RAEL Revista Electrónica de Lingüística Aplicada

RAEL: Revista Electrónica de Lingüística Aplicada
Vol./Núm.: 24/1
Enero-diciembre 2025
Páginas: 58-75
Artículo recibido: 13/08/2024
Artículo aceptado: 31/03/2025
Artículo publicado: 31/12/2025
Url: https://rael.aesla.org.es/index.php/RAEL/article/view/664
DOI: https://doi.org/10.58859/rael.v24i1.664

Communicating Climate Change: A Multimodal Comparative Study on Long and Short YouTube Videos

Comunicando el cambio climático: un estudio multimodal y comparativo entre vídeos largos y cortos (Shorts) de YouTube

Leticia Arrametapongsa-Brines
Universitat Jaume I

YouTube shorts have become popular since being launched in 2021 and climate change-oriented channels have subsequently made use of this setup to propagate their message. Three popular short videos concerning climate change alongside three long (regular) videos from the same science disseminators' channel were selected to conduct a comparative study on how climate change is communicated in the two YouTube formats. The quantitative analysis focused on the videos' modal density (Valeiras-Jurado & Bernad-Mechó, 2022), whereas modal coherence (Valeiras-Jurado, 2019) was analysed qualitatively. Findings showed that the shorts analysed exhibited higher modal density and greater modal coherence, reflecting the needs of their potentially distinct audiences. Given that many YouTube videos about climate change promote viewpoints that contradict the scientific consensus (Allgaier, 2019), understanding and sharing how effective climate change communication works across these different formats could help counter this trend.

Keywords: *multimodality; climate change dissemination; shorts; modal density; coherence.*

Desde su lanzamiento en 2021, los YouTube Shorts han ganado popularidad, y canales sobre el cambio climático han empezado a utilizarlos para difundir su mensaje. En este estudio, se seleccionaron tres vídeos cortos (shorts) y tres largos (normales) relacionados con el cambio climático de distintos canales de divulgación científica para realizar un análisis comparativo. El análisis cuantitativo evaluó la densidad modal (Valeiras-Jurado y Bernad-Mechó, 2022) mientras que el cualitativo se centró en la coherencia modal (Valeiras-Jurado, 2019). Los resultados revelaron que los vídeos cortos presentaban una mayor densidad y coherencia modal, potencialmente en consonancia con las necesidades de su público. Dado que la mayoría de los vídeos de YouTube sobre cambio climático promueven puntos de vista opuestos al criterio científico (Allgaier, 2019), examinar y difundir cómo se comunica eficazmente la realidad del cambio climático a través de estos formatos podría ayudar a revertir esta tendencia.

Palabras clave: *multimodalidad; divulgación del cambio climático; shorts; densidad modal; coherencia.*

Citar como: Arrametapongsa-Brines, L. Communicating Climate Change: A Multimodal Comparative Study on Long and Short YouTube Videos. *RAEL: Revista Electrónica de Lingüística Aplicada, 24*, 58-75. https://doi.org/10.58859/rael.v24i1.664

1. INTRODUCCIÓN

Be it in the form of a YouTube video or as another digital media product, the climate crisis has its singularities that set it apart from other online science dissemination topics. For one, these videos would encompass more than simply being *edutainment*, understood as the purposeful combination of educational and entertaining content to enhance learning (Aksakal, 2015), and would present differences in their underlying goals, such as promoting awareness or encouraging climate action. Secondly, due to its urgent condition, the need for climate communication to be understood across all types of audiences (such as specialist and non-specialist viewers) is a greatly pertinent matter.

The UN has acknowledged that "communicating on climate change is about educating and mobilising audiences to take action to confront the climate crisis" (United Nations, 2024). They refer to videos, podcasts, written articles, or graphics as "communication products" and encourage advocates "to make it a valuable, effective and reliable piece of content" (United Nations, 2024). Advice on storytelling, empowering people and focusing on solutions is given on their website. Many of these so-called *communication products* are YouTube videos created by scientists and communicators alike, given that this digital platform is one of the main drivers of science dissemination on the internet (Brennan, 2021).

However, concerns have been raised about the platform's effectiveness in supporting evidence-based climate communication. In their study on public responses to climate science on YouTube, Shapiro and Park (2015: 116) warned that "given polarised views on the science of climate change among the public, the likelihood of politicised science and the public's increased reliance on the Internet as a source of information, there is an urgent need to pay attention to these issues". More recently, Olausson and Berglez (2023) observed that digital platforms like YouTube function as "amplifiers of misinformation, especially when messages appeal to identity and grievance over facts" (p. 56). Similarly, Allgaier (2019) conducted an empirical analysis of climate change videos on YouTube and found that the majority of content promoted viewpoints opposing the scientific consensus.

As scholars have noted, climate change communication faces persistent challenges and remains an area that warrants further exploration, which may include examining how successful scientifically grounded messages are crafted and shared across the YouTube platform.

1.1. *Online science dissemination and YouTube*

In recent years, there has been a growing trend of science transcending the boundaries of academic institutions (Valeiras-Jurado & Bernad-Mechó, 2022). Online communication allows scientists and non-scientific intermediaries who act as experts to actively participate in the communication process between science and the public (Lobato, 2016; Welbourne & Grant, 2016). YouTube, in particular, has become a growing source of content for research dissemination, with some scholars describing it as "a powerful tool to communicate science and technology to the general public" (León & Bourk, 2020: 1).

