

Multimodal engagement strategies in science dissemination: A case study of TED talks and YouTube science videos

Discourse Studies

1–22

© The Author(s) 2023

Article reuse guidelines:

sagepub.com/journals-permissions

DOI: 10.1177/14614456231161755

journals.sagepub.com/home/dis**Edgar Bernad-Mechó**

Universitat Jaume I (UJI), Spain

Julia Valeiras-Jurado

Universitat Jaume I (UJI), Spain

Abstract

The growing interest on science dissemination offers new opportunities to communicate science openly to various audiences, but also brings on the challenge of adapting to an audience that does not share the same academic background. This adaptation has been referred to as *recontextualization*. In the case of the formats that concern this study, that is, TEDx Talks and YouTube science dissemination videos, their multimodal nature suggests that recontextualization, and therefore engagement as a crucial aspect of this process, is likely to go way beyond purely linguistic aspects. The aim of this study is to unveil how engagement strategies in two science dissemination formats (a face to face talk and an online video) are realized through complex multimodal ensembles, and to highlight differences across them. In order to fulfill this aim, two talks by the same presenter and dealing with similar content were selected for analysis: a TEDx talk and a YouTube science dissemination video from the channel PBS Space Time. The recordings were annotated using the software Multimodal Video Analysis. The annotation included engagement strategies; embodied modes, that is, modes carried out using the body; and, in the case of the YouTube video, filmic modes, that is, modes triggered by the editing process of the recorded video. Our results show that the role of both embodied and filmic modes is paramount in the realization of engagement strategies. Our findings also bring to the fore significant differences in the ways in which the two distinct audiences are engaged, concerning the frequency and use of both semiotic modes and engagement strategies.

Corresponding author:

Edgar Bernad-Mechó, Universitat Jaume I (UJI), Av Sos Baynat s/n, Castello de la Plana 12071, Spain.

Email: ebernad@uji.es

Keywords

Embodied modes, engagement strategies, filmic modes, multimodality, science dissemination

Introduction

Science popularization can be defined as the dissemination of scientific knowledge to a lay audience (Xia and Hafner, 2021). It is a trend that has attracted increasing scholarly interest. In particular, the recontextualization processes by which scientific content is made accessible and appealing to a non-specialized audience has been the object of recent studies (Calsamiglia and Van Dijk, 2004). These studies explore new ways of communicating science through genres such as TED talks (Scotto di Carlo, 2014, 2015; Xia and Hafner, 2021), research groups online science videos (Luzón, 2019), Three Minute Thesis presentations (Carter-Thomas and Rowley-Jolivet, 2020; Zou and Hyland, 2021), research pitches (Ruiz-Madrid, 2021) or online science dissemination videos (Valeiras-Jurado and Bernad-Mechó, 2022). What these genres seem to have in common is that they belong to a model of science popularization that brings scientists into direct contact with the general public, without the need of an intermediary such as a journalist (Xia and Hafner, 2021). The influence of online media is also very noticeable in these genres: some of them originate online (e.g. YouTube science videos), while others are migrating to such an online environment (e.g. TED talks, Three Minute Thesis). This shift brings along new possibilities, such as maximizing reach, but also new challenges, for example using the new affordances skillfully to bridge the knowledge gap, or to connect with an audience that is not always physically present.

These challenges can be overcome by means of different techniques. Previous studies have pointed out how the use of inclusive pronouns, emotionally-loaded adjectives, or the use of questions can greatly contribute to engaging the audience in science popularization genres (Caliendo, 2014; Hyland, 2010; Scotto di Carlo, 2014, 2015). However, recent research has noticed the need to account for recontextualization practices (and engagement strategies among them) that go beyond the merely linguistic and make use of varied semiotic modes, which calls for a multimodal approach to recontextualization. Luzón (2019), for example, identifies recontextualization strategies in online science videos that orchestrate speech, gestures and image. She groups them into four categories depending on their function: (i) build credibility; (ii) build persuasive arguments; (iii) tailor information to the assumed knowledge of the audience; and (iv) engage the audience. This last category is the one that concerns and informs the present study more directly. Xia and Hafner (2021) carry out a multimodal analysis of engagement strategies in TED talks. Their analysis shows that the engagement of online viewers is generally achieved by a combination of multiple semiotic resources, including, among others, visual aids, camera shot and gaze. Finally, Valeiras-Jurado and Bernad-Mechó (2022) further explore the complexity of multimodal ensembles in PBS online science videos, highlighting the crucial role of filmic modes that are added to the ensemble during montage. They find that these videos are tailor-made for online audiences and exploit the possibilities of the online medium successfully and coherently to recontextualize content. In particular, these videos show a high density of filmic modes in the form of visual prompts, visual and sound effects, coherent types of shots and cuts.

The present study focuses on a specific aspect of recontextualization: engagement. We argue that “being engaging” is one of the inherent characteristics of the formats object of this case study. As Bondi et al. (2015) put it, “recontextualization is aimed not only at making specialist knowledge available to the wider public [. . .], but also at making it relevant or interesting for the non-specialist reader” (p. viii). This paper aims to go beyond previous studies in science dissemination to describe how engagement is realized multimodally, comparing a single speaker sharing similar content in two main contexts: a face-to-face TEDx talk and an online YouTube science video. In order to fulfill this aim, the following research question was established:

- *What similarities and differences can be spotted across the two science dissemination formats under scrutiny in terms of multimodal engagement?*

The next sections explore the previous literature on engagement in science dissemination and the formats being studied to establish the starting point for the case study presented in this paper.

Engagement in science dissemination

Martin and White (2005) consider that all utterances have a dialogic nature and explore how speakers and writers engage with prior knowledge, acknowledge previous knowers and anticipate potential responses. Focusing on scientific communication, Hyland (2010) describes how engaging the readership in professional and popular science means acknowledging and connecting to others. In the same line, but focusing on discussion sessions in doctoral dissertations, for Loghmani et al. (2020), engagement in scientific communication means sourcing attitudes and stance in discourse. Although these three views on engagement differ in their foci (discourse in general, popular science and academic discourse respectively), they all highlight the relevance of starting a dialog with the audience and anticipating their reactions for successful communication.

