María-Ángeles Mestre-Segarra*

Universitat Jaume I, Spain
mamestre@uji.es

A MULTIMODAL RHETORICAL ANALYSIS OF VIDEO RESUMES

Abstract

Today's society is characterized by digital and multimodal technologies which allow the constant emergence of new digital genres, especially in professional contexts. One of these recent genres, the video resume (VR), is perceived as an innovative recruitment and selection tool attracting growing interest in business settings. However, very little research focuses on the definition of the VR as a genre and its multimodal nature in order to validate it as a professional digital genre. Therefore, the aim of this paper is twofold: a) to determine the multimodal rhetorical structure of the VR, and b) to assess the role multimodality plays in this digital genre. The dataset used for the analysis consists of 26 VRs, all of them in English, taken from the online platform YouTube. The methodological procedure is based on a move-and-step rhetorical analysis followed by a multimodal discourse analysis (MDA). The results show that in VRs meaning is conveyed through the interconnection of various modes (embodied, disembodied and filmic) and that the multimodal rhetorical functions identified in the analysis contributed significantly to the definition of the VR as a professional digital genre. Pedagogical implications are also presented, which demonstrate the potential of the VR as a teaching resource used in English for Specific Purposes (ESP) to enhance business students' multimodal skills and to improve the communicative competence much needed in today's business world.

Key words

* Corresponding address: María-Ángeles Mestre-Segarra, Departament d'Estudis Anglesos, Universitat Jaume I, Av. Sos Baynat s/n, E12071, Spain.

1. INTRODUCTION

In the 1990s the Internet spread globally, opening the door to the use of digital tools in the employee recruitment and selection process. According to Hiemstra and Derous (2015: 2), "[r]ecent technological developments and the increased use of digital applications have resulted in the emergence of so-called video résumés". They are described as "a short video-taped message in which applicants present themselves to potential employers on requested knowledge, skills, abilities, motivation and career objectives" (Hiemstra et al., 2012: 11). Research is needed to empirically examine the benefits of using digital tools such as the video resume (VR) in the recruitment and selection process, now that organizations can decide whether to use them and how best to standardize their use (Waung et al., 2014). Due to the potential impact of the VR on both the professional community and the general audience, it is considered to be of key importance for potential job applicants to acquire specific skills necessary to produce videos that make full use of the range of semiotic resources available in such media. In this paper, I present an empirical study investigating the multimodal rhetorical structure of the VR in order to provide insights into the systematic organization on which this genre is based.

VRs are increasingly gaining interest as they represent an opportunity to extensively study the ways in which first impressions are formed in an employment context (Nguyen & Gatica-Perez, 2016). Kemp et al. (2013) claim that a one-minute VR is an effective recruitment tool since it can convey applicants' desirable qualities, such as enthusiasm and verbal communication skills, as well as because it is a tool for effectively using the time employers spend screening videos. Waung et al. (2014: 238) argue that "video resumes are richer media than paper resumes, because multiple verbal and nonverbal cues are available along with language variety via captions, voice-overs, and speech". Hiemstra (2013: 4) points out that "the format of a video resume can vary from a video-taped message to a multimedia message, including animations and text".

As a result, the information presented in a VR becomes more dynamic compared to that of a paper resume because of the communicative skills and multimodal resources candidates use to show their potential aptitudes. The scope of the term *video resume* itself is broadened in this paper, as it does not merely refer to the digitized version of a paper-based resume. In fact, it shares significant features with job application letters (see Bhatia, 1993) and motivational letters (see Furka, 2008) in terms of content, register as well as rhetorical structure, which is the focus of the present study. Despite the similarities, these genres differ in their mode of transmission, written or oral. While job application and motivational letters emphasize language content and tend to be descriptive and impersonal, VRs convey a detailed description of the content from a more personal perspective and increase visibility, which stems from their multimodal nature.

As research into the structure of a relatively new genre such as the VR is still scarce, my study will examine the move patterns on which it tends to be structured.

350

The analysis follows Swales's (1990) rhetorical analysis in terms of moves and steps, Bhatia's (1993) systematic framework for analyzing another persuasive genre, job application letters, as well as O'Halloran's (2011) multimodal discourse analysis (MDA) approach. The combination of rhetorical and multimodal approaches will help me to establish the multimodal rhetorical structure of the VR and define it as a professional digital genre, which is the ultimate purpose of the study. I also focus on the combination of the concepts of genre and multimodality known as the Genre and Multimodality (GeM) model developed by Bateman (2014), in which socially recognizable categories and diverse expressive resources are mixed to form coherent messages (Bateman, 2008, 2017). When studying VRs, the nature of these messages requires the genre to be conceived of as "an interconnected, vibrant, and resilient social phenomenon" used to achieve "different communicative purposes under different social circumstances" (Xia, 2020: 145).
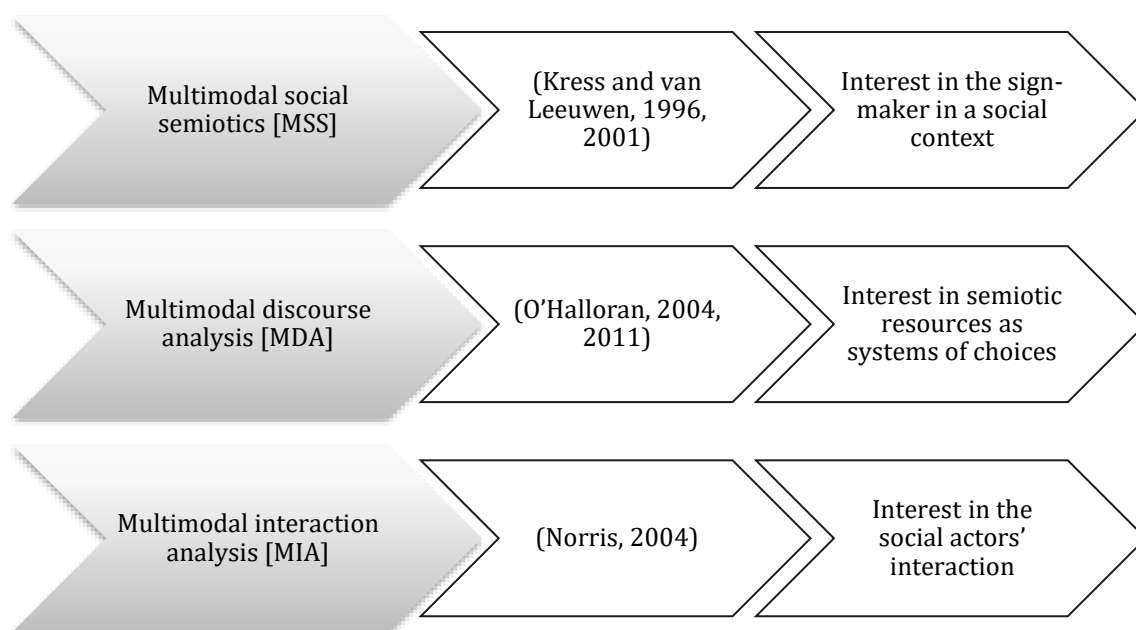
2. MULTIMODALITY IN DIGITAL GENRES APPLIED TO PROFESSIONAL CONTEXTS

Although there is a straightforward image of the VR as a potential selection and recruitment tool in the business world, there is also a lack of literature focusing on its definition as a professional digital genre from a multimodal perspective. Professional genre studies seem to have devoted more attention to researching the evolution of digital genres with the aim of illuminating their construction and rhetorical structure. As Pérez-Llantada (2016: 23) states, "the impact of digital technologies on genres has been investigated extensively in the context of professional communication". In this area of genre studies, the internalization of digitality is considered to be the first major development (Xia, 2020), the analysis of digital genres thus being inherently connected to the use of multiple communicative modes.