Videos on YouTube have the potential to engage and motivate learners, especially digital natives who prefer individualised and online learning experiences. Studies revolving around science communication in YouTube videos have examined participatory aspects, coverage of controversial issues, user comments, motivations for watching science videos and differences between user-generated and professionally generated content (Shapiro & Park, 2015; Welbourne & Grant, 2016; De Lara et al., 2017; Erviti et al., 2020). Some authors have attributed the success of knowledge dissemination to speakers' presenting science as a shared heritage

belonging to the entire community, rather than something distant and detached (Scotto di Carlo, 2015). Moreover, YouTube science videos have also been examined through their use of genre conventions (Muñoz Morcillo et al., 2016; Boy et al., 2020; Huang & Grant, 2020).

While these studies offer valuable insights, they focus primarily on general science communication and do not address the specific challenges of climate change. They also overlook short-form video formats, as the YouTube platform had not yet introduced this style of content at the time they were conducted.

1.2. *YouTube Shorts and Climate Communication*

YouTube Shorts is a new video format offered by this platform launched globally in 2021. They are designed to be up to a minute long and to be filmed in a vertical orientation, given they were conceived to be consumed on mobile phones. Users can engage with shorts through comments, likes and shares, as with regular YouTube videos. Moreover, YouTube's Shorts section often functions as a discovery mechanism, helping creators attract new viewers and convert them into subscribers for their longer-form content. For this reason, effectively capturing the viewer's attention becomes crucial. These short, instantaneous videos form an integral part of a creator's channel and remain accessible at any time through the Shorts category on the channel page. Chi and Park (2022) examined the characteristics of popular short-form videos and the respective roles of producers and viewers. Their findings indicate that the virality of long videos tends to depend primarily on their content, whereas the virality of short videos is driven more by community engagement. They further suggest that video language in shorts should be simple and intuitive, reflecting viewers' preference for easily consumable content.

Due to their scrolling nature, shorts may reach a wider and more varied audience than regular YouTube videos. As reported in the technology-news outlet TechCrunch, YouTube's product management lead for YouTube Shorts, Todd Sherman, stated that the "Shorts feed prioritises a more diverse feed because people are flipping through hundreds of videos versus maybe 10 or 20 in long-form" (Perez, 2023, para. 3). As Rajendran et al. (2024) observe, the expanded feed and scroll-based interface of short-form videos increases the pressure on creators to immediately capture viewers' attention. Even a brief lapse in engagement can prompt users to swipe away within seconds, a dynamic that helps explain the fast pace and intensity often seen in YouTube Shorts and that tends to heavily rely on edition.

Climate advocates have also engaged with the emerging shorts format. Consequently, this medium constitutes a relevant object of study for examining how climate-change communicators seek to convey their messages on YouTube. Given the inherently audiovisual and interactive character of YouTube videos, conventional linguistic analysis alone is insufficient to account for their full communicative complexity. As Jewitt and Price note (2012: 1), "multimodality offers a valuable approach for analysing video data, as it systematically attends to the interpretation of a wide range of communicational forms (e.g. gaze, posture, action, speech)". A multimodal analytical framework is therefore necessary and will be outlined in the following section.

2. THEORETICAL FRAMEWORK AND RESEARCH QUESTIONS

Multimodal Discourse Analysis (MDA) focuses on examining semiotic resources as structured systems of choices available to communicators, analysing how these choices combine to produce meaning. This approach considers language one of multiple legitimate

communicative modes. Modes can be conceived as semiotic resources. Van Leeuwen (2005: 285) defines these as:

> the actions, materials and artifacts we use for communicative purposes, whether produced physiologically—for example, with our vocal apparatus, the muscles we use to make facial expressions and gestures—or technologically—for example, with pen and ink, or computer hardware and software—together with the ways in which these resources can be organized.

The designed complex of different modes, referred to by Kress (2010) as a multimodal ensemble, may be defined as the purposeful integration of modes (e.g., speech, visuals, editing) to create cohesive communication. In the realm of science dissemination, YouTube science content has been regarded as a well-organised multimodal arrangement (Boy et al., 2020), and the incorporation of multimodal ensembles recognised as a crucial aspect of effective communication (Valeiras-Jurado & Bernad-Mechó, 2022).

Some previous studies have dealt with analysing YouTube science dissemination videos through the use of modes. Words, paralanguage, and kinesics have been investigated by different authors (Querol-Julián, 2011; Fortanet-Gómez & Ruiz-Madrid, 2014; Ruiz-Madrid & Fortanet-Gómez, 2019). Paralanguage and kinesics fall under what Norris (2004) defines as embodied modes—semiotic resources that involve the physical body in the process of communication and meaning-making. These include gestures, facial expressions, eye gaze, posture, and body movement, all of which contribute to how meaning is constructed and interpreted in interaction.

