Research Article

Inmaculada Fortanet-Gómez* and Vicent Beltrán-Palanques

Enhancing multimodal communicative competence in ESP: the case of job interviews

Abstract: ESP teaching has traditionally centred on the discourse practices of specific genres to develop students' discursive competence in the target language. However, with the advent of studies on multimodality, there has been a growing recognition that new pedagogical approaches are required in ESP teaching. In this study, we discuss a research-informed pedagogical proposal to deal with job interviews, a highly routinised spoken genre that is typically addressed in Business English courses. Drawing on previous literature (Cope, B. and Kalantzis, M. (2015). The things you do to know: an introduction to the pedagogy of multiliteracies. In: Cope, B. and Kalantzis, M. (Eds.), *A pedagogy of multiliteracies. Learning by design*. Palgrave MacMillan, Hampshire, pp. 1–36; Lim, F.V. (2018). Developing a systemic functional approach to teach multimodal literacy. Funct. Ling. 5: 1–17), we present a pedagogical proposal that aims to promote students' awareness of the multimodal nature of job interviews and develop their multimodal communicative competence through instruction, practice and feedback.

Keywords: ESP; job interviews; multimodal communicative competence; multimodal interactional analysis (MIA)

1 Introduction

Multimodality, understood as the representation and construction of meaning through varied communicative modes (e.g. gaze, gestures, speech) (Kress 2010), is a well-established research area that has attracted the attention of scholars in various fields (e.g. anthropology, linguistics). Furthermore, recent research has highlighted the importance of incorporating multimodality in ESP contexts from a genre perspective (e.g. Coccetta 2018; Crawford Camiciottoli 2019; Fortanet-Gómez and Bernad-Mechó 2019; Querol-Julián and Beltrán-Palanques 2021; Ruiz-Madrid and Valeiras-Jurado 2020). Although still emerging, this is a promising area that attempts to bridge the gap between multimodality and its pedagogical application to ESP teaching. Accordingly, this study addresses job interviews, a highly routinised genre that requires interlocutors to interact at a discursive and interpersonal level, drawing not only on their linguistic repertoire but also on other modes (e.g. gestures, facial expressions, intonation) to fulfil communicative purposes. A multimodal analysis of a job interview was carried out to identify the variety of semiotic modes interlocutors used to construct this communicative event. Taking this as a point of departure, we present a research-informed pedagogical proposal for teaching students how to conduct and perform well in job interviews. The proposal is framed within a genre-based approach and aims to develop students' multimodal communicative competence.

*Corresponding author: Inmaculada Fortanet-Gómez, Departament d'Estudis Anglesos, Facultat de Ciències Humanes i Socials, Universitat Jaume I, Castellón, Spain, E-mail: fortanet@uji.es. https://orcid.org/0000-0001-8981-565X
Vicent Beltrán-Palanques, Departament d'Estudis Anglesos, Facultat de Ciències Humanes i Socials, Universitat Jaume I, Castellón, Spain, E-mail: vbeltran@uji.es. https://orcid.org/0000-0003-3139-6629

2 Theoretical background

In the last few decades, research on communication has greatly evolved thanks to the advent of studies on multimodality in the mid-1990s. The acknowledgement of the multimodal nature of communication has contributed to understanding, for example, how speakers construct and shape communicative events and engage in interaction using varied communicative modes (e.g. visuals, gestures, intonation, gaze) (e.g. Jewitt et al. 2016; Norris 2004, 2016). Such knowledge is especially relevant in the case of spoken genres, and more specifically in communicative events involving interaction, such as research presentations, product pitches, meetings, negotiations, or job interviews as the domain of interest in this study. Due to their highly interactional nature, job interviews require interlocutors to employ a broad range of communicative modes, both verbal and non-verbal. Regarding this point, Lipovsky (2008) suggests that a multimodal analysis of candidates' gestures and facial expressions would provide further insights in the promotion of solidarity with the interviewers. In the same line, Jewitt et al. (2016), drawing on Campbell and Roberts (2007) study on job interviews, advocate for adopting a multimodal approach to explore the contribution of non-verbal resources (e.g. intonation, gaze, facial expressions, hand gestures) to the overall construction of meaning.

While multimodality is not a new research approach, its application to language learning/teaching in higher education contexts is relatively recent (e.g. Crawford Camiciottoli and Campoy-Cubillo 2018). The pedagogical proposal presented in this article relies mainly on Multimodal Interactional Analysis (MIA) (Norris 2004), which supports the idea that students should become aware of the multiple modes that intervene in the meaning making process. With specific reference to ESP contexts, a multimodal approach can provide learners with a wider set of semiotic resources other than speech to cope with the linguistic, discursive, and pragmatic challenges of domain-specific language (Crawford Camiciottoli 2019).

In a series of studies based on specialised language, Crawford Camiciottoli and Bonsignori (2015) and Bonsignori (2018) discussed how to increase students' awareness of the multiple semiotic resources involved in the meaning making process. Similarly, Coccetta (2018) proposed genre-focused activities to increase students' multimodal communicative competence in the language classroom. More recently, Ruiz-Madrid and Valeiras-Jurado (2020) described the application of multimodal discourse analysis to product and research pitches to foster students' multimodal literacy.