Professional communication is deemed to be multimodal in the sense that a wide range of semiotic modes, including writing, speech, images, animations, gestures, are orchestrated by speakers to make meaning (Jewitt & Kress, 2010). The role of language as a 'decentralized' element in the conveyance of meaning allows the analyst to view human communication from a holistic point of view where every single semiotic mode, verbal and non-verbal, is intentional to the same degree (Jewitt & Kress, 2010). In multimodal studies, *modes* are understood as semiotic systems with rules and regularities – such as images, gestures, speech, music, layout, writing, proxemics, and posture (Kress & Van Leeuwen, 2001). Norris (2004) distinguishes between embodied and disembodied modes where the former consist of resources produced by humans, while the latter refer to external elements which create an impact on the audience. Another key element in the analysis of digital genres is that of *semiotic resources*, defined by van Leeuwen (2005: 285) as "the actions, materials and artifacts we use for communicative purposes". Jones and

351

Hafner (2012) point out that *affordances* refer to what the resources enable us to do (for instance, the VR enables us to present information in a multimodal way and to disseminate the content to a wide range of audiences in a short period of time). The interrelationship between the modes (i.e., embodied and disembodied) is known as *modal density* (Norris, 2004), and the consistency in the use of these modes is widely known as *modal coherence* (Valeiras-Jurado, 2019).

Reflecting on the origins of multimodality, Jewitt (2014) distinguishes three main approaches to multimodal analysis (Figure 1):

Multimodal social semiotics [MSS]	(Kress and van Leeuwen, 1996, 2001)	Interest in the sign-maker in a social context
Multimodal discourse analysis [MDA]	(O'Halloran, 2004, 2011)	Interest in semiotic resources as systems of choices
Multimodal interaction analysis [MIA]	(Norris, 2004)	Interest in the social actors' interaction

Figure 1. Jewitt's (2014) three main approaches to multimodal analysis

The first approach, Multimodal Social Semiotics (MSS), focuses on how modes are used in specific social contexts on the basis of Halliday's (1978, 1985) Systemic Functional Linguistics (SFL) theories. Within this approach, "the emphasis is [...] on the sign-maker and the semiotic choices they make in a given context" (Bernad-Mechó, 2021: 181). The second approach, Multimodal Discourse Analysis (MDA), which also derives from Halliday's (1985) SFL theory, has a different focal point – the study of language as the result of a combination of semiotic resources to convey meaning (O'Halloran, 2011). Finally, the third approach is Multimodal (Inter)action Analysis (MIA), with the focus now on context and situated interaction. In other words, the purpose of this approach is to comprehend how participants in a communicative situation react to each other's discourse. Each approach comprises a wide range of features of multimodal communication, which may differ or coincide depending on the goals to be achieved.

352

Of these three distinct multimodal approaches, I opt here for a corpus-driven MDA approach since it explores the use of language in VRs in a holistic manner. My analysis is based on a comprehensive examination of excerpts from VRs with the aim of developing a framework for the multimodal analysis of VRs which will demonstrate how different semiotic resources interact in this genre. In fact, using videos to create multimodal content may allow students to learn to deploy a wide range of multisemiotic resources in the context of ESP, which entails promoting students' awareness of the multimodal perspective in any type of professional communication context (Ho, 2019).

Therefore, two research questions formulated to guide my study are the following:

RQ1. What is the multimodal rhetorical structure of the VR?
RQ2. What is the contribution of multimodality to this genre?

3. METHODOLOGY

3.1. The dataset

In order to validate the VR as a professional digital genre following a corpus-driven MDA approach, I compiled a set of 26 VRs, all of them in English, from the online platform YouTube. At an initial stage, I took a wide range of features into consideration when carrying out the data collection for the study (see Appendix 1). I then restricted the dataset to videos that met the following criteria: short duration (less than three minutes), presence of both male and female speakers; any type of professional or academic activity, the degree of impact of a VR expressed through the number of views, as well as its year of publication. The videos had to be edited so that they consisted solely of camera shots. The final dataset has a total duration of 52m 28s, which, based on previous, similar studies (Ruiz-Madrid, 2021), can be considered valid for qualitative analysis.

The dataset used for this study had a twofold purpose: (1) the 26 VRs were first analyzed to determine their rhetorical macrostructure in terms of moves and steps; and (2) on the basis of said analysis, a subcorpus of eight videos out of the 26 VRs (around a third of the total dataset) together with the eight excerpts from a particular move (Move 6), with a total length of 179 seconds, was created to examine their multimodal microstructure. This selection may appear modest, but again, it is valid for a fine-grained multimodal analysis as shown by previous studies (Ruiz-Madrid, 2021). The eight videos were selected according to the two distinct criteria: (1) constructing a gender-balanced dataset (four female and four male job applicants); and (2) ensuring interrater reliability, which was done by consulting two additional researchers from a similar research area until consensus was reached on the multimodal density these videos entailed. I opted for a particular

move (Move 6, *Using pressure tactics*) due to its persuasive communicative purpose and its multimodal nature. Move 6 has the function of exerting pressure on the addressee through various persuasive strategies with the aim of obtaining the desired job, its communicative goal essentially being to convince the recruiter. Additionally, this move can be expected to be modally dense, i.e., there is a high-frequency interrelationship among the modes, on account of its persuasive connotations. The Results section contains detailed information on time intervals and the excerpts from Move 6 to be discussed (Table 1). The whole process involves, first, the sequencing of structural elements, divided into moves and steps, each of which performs a specific rhetorical function, and then, the viewing of all the multisemiotic resources in increasing depth in order to identify key moments from a multimodal perspective (Jewitt, 2009).