Filmic modes, by contrast, refer to the layer of meaning introduced during the video's editing process (Valeiras-Jurado & Bernad-Mechó, 2022). In their study, the authors also develop the concept of modal density, defined as the intensity of modal use, measured by the frequency of mode occurrences per minute. Additionally, they examine modal coherence, which, according to Valeiras-Jurado (2019), describes the consistent alignment of modes to prevent contradictions and reinforce the intended message, such as synchronising gestures with visual prompts to create a unified communicative effect.

In the context of climate change communication, multimodal analysis has mainly focused on the use of visual metaphors. Schäfer and Yan (2023) examined the literature related to climate change media imagery, noting that while most studies emphasise the strong influence of visual content on audiences, certain imagery can also be problematic. Notably, one of the major shortcomings they identified was the limited scope of existing multimodal analyses, which tend to focus solely on imagery while overlooking the integration of written text, sound and particularly video content.

To begin addressing this gap, the present study conducts a comparative analysis of modal density and modal coherence in six popular climate change science dissemination videos on YouTube, comprising three long-form videos and their corresponding YouTube Shorts version. The objective is to explore how multimodal strategies function across these formats by identifying patterns in the deployment and interaction of all available meaning-making resources, with the aim of modestly contributing to a better understanding of effective climate change communication. To guide this exploration, the study poses the following research questions:

1) What are the similarities and differences in terms of modal density between short and long climate change dissemination videos?
2) Do the videos present modal coherence? To what extent and in what ways are the videos coherent?

3. METHOD

3.1. *Dataset*

This research aims to present a comparative analysis of successful short and long-term climate change videos to potentially bring forth the features that may be contributing to their effectiveness. In order to find experts and successful presenters on the matter, the words *scientist*, *climate* and *shorts* were typed into YouTube's search engine without logging into any account to counteract personal recommendations. To carry out the comparative analysis, the first three shorts with a presenter that had a long video tackling a similar topic in their YouTube channel were selected from the first results provided by the algorithm (which shows the most popular results), thus creating the three video pairs presented in Table 1. As shown in the table, only some excerpts from the long videos that correlated with the shorts video were analysed to obtain comparable data in the quantitative analysis. Henceforth, code video titles shown in Table 1 will be used to refer to each video.

Analysing specific sections of the long video related to the shorts instead of looking at the long video's full length was believed to offer more comparable results, as focusing on similar segments (for example, the explanation of the same phenomenon) in the long and short version would better reflect the differences in the communication style of the presenter in relation to each format, eliminating the interference of other factors that may have shaped the outcome. For instance, LV3 showcased a significantly long interview and a documentary-like, amateurish filming of laboratory installations, elements that are very rare in the channel and do not represent the general approach of this specific content creator. Had the video been analysed in its entirety, the results would not have reflected the communication style of the presenter nor would they have represented the average display of filmic modes of the channel. Thus, only some equivalent fragments from the long videos were selected.

Table 1: ***Corpus***

Channel	Code	video titles	Length	Analysed time	Word count analysed time
ClimateAdam	A1 LV1	One Million Years into Climate change	12 min	104 s	245
	A1 SV1	What does past #climate change teach us about the future? #shorts	57 s	57 s	148
Simon Clark	A2 LV2	Are humans really behind the extra CO_2 in the atmosphere?	10 min	129 s	246
	A2 SV2	Are humans really behind the extra CO_2? #Climatescience	59 s	59 s	211
Be Smart	A3 LV3	How Ancient Ice Proves Climate Change is Real	10 min	129 s	262
	A3 SV3	This PROVES that humans are causing climate change	59 s	59 s	194

A# = channel identifier; LV# = Long video identifier; SV# = Short video identifier

To further contextualise these materials, Table 2 presents basic information about the three selected YouTube channels and their presenters, all of whom are trained scientists working as science communicators. Adam Levy, who hosts *ClimateAdam*, is also an active

science journalist, whereas *Simon Clark*'s channel is known for his award-winning science communication work. *Be Smart*, hosted by Joe Hanson, forms part of PBS Digital Studios, and Hanson additionally co-created *Hot Mess*, a climate-focused educational series.

Table 2: ***Description of the three selected channels***

Channel	Subscribers (January 2024)	Creator / Host	Academic Background
ClimateAdam	48.5K	Adam Levy	PhD in Atmospheric Physics, University of Oxford
Simon Clark	524K	Simon Clark	PhD in Atmospheric Physics, University of Exeter
Be Smart	5.12M	Joe Hanson	PhD in Cell and Molecular Biology, University of Texas at Austin

3.2. *The multimodal annotation*

To answer the first research question, that of modal density, the six videos (the whole short videos and the long videos' fragments) were analysed using GRAPE MARS software (Ruiz-Madrid et al., 2023) (Figure 1). The programme is specifically designed to support in-depth analysis of videos, offering a quantitative analysis of the annotated modes and graphical representations of the data.

Figure 1: ***Screenshot of GRAPE-MARS software***

The modes were classified into embodied and filmic modes, as per Valerias-Jurado and Bernad-Mechó's (2022) distinction. This classification, reflected in Table 3, is adapted from their taxonomy, with several modifications made to fit the corpus. The paragraphs that follow discuss the annotation and sub-annotation scheme in detail, as well as any adjustments made to the original taxonomy.