In the present study, we departure from the notion of multimodal communicative competence, defined as "the ability to understand the combined potential of various modes for making meaning" (Royce 2002: 192), and draw on Lim's (2018) model to discuss our research-informed pedagogical proposal. Specifically, Lim (2018) develops a model to address multimodality in the classroom, which is in line with Cope and Kalantzis' (2015) framework for teaching multiliteracies. His model consists of a series of dimensions and knowledge processes students should go through to develop multimodal literacy, namely: 1) Situated Practice/Experiencing; 2) Overt Instruction/Conceptualising; 3) Critical Framing/Analysing, and 4) Transformed Practice/Applying.

The first dimension, *Situated practice,* involves learning through personal and prior knowledge (experiencing the known) and immersion in unfamiliar experiences (experiencing the new). Thus, students can make connections between what is known and what is unknown to build new knowledge, with special attention to the multimodal nature of the genres. In this dimension, the teacher offers students opportunities to be engaged in relevant educational experiences. The second dimension, *Overt instruction*, is characterised by teachers' explicit instruction, which provides students with opportunities to develop abstract, generalising concepts as well as theoretical grounds related to the new concepts. Therefore, teachers provide explicit information to help students expand their knowledge on particular aspects, with a special focus on the multimodal ensembles, and offer scaffolding when necessary. The third dimension, *Critical framing*, consists in "examining the function of a piece of knowledge, action, object or represented meaning" (Cope and Kalantzis 2015: 20). In the last dimension, *Transformed practice*, students learn by "applying experiential, conceptual or critical knowledge" (Cope and Kalantzis 2015: 21) and are required to make their own constructions as a way to demonstrate their learning. This dimension consists of the processes of

applying knowledge appropriately and creatively, including the suitable combination of modes. The former refers to the knowledge that is built in a predictable manner, while the latter is concerned with transferring the knowledge and capabilities from one particular setting to another.

The framework described above can provide a basis for addressing the semiotic affordances in specific genres in the ESP classroom from a learner-centred perspective and for making connections between academic and real-life practices. In the following section, we describe the model that we adapted from this framework to teach students how to conduct and perform well in job interviews.

3 Multimodal genre-based model for job interviews

The model we propose was designed for a Business English course taught in the degree of Business Administration at our university. It focuses on the genre of the job interview as a key communicative event that students need to learn due to its relevance in their future professional careers (Ho 2019), especially in globalised contexts that require competence in English. As a form of institutional discourse, a job interview involves at least two interlocutors, namely interviewers and the interviewees or candidates (Scheuer 2001), who engage in a series of questions-answers to fulfil particular communicative purposes. This type of conventionalised interaction can be very demanding as it has its own specific discourse features. Furthermore, in job interviews, as in any other type of interaction, interlocutors may make use of paralinguistic and extralinguistic resources (see Jewitt et al. 2016). For this research, we have selected intonation, pitch, pauses, hand gestures, gaze, and facial expressions, following previous multimodal studies (e.g. Beltrán-Palanques and Querol-Julián 2018; Bonsignori 2018; Crawford Camiciottoli 2015; Morell and Pastor 2018; Querol-Julián and Fortanet-Gómez 2012; Ruiz-Madrid and Valeiras-Jurado 2020). Both the interviewers and candidates may employ ensembles of communicative modes to exchange information and co-construct interpersonal meanings in this communicative event.

To teach multimodal communicative competence, it is necessary to first carry out a multimodal analysis of a sample of the genre selected. We believe that Norris's (2004) approach to multimodal analysis (i.e. MIA) can be usefully applied in pedagogical contexts to explore how interaction is constructed in job interviews. More specifically, this analytical approach can serve to examine how varied communicative modes "are brought into and are constitutive of social interaction, identities and relations" (Jewitt et al. 2016: 132). Thus, MIA broadens the scope of interaction and considers how people make use of communicative modes to mediate interaction (Jewitt 2014). Within the MIA framework, Norris (2004) distinguishes between lower-level actions (LLAs) and higher-level actions (HLAs). While lower-level actions represent the smallest pragmatic unit such as gestures, utterances, and gaze, higher-level actions are actions that social actors intend to perform and pay attention to Norris (2016), such as greeting somebody or holding a business meeting, and consist of a "multiplicity of chained lower-level actions" (2004: 13). To analyse HLAs, Norris (2004) introduces the modal density foreground-background continuum, which can be explored through modal intensity, defined as "the intensity or weight that a mode carries in the construction of a higher-level action" (90) and modal complexity which refers to the interrelation of modes (see also Jewitt et al. 2016). In this paper, we present a simplified analysis of the HLAs in a fragment of a sample job interview that centres only on the modes most often used (i.e. modal intensity) by the interlocutors. This type of analysis will be later replicated by the students in a different interview fragment to critically analyse multimodal interaction. The purpose of this task is to raise students' multimodal awareness as it will draw their attention to how various modes intervene in the construction of discourse. It will be carried out in the *Critical framing* stage in which students are provided with opportunities to examine the communicative function of a particular piece of knowledge, before proceeding to *Transformed Practice*, in which they have to produce their own interview (Cope and Kalantzis 2015).

