video recordings of their presentations. The studies reviewed in this sub-section suggest that providing L2 learners with opportunities to consciously practice and evaluate their own non-verbal competence can help them become not only more proficient but also more confident speakers of English.

3. The contributions to this Special Issue

As previously suggested, multimodal texts and the corresponding learning activities that may be associated with them are often complex, with various modes coming into play in more or less prominent ways. This means that a multimodal approach to teaching English language competence cannot always be strictly matched to only one of the four core skills. In the following paragraphs, we have attempted to present the contributions to this Special Issue according to the skill which seems to be most foregrounded in the learning process.

Two papers offer new insights into how multimodality can be used to enhance reading skills. In the first paper, Wawra focuses on the potential of multimodal texts to help English language learners become more critical readers. She shows how political cartoons can be used in the classroom to foster and enhance multimodal literacy. In particular, as multimodal artefacts, the cartoons provided students with the opportunity to analyse and interpret ideological meanings and key political themes and narratives, as well as stylistic devices such as metaphor, metonymy, personification, and blending, which all contribute to multimodal meaning construction. The aim was to improve students' understanding of how different modes interact to make meaning in ideological and creative ways, in order to critically evaluate the message of the cartoons. In line with Royce (2015), who also analysed legal cartoons following a systemic functional approach and a social-semiotic perspective, Wawra's paper puts this framework into action, taking it to the language classroom and carefully selecting a number of political cartoons that the students can easily relate to. The model developed by Wawra in a learning environment also illustrates the concept of "visually encoded and contextually loaded *intersemiotic interaction*" identified by Royce (2015, p. 721), and is relevant for independent and proficient users (B1 to C2 in the CEFR).

In the second paper, Coccetta illustrates a university English language course aiming to help students perform text analysis with tools that allow them to identify the various semiotic resources that contribute to meaning and to develop appropriate metalanguage for describing this phenomenon. The course thus reflects a central idea for the teaching of multimodal frameworks in higher education context, i.e., awareness raising. Coccetta's systemic functional approach to genre

analysis for instructional and procedural text types focuses on raising students' awareness of the different semiotic resources employed in the selected text types. The author works with two different texts: one contains animated images while the other has static visual information. For the latter, students are provided with a multimodal version and annotations, as well as the same texts deprived of their visual mode information. The author shows how these two strategies were valuable for creating an awareness of how visual modes are employed and how the distinct resources are selected to create meaning. Regarding the animated text, a multimodal transcription of the video was used for the students to process and interpret how the different semiotic resources interact to create meaning. By providing the students with the multimodal transcription of the video, they could discuss the different agents participating in the process and how they are represented both verbally (soundtrack) and non-verbally (animated images). The study offers a multimodal systemic functional approach for the development of multimodal receptive skills, where students reported on how their multimodal competence improved through the development of multimodal analytical strategies to observe the effects that mode choice and mode combination may have for specific contexts.

Other important issues connected to the awareness-raising process are content and background knowledge. While Coccetta's awareness-raising proposal focuses on receptive multimodal skills for reading, Bellés-Fortuño's paper shows how medical students apply multimodal reading as part of a three-phase process aimed at producing a multimodal composition (a medical informative leaflet), reflecting a topic that is relevant to their future profession. In the first phase, multimodal reading skills come into play when students use their content and background knowledge related to medical topics as a starting point to become more aware of how medical leaflets are composed. Indeed, knowledge of text type and self-learning strategies to read and analyse text format and use of communicative modes are part of the awareness-raising phase. In the second phase, students work in groups to become engaged in the development of their productive skills to generate a multimodal medical leaflet. The final stage of their project involves

the oral presentation of their leaflets and provides them with the opportunity to show the effectiveness of their creation, since it is through their presentation of the different modes employed in the design of the medical leaflet that the appropriateness of their choices (or lack of it) is made visible.