3.2. Annotation and analysis

Fully automated software tools that identify and annotate functional units in multimodal videos have not yet been fully developed. Nonetheless, the analytic strategies I used in this study rely on two specialized software packages designed to perform computer-assisted rhetorical and multimodal analyses, supplemented by manual analysis. The first package is ATLAS.ti (https://atlasti.com/), defined as qualitative data analysis (QDA) software, which was used here for the move-step rhetorical analysis of the VR. Appendix 2 shows the interface of the operating software for rhetorical transcription annotations. In its left corner there is a video window where the clips can be played; in the right corner of the interface there is space for verbal transcriptions as well as a list of strips where choices are annotated in time. I created these strips manually with the aim of annotating all relevant moves and steps used in the selected excerpts. ATLAS.ti provides functional analytic tools in academic and professional research, particularly in the social science disciplines (Hwang, 2008). The same author states that employing QDA software can be beneficial because it makes the process more transparent and replicable. Additionally, given the increased credibility it brings, it can be time-saving and more effective in terms of project management. Hwang (2008) also argues that one of the main reasons for using ATLAS.ti in methodological development is its ability to efficiently handle digital media formats such as video.

Move analysis, i.e., the discourse structuring of a communicative event, is a procedure in genre studies based on the notion that moves are "semantic and functional units of texts which can be identified because of their communicative purposes and linguistic boundaries" (Ding, 2007: 370). A *step* is a smaller unit than a move and it is, therefore, placed on a subordinate level. Accordingly, I divided the videos into the components of their systematic structure (moves and steps), and the sections thus identified were treated as the basic units of analysis and as units involved in a goal-directed communicative event. When categorizing the moves and

354

steps identified in the VR dataset, I took into account the methods and results used in previous studies on professional genres, such as Bhatia's (1996), Furka's (2008), or Wang's (2005) rhetorical analysis of job application letters, as well as on academic genres, such as Hyland's (2004) analysis of research articles and Plastina's (2017) work on video abstracts.

According to Valeiras-Jurado et al. (2018: 99), "an MDA approach requires the use of different specialized software packages to look into the data". Therefore, the second software package I used for this study is the multimodal annotation software application, known as Multimodal Analysis-Video (MMA-Video) (O'Halloran et al., 2012), which was used to identify verbal and non-verbal cues conveying meaning in VRs. O'Halloran et al. (2017: 22) argue that MMA-Video "provides the necessary tools for investigating the use of semiotic resources and the ways in which semiotic choices interact to fulfill particular objectives in a multimodal video". This tool was used to address the second research question – the contribution of multimodality to the development of the genre – and was applied to the eight VRs selected. Its interface is shown in Appendix 3. These eight VRs were viewed several times with the aim of making preliminary notes on the types of embodied and disembodied modes (Norris, 2004) present in them and on significant elements added in the postproduction process, or "filmic modes" (Valeiras-Jurado & Bernad-Mechó, 2022).

I examined embodied, disembodied and filmic modes, because meaning is sometimes based on the interrelation of specific modes, or "multimodal ensembles" (Ruiz-Madrid, 2021). Speakers make use of them because videos "fulfill their respective communicative aims and functions through various combinations of semiotic choices in their organizational structure, functional stages and properties" (O'Halloran et al., 2017). I therefore paid particular attention to the nature of the multimodal ensembles used by speakers in VRs drawing on previous research on the use of semiotic resources in other formats (Bernad-Mechó, 2022; Ruiz-Madrid, 2021). I performed my analysis only after all instances occurring for each of the modes had been annotated in the program.

355

4. RESULTS AND DISCUSSION

As the main purpose of this study is to validate the VR as a professional digital genre, in this section I present the results obtained after conducting a specific type of genre analysis determining the multimodal rhetorical structure of VRs and discuss the most relevant findings.

4.1. Identification and analysis of the VRs' rhetorical structure

In response to RQ1 – What is the multimodal rhetorical structure of the VR? – I conducted a move-and-step rhetorical analysis. Based on Bhatia's (1993) seminal

study of job application letters in terms of genre analysis, I devised a more comprehensive version of the move-and-step rhetorical organization for VRs, as presented in Figure 2.
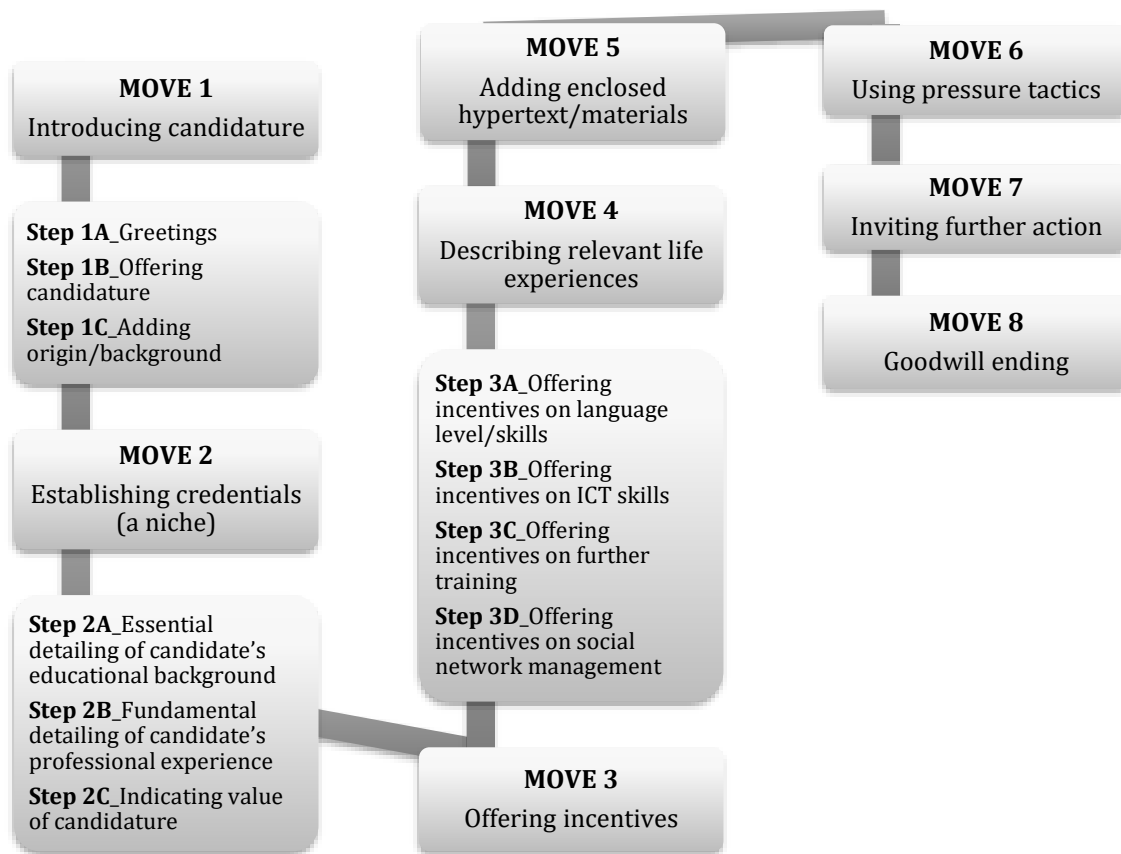
Figure 2. Taxonomy of video resumes' rhetorical organization