Thus, this study adopts a model that engages students in a series of stages aiming to develop multimodal communicative competence while addressing the genre of job interviews. Figure 1 illustrates the teaching/

learning cycle designed to teach students how to conduct and perform well in job interviews from a multimodal perspective.

As shown in Figure 1, the pedagogical procedure adopted in this study entails four dimensions. Accordingly, this approach aims to provide students opportunities to become aware of the multimodal nature of the genre through experience, instruction, and sample analysis to finally apply the knowledge acquired to construct their own texts. Specifically, the procedure followed consists in:

(1) Experiencing a new situation based on the students' previous knowledge from the everyday world, in this case by sharing some students' experience of job interviews as multimodal communicative events (*Situated practice*),

(2) Learning from abstract generalising terms, by drawing distinctions, identifying similarities and differences, and connecting concepts with language and visuals, in this case through the teacher's explanations of how job interviews are structured and of the multimodal affordances they have (*Overt instruction*),

(3) Analysing the sample of a genre, examining structure and function, as well as patterns, and trying to relate multimodal ensembles to the need for engagement in a job interview (*Critical framing*), and

(4) Applying the knowledge acquired in previous phases to reproduce a job interview adapting it to a given situation (*Transformed practice*).

In the next section, we present the adaptation of this model to a specific setting: a group of Business English students at Authors' university.

4 Research-informed pedagogical proposal

The pedagogical proposal described here was partially implemented in the Business English course during the first semester in the academic year 2020–2021. Due to the COVID-19 circumstances, all teaching in the Business degree was delivered in a hybrid mode. Therefore, the proposal, originally designed for face-to-face classroom teaching, could not be implemented as planned. Notwithstanding this, the partial implementation of the proposal provided valuable information towards improving it for incorporation into the course syllabus in the next few years. Two main aspects were found that needed reinforcement or modification. The first involved the focus of the multimodal analysis. Specifically, intonation, which needs to build on an improvement in pronunciation, together with facial expression, were identified as the modes that require the most attention, due to their importance in interaction and the difficulty they entail for students. Secondly, the interviews
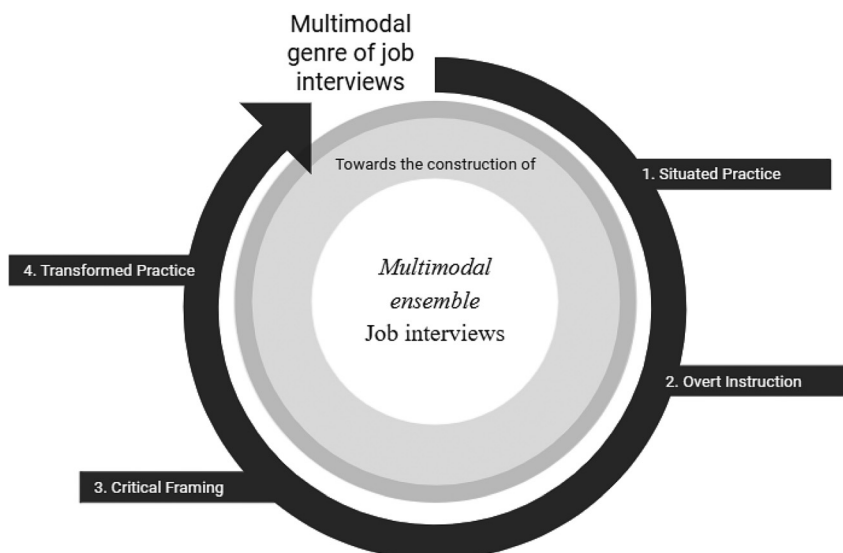
Figure 1: Pedagogical steps to teaching the multimodal genre of job interviews (adapted from Querol-Julián and Fortanet-Gómez 2019).

produced by students need to be as spontaneous as possible, rather than relying too much on written scripts, since in authentic situations, they will not have the chance to prepare complete scripts for these communicative events.

4.1 Students' profile

The students involved in the ESP course (Business English) belong to three different bachelor's degree courses (i.e. Business Administration, Finance and Accounting, and Economics) taught at Authors' university. The same ESP course is taught in five different groups and about 300 students are usually enrolled. In this study, a proficiency self-assessment English test was distributed to a sample of 100 students attending the course in the academic year 2020–2021. The test, run by Cambridge Assessment English (https://www.cambridgeenglish. org/test-your-english/), served to indicate students' proficiency level according to the Common European Framework of Reference for Languages (Council of Europe 2020). Although the participants' proficiency level differed and ranged from A1 to C1, most of them were between the threshold levels of A2 (54%) and B1 (26%). In light of this, all activities were adapted to these levels. Students were also asked about their previous experience in job interviews. This questionnaire revealed that 33% of the students had participated in a job interview, even though it was usually in their own language and as candidates, but not as interviewers.

4.2 The course

The Business English course addressed in this study is taught in the third academic year. Although a B1 level of English language proficiency (Council of Europe 2020) is recommended, as shown above, the students do not always demonstrate this level. The class sessions are 4 h long and held only once a week to complete six credits (approximately 60 h of teaching) in the first semester of the academic year.