The concept of multimodal composition is also prominent in two other contributions that focus mainly on writing skills. In the first paper, DePalma and Poe Alexander address the challenges of instructors who teach multimodal composition in First Year Writing courses in universities by proposing a pedagogical approach which they call *distributed collaboration*. This approach aims to help instructors cope with the multi-faceted nature of the technologically advanced texts that students are now called upon to create by enlisting the collaboration of specialists with targeted expertise, for example, new media specialists, software specialists, and local professionals. Such collaboration not only facilitates and increases the efficacy of the instructor's task, but more importantly, it allows students to enhance their multimodal literacy while producing high-quality multimodal artefacts at the same time. The paper describes a graduate level pedagogy course to illustrate how students made use of distributed collaboration in a team project that required them to compose a multimodal text (i.e., a short video profile of writing courses offered at their university), and how this approach helped them to improve their multimodal writing skills and gain a better understanding of the potential of different semiotic modes in the meaning-making process. The authors then show how the distributed collaboration approach can be adapted to English language teaching, specifically to give learners opportunities to 1) practice speaking and listening skills when interacting with the enlisted expert collaborators, 2) engage in critical reading as they research sources and materials for the project, and 3) produce written output for the student-produced video.

In the second paper, Chen illustrates how students were involved in a multimodal composition project to produce a video on the topic of cyberbullying. The aim of the project was for students to create a multimodal artefact that would also help them develop an awareness of *digital*

empathy, which refers to the ability (both cognitive and emotional) to engage in reflection and behave in socially responsible ways when using digital media. The author points out that teaching digital empathy has become an important issue in today's technology-driven world where a growing use of digital instruments for communication increases vulnerability, especially among young people. From the language learning perspective, the project was designed so that students were involved in activities that allowed them to practice their writing skills, but also listening and speaking skills. Chen carried out a post-project questionnaire to explore students' perceptions about the video project experience and what they had learned from it. The results indicated that the students had become more cognizant of the importance of digital empathy and that the video project had provided them with a new learning opportunity. They also expressed an awareness of how semiotic resources such as images, sound effects, and music made important contributions to the message that they sought to convey. The study suggests that the digital empathy video project was useful not only to enhance language skills and multimodal literacy, but also to expose students to important social issues.

The following two papers address listening skills and interactive skills. In the paper by Bonsignori the focus is on listening skills and their use in films to promote the learning of genre specific characteristics in specialized professional domains. The second paper presented by Morell discusses an interactive event in a university classroom and pays particular attention to the speaking component of the interaction.

Research conducted by Bonsignori exemplifies the use of films as a multimodal resource to teach ESP, most notably, on how they are an effective multimodal artefact for teaching language and culture. The films are part of a multimodal corpus compiled at the Language Center at the University of Pisa that includes searchable linguistic, pragmatic and cultural annotations elaborated with the ELAN software (Wittenburg et al., 2006). The added value of this paper is that it can easily be related to the 2018 CEFR audiovisual reception descriptors in the "watching TV, film and video" scale. Not only that, it also shows how this can be taken to the ESP arena in three specialized

domains (politics, law, and tourism). While other descriptors such as those included in the scale for "listening as a member of a live audience" (CEFR 2018, p. 59) include reference to specialized and professional domains (e.g., B2 - Can understand the speaker's point of view on topics that are of current interest or that relate to his/her specialized field, provided that the talk is delivered in standard spoken language), the audio-visual reception descriptors do not include such references. Bonsignori shows how this may be done and how aspects such as specialized lexis, bodily behavior in political discourse, specific legal contexts and situations, or soundscape and visual cues in tourism clips can be analyzed and introduced in the ESP classroom.

Morell illustrates an experience in an advanced proficiency level classroom where the pedagogical functions of semiotic resources used by a university lecturer are presented as part of the lecturer's instructional strategies. In such a situation, students as lecture listeners will need to interpret such communicative devices adequately. The author studies the instructor's multimodal competences in an English as a Medium of Instruction (EMI) lecture. She pays attention to how the instructor sequences multimodal resources to make students understand a key concept in the lecture. Specifically, she explains how the EMI instructor works with multimodal ensembles to instantiate the setting up, supervising and eliciting moves of the activity. This study includes the analysis of temporal and spatial elements and their relationship with gaze, gestures, and language use. The interesting aspect of this paper is that it focuses on instruction as a multimodal event. Instruction is a relevant aspect of language teaching and a multimodal approach to language teaching cannot be understood unless learners' input and the instruction they receive belong to the same multimodal agenda. In this case, the study reveals how ensembles of three to four modes seem to be useful in order to achieve the desired effects to involve students in the understanding of conceptual meaning.