In the case of science dissemination genres, in which the audience is no longer a specialized one, this task is further challenged by a greater knowledge and interest gap. Therefore, engagement strategies can be assumed to be particularly rich in these genres, and authors are expected to master a wide array of semiotic resources to engage their audiences. In fact, engagement strategies have recently attracted the interest of researchers involved in the study of academic and scientific discourse.

The notion of engagement has been adapted to studies on science communication in various ways to describe how engagement strategies can be realized multimodally, that is, employing a variety of semiotic modes. Carter-Thomas and Rowley-Jolivet (2020), for instance, have investigated engagement strategies in the genre of Three Minute Theses (3MT) presentations. The main strategies they observe in their data are catchy titles, visual impact, various personalization devices, questions, humor and “street cred” (i.e. a common framework based on shared cultural values rather than scientific know-how). In addition, they note that speakers make an effort to present scientific content in the most appealing and stimulating way possible. To that aim, they portray themselves as approachable people who share concerns with the audience, and even exploit their own

personalities, smiling, joking and recounting details of their lives, which is in sharp contrast with traditional academic genres like the lecture. Similarly, Jiang and Qiu (2022) have also explored engagement in 3MTs presentations and the process to communicate disciplinary knowledge to wider audiences, and describe orientation, rationale, purpose, methods and results as the five moves that every presentation must have. Although disciplinary differences are found, these researchers argue that orienting the topic toward the audience's interests, and establishing common ground and connections between specialized knowledge and non-specialized audiences are key engaging strategies.

The study by Xia and Hafner (2021) on the other hand, approaches engagement in TED talks from a multimodal perspective. Their results reveal five strategic multimodal configurations employed in TED talks to achieve viewer engagement, with different modes playing the leading role in each case: visual aids, a long shot, gaze, questions and reference to personal emotions.

Valeiras-Jurado et al. (2018) suggest a taxonomy of persuasive strategies found in a corpus of conference presentations and research dissemination talks. Some of the strategies they identify are very much aligned with the views on engagement offered in the previous paragraphs. For example, with the strategy *anticipation and control of responses* speakers try to predict potential responses from the audience and prompt desirable ones. Speakers also raise and maintain the interest of the audience (i.e. *attention getting*), and highlight parts of the message to make them more salient through *emphasis*.

As the literature review presented in the previous paragraph shows, previous studies on engagement in science dissemination are mostly rooted on linguistic analyses. When non-verbal semiotic resources are considered, language is still the main meaning carrier. In addition, most of these studies are corpus-driven and suggest their own taxonomies or categories without taking into account previous ones. Our study, on the contrary, tries to find areas of agreement between previous taxonomies and looks at science dissemination from a multimodal perspective, considering language just as one element within the multimodal ensembles. Thus, for the purposes of this paper, we combine engagement strategies found in the literature discussed in the previous paragraphs and distil the following taxonomy of multimodal engagement strategies:

- Emphasis: speakers highlight specific parts of the message.
- Attention getting: speakers use devices to obtain and maintain the audience's attention.
- Dialogic involvement: speakers start a fictional (or real) dialog with the audience by directly referring to them or including them within the contents of their speech.
- Humor: speakers use jokes, irony, and other humoristic devices.
- Control of responses: speakers try to predict potential responses and trigger the ones that are in line with their communicative intentions.

The relevance of engagement in science popularization and the multimodal nature of science dissemination genres have been acknowledged in previous literature. However, to our knowledge, no studies have yet adopted a comparative approach to multimodal engagement strategies in science dissemination formats. In order to fill this gap, we combine quantitative and qualitative methods to shed further light into the use of engagement

in two science popularization formats (a face-to-face and an online event): a TEDx talk and a PBS online science video by the same speaker and dealing with similar content.

Background: TEDx talks and PBS online science videos

With the advance of information and communication technologies in the last decades, new online platforms have arisen, thus widening the scope of possibilities for the creation of new genres (Maingueneau, 2010). In fact, genre and discourse analysis have required a paradigm shift to adapt to the analyses of these new media. Shepherd and Watters (1998) state that a continuum may be established to describe the evolution of traditional genres into cybergenres, ranging from copied or replicated genres, through variant genres –adaptations of previously existing genres–, to novel genres, which sprout up directly on the Internet. This paper focuses on two main science dissemination formats: a TEDx talk and a PBS science dissemination video. The former could be classified as *traditional*, and is arguably becoming replicated as it migrates to online platforms. The latter can be considered novel.

TED talks are science dissemination events (Sugimoto and Thelwall, 2013) defined as “a series of short popularizing talks (of approximately 20 minute), addressing a mass audience and delivered by top-level experts in a wide variety of domains” (Caliendo and Compagnone, 2014: 105). Apart from the main TED talks, a more modest group of talks arose to reach wider local audiences; these events are referred to as TEDx talks and are independently organized by universities and other entities. One of these talks will be studied in this paper. TED conferences have become popular in the last decade and have been the center of much research. From a multimodal point of view, TED talks have been described from verbal and non-verbal perspectives to be used as a tool to attract and persuade the audience in Technological English courses (García Pinar and Pallejá, 2018). Similarly, Renuka Devi and Srinivasa (2022), identify content, voice, and facial and bodily expressions as elements contributing to oratory performance, and establish a series of recommendations for science communicators.