The genre of job interviews is covered after a unit entitled "Job ads, application letters and CVs". We believe that both units must be closely interconnected for students. That is why they are asked to form teams of five or six students, in which two students represent the recruiting company and the rest are candidates. The recruiting company creates and publishes a job ad, and candidates apply for the job by sending a letter and a CV or résumé. Then, they receive a reply inviting them for an interview, which will be held by the representatives of the recruiting company and each of the candidates.

The instructions that students have traditionally received on job interviews are based on the vocabulary and the phrases that make the interview effective. Only some indications such as the following refer to non-verbal elements in a very general manner: "dress for the situation" or "you must project yourself in a positive, enthusiastic manner". Though these may be valuable pieces of advice, how can candidates and interviewers project themselves in a positive, enthusiastic manner? There is a need for more specific indications on how to achieve this. We tried to fulfil this need by adding some activities which would highlight the relevance of the use of multimodal ensembles in job interviews, following the model described in Section 3.

4.3 Proposed activities

4.3.1 Situated practice

Before starting with the teacher's explanation about the materials in the unit, our proposal includes *Situated practice* as an activity. Considering that about 30% of the students have some previous experience with job interviews, we will invite those students to share their experience with the rest, asking them to reflect on the several modes that can be found in this communicative event. The activity is described in Table 1.

Table 1: Activity for *Situated practice*.

Dimension	Activity	Description
Situated practice	Discussion of what students know about job interviews	The teacher provides 6 questions to guide the discussion. All students can contribute by explaining their experiences and asking and answering questions.
		Guiding questions:
		(1) How did you prepare for the interview?
		(2) What was more challenging for you before the interview?
		(3) How did you feel during the job interview?
		(4) What did you find more challenging during the interview?
		(5) What was the result? How did you feel about it?
		(6) What could you have improved?

4.3.2 Overt instruction

After the *Situated practice*, students are provided with explicit information on how to communicate effectively in job interviews (*Overt instruction*). For this purpose, the teacher uses published materials for teaching Business English; in this case the book *English Communication for International Business* (Ruiz-Garrido and Fortanet-Gómez 2006) as well as extra materials purposefully designed to deal with the genre, available to students in the Virtual Classroom. Broadly speaking, these materials address the main features and stages of a job interview, while also providing guidelines for effective communication, and information about how job interview discourse is shaped, both for the candidate and for the interviewer.

To complement this information, we propose specific instruction on multimodal ensembles. To that end, a sample of a job interview was selected, in which an experienced candidate is interviewed for the position of Human Resources Manager. It is a short realistic simulated interview (it lasts 3 min and 51 s), easily understandable by the students, and freely available on YouTube at https://www.youtube.com/watch?v=SieNfciN274.

To begin this activity, some specific information is provided about the sample job interview such as context of the interview and characteristics of the position advertised. Next, the teacher opens a discussion about what a job interview is like (*Situated Practice*). Then, students are instructed on the structure of the job interview by means of a multimodal interaction analysis, a simplification of a previous study carried out by the teacher. Following Norris's (2004), 11 main Higher-Level Actions (HLAs) can be found in this job interview. In the MIA approach, the actors are in focus rather than the communicative event. That is why HLAs are different for each speaker. The job interview is led by the interviewer and she initiates and ends turns. In HLA 1, 3, 5, 7, 8, 9 and 10, the speaker is the interviewer; therefore, these HLAs are foregrounded, although the candidate remains as an active receiver with LLAs such as smiling, nodding, and looking at the interviewer. In HLA 2, 4, 6, and 11, the speaker is the candidate, and her HLAs are in the foreground, while it is the interviewer who listens and produces backchanneling, usually through facial expressions and, incidentally, with a short verbal reply. We could have included the HLAs dealing with the function of listening. However, because the videorecording mainly shows the speaker and rarely the listener, we decided not to focus explicitly on these HLAs, although the listeners are sometimes referred to in the analysis when it is relevant. Table 2 summarizes the HLAs in a job interview.

Each HLA shows a clear beginning and a clear end, marked by verbal and non-verbal modes. In job interviews, HLAs come one after the other in the foreground. Moreover, HLA 3 and HLA 4 are constituted by several other embedded HLAs, as the interviewer asks the candidate seven questions. However, as these two interlocutors are not just exchanging information, there are other HLAs that remain in the midground, as shown in Table 3.

Table 2: HLAs in the foreground in the sample job interview.

HLA 1: Salutations and welcome	HLA 2: Reply to salutations
HLA 3: Interviewer asks the candidate questions (HLA 3.1–3.7)	HLA 4: Candidate answers questions (HLA 4.1–4.7)
HLA 5: Interviewer asks the candidate to ask questions (transition)	
HLA 7: Interviewer answer the questions	HLA 6: Candidate asks the interviewer one question
HLA 8: Interviewer ends up question time	
HLA 9: Closing by interviewer	
HLA 10: Farewell	HLA 11: Answer to farewell

Table 3: HLAs in the midground in the sample job interview.