Paralanguage examined the presenter's use of pauses, prominence, such as emphasised strands of speech, and tempo. Gaze assessed the presenter's eye direction, which in this analysis was further marked towards the camera/audience, at a visual prompt or other (accounting for the host looking up, down, left or right). Gestures, conceived as arm, hand and shoulder movement (the latter being an addition to the original taxonomy) were broken down into iconic, metaphoric, deictic and beats. Gesture phrases were classified as one occurrence (i.e., if a series of repetitive beats were accompanying a strand of speech, they would be counted as 1 beat), whilst facial expressions were dissected into eyebrow rising, frown, smile with two more categories, grimace and eye squinting, inserted to better suit this corpus. Head movements were divided into tilts, nods and shakes.

Table 3: ***The quantitatively analysed modes adapted from Valerias-Jurado and Bernad-Mechó (2022)***

Modes		**Measured in**	
Embodied	Paralanguage	#/min. of total time words/min.	prominence other tempo
	Gaze	% of presenter's time	camera/audience visual prompt other
	Facial expressions	#/min. of presenter's time	eyebrow rising frown smile eye squinting grimace other
	Gestures	#/min. of presenter's time	iconic metaphoric deictics beats
	Head movements	#/min. of presenter's time	tilt nod shake
Filmic	Scene change	#/min. of total time	
	Visual prompts	#/min. of total time	image graphs text other
	Sound effects	#/min. of total time	
	Music	% of total time	music silence
	Type of shot	%/total time	close-up medium-close up medium
	Zoom	#/min. of total time	
	Cuts	#/min. of total time	
	Subtitles	%/total time	regular emphasised

Filmic modes dealt with the video's post-production. The scene change annotation was specifically incorporated into the taxonomy to examine all the digital editions that affected how objects (including the presenter) appeared on screen. Moreover, visual prompts registered not only the use of images as in the reference study but also graphs and text. Sound effects were left as in the original taxonomy, without further classifications, whereas the music annotation incorporated silence and music change sub-annotations. The type of shot considered was more restricted than that in Valerias-Jurado and Bernad-Mechó's proposal, as these videos just used close-ups, medium close-ups and medium shots to frame their presenters. For the sake of comprehensive analysis, a subtitles annotation had to be added and further split into regular or emphasised sub-annotation categories. Similarly, the zoom annotation was added to account for zooming in and out instances. This taxonomy, which was used to analyse quantitatively the videos' modal density, also informed the qualitative analysis regarding modal coherence.

3.3. *Analytical procedures*

To address the first research question on modal density, the videos were classified and analysed with the GRAPE MARS software, whose Analytics section provided the frequency and time-percentage data used to discuss the modal density results. To account for it, the annotations of paralanguage, facial expressions, gestures, head movements, scene change, visual prompts, types of shot, cuts, sound effects and zooming were reviewed in terms of occurrences, as previously shown in Table 3, where these modes are presented as average occurrences per minute to facilitate the comparative analysis between the long and short videos. With respect to the embodied modes that required the visual presence of the presenter, namely, gestures, facial expressions and head movements, further conversions were required as the presenter was not constantly on screen. Therefore, to convey reliable results that reflected the presenter's communication style in the long and short videos, the data was exhibited agreeing with the presenter's screentime (i.e., if the presenter only appeared 89s out of the 129s examined in LV2, long video number two, the head movements' average number of occurrences per minute was calculated using the 89s reference). Table 4 shows the screentime percentage share that was used to calculate the averages for embodied modes requiring the presenter's on-screen presence (i.e., gaze, gestures, facial expressions and head movements) that also supported the coherence analysis.

Table 4: ***Screentime analysis of the 6 videos***

	measured in		**LV1**	**SV1**	**LV2**	**SV2**	**LV3**	**SV3**
Screentime	**% of total time**	**presenter**	**28**	**0**	**14**	**34**	**13**	**30**
		presenter with visual prompt	58	79	55	48	0	0
		video clip	10	18	14	6	53	9
		animation	4	3	2	12	26	54
		other	0	0	15	0	8	7

Furthermore, the paralanguage figures for LV3 required additional adjustment, since the channel's host, as noted in the previous section, was not the only speaker in the video, which featured an interview with a palaeoclimatologist. Consequently, the number of occurrences per minute of the paralinguistic features was calculated according to the narrator/presenter's speech time, which amounted, in this case, to 24s out of the total 129s of the excerpts' combined duration. Conversely, elements such as gaze, type of shot, music, and subtitles were evaluated based on their average percentage of total screen time, as these features are more accurately

represented through duration rather than frequency. Unlike discrete actions such as hand gestures or spoken words, most of these elements are continuous or fluid in nature. For example, it is difficult to determine what constitutes a single gaze: does it end when the person blinks or shifts attention? Likewise, the communicative impact of various shot types depends not only on their occurrence but also on the duration for which they remain onscreen; a medium shot held for most of a video, for instance, can create a sense of neutrality or emotional distance. As per paralanguage's tempo sub-annotation, it had to be examined outside the software. The number of words uttered by the presenter in the short videos and the long videos' combined fragments were recorded and the results were then translated to average words per minute.