The last two papers address mainly speaking skills. In the first paper, Beltrán-Palanques and Querol-Julián tackle the issue of interlanguage pragmatic competence from a multimodal perspective. Using innovative software that enables the annotation and elaboration of multimodal input, they provide an in-depth linguistic and extralinguistic analysis of video-recordings involving

learners of English as an additional language at two distinct proficiency levels: lower intermediate and upper intermediate. The learners who participated in the study engaged in face-to-face interaction, specifically in a role-play task that focused on the speech act of complaining as a highly complex face-threatening act that typically involves extended sequences of discourse, rather than the more restricted adjacency structure that is typical of most other speech acts. Thanks to the software that was applied, the authors were able to identify a whole range of semiotic resources that accompanied the verbal message (i.e., gestures, facial expression, head movement, and gaze direction) and measure their relative contribution in terms of modal density which, as explained in section 1, encompasses both modal intensity (the weight carried by a particular mode) and modal complexity (the intertwining of multiple modes). The findings revealed that the verbal mode was only one of the many modes employed by the participants during a complaint sequence and was not necessarily the most prominent one. The implications are that our interpretation of pragmatic competence needs to be expanded to include the multimodal dimension, which should also be reflected in the teaching and assessment practices of the English language classroom.

In the final paper of the Special Issue, Franceschi proposes an innovative method to help future physicians improve their performance during physician-patient communication, which comprises the awareness and application of multiple semiotic modes beyond verbal language. The author includes a review of existing materials for teaching medical English, pointing out that the focus is almost exclusively on verbal meanings, with activities designed to build vocabulary and phraseological structures associated with the medical profession. This verbal focus results in a lack of attention to communication as an embodied phenomenon, which also entails non-verbal elements such as facial expressions, gestures and body movements. In physician-patient interactions, such non-verbal elements can greatly contribute to establishing trust and encouraging patients to comply with physicians' recommendations. Franceschi's approach to teaching students how to communicate effectively with patients is inspired by Gestalt psychotherapy/counselling, incorporating four methods used in this context: the empty chair technique, the making-the-rounds

exercise, the exaggeration task, and empathic listening. Integrating such activities into the medical English classroom can result in a more humanistic approach, as well as more emotional awareness during medical training, which can ultimately enhance the quality of healthcare. These Gestalt methods also provide medical English learners with ample opportunities to practice their speaking skills through various types of exercises based on interviewing, role-playing, and social conversations.

This paper provides an illustration of a holistic approach to developing medical English students' speaking skills, but with particular attention to emotional awareness training and an emphasis on effective relation-building communication.

To conclude, the papers that have been selected for this Special Issue offer numerous new insights into how English language researchers and practitioners can draw on the interrelated concepts of multimodality and multimodal literacy in the context of English language teaching in the higher education. It is our sincere hope that this new knowledge will be fruitfully applied in English language classrooms in universities around the world.

References

Abrams, Z. I. (2003). The effect of synchronous and asynchronous CMC on oral performance in German. *The Modern Language Journal*, *87*(2), 157–167.

Bowen, T., & Whithaus, C. (2013). "What else is possible": Multimodal composing and genre in the teaching of writing. In In T. Bowen & C. Whithaus (Eds.), *Multimodal literacies and emerging genres*, (pp. 1–12). Pittsburgh, PA: University of Pittsburgh Press.

Busà, M. G. (2010). Sounding natural: Improving oral presentation skills. *Language Value*, *2*(1) 51–67. Retrieved from http://www.e-revistes.uji.es/languagevalue.

Busà M. G. (2015). Teaching learners to communicate effectively in the L2: Integrating body language in the students' syllabus. *Lingue e Linguaggi*, *15*, 83–98.