As for PBS science dissemination videos, these are short clips (7–12 minutes) in which experts disseminate science in various fields (physics, chemistry, biology, anthropology, etc.). They are produced by PBS (Public Broadcasting Service), an American non-profit entity that produces content for both TV and the Internet. Although this is a rather new format, it could be included within “scifotainment” or “edutainment” genres (Pérez-Llantada, 2021). These types of videos are designed specifically for YouTube, which has become a platform to provide wide visibility to science communication (Allgaier, 2020; Geipel, 2018; Osterrieder, 2013) and fosters the participation of wider audiences for educational purposes (Burgess and Green, 2009; Soukup, 2014). The increasing popularity of the platform has led to a wider offer of contents, which implies a stronger competition to attract and maintain the audiences. In this sense, Khan (2017) reflects on the importance of creating entertaining materials for YouTube that would trigger more interaction with the audience. In fact, engaging the audience right from the start through catchy introductions is essential for the success of YouTube popular science videos (Munoz Morcillo et al., 2016). For these authors, the relationship with the audiences becomes one of the main reasons for success in these videos. In particular, they

establish charisma, a direct view to the camera, and the use of dramatic means as engagement strategies. In another study, Welbourne and Grant (2016) analyze 390 research dissemination YouTube videos and identify engagement strategies such as the type of production, the continuity, the ability of the speakers to communicate authenticity, the creation of medium-length videos, and fast speech rates. In short, YouTube videos are particularly multimodal and are made of complex systems of semiotic resources (Boy et al., 2020). In this type of videos, filmic modes contribute to exploiting the possibilities of the digital medium.

Method

The dataset

To carry out this study, two science dissemination videos by the same scientist were selected: a PBS Space Time video and a TEDx talk. The former is entitled “Why haven’t we found alien life?” and it contains a talk by Dr. Matthew O’Dowd, an astrophysicist who explores the possible reasons why alien life has never been spotted before. It was created directly for YouTube and uploaded on November 5 2015. The latter is entitled “Fermi’s Paradox and the Psychology of Galactic Empires.” Like the previous one, this TEDx talk is given by Dr. O’Dowd and it discusses, to a different extent, the same issue: the reasons why alien life is not evident. This talk took place at the University of Wien on May 26 2017 and its recording was later uploaded to YouTube on August 9 2017. As indicated in the requirements to license TEDx events,¹ speakers should receive broad instructions to prepare their presentations, but unlike in TED talks, there is no briefing program as such. These videos were chosen as part of a larger project to analyze science communication in digital genres (cf. Bernad-Mechó and Girón-García, in press; Valeiras-Jurado and Bernad-Mechó, 2022), and based on the coincidence in speaker and content across the two videos. After an initial viewing, six pairs of shorter excerpts were selected from each of the videos for further comparative analyses, based on the similarities in their content. Table 1 below summarizes the dataset and the contents of each clip.

As argued above, the main reason why these fragments were selected was the similarity of their content (the pairs of clips portray the same speaker talking about similar topics), albeit in two different contexts: the PBS Space Time video was designed to be an online-only video, while the TEDx talk constitutes a live event that was videorecorded and then uploaded online. Although it is evident that both videos are, as of today, solely watched by online viewers (the live TEDx talk was a one-off event and nowadays both videos are available on YouTube), we analyze the TEDx talk as it was originally intended: as a face-to-face event. A second reason for choosing these videos was their popularity in YouTube, which would presuppose engagement. The PBS Space Time video had, at the time this paper was written, over 3 million views and 36k likes; and the TEDx talk had 1.1 million views and over 12k likes. Furthermore, vidIQ (a software that analyzes metrics for YouTube videos) was run on both videos showing a popularity score of 71.5/100 in the case of the PBS video and 60.3/100 in the case of the TEDx talk based on engagement and views relative to other videos in their channels. These scores indicate a

Table 1. Description of the dataset.

Code	Content	Extracted from	Duration	No. of words
PBS_1	Introduction. Why don't we see alien civilizations?	00:03-00:10 and 00:15-00:26	0' 18"	70
TED_1	We don't see alien civilizations.	07:17-00:47	0' 30"	48
PBS_2	The Kepler space observatory. Number of planets.	00:26-00:46	0' 20"	61
TED_2	The Kepler space telescope. Number of planets.	03:57-00:48	0' 52"	95
PBS_3	The Fermi Paradox. The Great Filter. Personal opinion.	00:46-01:22	0' 36"	119
TED_3	The Fermi Paradox. Possible civilizations and implications. Where are we on The Great Filter?	08:38-12:10	3' 32"	427
PBS_4	Humans might be first. Earth as the one instance to analyze.	01:23-02:01	0' 38"	122
TED_4	Humans as a proto-galactic civilization.	12:11-13:09 and 14:12-14:52	1' 38"	195
PBS_5	The origins of life. The late heavy bombardment. Early fossils.	03:22-06:10	2' 48"	489
TED_5	Early fossils. The origin of life.	05:17-07:16	2'	247
PBS_6	Answers to The Great Filter: "Emergence of intelligence," "What if we are early?"	08:09-08:46 and 09:00-10:24	2' 01"	359
TED_6	An answer to The Great Filter: it is behind us. "What if we are early?"	21:15-23:44	2' 30"	274

high engagement rate. Unfortunately, data on viewer retention for the videos could not be obtained as this information may only be accessed by the owners of the YouTube channels. Furthermore, vidIQ indicated a ratio of 54.7 views per hour for the PBS Space Time video and 28.2 for the TEDx talk.

The analysis

After choosing the fragments for analysis, a multimodal annotation was conducted using *Multimodal Analysis Video (MAV)* (O'Halloran et al., 2012), a software that allows for the creation of multiple layers of analysis. For both, the TEDx talk and the PBS Space Time video, and following previous research on the multimodal analysis of oral academic discourse (Bernad-Mechó, 2021), seven modes were selected for their annotation: spoken language, that is, the contents of the speaker's utterances; paralinguistic, which included pauses that were longer than 0.5 seconds, prominent parts of speech (e.g. when the use of a specific intonation or stress would make a fragment particularly salient), and tempo, measured in words per second; gestures, classified as iconic, metaphoric, deictic and beats following McNeill's (1992) taxonomy; direction of gaze; proxemics, or the use of physical space on stage; head movements, identifying nods, shakes and tilts; and facial expression, which described eyebrow raisings, frowns and smiles. These modes are referred to as *embodied* as they were controlled by the speaker's body. On top of that, *filmic* modes were also considered following (Valeiras-Jurado and Bernad-Mechó, 2022) taxonomy, especially in the case of the PBS Space Time video. These modes refer to semiotic resources that come about through the production and editing processes of videos, including type of shot (close-up, medium shot, full shot, etc.); angle (frontal or lateral); *mise-en-scène*, or background; use of cuts; music; visual prompts, either using image or text; sound effects; and visual effects.