The second question looked at coherence, which was analysed using a qualitative approach. The six YouTube videos were reviewed paying attention to highly packed modal instances. Borrowing from Valerias-Jurado's (2019: 95) conception: "(…) modal coherence means that the modes are being used in a consistent way and do not contradict each other". For instance, in LV1 (long video one), the presenter showcases a visual prompt, a rising graph line added during editing, while simultaneously producing a metaphorical gesture with his left hand to represent a decrease in temperature, moving in the opposite direction of the graph. This would be counted as an instance of incoherence as the direction that both modes are projecting contradict each other. In order to display the results, it was decided that a brief comment on each video pair would be procured to illustrate the findings. The modal density table was used to supplement the qualitative analysis, as it aided in guiding the observations.

Additionally, the second question was not only related to modal coherence, but also to coherence in a broader sense. This analysis examined whether the creators had adapted their content and style to the different formats, which potentially attract different audiences. Long videos, which viewers typically click on, were presumed to attract a more specialised or already interested audience, likely more engaged with climate-related issues and science dissemination. In contrast, YouTube Shorts' scrolling format and preference towards a varied feed suggested a broader audience, one likely to be less familiar with or less invested in the topic of climate change.

4. RESULTS

4.1. *Modal Density*

The first research question revolved around the similarities and differences in terms of modal density between short and long climate change dissemination videos. The results of this analysis are presented in Table 5 for the embodied modes and in Table 6 for the filmic modes.

Table 5: ***Results of the embodied modes modal density analysis***

	Measured in		**LV1**	**SV1**	**LV2**	**SV2**	**LV3**	**SV3**
Paralanguage	#/ min. of total time	-	9.231	12	13.95	23	17.65*	20
		prominence (speech)	7.500	12	12.64	19	17.65	16
		prominence pauses	0.577	0	1.40	4	0	0
		other	1.73	0	0	3	0	4
	words/min	tempo	141.35	148	114.42	211	121.74	194

Table 5, continued on next page.

Table 5, continued from previous page.

	Measured in		**LV1**	**SV1**	**LV2**	**SV2**	**LV3**	**SV3**
Gaze	% of presenter's time	camera /audiance	92	83	96	97	100	100
		visual prompt	2	17	0	0	0	0
		other	6	0	4	3	0	0
Facial expressions	#/min. of presenter's time	-	27.30	41.73	23.59	28	35	36.66
		eyebrow rising	12.36	24.78	20.89	20	25	20
		frown	12.36	15.65	2.02	8	5	3.33
		smile	1.35	0	0.67	0	0	0
		eye squinting	1.35	0	0	1.33	0	0
		grimace	0	0	0	0	0	13.33
		other	0.67	0	0	0	5	0
Gestures	#/min. of presenter's time	-	19.55	23.47	4.75	16	20	26.66
		iconic	0	2.60	1.35	0	5	0
		metaphoric	4.04	11.73	2.03	2.66	15	3.33
		deictics	3.37	1.30	0.68	6.66	0	23.33
		beats	10.17	7.82	0.68	6.66	0	0
Head movements	#/min. of presenter's time	-	21.35	28.69	35.84	41.33	30	50
		tilt	8.09	5,21	7.87	12	20	13.33
Head	#/min. of presenter's time	nod	10.11	18.26	23.96	25.33	10	33.33
		shake	2.69	3.91	4.04	2.66	0	3.33

*#/ presenter's speech time

After processing the results, it was noted that, on average, the three shorts, SVI, SV2 and SV3 presented a higher modal density than their long counterparts. Some of the most striking figures in relation to the differences between the two formats are reported below.

Regarding *embodied modes*, a faster tempo was registered in the shorts, measured by average number of words per minute (LV1=141.35 vs. SV1=148, LV2=114.42 vs. SV2= 211, LV3= 121.74 vs. SV3=194). It should be noted that this acceleration in shorts may reflect not only embodied delivery but also post-production editing (e.g., artificial speed adjustments), which could align tempo with filmic modes too.

This pattern of higher occurrences in shorts was repeated through the paralanguage results that looked at *prominence* and *other* which, in combination, resulted in: LV1=9.23 vs. SV1=12, LV2=13.95 vs. SV2=23, LV3=17.65 vs. SV3=20. Regarding the *other* subcategory, it was added for the sake of covering any unexpected factor, and in the case of paralanguage features, it encompassed the strands of speech that were altered for humoristic purposes such as shifts in intonation or deliberate mumbling.

With respect to *facial expressions* and, again, in alignment with the more intense and fast-paced nature of shorts, the three shorts significantly displayed more occurrences per minute (LV1=27.30 vs. SV1=41.73, LV2=23.59 vs. SV2=28, LV3=30 vs. SV3=36.66).

The same may be applied to *gestures* (LV1=19.55 vs SV1=23.47, LV2=4.75 vs. SV2=16, LV3=20 vs SV3=26.66). The strikingly lower number of gestures in LV2 compared to SV2 had to do with the type of shot used to encase the presenter, as most gestures were left out of the frame. The *filmic modes* results, displayed in Table 6 are discussed in the paragraph below.