Campoy-Cubillo, M.C. & Querol-Julián, M. (2015). Assessing multimodal listening. In B. Crawford, & I. Fortanet-Gómez (Eds.), *Multimodal Analysis in Academic Setting: From Research to Teaching* (pp. 193–212). London: Routledge.

Celce-Murcia, M. (2008). Rethinking the role of communicative competence in language teaching. In E. Alcón Soler & M. Safont Jordà (Eds.), *Intercultural language use and language learning* (pp. 41–57). Dordrecht, The Netherlands: Springer.

Council of Europe (2001). *Common European Framework of Reference for Languages: Learning, teaching, assessment*. Cambridge: Cambridge University Press.

Council of Europe (2017). *Common European Framework of Reference for Languages: Learning, Teaching, Assessment. Companion Volume with New Descriptors*. Language Policy Programme. Education Policy Division. Education Department. Provisional Edition September 2017.

Council of Europe (2018). *Common European Framework of Reference for Languages: Learning, teaching, assessment. Companion volume with new descriptors*. Cambridge: Cambridge University Press.

Francesconi, S. (2016). Green hypermodality: A social semiotic multimodal analysis of the 'Green your house' webpage. In Palusci, O. (Ed.), *Green Canada. Series: Études canadiennes – Canadian Studies*. (pp. 251-262) Bern, Switzerland: Peter Lang.

Gee, J. P. (1996). *Sociolinguistics and literacies: Ideology in discourses*, (2nd ed.). London: Taylor & Francis.

Goman, C. K. (2008). *The nonverbal advantage: Secrets and science of body language at work*. San

Francisco: Berrett-Koehler.

Guichon, N. & McLornan, S. (2008). The effects of multimodality on L2 learners: Implications for

CALL resource design. *System, 36*(1), 85–93.

Hafner, C. (2014). Embedding digital literacies in English language teaching: Students' digital

video projects as multimodal ensembles. *TESOL Quarterly*, *48*(4), 655–685.

Itani, O. S., & Inyang, A. E. (2015).The effects of empathy and listening of salespeople on

relationship quality in the retail banking industry: The moderating role of felt stress.

International Journal of Bank Marketing, *33*(6), 692–716.

Jewitt, C., & Kress, G. (eds.) (2003). *Multimodal literacy*. New York: Peter Lang.

Jewitt, C. (2009*). Routledge handbook of multimodal analysis*. London: Routledge.

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.

Kress, G. (2011). Multimodal discourse analysis. In J. P. Gee & M. Handford (Eds). *The Routledge

handbook of discourse analysis*. Routledge Handbooks Online (pp. 35–50). Retrieved from

https://www.routledgehandbooks.com/doi/10.4324/9780203809068.ch3

Kress, G., & van Leeuwen, T. (1996). *Reading images. The grammar of visual design*. London/New

York: Routledge.

Lemke, J. L. (1998). Multiplying meaning: Visual and verbal semiotics in scientific text. In J. R.

Martin & R. Veel (Eds.), *Reading science: Critical and functional perspectives on discourses of

science* (pp. 87-113). London: Routledge.

Lemke, J. L. (2002). Travels in hypermodality. *Visual Communicaton*, *1*(3), 299–325

Liang, M. Y. (2010). Using synchronous online peer response groups in EFL writing: Revision-

related discourse. *Language Learning & Technology, 10*(2), 56–76.

Liu, J. (2013). Visual images interpretive strategies in multimodal texts. *Journal of Language

Teaching & Research*, *4*(6), 1259–1263.

Martinec. R. (2000). Types of processes in action. *Semiotica*, *130*(3–4), 243–268.

Munter, M. (2009). *Guide to managerial communication. Effective business writing and speaking*. Upper Saddle River: Prentice Hall.

Nelson, M. E. (2006). Mode, meaning, and synaesthesia in multimedia L2 Writing. *ICFAI journal of English Studies*, 2(1), 69–91.

New London Group. (1996). A pedagogy of multiliteracies: Designing social futures. *Harvard Educational Review 66*(1), 60–92.