Of the eight moves identified, seven are consistently apparent in the analysis as core moves, and only one move (Move 4) is optional. The first move (Introducing candidature) allows the candidate to introduce themselves, and it is divided into three main steps (Greetings, Offering candidature and Adding origin/background information) in which the candidate greets the viewer, mentions the job offer they are referring to and talk about their origin/background. This type of information can be illustrated as follows:

(1) Hi, my name is XX, and I am a front-end web developer. I'm looking for a job and instead of just sending out the usual cover letter resume, I'm making this quick video so that you can more easily learn a little about me. (VR21)

In this specific case, the speaker is already using a discursive strategy to distinguish herself from the candidate who presents the traditional paper-based resume.

In Move 2, Establishing credentials (a niche), the candidate establishes their credentials as pertaining to someone who can undoubtedly fill the job niche. Three different ways of supplying the information were identified: presenting the candidate's essential educational background details (e.g., "In November 2015 I graduated from the Iron Yard" [VR21]), reviewing the candidate's primary professional experiences, and evaluating themselves for the desired job. This move is strongly connected with Move 3, where the candidate offers their personal strengths on the most relevant aspects. Four different types of personal strengths were identified: foreign language level/skills, ICT skills, further training, and social network management. This is clearly illustrated in the following excerpt:

(2) I've been creating custom web layouts, navigations and smooth user experience design [...] I've been increasing my knowledge in Photoshop's frameworks and JavaScript library. (VR21)

Move 4 is regarded as optional because the information given is less relevant to the speaker's overall goal than in other moves, and allows the speaker to describe life experiences in terms of hobbies and interests:

(3) Apart from sunsets and long walks on the beach, I also happen to enjoy just working out mixed martial arts, golfing, guitar, acting, photography and blogging, traveling and socializing. (VR20)

357

Move 5 (Adding enclosed/hypertext materials) is one of the most frequent moves in the VRs due to the limited time length which this video format entails. To compensate for this, the candidate provides additional information through links shared on the screen to complement the VR.

(4) For those of you who are unfamiliar with the Iron art program you can click 'here' to learn about the training we provide ... You can check out my portfolio 'here' or you can contact me directly 'here'. (VR21)

In Move 6 (Using pressure tactics), the candidate exerts pressure on the addressee through persuasive strategies used to obtain the desired job. As explained in the previous sections, this move is the one selected for the multimodal analysis, since it is here that the speaker emphasizes his/her strengths and potential aptitudes so as to be considered the most suitable candidate for the job position.

(5) I'm an excellent candidate for a company or team looking to capitalize a lot ... I'm always up for a challenge, and I'm looking for a place where I can continue to grow and learn. (VR21)

In Move 7 (Inviting further action), the candidate closes his/her speech persuasively by suggesting that the addressee should take action:

(6) I'd love to hear from you; hopefully we'll be working together soon. (VR21)

The last move, Move 8 (Goodwill ending), allows the candidate to end the VR politely using a short affirmative sentence, e.g. "Thank you for your time" (VR21).

With the aim of complementing the rhetorical structures with the multisemiotic resources, which are present in audiovisual formats, I employed an MDA approach to analyze Move 6 (Using pressure tactics).

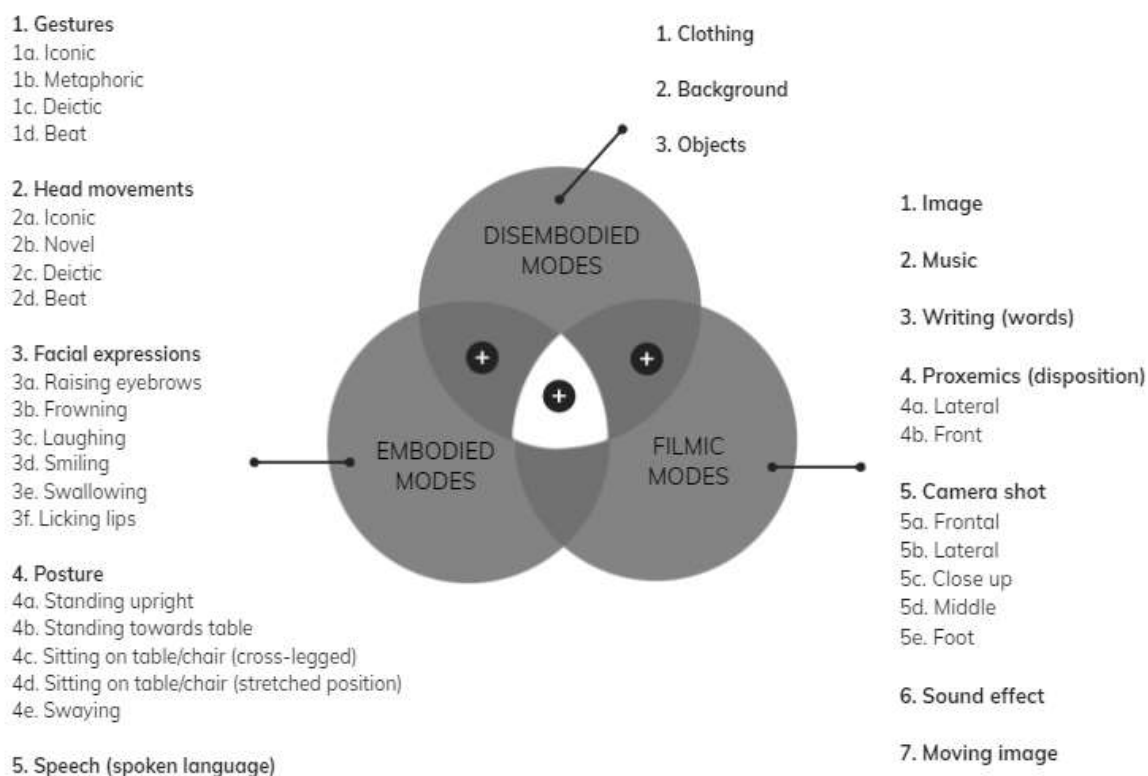## 4.2. The MDA approach: multimodal ensembles in Move 6

In response to RQ2 – What is the contribution of multimodality to this genre? – the analytic focus now shifts to the role of multimodality in the digital genre under study. As explained earlier, Move 6 (i.e., Using pressure tactics) was selected from a set of eight VRs for a specific MDA in terms of presumed multimodal density and communicative aim. The initial hypothesis is based on the idea that Move 6 should be modally dense in order to successfully convince the addressee and obtain the desired job position. Table 1 shows the exact time intervals as well as the excerpts from Move 6 identified in the subcorpus of this study.