Table 6: ***Results of the filmic modes modal density analysis***

	Measured in		**LV1**	**SV1**	**LV2**	**SV2**	**LV3**	**SV3**
Scene change	#/min. of total time		5.19	9	5.58	15	4.61	10
Visual prompts	#/min. of total time	-	5.77	13	4.19	13	1.98	6
		image (fixed+dynamic)	1.73	2	1.40	6	1.32	4
		graphs (fixed+dynamic)	2.88	7	1.86	7	0.66	2
		text	1.15	2	0	0	0	0
		other	0	2	0	0	0	0
Sound effects	#/min. of total time		12.11	9	0	0	0	5
Music	% of total time	music	92	97	100	100	100	8
		silence	8	3	0	0	0	92
Zooming	#/min. of total time		86	78	69	80	13	30
		close-up	86	0	0	0	13	4
		medium-close up	0	38	69	52	0	26
		medium	0	40	0	28	0	0
Cuts	#/min. of total time		0.57	2	0	0	4.61	2

Focussing on filmic modes and concerning *scene changes*, all the analyses depicted greater alternation in scenes in their summarised adaptations (LV1= 5.19 vs. SV1=9, LV2=5.58 vs. SV2=15, LV3=4.61 vs. SV3=10).

As regards *visual prompts*, the figures revealed that the shorts had a higher number appearing per minute; LV1=5.77 vs SV1=13, LV2=4.19 vs SV2=13, LV3=1.98 vs SV3=6. As for the type of shot encasing the presenters, only one type, be it close-up (LV1 and LV3) or medium close-up (LV2) was used in the regular videos, whereas shorts resorted to alternating between two different types (SV1 and SV2 used medium-close up and medium whereas SV3 relied on close-up and medium close-up).

Regarding the *zoom* annotation, LV1 showcased two zooming actions, whereas the short shows twice that amount. Similarly, SV3 showcased two cuts that acted as abrupt zooming-ins towards the presenter, whilst its long version has no zooming in or out on the presenter. Film modes related to *framing* and *virtual proxemics* were denser in the corresponding shorts, too. Finally, none of the long videos presented *fixed subtitles*, in contrast to their short counterparts.

4.2. *Coherence*

The second research question concerned assessing the long and short videos' display of coherence. Modal coherence examined the multimodal ensembles, while the wider framework of coherence examined whether the content and delivery of the videos catered to potentially different audiences in the long versus the short format.

Concerning modal coherence, the long videos and their compact counterparts depicted scarce inconsistent instances. Overall, a high coherence between embodied and editing modes dominated the films, as the montage accompanying the presenter's fluid scientific explanations in both cases was swift, reflecting the level of professionalism of their post-production. The analysis revealed that, in general, shorts exhibited more complex coherence instances across their embodied and filmic layers. A few remarkable episodes in each short video is subsequently examined.

Figure 2: ***SV1 screenshots of instances with significant gestures and gaze correlating with other modes***

In A1 (channel one's videos), the *gaze* annotation in the long video only registered the presenter looking at the camera/audience. However, in the short version, the youtuber's gaze is directed at a visual prompt, specifically a graph, 17% of the time, as showcased in Table 5. Furthermore, the graph's materialisation is also accompanied not only by the presenter's gaze, but by a sound effect and a metaphorical gesture that frames the graph's line, shown in Figure 2.

The data reflects this in that the main category in the short regarding gestures is composed of metaphoric movements that mimic the graph's appearance or its oscillations, which occurs nine times. On the other hand, the extended video's main category is *beats*, with 12 occurrences per minute. The metaphoric gestures, even though employed in the same way as in the shorts (albeit less consistently), lack the precision in execution that the shorts' gestures display, sometimes generating incoherences in relation to the visual prompt.

A2 (channel two's videos) had, out of the three pairs examined, the most substantial disparity in the number of scene changes per minute between its long and short analysis. In this case, more scene variations brought about a higher number of visual prompts (each scene often had a different visual prompt as background) and were also linked to the presenter's discourse.

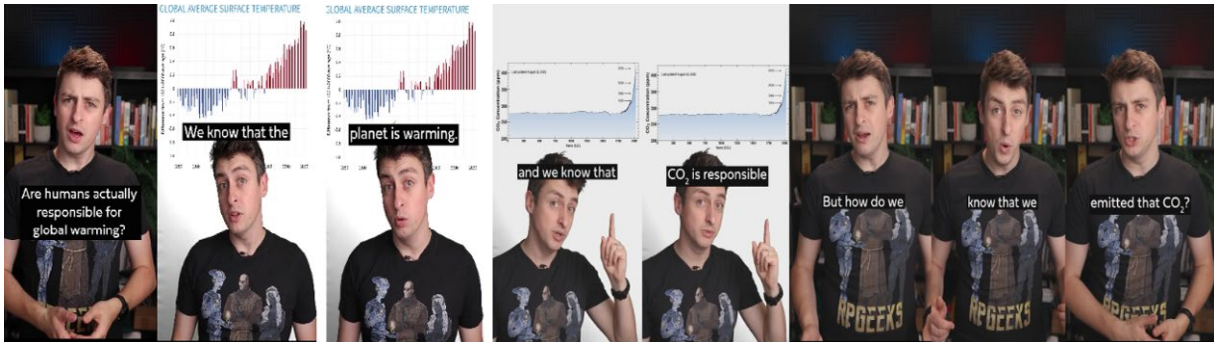

Figure 3: ***SV2 screenshot sequence: correlation between statements, scene changes and subtitle arrangements***