HLA 12: Establish rapport	
HLA 13: Find out attitudes	HLA 14: Show best attitudes

These HLAs are not so evident and their boundaries are not so clearly established. Nevertheless, there are some modes that are especially intense within them, especially in order to establish rapport (HLA 12), such as the use of polite language, eye contact accompanied by a smile, nodding and complying back-channeling, both by the interviewer and the candidate. One of the ultimate aims of the interviewer is to find out the candidate's attitudes (HLA 13), and the candidate needs to show her best attitudes to comply with the interviewer's expectations (HLA 14). Mid- and foregrounded HLAs fluctuate and change as communication unfolds, and these HLAs are foregrounded especially when some modes gain intensity by means of LLAs. They are important because attitudes and reactions are often what most interviewers seek in a face-to-face job interview in order to find out about the candidate's personality.

In the teacher's *Overt instruction*, students need to learn about the foregrounded and midgrounded HLAs, since all of them can be relevant for the result of the interview, though this explanation needs to be adapted to the profile of the students. Among the HLAs, two have been selected to be analysed by the teacher as *Overt instruction* in this proposal. They are HLA 3.6, the sixth question by the interviewer and the answer by the candidate, HLA 4.6. The question asked is about the weaknesses that the candidate, Tamra, thinks she may have in doing her job. Her answer deals with the relationship between employees and how to teach them to improve these relationships. We selected these HLAs because they require an extra effort by interviewers and interviewees, as they can be face threatening. In addition, they may be accompanied by midgrounded HLAs in which the interviewer intends to find the interviewee's best attitudes, while the latter tries to show these best attitudes (HLA 13 and 14), as explained above.

Students then watch the video, and the teacher produces an analysis of the multimodal ensembles in the actions. This is based on the analysis of the modal density found in these HLAs (Norris 2004): language (with the highest intensity, as it contextualises the content), intonation, pitch, personal appearance, facial expression and eye contact. For the analysis of language and intonation, the teacher uses the transcript, also provided together with the video on the YouTube website. Figure 2 shows how this would appear on a slide.

In this question, the candidate, who already has experience as a Human Resources Manager, is asked what she thinks she needs to improve (HLA 3.6). She answers that for her, it is difficult to discern whether she needs to be strict or not when somebody makes a mistake, as it may be more important to find out why that employee made a mistake in order to prevent it from happening again (HLA 4.6).

Regarding language, the teacher explains the use of polite formulae by means of the conditional auxiliary verb "would" and a hedge such as "a little better". Another interesting feature in HLA 3.6 is the combination of

SCRIPT- LANGUAGE FEATURES

02:06-02:33
I = what would they say is the thing that Tamra needs to do a little better? we all have those things we can improve upon. What's the thing that you're working on?

C = Ehhh, coming off or not coming off so harsh with an individual (ehh, individual) things like oh if you make a mistake, you know why did you make the mistake and eh try to find out the root problem of it and then teach them on it, so it doesn't happen again.

SCRIPT- PITCH AND INTONATION

02:06-02:33
I = what would they say is the thing that Tamra **needs to do** a little better? we all have those things we can **improve upon.** What's the thing that you're **working on?**

C = Ehhh, **coming off** or **not coming off** so harsh with an individual (ehh, individual) things like oh if you **make a mistake** you know **why** did you make the mistake and eh try to find out the **root problem** of it and then **teach** them on it so it doesn't **happen again**

Figure 2: Analysis of language features, pitch and intonation in HLAs 3.6 and 4.6.

subjects, such as "Tamra" to refer to the candidate in the third person and "you" to also refer to her. There is also an inclusive "we" in the interviewer's contribution, in an effort to establish a good rapport with the candidate, especially with this type of question which is not comfortable to answer. On the other hand, in HLA 4.6, the candidate responds using "individual" at the start, to whom she refers later indirectly using the impersonal "you" and, later, "them", in what seems to be an intention to detach herself from the employees she has to deal with in her job.

As for pitch and intonation, students are asked to listen to these HLAs several times without seeing the image in order to discern differences, and after that, the information is shown to students, where they can check the stressed words and expressions marked in bold. They are mainly verbs and some subjects and objects. The intonation helps the hearer to distinguish the various questions that the interviewer is asking, as well as the different parts of the answer.

To analyse other non-verbal expressions, students watch the HLAs without sound. By doing this, students can observe the personal appearance of both the interviewer and candidate, their facial expressions and the use of smiles and eye contact. Finally, students are asked to watch the complete video sequence in order to appreciate how the multimodal ensemble provides the meaning to the scene. Table 4 shows this analysis.

After analysing HLA 3.6 and HLA 4.6, students are asked about the midgrounded HLAs, especially about the presence of HLA 13, *Find out attitudes* and HLA 14, *Show best attitudes*. The former is addressed by the interviewer and the latter by the candidate in the interactive roles they play. HLA 13 is represented in the way the interviewer asks the question. It is a difficult question to answer, so she uses two questions with a statement in between them (Figure 2), which makes it very clear the candidate has to address her weaknesses. The most intense mode in this HLA is verbal discourse. Regarding HLA 14, the candidate shows her uncertainty with several crutches and by changing the way she refers to the employees in the company. The way she utters "coming off or not coming off so harsh" at the start, with a stress on "coming off" and the addition of "not coming off", reinforced by her facial expression, is the most significant part with the joint intensity of the language, the stress, and the facial expression, in an effort to show her professionalism as well as her understanding when an employee makes a mistake. Table 5 summarizes the activity in this dimension.