8 VRs (SUBCORPUS)	TIME INTERVALS	TOTAL OF SECONDS	EXCERPTS FROM MOVE 6
VR19	1'27''- 1'56''	29 sec	*Why should you pick me out of seven billion people living on this planet? Well, as a strong communicator and negotiator who can build effective relationships, I specialize in finding creative and innovative solutions to the toughest problems. I also happen to speak English, French and Spanish and have been involved in very exciting stuff for the last few years. I truly believe that I can bring a real and positive impact to the organization from day one. I'd love to join a creative agency, working for my skills and experience to good use and apply my entrepreneurial principles to foster social transformation.*
VR20	1'46''- 2'12''	66 sec	*Why should you pick me?* *Well, I'm very interested in digital trends, new ways to sell products and services and how to build strong marketing strategies for companies around the world. I'm also able to speak English, French, Arabic and Spanish. I'm a fully creative, strong communicator and extremely motivated to build effective relationships with the ability to fit in with any group, any service, any company. I'm sure I can bring a really positive energy to the organization from day one.*
VR21	1'18-1'32''	14 sec	*As a recent graduate from the Iron Yard, I'm an excellent candidate for a company or team looking to capitalize a lot I currently know and shape me into what they specifically need. I'm always up for a challenge, and I'm looking for a place where I can continue to grow and learn.*

VR22	1'21''-1'29''	8 sec	*I believe I am your perfect candidate for your company and will become a great future employee.*
VR23	2'18''-2'35''	17 sec	*I'm looking for a company that believes in getting stuff done, low bureaucracy, and empowering the staff to use their skills and talents for the better at a business. I'd prefer to be used for my brainpower and skills, but I'm willing to crush whatever is put in front of me. I hope that this is interesting to you or someone that you might know.*
VR24	2'34''-2'44''	10 sec	*I am convinced that you are looking for a young, creative, hard-working, reliable and committed employee who will do their best. I'm one of those!*
VR25	1'26''-1'56''	30 sec	*Why should a company employ me?* *I have all these skills that I feel are very transferable: soft skills that are necessary in job roles when I apply for jobs. I feel like companies are looking for this. Alongside being very good at communicating, a good problem solver. I find that I've learned from my project management how to be very organized, how to plan better, and I'm just good at time management as well, and I think this works in the role I'm looking for.*
VR26	0'41''-0'46''	5 sec	*If you're looking to increase sales production, develop your employees and retain them, I'm the person for the job.*

Table 1. Time intervals and excerpts from Move 6

359

In order to examine the correlations between the different semiotic modes from the extracts selected, a detailed framework for the multimodal annotation was established, considering previous studies on multimodal analysis (e.g., Bernad-Mechó, 2022), as shown in Figure 3 below. This framework is used to evaluate the figure of the presenter, that is, how the presenter communicates, or what multimodal resources they use to convey the message. To provide a detailed description of the most salient semiotic resources used by the presenters in their videos, six Figures are supplied below (Figures 4, 5, 6, 7, 8 and 9). Each of these VRs deploys a coherent multimodal ensemble in which the most frequent modes interact, with the aim of conveying the rhetorical function of Move 6 (i.e., Using pressure tactics) effectively. Consequently, meaning is conveyed by a multiplicity of modes that are set in motion when contents and engagement features are directed toward the audience, although meaning is also expressed by verbal and non-verbal resources and other supporting elements.

1. Gestures
1a. Iconic
1b. Metaphoric
1c. Deictic
1d. Beat

2. Head movements
2a. Iconic
2b. Novel
2c. Deictic
2d. Beat

3. Facial expressions
3a. Raising eyebrows
3b. Frowning
3c. Laughing
3d. Smiling
3e. Swallowing
3f. Licking lips

4. Posture
4a. Standing upright
4b. Standing towards table
4c. Sitting on table/chair (cross-legged)
4d. Sitting on table/chair (stretched position)
4e. Swaying

5. Speech (spoken language)

1. Clothing

2. Background

3. Objects

DISEMBODIED MODES

EMBODIED MODES

FILMIC MODES

1. Image

2. Music

3. Writing (words)

4. Proxemics (disposition)
4a. Lateral
4b. Front

5. Camera shot
5a. Frontal
5b. Lateral
5c. Close up
5d. Middle
5e. Foot

6. Sound effect

7. Moving image

Figure 3. Framework for the multimodal analysis of VRs

Figure 4 displays a coherent multimodal ensemble that allows the speaker to draw the audience's attention to himself as a suitable candidate for the job position. He employs both verbal and non-verbal semiotic resources to introduce the segment: the former are represented by a direct question (i.e., "Why should you pick me out of seven billion people living on this planet?" [VR19]), which is in turn complemented by a metaphoric gesture (i.e., open hand supine gesture) suggesting openness and honesty and creating rapport. Moreover, he appears to use different embodied modes intentionally, with the aim of keeping the audience's attention. For instance, he frequently uses head movements (i.e., head beat movement when turning his head to the left side of the frame) as well as various facial expressions to emphasize his explanation. This is mainly achieved with an eyebrow raise gesture, which co-occurs with the phrase "I truly believe that I can bring a real and positive impact to the organization from day one". These semiotic resources combine with various filmic modes to enrich the candidate's speech. For example, the speaker starts the section in a front disposition (i.e., proxemics) and moves into a lateral disposition to change the topic. Moreover, there is a variation in the camera shot, as the excerpt starts with the speaker in a middle frame, then moves to a close-up camera shot to reinforce his words. In short, this speaker utilizes a complex multimodal ensemble.
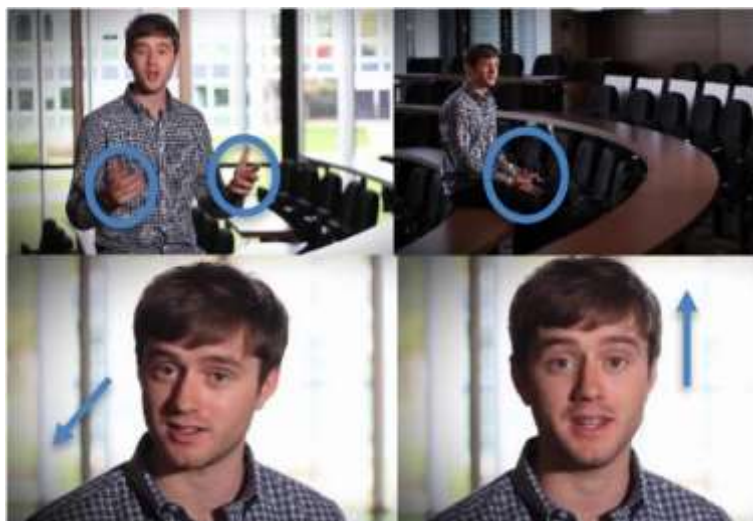
Figure 4. Semiotic resources from VR19

The speaker in Figure 5 is also multimodally coherent, and his presentation is mostly based on embodied modes to achieve the communicative aim of Move 6. For example, variation in performing gestures may be noticed, especially when used metaphorically when the speaker puts his index finger up. He aims to emphasize information he considers relevant, since immediately after this gesture the speaker says: "I'm sure I can bring a really positive energy to the organization from day one" (VR20). In terms of posture, during the whole excerpt he remains seated on a chair in a stretched-out position. In addition, the camera shoots from a middle distance and a front disposition.