Once the annotation of embodied and filmic modes was completed, engagement strategies being used in the clips were also examined (emphasis, attention getting, dialogic involvement, humor, and control of responses). Thus, an extra layer of analysis was included in MAV to code instances of engagement. The annotations of engagement strategies were conducted individually by each of the researchers and agreement was reached for the cases showing discrepancies in their individual analyses. Table 2 summarizes the annotation framework for the multimodal analysis conducted in this paper.

Finally, once the dataset was fully annotated and explored for the use of engagement strategies, two main types of analyses were carried out. First, using the tool *State Machine* in MAV, a quantitative inspection of the use of engagement strategies was run to discern the percentages of use of each strategy over the total duration of each set of clips. A second part of the exploration of the data was directed toward qualitatively detecting how these engagement strategies occurred from a multimodal point of view; in other words, the main focus was to describe the specific multimodal ensembles that were orchestrated to perform each of the engagement techniques. Lastly, both datasets (the TEDx talk clips and the PBS Space Time video ones) were compared to account for any possible similarities and differences in the quantitative and qualitative use of engagement strategies.

Table 2. Framework for the multimodal analysis of engagement strategies.

Engagement strategies	Emphasis	
	Attention getting	
	Dialogic involvement	
	Humor	
	Control of responses	
Embodied modes	Spoken language	
	Paralanguage	Pauses
		Prominences
		Tempo (words/m)
	Gestures	Iconic
		Metaphoric
		Deictic
		Beats
	Gaze	
	Proxemics	Standing
		Pacing
	Head movements	Nod
		Shake
		Tilt
	Facial expression	Eyebrow raising
		Frown
		Smile
Filmic modes	Type of shot	Close-up
		Medium close-up
		Medium shot
		Cowboy shot
		Medium-full shot
		Full shot
		Long shot
	Angle	Frontal
		Lateral
	Mise-en-scène	
	Cuts	
	Music	
	Visual prompts	Image
		Text
	Sound effects	
	Visual effects	

Source: Adapted from Valeiras-Jurado and Bernad-Mechó, 2022.

Results

In this section we provide the results of the quantitative analysis, which sheds light into how much each engagement strategy is used (i.e. how long they are used in relation to the full length of the video), and the qualitative analysis, which illustrates how

Table 3. Average use of engagement strategies.

		TEDx average	PBS average
Engagement strategies	Emphasis (%)	17.26	58.76
	Attention getting (%)	8.06	52.30
	Dialogic involvement (%)	18.60	3.33
	Control of responses	27.7	46.19
	Humor (%)	0.68	11.07

multimodal ensembles are orchestrated to realize these strategies. A discussion of these results is provided in a subsequent section.

Quantitative analysis

A quantitative analysis of the use of engagement strategies was conducted using State Machine in MAV. To do so, the fragments were explored and annotated manually. Table 3 below provides data regarding the presence of engagement strategies in both, PBS Space Time and TEDx videos. The data is shown in percentages of use over the total duration of the clips, that is, it reflects the amount of time in which certain strategies are being used. It is important to remark that overlap may occur among individual strategies (e.g. when a given moment is both emphatic and humoristic). This accounts for the fact that figures for each strategy do not add up to 100%. Moreover, the results show the average of all short clips for each dataset. These results have been computed proportionally; in other words, values have been weighted according to the duration of each of the clips.

As shown in Table 3, all strategies are significantly more common in the PBS Space Time video, except for dialogic involvement. There is 20 times more humor in PBS Space Time than in TEDx, which is to be expected given the higher level of formality in the latter. Interestingly, the use of dialogic involvement is greater in TEDx than in PBS Space Time (18.60% vs 3.33%). This is not surprising if we take into account that there is a live audience in TEDx, and that interacting with online audiences is challenging for online speakers (Querol-Julián, 2023). An equally interesting finding is that there is significantly more attention getting in PBS Space Time (52.30%) than in TEDx (8.06%). This can be accounted for by the need to retain the online audience until the end of the video (Munoz Morcillo et al., 2016), a pressure that is inexistent in the live TEDx talk. The greater presence of emphasis in PBS is also in line with this trend. In fact, emphasis is the most frequent strategy in PBS, while in the TEDx talk the speaker resorts to control of responses more frequently. This suggests that while the speaker in PBS tries to retain viewer attention, the same speaker in TEDx tries to monitor the audience that is physically present and obtain a desirable response from them.

Qualitative analysis

A qualitative analysis of four selected excerpts is provided in this section to illustrate the combined use of engagement strategies and the multimodal ensembles that make them





1		<p>Why don't we see alien civilizations? We've asked this question before here on Space Time. And if you haven't seen that episode, you should check it out.</p>
2		<p>But today (PAUSE), we'd like to go deeper, because</p>
3		<p>it really does seem like</p>
4		<p>there SHOULD, BE, ALIENS.</p>

Figure 1. Multimodal analysis of Example 1.

possible. The excerpts were chosen because the quantitative analysis revealed them as particularly rich in engagement strategies. In addition, the excerpts illustrate the five engagement strategies distilled from previous literature in each of the formats. The first two examples correspond to the PBS Space Time video. Examples 3 and 4 focus on the TEDx talk.

Example 1

Figure 1 below describes a section from the introduction to the PBS Space Time video. In it, the speaker introduces the main topic for the video, that is, the possible reasons why humanity has never seen aliens. As argued by Saurabh and Gautam (2019), the introductions to YouTube videos need to be as engaging as possible to achieve audience retention. In the multimodal analysis of this excerpt, engagement is achieved through the use

of four strategies: emphasis, attention getting, dialogic involvement and control of responses.

The excerpt begins with the speaker putting forward a question to the audience and referring to a previous episode (Image 1). These utterances entail examples of dialogic involvement in which the speaker addresses the audience explicitly offering advice (“if you haven’t seen that episode, you should check it out”) or including himself as part of the audience through the use of the plural first-person pronoun (“Why don’t we see alien civilizations?”). This excerpt is also rich in the use of modes. For instance, the speaker performs a deictic gesture with both hands as he points to the hyperlink connecting to the previous episode. On top of that, the use of visual prompts that enter and exit the screen contributes to getting the attention of the audience and emphasizing the message.