Table 4: Analysis of nonverbal LLAs in HLA 3.6 and 4.6.

Transcript	Rapport, smiles, facial expression, gestures	Function
I = what would they say is the thing that Tamra needs to do a little better?	Smile, nose wrinkle with the word "needs", and head tilt to the left	Asking an uncomfortable question
We all have those things we can improve upon.	Head tilt to the right and nod	Reinforcing hedge
What's the thing that you're working on?	Hand gesture pointing at candidate and open smile	Trying to look sympathetic
C = Ehhh, coming off or not coming off so harsh with an individual (ehh, individual)	Pause, slight smile, head tilt and nod, closing eyes while saying "not coming off"	Hedging her position in a difficult question to answer by adding a contradiction
Things like oh if you make a mistake,	Looking sidewards and smile	Thinking aloud
You know why did you make the mistake and eh try to find out the root problem of it	Nodding, eye rapport	Reinforcing words, she is more confident with what she says
And then teach them on it, so it doesn't happen again.	Eye rapport, nodding and momentarily closing her eyes	Condescending feeling

Table 5: Activity for *Overt instruction*.

Dimension	Activity	Description
Overt instruction	Description of HLA, explanation of modes and multimodal analysis of HLA 3.6 and 4.6	The teacher explains the structure of the job interview and contextualises HLA 3.6 and 4.6. She explains the various modes, with the collaboration of the students, who are asked to observe the transcript, the video without image, the video without sound and then, image and sound, to appreciate the multimodal ensemble. Then, with the collaboration of the students, the teacher unveils the presence of HLA 13 and HLA 14.

4.3.3 Critical framing

In the third dimension, *Critical framing*, students are asked to do the same analysis with two other HLAs, in this case HLA 3.7. and 4.7., in which the interviewer asks the candidate's opinion on what it is like to be a good employee. Students are provided with the transcript to facilitate this analysis. Figure 3 shows the exercise as presented to the students for language, pitch and intonation. Figure 4 shows a fill-in table for eye rapport, smiles and facial expression.

After this analysis, students are asked about the presence of HLA 13 and 14 in HLA 3.7 and 4.7. It is HLA 14 that is more evident here, as the candidate speaks most of the time. Then, students are asked how meaning has been created and how the multimodal ensemble has contributed to establishing a relationship between interviewer and candidate. Table 6 summarizes this dimension.

4.3.4 Transformed practice

The fourth dimension of the model requires the students' production of a sample of the genre, in what Lim (2018) calls *Transformed practice*. We have divided it into three parts: preparation, rehearsal and performance. Students are requested to do a role play activity in which they have to perform a job interview in groups of six students, two representing the recruiting company and the other four as candidates for the job. As preparation, the company representatives have produced job ads for real companies that can be found on the

1. Please mark the most representative language features in this job interview Action.

    I = What do you think it takes to be a good employee?

    C = A good employee, showing to work on time, being there as a team member, knowing your job, knowing your job

    I = Knowing your job

    C = If you don't know the job, ask questions, don't sit there like a knot on a log and not ask the question and I truly believe in taking self-initiative, you know, you bring something to the table.

2. Now mark the pitch (underline) and intonation (draw arrows).

    I = What do you think it takes to be a good employee?

    C = A good employee, showing to work on time, being there as a team member, knowing your job, knowing your job

    I = Knowing your job

    C = If you don't know the job, ask questions, don't sit there like a knot on a log and not ask the question and I truly believe in taking self-initiative, you know, you bring something to the table.

Figure 3: Activity on the analysis of language features, pitch and intonation in HLA 3.7 and 4.7.

3. Please watch the video and describe the use of eye rapport, smiles and face expression as the interviewer and the candidate speak. Why do they use them?

TRANSCRIPT	RAPPORT, SMILES, FACIAL EXPRESSION	FUNCTION
I = What do you think it takes to be a good employee?		
C = A good employee, showing to work on time,		
being there as a team member,		
knowing your job, knowing your job		
I = Knowing your job		
C = If you don't know the job, ask questions,		
don't sit there like a knot on a log and not ask the question		
and I truly believe in taking self-initiative, you know, you bring something to the table.		

Figure 4: Activity on the analysis of eye rapport, smiles and facial expression in HLA 3.7 and 4.7.

Table 6: Activities for *Critical framing*.

Dimension	Activity	Description
Critical framing	Analysis by the students of HLA 3.7 and 4.7 in order to show their multimodal awareness	1. Students work in teams in order to analyse the sequence and fill in a table in which they are asked to observe several of the modes previously analysed by the teacher. 2. In a class general discussion, they will be asked to answer how the multimodal ensemble contributes to creating a relationship between the interviewer and the candidate (HLA 12). The teacher will supervise work in teams in class and will chair the discussion.

Internet, and the candidates have submitted their application letters and CVs or resumes. Next, the company representatives invite candidates for an interview. Interviewers have prepared the most adequate questions which they will not share with the candidates, who would have done some research on the company and prepared their answers for the questions the interviewers may pose. Then, in a second part of this activity, each company representative will interview two of the candidates in his or her team in class as a rehearsal, while the other company representative and the other two candidates observe and take notes with the aid of a simplified checklist in order to provide feedback. The teacher will supervise and take notes on the whole process as part of her assessment.