361

Figure 5. Semiotic resources from VR20

Figure 6. Semiotic resources from VR23

Similarities in performance to the speaker from VR23 (Figure 6) can be observed, as the embodied modes are the most widely used. In particular, the metaphorical gesture is repeated along with the eyebrow raise (i.e., facial

expression) representing the action of thinking. Nevertheless, it should also be noted that he closes his eyes momentarily to pause and consider his next words, which can be accounted for as a metaphoric representation of silence. Three speakers (VR19, VR20 and VR23) wear a light-color shirt, which can be regarded as a disembodied mode intended to show seriousness and elegance. In fact, as in any online performance, it is vital to choose an optimal outfit that will make a strong impression while better defining the individual's personality (Palmer-Silveira & Ruiz-Garrido, in press).

However, one of the most outstanding disembodied modes observed is the university T-shirt (i.e., clothing) in VR21 (Figure 7), which the speaker appears to wear intentionally in the video in order to show that she takes pride in her university. This co-occurs with the phrase "As a recent graduate from the Iron Yard, I'm an excellent candidate for a company". Furthermore, she is fairly expressive and continually communicates through facial expressions. The most recurrent of these are eyebrow raises and smiles. A similar instance can be identified in VR24 (Figure 8). The speaker wears a bow tie (i.e., object or clothing as a disembodied mode) as a way of distinguishing herself from other candidates. She has an eloquent manner, which may explain her use of facial expressions throughout the video (i.e., eyebrow-raises and smiles). She emphasizes certain parts of her message over others to attract more attention. This is accomplished by the co-occurrence of two embodied modes: gestures (i.e., iconic and beat) and speech (i.e., "I'm one of those!"). She also uses her fingers in a coherent way to establish the order or appearance of concepts at a specific moment of the talk while listing some of her strengths ("I am convinced that you are looking for a young, creative, hard-working, reliable and committed employee who will do their best").

362

Figure 7. Semiotic resources from VR21

Figure 8. Semiotic resources from VR24

In VR25 (Figure 9) the speaker makes use of one of the most recurrent filmic modes identified in this study: words superimposed on video. In this particular case, the speaker uses the direct question "Why should a company employ me?" to support her speech and provide the audience with visual guidance. She also delivers her speech in a somewhat impassive way, although certain facial expressions, such as eyebrow raises, a shy smile and closed eyes during pauses, can also be observed.

Figure 9. Semiotic resources from VR25

In sum, all the multimodal ensembles orchestrated by the speakers help to engage the audience, making their message relevant and decisive and easy to understand.

5. CONCLUSION

This research has drawn attention to the growing use of digital technologies in business communication contexts of which VR is a prime example. Such technologies can better showcase job candidates' potential and make them more aware of the strengths they can offer potential employers through a digital medium. The study, therefore, has attempted to contribute to raising awareness and the understanding of professional genres in the current digital era. As Pérez-Llantada (2016) points out, certain features of multimodality, such as those analyzed in this study (embodied, disembodied and filmic modes), act jointly with those characteristic of stabilized and traditional genres, indicating that new forms and practices are being used.

The multisemiotic perspective taken in this study can also lead to a reflection on the use of various semiotic resources in professional communication. In line with previous research (Ruiz-Madrid, 2021), genre analysis is perceived in the present research from a wider perspective. It goes beyond the traditional Swalesean approach (Swales, 1990), where language is the only mode taken into consideration. It has defined the VR as a professional digital genre using two analytical approaches, the rhetorical and multimodal, and it has determined the multimodal rhetorical structure of the VR.

The findings of this study can provide several pedagogical implications. Analyzing VRs using multimodal semiotic resources can contribute to the teaching of communication skills in higher education, and help to develop students' multimodal understanding required for producing VRs, due to their strategic orchestration of multiple semiotic modes (Ho, 2019). More specifically, "the creation of authentic materials can increase students' motivation and expose them to real language and cultures as well as to the different genres of the professional community to which they aspire" (García-Ostbye & Martínez-Sáez, 2023: 56).

As ESP analysts, we need to develop new methods of analyzing VRs and similar genres for both research and teaching purposes (Coccetta, 2018). Speakers need to acquire multimodal skills in order to be able to communicate effectively with their audience, regardless of their level of expertise. From a practical standpoint, ESP learners should develop their multimodal skills by constructing meaning from the multiple semiotic resources that characterize these digital genres. The genre analysis I present in this paper offers insights that can be applied to the teaching of ESP courses, in which the VR would act as a teaching resource to develop multimodal skills for communication in professional contexts. Encouraging students to create their own VRs by using a comprehensive teaching proposal with a well-founded move structure framework will allow them to be better prepared for their professional life, create a long-lasting effect when applying for a vacant position and generally enhance their multimodal communication skills. In sum, the teaching of VRs seems pedagogically advisable, as it can help refine the curricular design of oral and digital genres in higher education (O'Halloran et al., 2016).

Finally, as in any empirical research, some limitations need to be pointed out. First, there is a significant difference in the size of the datasets used for the two analyses carried out in this study. Although multimodal studies are usually based on small datasets due to the fine-grained analysis they perform, the total duration of my multimodal subcorpus was only 179 seconds in comparison to the dataset used for the rhetorical analysis of the 26 VRs, which amounted to a total of 52m 28s. Furthermore, the dataset of the multimodal analysis only included one move (Move 6). Therefore, further research on VRs could draw on larger corpora and more moves to develop a deeper understanding of the genre. It would also be worthwhile to replicate the study in a wider range of written and spoken professional genres, such as job applications, motivational letters and cover letters, or job interviews, to delve deeper into potential signs of hybridity. We hope that this paper provides a

364

flexible, easily adaptable and data-driven analytical framework. Furthermore, the taxonomy of VRs' rhetorical organization and the multimodal annotation framework presented here can serve as a basis for digital genre analysis and MDA research in general.