After referring to the previous episode, the speaker establishes the differences between that episode and the present one. He suggests that there is something odd about the fact that aliens have never been spotted (Images 2, 3, and 4). In this section, the speaker uses facial expression to show disbelief and surprise. He does this by averting his gaze to simulate he is reflecting (Image 2) and then gazing back to the camera as he frowns (Images 3 and 4). This surprise is also shown through the use of prominence as he utters the words “should be aliens” marking them as he speaks. So far, the speaker is using the strategy of control of responses in the multimodal ensemble to convey genuine surprise and to create intrigue, which will contribute to retaining viewers who might watch the full video. As for the use of filmic modes, apart from the thumbnail referring to the previous episode, visual prompts and visual and sound effects are employed as they co-occur with the verbal message (e.g. “there should be aliens”) to emphasize and get the attention of the viewers. Finally, the choice of a medium shot allows the viewer to notice both the speaker’s use of embodied modes and the visuals utilized. All in all, the speaker creates a fictional dialog with the viewer and controls responses by presenting content in an intriguing manner. At the same time, the message is emphasized through the use of visual prompts. All this is done coherently as an attempt to attract and retain attention.

Example 2

The second example from PBS Space Time corresponds to an excerpt in which three engagement strategies are being used: emphasis, humor, and control of responses. Although, as shown in Table 3, humor is less frequent in terms of percentages of use, qualitatively, it seems to play a more central role: humor is very explicit for the audience within the development of the explanations, as it interrupts more formal strains of content. In particular, humor is introduced in this excerpt as the speaker presents a series of possibilities for what “The Great Filter” could be. The Great Filter is described as an event or a point in the development of any civilization that prevents their expansion to the rest of the galaxy. To exemplify this, the speaker explores three possibilities: a nuclear war, an environmental catastrophe, and the accidental creation of a black hole that would consume the planet (see Figure 2).

The latter example is, clearly, far-fetched and is added as an engaging humoristic device. This is shown both in the embodied and filmic modes in the sequence together with the rest of engagement strategies. First, the speaker presents two possibilities for the

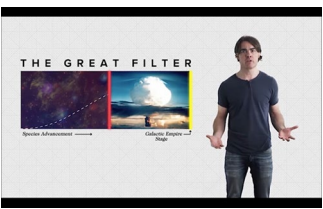
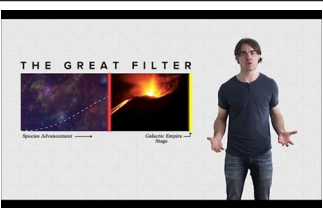
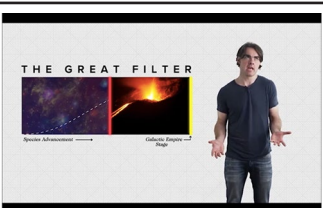
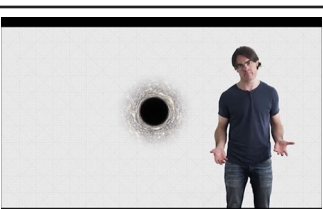
1		[...] whether by nuclear war,
2		environmental catastrophe,
3		(PAUSE) accidentally making a black hole that swallows the planet,
4		(PAUSE) etcetera.

Figure 2. Multimodal analysis of Example 2.

Great Filter that seem to be understood as plausible: a nuclear war and an environmental catastrophe. This is shown through a serious facial expression and a direct gaze to the camera (Images 1 and 2). At the same time, visual prompts of a nuclear explosion and the eruption of a volcano are added as filmic modes emphasizing the verbal message. Next, the third possibility is introduced. At this point, the speaker's expression shifts as he gazes upwards and slightly tilts his head (Image 3), portraying a facial expression of disbelief. Furthermore, his verbal utterance (“accidentally making a black hole that swallows the planet”) is both preceded and followed by short pauses. The use of pauses has been demonstrated to be used to encapsulate humoristic expressions (Bernad-Mechó and Girón-García, in press; Fortanet-Gómez and Ruiz-Madrid, 2016; Kyratzis, 2003). Finally, in Image 4, the speaker looks back at the camera, frowns, and tilts his head to the other side as he sentences “etcetera”. Within filmic modes, a visual effect of a swirling

black hole dissolves the previous visual content. After this, there is a cut and the speaker moves on to a more formal tone. By looking back to the audience and changing his facial expression, in combination with the visual effects that put an end to this short sequence, the speaker appears to be indirectly reassuring the audience that the last example is, indeed, an implausible one. All in all, Images 3 and 4 show the use of humor as well as control of responses.



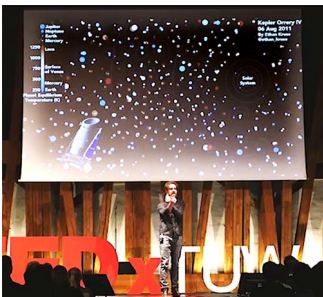

Example 3

In the following excerpt from the TEDx talk the speaker is discussing advances in our knowledge of the universe. He is arguing that the more we know about the universe, the more likely it seems that there should be aliens, and the more puzzling the fact that we cannot see them (Figure 3). As he speaks he is standing and mostly gazing at the audience. The projection behind him shows images from the galaxy captured by the Kepler telescope (Images 1 and 2). He is also gesturing throughout the excerpt. In addition to beats that mark the rhythm of his discourse he also makes an iconic gesture that represents the orbit of a planet around its star (Images 3 and 4), and two metaphoric gestures, one moving both hands outwards with palms up, as if offering to the audience the knowledge that we now have validated about our universe (Image 2), and another one with both hands making a circular movement that represents the scope of “every star” (Image 5). A head nod also reinforces the extreme quantification “almost all” (Figure 1), and can be interpreted as preventing a potential counterargument (as if saying “believe me, really almost all stars have orbiting planets”). Simultaneously, prominence in key words like “planetary systems,” “discovered” and “every star” highlight important aspects of his message. The use of inclusive “we” is also noticeable throughout the excerpt. With this multimodal ensemble the speaker achieves the following engagement strategies:

- i. he attracts the attention of the audience and emphasizes relevant parts of the talk (strategies attention getting and emphasis),
- ii. he builds dialogic involvement and engages the audience in a fictional dialog (strategy dialogic involvement),
- iii. he prompts desirable responses from the audience by presenting the content he is presenting as well-established knowledge, not questionable (strategy anticipation and control of responses).