Table 7: Activities for *Transformed practice*.

Dimension	Activity	Description
Transformed practice	Students apply what they have learned replicating a job interview, adapted to a given situation	1. Students prepare for the job interview. Interviewers select the questions according to the candidates' CVs. Candidates research on the company and prepare for expected questions. 2. Students rehearse the role play in front of their teammates and receive feedback 3. Students perform the role play of the job interview face to face and in front of the teacher. It is also video recorded.

The third part of this dimension is performance, in which one interviewer in each team will interview two candidates, and the other one will interview the other two. Then, they will meet and decide on the best candidate by filling in a table with the requirements and merits of the job. The interviews will be face to face in front of the teacher as a role play, that is, the students need to perform the situation. They will also be recorded for later assessment. Table 7 sums up this dimension.

4.4 Instruments to assess the validity of the pedagogical proposal

To validate the effectiveness of the pedagogical proposal, we suggest the use of the following instruments: a pre-test questionnaire, a post-test questionnaire and a final test, which may be powered by Google Forms. The pre-test and post-test need to be administered to the experimental group and to at least one control group with similar characteristics but without receiving the *Overt instruction* on multimodal communicative competence.

The pre-test will be used to gather data about students' prior experience and knowledge regarding job interviews. In addition, it may include a short 4-min video of a job interview about which students will have to identify the foregrounded HLA and also if there are elements that identify one of the midgrounded HLAs. Then, they will be asked to rank the relevance given to specific items (ranking scale). The items to be ranked will deal with verbal discourse (the most intense mode), pronunciation, intonation, and stress on relevant parts of discourse, eye rapport, facial expression, posture, and hand gestures. This pre-test will be administered to the experimental group as well as a control group of students.

The post-test, also administered to the experimental and control groups, will be used to explore the participants' improvement after the pedagogical treatment. It will consist of only two items: the same video with identification of the foregrounded and midgrounded HLAs and the ranking scale used in the pre-test.

The third questionnaire will only be administered to the experimental group and will include a ranking scale to measure students' satisfaction and five-item open-ended questions to explore their opinion on the content of the unit, on the usefulness of the feedback received, and on the relevance of the unit for their future professional career.

5 Conclusions

Multimodal competence is one of the most important abilities that ESP students should acquire to effectively construct meanings in spoken genres. The objective of this article was to present a research-informed pedagogical proposal for the teaching of one of those spoken genres, namely, job interviews, from a multimodal perspective.

The pedagogical proposal we present is based on Norris's (2004) MIA approach and aims to introduce multimodal interactional analysis in the classroom. Because not much research has focused on the teaching of multimodal communicative competence in ESP, our proposal takes Lim's (2018) model with four dimensions as a point of departure: *Situated practice*, *Overt instruction*, *Critical framing* and *Transformed practice*. The objective of the proposal is to make students aware of the relevance of multimodal affordances as they learn how to effectively conduct and perform in a job interview in English.

Although this proposal was partially implemented during the academic year 2020–2021, it was not possible to introduce it as stated previously due to the COVID-19 circumstances. Even so, we obtained sufficient useful insights to inform the proposal in order to adapt it better to the profile of the students and to the teaching situation. However, we acknowledge that this is also a limitation and that the proposal should be implemented in order to assess its potential. Moreover, because the job interview is only one genre within the genre chain (Paltridge 2012; Räisänen 1999; Swales 2004) of the job application process, this multimodal approach should be extended to its other genres (e.g. job advertisement, application letter, CV or resume, offer of appointment). In this way, it would be seen by students as a holistic approach that can be followed for all kinds of communicative events. In addition, in this study we have mainly focused on modal intensity to purposely illustrate the modes that are more prevalent in the interaction. Further research could also introduce modal complexity to explore, along with modal intensity, how varied communicative modes are involved and interrelated.

This article argues that there is a need to equip Business English students with multimodal literacy so that they may acquire multimodal communicative competence. We claim that if all communication is multimodal, teaching how to understand and how to produce business genres should be based on this premise.

Appendix 1

RUBRIC FOR JOB INTERVIEWS
    GROUP:
    INTERVIEWER:
    CANDIDATE:
    JOB DESCRIPTION:
    VALUE: 1 (VERY LOW) AND 5 (VERY HIGH)

GENERAL	CANDIDATE	INTERVIEWER
1 Is the general appearance appropriate?		
2 How would you rate the candidate and the interviewers?		
LANGUAGE		
3 Is the language correct?		
4 Is the language appropriate for the context?		
5 Does the interview follow the established structure?		
6 Is there interaction between speakers? Ex. turn-taking respect, and asking and giving the floor		
POSTURE AND GESTURES		
7 Is the posture and attitude shown adequate?		
8 Are gestures used appropriately?		
FACIAL EXPRESSION		
9 Is smile and/or serious expressions used to accompany the discourse?		
10 Do they use other expressions to accompany their speech, such as frowning eyebrows, opening eyes widely or closing eyes?		
EYE RAPPORT		
11 Do the speakers look at their interlocutor?		
PAUSES, INTONATION AND STRESS		
12 Is the pronunciation correct and clear?		
13 Are there strategic pauses used?		
14 Is there a suitable intonation?		
15 Are strategic words and phrases emphasized?		
TOTAL		
COMMENTS:		

References

Beltrán-Palanques, V. and Querol-Julián, M. (2018). English language learners' spoken interaction: what a multimodal perspective reveals about pragmatic competence. System 77: 80–90.