References

Bateman, J. A. (2008). *Multimodality and genre: A foundation for the systematic analysis of multimodal documents.* Palgrave Macmillan. http://dx.doi.org/10.1057/9780230582323

Bateman, J. A. (2014). Developing a GeM (genre and multimodality) model. In S. Norris & C. D. Maier (Eds.), *Interactions, images and texts: A reader in multimodality* (pp. 25–36). De Gruyter Mouton. https://doi.org/10.1515/9781614511175.25

Bateman, J. A. (2017). Multimodality and genre. In A. Black, P. Luna, O. Lund, & S. Walker (Eds.), *Information design: Research and practice* (pp. 237–258). Routledge. https://doi.org/10.4324/9781315585680

Bernad-Mechó, E. (2021). Combining multimodal techniques to approach the study of academic lectures: A methodological reflection. *Atlantis: Journal of the Spanish Association for Anglo-American Studies*, *43*(1), 178–198. https://doi.org/10.28914/Atlantis-2021-43.1.10

Bernad-Mechó, E. (2022). Modal density in structuring segments containing organizational metadiscourse versus content sequences. *ESP Today*, *10*(1), 2–21. https://doi.org/10.18485/esptoday.2022.10.1.1.

Bhatia, V. K. (1993). *Analyzing genre: Language use in professional settings.* Longman.

Bhatia, V. K. (1996). Methodological issues in genre analysis. *HERMES: Journal of Language and Communication in Business*, *9*(16), 39–59. https://doi.org/10.7146/hjlcb.v9i16.25383

Coccetta, F. (2018). Developing university students' multimodal communicative competence: Field research into multimodal text studies in English. *System*, *77*, 19–27. https://doi.org/10.1016/j.system.2018.01.004

Ding, H. (2007). Genre analysis of personal statements: Analysis of moves in application essays to medical and dental schools. *English for Specific Purposes*, *26*(3), 368–392. https://doi.org/10.1016/j.esp.2006.09.004

Furka, I. (2008). The curriculum vitae and the motivational letter: A rhetorical and cultural analysis. *WoPaLP: Working Papers in Language Pedagogy*, *2*, 18–37.

García-Ostbye, I. C., & Martínez-Sáez, A. (2023). Reading challenges in higher education: How suitable are online genres in English for medical purposes? *ESP Today*, *11*(1), 53–74. https://doi.org/10.18485/esptoday.2023.11.1.3

Halliday, M. A. K. (1978). *Language as a social semiotic: The social interpretation of language and meaning*. Edward Arnold.

Halliday, M. A. K. (1985). *An introduction to functional grammar*. Edward Arnold.

Hiemstra, A. M. F. (2013). *Fairness in paper and video resume screening* [Unpublished doctoral dissertation]. Erasmus University Rotterdam.

365

Hiemstra, A. M. F., & Derous, E. (2015). Video résumés portrayed: Findings and challenges. In I. Nikolaou & J. K. Oostrom (Eds.), *Employee recruitment, selection, and assessment: Contemporary issues for theory and practice* (pp. 45–60). Routledge/Taylor & Francis Group.

Hiemstra, A. M. F., Derous, E., Serlie, A. W., & Born, M. P. (2012). Fairness perceptions of video resumes among ethnically diverse applicants. *International Journal of Selection and Assessment*, *20*(4), 423–433. https://doi.org/10.1111/ijsa.12005

Ho, W. Y. J. (2019). Developing professional communication: The construction of a multimodal understanding of job interviews. *Languages*, *4*(1), 1–12. https://doi.org/10.3390/languages4010005

Hwang, S. (2008). Utilizing qualitative data analysis software: A review of Atlas.ti. *Social Science Computer Review*, *26*(4), 519–527. https://doi.org/10.1177/0894439307312485

Hyland, K. (2004). *Genre and second language writing*. The University of Michigan Press.

Jewitt, C. (2009). An introduction to multimodality. In C. Jewitt (Ed.), *The Routledge handbook of multimodal analysis* (pp. 14–27). Routledge.

Jewitt, C. (2014). Different approaches to multimodality. In C. Jewitt (Ed.), *The Routledge handbook of multimodal analysis* (2nd ed., pp. 28–39). Routledge.

Jewitt, C., & Kress, G. (2010). Multimodality, literacy and school English. In D. Wyse, R. Andrews, & J. Hoffman (Eds.), *The Routledge international handbook of English, language and literacy teachin*g (pp. 342–352). Routledge.

Jones, R., & Hafner, C. A. (2012). *Understanding digital literacies: A practical introduction*. Routledge. https://doi.org/10.4324/9780203095317

Kemp, K. J., Bobbitt, L. M., Beauchamp, M. B., & Peyton, E. A. (2013). Using one-minute video résumés as a screening tool for sales applicants. *Journal of Marketing Development and Competitiveness*, *7*(1), 84–92.

Kress, G., & Van Leeuwen, T. (2001). *Multimodal discourse: The modes and media of contemporary communication.* Edward Arnold.

Nguyen, L. S., & Gatica-Perez, D. (2016). Hirability in the wild: Analysis of online conversational video resumes. *IEEE Transactions on Multimedia*, *18*(7), 1422–1437. https://doi.org/10.1109/Tmm.2016.2557058

Norris, S. (2004). *Analysing multimodal interaction: A methodological framework*. Routledge. https://doi.org/10.4324/9780203379493

O'Halloran, K. (2011). Multimodal discourse analysis. In K. Hyland & B. Paltridge (Eds.), *Continuum companion to discourse analysis* (pp. 120–137). Continuum International Publishing Group.

O'Halloran, K. L., Tan, S., & E, M. K. L. (2017). Multimodal analysis for critical thinking. *Learning, Media and Technology*, *42*(2), 147–170. https://doi.org/10.1080/17439884.2016.1101003

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). Routledge.

O'Halloran, K. L., Podlasov, A., Chua, A., & E, M. K. L. (2012). Interactive software for multimodal analysis. *Visual Communication*, *11*(3), 363–381. https://doi.org/10.1177/1470357212446414

Palmer-Silveira, J. C., & Ruiz-Garrido, M. F. (in press). Introducing science to the public in 3-minute talks: Verbal and non-verbal engagement strategies. In R. Plo-Alastrué, R. & I. Corona-Marzol (Eds.), *Digital scientific communication: Identity and visibility in research dissemination.* Palgrave Macmillan.

Pérez-Llantada, C. (2016). How is the digital medium shaping research genres? Some cross-disciplinary trends. *ESP Today*, *4*(1), 22–42. https://www.esptodayjournal.org/pdf/current_issue/3.6.2016/CARMEN%20P%C3%89REZ-LLANTADA-full-text.pdf

Plastina, A. F. (2017). Professional discourse in video abstracts: Re-articulating the meaning of written research article abstracts. In G. Garzone, P. Catenaccio, K. Grego, & R. Doerr (Eds.), *Specialized and professional discourse across media and genres* (pp. 57–73). Ledizioni.