Example 4

In the next example from the TED Talk the speaker is explaining what is special about humans that makes us so close to meet other alien civilizations (Figure 4). In the background, the screen shows a time scale and a question mark. The speaker alternates standing and pacing, and keeps eye contact with the audience in the room. His register becomes less formal in this excerpt, and includes slang like “cocky.” He also addresses the audience directly using an imperative (“don’t get cocky”). Gestures, face expression and head movements also play an important role in this excerpt. In addition to beats that once more mark the rhythm of discourse, there is a metaphoric gesture with palms down that

1		<p>Those stars, almost all have planetary systems</p>
2		<p>as discovered by the Kepler space telescope</p>
3		<p>watching the miniscule dimming of light</p>
4		<p>as distant alien worlds pass in front of their parents star</p>

(Continued)

Figure 3. (Continued)



5		we've now determined that almost every star in the Milky Way
6		has a planetary system

Figure 3. Multimodal analysis of Example 3.

seems to call for calmness as he says “don’t get cocky” (Image 1), and an iconic gesture as if discarding something is synchronous with the expression “way back” (Image 2), which is also made prominent through intonation. “Special,” in “something special about humanity” is accompanied by a reaffirming nod that prevents counterclaims, and a shake also adds to the uncertainty introduced by “maybe” in “maybe that chance fusion” (Image 3). This hedging is further reinforced by the eyebrow raising conveying surprise that is synchronous with “chance” and an open hand palms up gesture (Image 4) synchronous with the estimation (“one in trillions”) that seems to present it as an educated guess. This multimodal ensemble has the effect of getting the attention of the audience (strategy attention getting) and engaging in a dialog with them (strategy dialogic involvement). The speaker also introduces a humorous pun (strategy humor), and through this humor and the hedging devices he presents the idea that “there’s something special about humanity” as something agreed upon, in contrast to the potential explanations, which are presented as tentative (strategy anticipation and control of responses).

Discussion

In this section we further discuss the results of our quantitative and qualitative analysis so as to provide an answer to our research question.



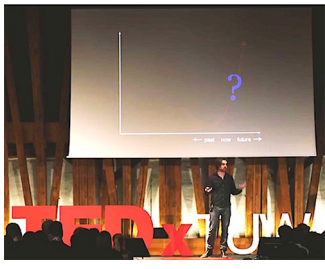
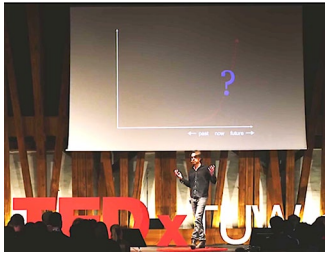
1		<p>there's something special about humanity, don't get cocky</p>
2		<p>it could have been way back in the evolutionary tree</p>
3		<p>maybe that chance fusion of proto eukaryote proto mitochondria to forge the modern complex cell</p>
4		<p>(PAUSE) etcetera.</p>

Figure 4. Multimodal analysis of Example 4.

- *What similarities and differences can be spotted across the two science dissemination formats under scrutiny in terms of multimodal engagement?*

Regarding similarities, the two events feature instances of the five engagement strategies suggested in previous literature. In addition, we have found that their realization goes beyond the merely linguistic, and involves a variety of modes. So far, previous studies have approached engagement in science dissemination applying taxonomies that are rooted in linguistic-based discourse analyses. Even if some of these studies acknowledge the role of other semiotic modes, they still attach a dominant role to language. Alternatively, when engagement is approached from a multimodal perspective, a corpus-driven approach tends to be adopted and existing taxonomies are not considered. We believe that our results show areas of convergence between existing classifications of engagement techniques, which is a step forward in a systematic study of engagement. In addition, they reveal how it is ultimately the effect of a well-orchestrated multimodal ensemble that makes these techniques effective.

The examples analyzed show that the strategy “Control of responses” relies heavily on embodied modes regardless of the format (online video or live TEDx talk), and is realized through intonation, face expression, gaze, head movements and gestures. Likewise, humor seems to be very dependent on spoken language in both formats, and the other modes contributing to it are also embodied (intonation, head movements, face expression and gaze).

On the other hand, remarkable differences across formats can be spotted in the use of modes. In the case of the PBS video, since it undergoes an editing process, the multimodal ensemble includes filmic modes such as cuts, visual effects, sound effects and visual prompts, which are inexistent in the live TEDx talk and used in the video version of the talk only to a very limited extent. Another difference is the use of pacing, which is frequent in TEDx and inexistent in PBS. We believe it can contribute to engagement in the case of the live TEDx event, because it helps the speaker to include the whole audience and not just one sector.

In addition, although the two formats use the five engagement strategies in our taxonomy, there are interesting differences in their use. Our quantitative findings show that, taking into account all engagement strategies together, the PBS Space Time video features considerably more engagement than the TEDx video. The PBS video relies on emphasis and attention getting, while the TEDx talk shows a preference for control of responses and dialogic involvement. We believe this is a result of the need to adapt to the two diverse audiences: an online audience that can freely stop watching the video at any time versus a physically present audience that the speaker can monitor as he is able to see their reactions.

In addition, the strategies attention getting and emphasis are realized through gestures and paralinguistic in TEDx, while PBS Space Time uses filmic modes such as visual prompts, visual effects, sound effects and cuts to that aim – a trend previously noted by Munoz Morcillo et al. (2016), and which is facilitated by the affordances of the online media.