Bonsignori, V. (2018). Using films and TV series for ESP teaching: a multimodal perspective. System 77: 58–69.

Campbell, S. and Roberts, C. (2007). Migration, ethnicity and competing discourses in the job interview: synthesizing the institutional and personal. Discourse Soc. 18: 243–271.

Coccetta, F. (2018). Developing university students' multimodal communicative competence: field research into multimodal text studies in English. System 77: 19–27.

Cope, B. and Kalantzis, M. (2015). The things you do to know: an introduction to the pedagogy of multiliteracies. In: Cope, B. and Kalantzis, M. (Eds.), *A pedagogy of multiliteracies. Learning by design*. Palgrave MacMillan, Hampshire, pp. 1–36.

Council of Europe. (2020). Common European framework of reference for languages: learning, teaching, assessment, Companion volume, https://rm.coe.int/common-european-framework-of-reference-for-languages-learning-teaching/16809ea0d4 (retrieved 10 06 2021).

Crawford Camiciottoli, B. (2015). Elaborating explanations during OpenCourseWare humanities lectures: the interplay of verbal and nonverbal strategies. The interplay of verbal and nonverbal strategies. In: *Multimodal analysis in academic settings. From research to teach*. Routledge, London, pp. 144–163.

Crawford Camiciottoli, B. (2019). Harnessing multimodal literacy for knowledge dissemination in ESP teaching. In: Bonsignori, V., Cappelli, G., and Mattiello, E. (Eds.), *World of words: complexity, creativity, and conventionality in English language, literature and culture*, Vol. 1. Language, Pisa University Press, pp. 47–52.

Crawford Camiciottoli, B. and Bonsignori, V. (2015). The Pisa Audiovisual Corpus Project: a multimodal approach to ESP research and teaching. ESP Today 3: 139–159.

Crawford Camiciottoli, B. and Campoy-Cubillo, M. (2018). Introduction: the nexus of multimodality, multimodal literacy, and English language teaching in research and practice in higher education settings. System 77: 1–9.

Fortanet-Gómez, I. and Bernad-Mechó, E. (2019) Análisis multimodales del discurso específico. Special issue. *Ibérica*. Available at: http://www.aelfe.org/?s=revista&veure=37.

Ho, W.Y.J. (2019). Developing professional communication: the construction of a multimodal understanding of job interviews. Languages 4: 2–12.

Jewitt, C. (2014). Different approaches to multimodality. In: Jewitt, C. (Ed.), *The Routledge handbook of multimodal analysis*, 2nd ed. Routledge, London, pp. 31–43.

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

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

Lim, F.V. (2018). Developing a systemic functional approach to teach multimodal literacy. Funct. Ling. 5: 1–17.

Lipovsky, C. (2008). Constructing affiliation and solidarity in job interviews. Discourse Commun. 2: 411–432.

Morell, T. and Pastor, S. (2018). Multimodal communication in academic oral presentations by L2 Spanish students. J. Spanish Lang. Teach. 5: 125–138.

Norris, S. (2004). *Analyzing multimodal interaction: a methodological framework*. Routledge, London.

Norris, S. (2016). Concepts in multimodal discourse analysis with examples from video conferencing. Yearb. Pozn. Linguist. Meet. 2: 141–165.

Paltridge, B. (2012). *Discourse analysis. An introduction*, 2nd ed. Bloomsbury, London.

Querol-Julián, M. and Beltrán-Palanques, V. (2021). PechaKucha presentations to develop multimodal communicative competence in ESP and EMI live online lectures: a team-teaching proposal. Comput. Assist. Lang. Learn. 22: 73–90.

Querol-Julián, M. and Fortanet-Gómez, I. (2012). Multimodal evaluation in academic discussion sessions: how do presenters act and react? English Specif. Purp. 31: 271–283.

Querol-Julián, M. and Fortanet-Gómez, I. (2019). Collaborative teaching and learning of interactive multimodal spoken academic genres for doctoral students. Int. J. Engl. Stud. 19: 61–82.

Räisänen, C. (1999). *The conference forum as a system of genres: a sociocultural study of academic conference practices in automotive engineering*. Acta Universitatis Gothoburgensis, Gothenberg.

Royce, T. (2002). Multimodality in the TESOL classroom: exploring visual-verbal synergy. Tesol Q. 36: 191–205.

Ruiz-Garrido, M.F. and Fortanet-Gómez, I. (2006). English Communication for International Business I. Universitat Jaume I, Castelló.

Ruiz-Madrid, N. and Valeiras-Jurado, J. (2020). Developing multimodal communicative competence in emerging academic and professional genres. Int. J. Engl. Stud. 20: 27–50.

Scheuer, J. (2001). Recontextualization and communicative styles in job interviews. Discourse Stud. 3: 223–248.

Swales, J.M. (2004). *Research genres. Explorations and applications*. Cambridge University Press, Cambridge.