Ruiz-Madrid, M. N. (2021). A multimodal discourse approach to research pitches. *Journal of English for Academic Purposes*, *52*, Article 101003. https://doi.org/10.1016/j.jeap.2021.101003

Swales, J. (1990). *Genre analysis: English in academic and research settings.* Cambridge University Press.

Valeiras-Jurado, J. (2019). Modal coherence in specialised discourse: A case study of persuasive oral presentations in business and academia. *Ibérica*, *37*, 87–114. https://www.revistaiberica.org/index.php/iberica/article/view/111

Valeiras-Jurado, J., & Bernad-Mechó, E. (2022). Modal density and coherence in science dissemination: Orchestrating multimodal ensembles in online TED talks and youtube science videos. *Journal of English for Academic Purposes*, *58*, Article 101118. https://doi.org/10.1016/j.jeap.2022.101118

Valeiras-Jurado, J., Ruiz-Madrid, N., & Jacobs, G. (2018). Revisiting persuasion in oral academic and professional genres: Towards a methodological framework for multimodal discourse analysis of research dissemination talks. *Ibérica*, *35*, 93–118. https://revistaiberica.org/index.php/iberica/article/view/136

van Leeuwen, T. (2005). *Introducing social semiotics: An introductory textbook.* Routledge.

Wang, H. (2005). A pragmatic genre analysis of job application letters. *Sino-US English Teaching*, *2*(12), 76–81.

Waung, M., Hymes, R. W., & Beatty, J. E. (2014). The effects of video and paper resumes on assessments of personality, applied social skills, mental capability, and resume outcomes. *Basic and Applied Social Psychology*, *36*(3), 238–251. https://doi.org/10.1080/01973533.2014.894477

Xia, S. A. (2020). Genre analysis in the digital era: Developments and challenges. *ESP Today*, *8*(1), 141–159. https://doi.org/10.18485/esptoday.2020.8.1.7

367

MARÍA-ÁNGELES MESTRE-SEGARRA is a PhD student at Universitat Jaume I (UJI), Spain. Within the Group for Research on Academic and Professional English (GRAPE), she studies the video resume (VR) as a professional digital genre. Her research interests are related to digital genres, multimodality and ESP studies. Her most recent publication has appeared in *System*.

Appendix 1. Description of the dataset

VRs	Duration	PROFESSIONAL/ ACADEMIC POSITION	Gender	IMPACT (Views)	YEAR OF PUBLICATION	VIDEO TYPE (Edited video)	CAMERA SHOT
VR1	2' 16''	Translator	F	228,234	2013	image + music + person	half body
VR2	2' 31''	Student	M	113,229	2015	image + music + audio + person	face, half body, full body
VR3	1' 22''	Visual Designer	M	30,419	2018	image + music + audio + person	half body
VR4	1' 44''	Art director	M	62,973	2017	image + music + audio + person	half body
VR5	2' 02''	Lawyer	M	65,732	2017	image + music + audio + person	half body
VR6	2' 05''	Social media specialist	F	18,727	2018	image + music + audio + person	face, half body
VR7	1' 36''	Publicist	M	34,922	2020	image + music + audio + person	half body
VR8	2' 25''	Student	F	79,125	2020	image + music + audio + person	half body
VR9	2' 12''	Marketing professional	F	87,135	2019	image + music + audio + person	face
VR10	1' 47''	Advertisement	F	15,445	2020	image + music + audio + person	half body
VR11	2' 04''	Student	F	109,340	2015	image + music + audio + person	half body
VR12	1' 50''	Administrative	M	45,617	2010	image + music + audio + person	face, half body, full body
VR13	2' 46''	Student	F	369,110	2016	image + music + audio + person	face
VR14	1' 02''	PR practitioner	M	663,528	2010	person + audio	half body
VR15	2' 16''	Sales and marketing	F	335,704	2012	image + music + audio + person	half body
VR16	2' 02''	Sales and business development	M	1,488	2019	person + audio	half body

VR17	2' 41''	Filmmaker	M	67,432	2016	image + music + audio + person	half body
VR18	0' 58''	Sales manager	F	149,173	2019	person + audio	face
VR19	2' 10''	Publicist	M	977,826	2012	person + audio	face, half body, full body
VR20	2' 34''	Salesman	M	16,652	2019	image + music + audio + person	half body
VR21	1' 52''	Web developer	F	124,584	2016	image + music + audio + person	half body
VR22	1' 38''	Student	M	22,871	2020	image + music + audio + person	half body
VR23	2' 50''	Travel manager	M	5,655	2020	image + music + audio + person	half body
VR24	2' 57''	Student	F	293,152	2014	image + music + audio + person	half body
VR25	2' 02''	Project manager	F	865	2019	person + audio	half body
VR26	0' 46''	Sales manager	F	202,454	2010	person + audio	half body

369

Appendix 2. ATLAS.ti interface for rhetorical transcription annotations
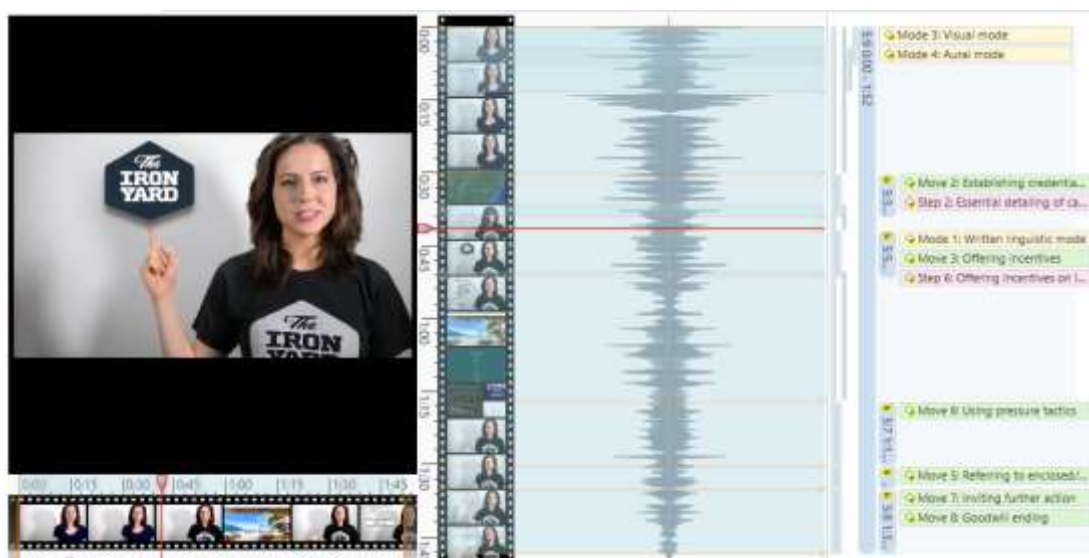
Appendix 3. MMA-Video interface for multimodal transcription annotations