In short, the quantitative and qualitative results presented in this study show that the speaker in the TEDx talk relies on embodied modes such as gestures and paralinguistic

to engage the audience as the physical presence of an audience enables the speaker to actually see their reactions. This provides immediate feedback, something that is not possible in the online setting of a PBS video, and is therefore compensated by a more extensive use of filmic modes and a clear focus on retaining the audience's attention.

Conclusions

In this paper we have presented a multimodal discourse analysis of engagement strategies in two science dissemination formats: TEDx talks and PBS science dissemination videos. Our results further bring to the fore the multimodal nature of these formats, as suggested in previous literature (Valeiras-Jurado and Bernad-Mechó, 2022; Luzón, 2019; Xia and Hafner, 2021). The study has also shed light on the complex multimodal ensembles that are orchestrated in them, involving embodied and filmic modes, in order to engage their audiences. Our findings highlight areas of agreement between previous classifications of engagement strategies or techniques, together with their potential modal realizations, which we believe is a step forward in a comprehensive, systematic analysis of engagement in science dissemination.

In addition, both similarities and differences across the formats analyzed have emerged regarding their use of engagement. Mainly, these can be summarized as follows: (i) the strategies dialogic involvement and humor are realized through embodied modes in both events; (ii) filmic modes are used more extensively in PBS, especially for emphasis and attention getting; (iii) emphasis and attention getting are used more extensively in PBS and (iv) control of responses and dialogic involvement become more relevant in TEDx.

These differences can be accounted for by two crucial aspects: the presence of a physical audience and the importance attached to montage and editing. Regarding the former, our results suggest that engagement is more effective when there is one clear audience in mind. They also reveal the importance of constantly monitoring the audience. Speakers in TEDx talks can exploit the shared physical space to this aim; online presenters can rely on filmic modes to obtain and retain attention. In other words, what contributes to engagement in a live event might not work for the online audience and vice versa.

Concerning the latter aspect, that is, editing, we have found a correlation between the importance attached to it and the extent and mastery to which filmic modes are used. In the case of the live TEDx talk, there is no editing as such, but rather prior preparation of the event. Except for slides projected on a screen, the speaker relies on his body and the space provided by the stage to engage with the audience. The PBS video, on the other hand, exploits filmic modes to the fullest to enact engagement strategies such as emphasis, or attention getting.

It is important to note that filmic modes are not controlled by the speaker. This means that the PBS video is co-authored between the speaker and the editing team. We believe our results show the importance of a smooth coordination between both so as not to jeopardize engagement.

The case study discussed in this paper presents certain limitations. One of them is the lack of focus on the audience's reaction, which is, in a way, looking at just one side of the coin. Further studies could move beyond speaker's intent and investigate the audience's

uptake. In addition, these types of studies could benefit from the insight of ethnographic interviews with speakers and editors. Finally, we also need to acknowledge the limited generalizability of our results, since the case study presented in this paper does not allow for establishing traits. Further studies comparing more videos from the same speaker, videos from different speakers, or delving deeper into qualitative variation with focus on individual strategies would yield promising findings that can help us better understand the complex ways in which scientists manage to bring science closer to the general public and turn scientific content into something accessible, engaging and enjoyable.

Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This study is encompassed within a research project funded by the Ministerio de Ciencia, Innovación y Universidades, Spain (Ref: PID2021-127827NB-I00) and entitled *Cybergenres and English as a Medium of Instruction. Multimodal analysis of digital academic genres and their pedagogical implications in Higher Education EMI contexts*; and a research project funded by Generalitat Valenciana – Conselleria d’Innovació, Universitats, Ciència i Societat Digital, Spain (Ref: CIAICO/2021/069) entitled *Análisis y descripción de la naturaleza multimodal y multimedial de los géneros académicos en la educación superior. La comunicación académica digital en inglés*.

Note

1. <https://www.ted.com/participate/organize-a-local-tedx-event/tedx-organizer-guide/speakers-program/prepare-your-speaker>

References

- Allgaier J (2020) Science and medicine on YouTube. In: Husinger J and Klasturp MML, (eds) *Second International Handbook of Internet Research*. Berlin: Springer, pp.7–27.
- 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.
- Bernad-Mechó E and Girón-García C (in press) A multimodal analysis of humour as an engagement strategy in YouTube research dissemination videos. *The European Journal of Humor Research*.
- Bondi M, Cacchiani S and Mazzi D (2015) *Discourse in and through the media: Recontextualizing and reconceptualizing expert discourse*. Newcastle upon Tyne: Cambridge Scholars Publishing.
- Boy B, Bucher H-J and Christ K (2020) Audiovisual science communication on TV and youtube. How recipients understand and evaluate science videos. *Frontiers in Communication* 5: 1–18.
- Burgess J and Green J (2009) *YouTube: Online Video and Participatory Culture*. New York, NY: Wiley.
- Caliendo G (2014) The popularisation of science in Web-Based genres. In: Caliendo G and Bongo G (eds) *The Language of Popularisation: Theoretical and Descriptive Models*. Bern: Peter Lang, pp.101–132.