As Figure 3 illustrates, every time there was a new statement, there was a scene change, contributing to the overall higher intensity of the video. This tight-knit ensemble was further complemented by the incorporation of subtitles. Since this short video had constant scene changes where the presenter was moved around the screen, subtitles could not be set in the same place throughout its running, as they would have inevitably interfered with the presenter's face and the facial expression modes. Consequently, editions were made to accommodate the video's subtitles every time there was a change in frame.

Figure 4: ***SV3 screenshots showing subtitle emphasis that accompanies speech prominence***

As outlined in previous sections, no subtitles are incorporated in the long videos, yet they are present in all 3 shorts. In SV3, subtitles are emphasised using capital letters on three occasions. The first, shown in Figure 4, mirrors the video's title, which also capitalises the word *PROVES*. This textual emphasis is replicated in the YouTuber's spoken language, as its enunciation is marked in prominence. The other two moments ("total carbon has gone WAY up" and "the amount of carbon-13 has gone WAY down") once more attempt to correlate with key information and a speech prominence instance. It should be noted that the last capitalised

word does not coincide with the prominence in the narrator's spoken language as "down" is the most emphasised unit of the word cluster. Yet, it manifests the creator's effort to establish correlations between the embodied and filmic modes and to convey meaning through the latter.

The last part of this second research question had to do with analysing coherence in a broader sense. The assessment considered the creator's distinct approaches to the same topic across the two YouTube formats and their potential different audiences.

A1's videos focus on climate variations in the last million years and the relationship between temperature and carbon dioxide, emphasising the impact of human activities on changing temperatures and the potentially devastating consequences to humankind and the planet. The long video merely refers to possible scientific solutions without explicitly showcasing them, while the short video immediately introduces mitigation and potential solutions after hinting at the catastrophe.

A2's videos aim to demonstrate through deductive reasoning that the additional CO_2 in the atmosphere is a result of human activity. The long video provides three different methods for arriving at this conclusion, while the short video only displays the latest method. A3 (channel three's vidoes) also strives to prove that climate change is due to human causation. The long video, created in 2019, resembles an amateur documentary wherein the presenter visits the Scripps Institution of Oceanography, while the short version from 2023 makes extensive use of the medium's editing affordances, as reflected in the modal density table in regard to the filmic layer.

The implications of these differing approaches, in relation to the affordances of each video format and its potential audience, are examined in the coherence subsection of the discussion.

5. DISCUSSION

5.1. *Modal density*

Broadly speaking, it was observed that shorts exhibited a greater modal density compared to their longer counterparts. These higher occurrences manifested through both the embodied modes and the filming layer, aligning with Valeiras-Jurado and Bernad-Mechó's observations in their 2022 paper. The significant editing of the videos follows the standards of effective popular science videos on YouTube (Muñoz Morcillo et al., 2016), which would explain why it is replicated and maximised in the shorts. The shorts' higher modal density could reflect a calculated adaptation to the platform's format, where creators must weaponise brevity and attention (Rajendran et al., 2024) to survive the swipe and potentially convert fleeting views into subscriptions.

5.2. *Coherence*

Across the corpus, both long and short videos displayed high modal coherence, but the shorts consistently showed more intricate coordination across embodied and filmic modes. This heightened complexity is attuned to the format's scrolling environment, where visually rich modal ensembles may help capture and retain viewer attention, a requirement for success in the shorts medium (Rajendran et al., 2024).

With respect to the wider coherence analysis, the results point to clear distinctions between formats that could be justified by their potential different audience profile. The fact that A1's short uses more mitigation when voicing the uncomfortable reality of climate change aligns with a discourse more suitable for the masses. This is consistent with the platform

dynamics previously discussed: YouTube's Shorts algorithm tends to deliver content to a wider audience, often reached through passive scrolling, whereas long videos tend to be clicked on by viewers and subscribers, showcasing a higher interest in the topic. The long video in A1 offers fewer reassurances and does not focus on solutions, further reinforcing its orientation toward a more specialised public prepared to confront the grave implications of the subject matter. Similarly, in A2 the long video presents multiple methods and a more detailed scientific explanation of its topic, while the short simplifies the content to a single, brief method, hinting at the uploader's desire to appeal to a less knowledgeable public. In A3, the short extensively leverages the medium's editing affordances likely in hopes of capturing transient viewers, while the long version, as with the other long videos, employs a much looser editing, which may suggest that the presenters place greater trust in the viewers' interest in the topic. Moreover, the three shorts displayed a simpler and less nuanced discourse style, consistent with Chi and Park's (2022) observation that the language used in shorts tends to be simple and intuitive to meet audiences' preference for easy consumption. Overall, the corpus suggests that the content and style of the videos were informed by their medium and the types of audiences they are likely to reach, making the six videos highly coherent.