Norris, S. (2004). *Analyzing multimodal interaction: A methodological framework*. New York, NY: Routledge.

O'Halloran, K. L., & Smith, B. A. (2011). *Multimodal studies: Exploring issues and domains*. London: Routledge.

O'Halloran, K. L., Tan, S. & Smith, B. A. (2016). Multimodal approaches to English for Academic Purposes. In K. Hyland & P. Shaw (Eds.), *The Routledge handbook of English for academic purposes* (pp. 256–269). London & New York: Routledge.

Palmer-Silveira, J. C. (2015). Multimodality in business communication. Body language as a visual aid in student presentations. In B. Crawford Camiciottoli & I. Fortanet Gómez (Eds.), *Multimodal analysis in academic settings: From research to teaching* (pp. 171–192). New York: Routledge.

Rance-Roney, J. A. (2010). Reconceptualizing interactional groups: Grouping schemes for maximizing language learning. *English Teaching Forum*, *48*(1), 20-26.

Roth, W. M. (1994). Thinking with hands, eyes, and signs: Multimodal science talk in a grade 6/7 unit on simple machines. *Interactive Learning Environments, 4*, 170-187.

Royce, T. (2002). Multimodality in the TESOL classroom: Exploring visual - verbal synergy. *TESOL Quarterly 36*, 191–205.

Royce, T. D. (2007a). Intersemiotic complementarity: A framework for multimodal discourse analysis." In T. D. Royce & W. L. Bowcher (Eds.), *New directions in the analysis of multimodal discourse* (pp. 63–109). Mahwah, NJ: Lawrence Erlbaum.

Royce, T.D. (2007b). In T. D. Royce, & W. L. Bowcher, (eds.). In T. D. Royce & W. L. Bowcher (Eds.), *New directions in the analysis of multimodal discourse* (pp. 361–390). Mahwah, NJ: Lawrence Erlbaum.

Royce, T. D. (2015). Intersemiotic complementarity in legal cartoons: An ideational multimodal analysis. *International Journal for the Semiotics of Law-Revue internationale de Sémiotique juridique*, *28*(4), 719–744.

Selfe, C. L., ed. (2007). *Multimodal composition: Resources for teachers*. Cresskill: Hampton Press.

Serafini, F., & Clausen, J. (2012). Typography as semiotic resource. *Journal of Visual Literacy*, *31*(2), 1–16.

Shih, Y. C. (2014). Communication strategies in a multimodal virtual communication context. *System*, *42*, 34–47.

Shumin, K. 2002. Factors to Consider. Developing Adult EFL Students' Speaking Abilities. In J. Richards & W. Renandya (Eds.) *Methodology in language teaching. An anthology of current practice*. (pp. 204–211). Cambridge: Cambridge University Press.

Stein, P. (2000). Rethinking resources: Multimodal pedagogies in the ESL classroom. *TESOL Quarterly*, *34*(2), 333–336.

Street, B., Pahl, K., & Rowsell, J. (2011). Multimodality and new literacy studies. In C. Jewitt (Ed.), *The Routledge handbook of multimodal analysis*, 191–200, London: Routledge.

Sueyoshi, A. & Hardison, D. M. (2005). The role of gestures and facial cues in second language listening comprehension. *Language Learning*, *55*, 661–699.

van Leeuwen, T. (1999): *Speech, music, sound*. London: Macmillan.

Van Slyke, E. (1999). *Listening to conflict: finding constructive solutions to workplace disputes*. New York: AMACOM.

Walsh, M. (2010). Multimodal literacy: What does it mean for classroom practice? *Australian Journal of Language and Literacy*, *3*, 211–223.

Wittenburg, P., Brugman, H., Russel, A., Klassmann, A., & Sloetjes, H. (2006). ELAN: A professional framework for multimodality research. In *Proceedings of LREC 2006, Fifth International Conference on Language Resources and Evaluation*, 1556–1559.

Yang, Y. F. (2012). Multimodal Composing in Digital Storytelling. *Computers and Composition 29*, 221–238.