- Caliendo G and Compagnone A (2014) Expressing epistemic stance in university lectures and TED talks: A contrastive corpus-based analysis. *Lingue e Linguaggi* 11: 105–122.
- Calsamiglia H and Van Dijk TA (2004) Popularization discourse and knowledge about the genome. *Discourse & Society* 15(4): 369–389.
- Carter-Thomas S and Rowley-Jolivet E (2020) Three minute thesis presentations: Recontextualisation strategies in doctoral research. *Journal of English for Academic Purposes* 48: 1–14.
- Fortanet-Gómez I and Ruiz-Madrid N (2016) Multimodal humor in plenary lectures in English and in Spanish. *Multimodal Communication* 5(1): 55–69.
- García Pinar A and Pallejá C (2018) TED talks: A multimodal tool for students of technological English. *Docencia e Investigación* 29: 6–24.
- Geipel A (2018) Wissenschaft@YouTube. Plattformspezifische Formen von Wissenschaftskommunikation. In: Lettkemann E, Wilke R and Knoblauch H (eds) *Knowledge in Action. Neue Formen der Kommunikation in der Wissensgesellschaft*. Berlin: Springer, pp.137–163.
- Hyland K (2010) Constructing proximity: Relating to readers in popular and professional science. *Journal of English for Academic Purposes* 9: 116–127.
- Jiang F and Qiu X (2022) Communicating disciplinary knowledge to a wide audience in 3MT presentations: How students engage with popularization of science. *Discourse Studies* 24(1): 115–134.
- Khan ML (2017) Social media engagement: What motivates user participation and consumption on YouTube? *Computers in Human Behavior* 66: 236–247.
- Kyrtzidis S (2003) Laughing metaphorically: metaphor and humor in discourse. In: *8th international cognitive linguistics conference*, University of La Rioja, Spain, 20–25 July. Theme Session: Cognitive-Linguistic Approaches to Humor. <https://doi.org/10.1.1.132.9689>
- Loghmani Z, Ghonsooly B and Ghazanfari M (2020) Engagement in doctoral dissertation discussion sections written by English native speakers. *Journal of English for Academic Purposes* 45(1): 100851.
- Luzón MJ (2019) Bridging the gap between experts and publics: The role of multimodality in disseminating research in online videos. *Ibérica* 37: 167–192.
- Maingueneau D (2010) Types of genres, hypergenre and internet. In: Luzón MJ, Ruiz-Madrid ML and Villanueva (eds) *Digital Genres, New Literacies and Autonomy in Language Learning*. Newcastle upon Tyne: Cambridge Scholars Publishing, pp.23–42.
- Martin JR and White P (2005) *The Language of Evaluation: Appraisal in English*. London: Palgrave Macmillan.
- McNeill D (1992) *Hand and Mind: What Gestures Reveal About Thought*. Chicago, IL: The University of Chicago Press.
- Munoz Morcillo J, Czurda K and Robertson-von Trotha CY (2016) Typologies of the popular science web video. *Journal of Communication Science* 15(04): 1–32.
- Osterrieder A (2013) The value and use of social media as communication tool in the plant sciences. *Plant Methods* 9: 26.
- O'Halloran KL, Podlasov A, Chua A, et al. (2012) Interactive software for multimodal analysis. *Visual Communication* 11(3): 363–381.
- Pérez-Llantada C (2021) Genres and languages in science communication: The multiple dimensions of the science-policy interface. *Language & Communication* 78: 65–76.
- Querol-Julián M (2023) Multimodal interaction in English-medium instruction: How does a lecturer promote and enhance students' participation in a live online lecture? *Journal of English for Academic Purposes* 61: 101207.
- Renuka Devi MN and Srinivasa G (2022) Multimodal analysis of TED talks toward designing a recommendation system. In R. P. Mahapatra, S. K. Peddoju, S. Roy, P. Parwekar, & L. Goel

- (Eds.), *Proceedings of International Conference on Recent Trends in Computing. Lecture Notes in Networks and Systems*, vol 341. Berlin: Springer, pp.455–467.
- Ruiz-Madrid MN (2021) A multimodal discourse approach to research pitches. *Journal of English for Academic Purposes* 52: 101003.
- Saurabh S and Gautam S (2019) Modelling and statistical analysis of YouTube's educational videos: A channel owner's perspective. *Computers & Education* 128: 145–158.
- Scotto di Carlo G (2014) Ethos in TED talks: The role of credibility in popularised texts. *Facta Universitatis. Series: Linguistics and Literature* 12(2): 81–91.
- Scotto di Carlo G (2015) Stance in TED talks: Strategic use of subjective adjectives in online popularisation. *Ibérica* 29: 201–221.
- Shepherd M and Watters C (1998) The evolution of cybergenres. In: *Proceeding HICSS '98. Proceedings of the thirty-first annual Hawaii international conference on system sciences*, vol 2, The Institute of Electrical and Electronics Engineers, Inc., Los Alamitos, CA, USA, 6–9 January 1998, pp.97–109.
- Soukup PA (2014) Looking at, with, and through YouTube. *Communication Research Trends* 33(3): 3–34.
- Sugimoto CR and Thelwall M (2013) Scholars on soap boxes: Science communication and dissemination in TED videos. *Journal of the American Society for Information Science and Technology* 64(4): 663–674.
- Valeiras-Jurado J and 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: 101118.
- Valeiras-Jurado J, Ruiz-Madrid N and 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.
- Welbourne DJ and Grant WJ (2016) Science communication on YouTube: Factors that affect channel and video popularity. *Public Understanding of Science* 25(6): 706–718.
- Xia SA and Hafner CA (2021) Engaging the online audience in the digital era: A multimodal analysis of engagement strategies in TED talk videos. *Ibérica. Revista de La Asociación Europea de Lenguas Para Fines Específicos (AELFE)*, 42, 33–58.
- Zou H and Hyland K (2021) A tale of two genres: Engaging audiences in academic blogs and three minute thesis presentations. *Australian Journal of Linguistics* 41(2): 131–151.

Author biographies

Dr. Edgar Bernad-Mechó is Assistant Professor at Universitat Jaume I, Spain, in the Department of English Studies. He specializes in the use of metadiscourse and multimodality in academic contexts and for science dissemination. He has recently published a paper in the *Journal of English for Academic Purposes* looking at modal density and coherence of YouTube videos for science dissemination. He has co-edited a special issue on multimodality for *Ibérica* and a volume on English as a Medium of Instruction and multimodality for Routledge that is expected to be released in 2023.

Dr. Julia Valeiras-Jurado is Assistant Professor at Universitat Jaume I, Spain, Department of English Studies. Her main research interests are Multimodal Discourse Analysis (MDA) and English for Specific and Academic Purposes (ESP/EAP). In particular, she has published MDA studies on academic and business discourse, focusing on the multimodal nature of academic, scientific and business genres, and on how they achieve persuasion and engagement multimodally. Her publications have appeared in *Iberica*, *Text&Talk*, *International Journal of English Studies*, *Language and Communication*, *International Journal of Applied Linguistics*, *Discourse Studies*, and *Journal of English for Academic Purposes*.