6. CONCLUSIONS

YouTube, which could be considered part of the social media ecosystem (van Dijck, 2013), stands out as a prominent digital space for disseminating scientific knowledge (Hill et al., 2022). This article has presented a comparative analysis examining how climate change is portrayed in both regular and short YouTube videos. The study focused on assessing modal density and coherence relationships within six videos, aiming to provide insights into how climate change communicators adapt their discourse to effectively convey their message on YouTube taking the two formats into account.

In sum, both filmic and embodied modes presented, overall, a higher modal density in the short videos. As per coherence, the examination of modal density results combined with the qualitative analysis indicated a more complex modal interplay between filmic and embodied modes in the short products. The creators' different approaches to the same issue in long and short videos may be justified by their knowledge of their medium and potential audience profile in, making the videos highly coherent.

The limitations of this study were the restricted scope of the corpus, which encompassed three YouTube channels of climate change advocates and scientists from which only six videos were analysed. Future research could enlarge the scale of the corpus and expand on the analyses made in this paper. For instance, the language employed in the videos could be quantitively analysed to confirm the impressions related in the qualitative analysis. Similarly, the abundant differences found between the two video formats regarding modal density, modal coherence and their distinct discursive approach, potentially tailored to different audiences, could indicate that the shorts and the long videos correspond to two different digital genres. On the subject of climate change communication, further research could explore the use of mitigation and omission of the harsher truths of the phenomenon in the long and short formats. In addition, and building on this paper's findings regarding modal density and coherence, a multimodal guideline for effectively communicating climate change on YouTube could be developed to offer deeper insights for online advocates. This guideline could even be expanded to include other short-form video formats, such as TikToks and Instagram Reels.

Finally, given the urgency of the climate crisis, there is a critical need for further research to aid communicators in effectively disseminating their messages, thereby combating the pervasive online misinformation. As highlighted in a prior section, Allgaier's (2019) study revealed that a significant portion of YouTube content on climate change perpetuates false information. Climate change is an imminent threat with far-reaching implications, necessitating contributions from every academic discipline to mitigate its rapid progression.

REFERENCES

Aksakal, N. (2015). Theoretical view to the approach of the edutainment. *Procedia - Social and Behavioral Sciences*, *186*, 1232–1239. https://doi.org/10.1016/j.sbspro.2015.04.222

Allgaier, J. (2019). Science and environmental communication on YouTube: Strategically distorted communications in online videos on climate change and climate engineering. *Frontiers in Communication, 4*. https://doi.org/10.3389/fcomm.2019.00036

Boy, B., Bucher, H.J., & Christ, K. (2020). Audiovisual science communication on TV and YouTube: How recipients understand and evaluate science videos. *Frontiers in Communication*, *5*. https://doi.org/10.3389/fcomm.2020.608620

Brennan, E.B. (2021). Why should scientists be on YouTube? It's all about bamboo, oil and ice cream. *Frontiers in Communication*, *6*. https://doi.org/10.3389/fcomm.2021.586297

Chi, Y., & Park, E. (2022). Counterattacking long videos: Exploring the characteristics of popular instant videos and roles of producers and viewers. *Library Hi Tech*, *41*(3), 694-710. https://doi.org/10.1108/LHT-05-2022-0230

De Lara, A., García-Avilés, J.A., & Revuelta, G. (2017). Online video on climate change: a comparison between television and web formats. *Journal of Science Communication*, *16*(01), A04. https://doi.org/10.22323/2.16010204

Erviti, M., Codina, M., & León, B. (2020). Pro-science, anti-science and neutral science in online videos on climate change, vaccines and nanotechnology. *Media and Communication*, *8*(2), 329-338. https://doi.org/10.17645/mac.v8i2.2937

Fortanet-Gómez, I., & Ruiz-Madrid, M. (2014). Multimodality for comprehensive communication in the classroom: Questions in guest lectures. *Ibérica*, *28*, 203-224.

Hill, V.M., Grant, W.J., McMahon, M.L., & Singhal, I. (2022). How prominent science communicators on YouTube understand the impact of their work. *Frontiers in Communication*, *7*. https://doi.org/10.3389/fcomm.2022.1014477

Huang, T., & Grant, W.J. (2020). A good story well told: Storytelling components that impact science video popularity on YouTube. *Frontiers in Communication*, *5*. https://doi.org/10.3389/fcomm.2020.581349

Kress, G. (2010). *Multimodality: A social semiotic approach to contemporary communication.* London: Routledge. ISBN 9780415320610

León, B., & Bourk, M. (Eds.). (2020). *Communicating science and technology through online video.* London: Routledge.

Lobato, R. (2016). The cultural logic of digital intermediaries: YouTube multichannel networks. *Convergence*, *22*(4), 348-360. https://doi.org/10.1177/1354856516641628

Jewitt, C., & Price, S. (2012, September). *Multimodal approaches to video analysis of digital learning environments* In *Proceedings of BCS HCI 2012 Workshops: Video Analysis Techniques for Human–Computer Interaction*. BCS. https://doi.org/10.14236/ewic/HCI2012.105

Muñoz Morcillo, J., Czurda, K., & Robertson-von Trotha, C.Y. (2016). Typologies of the popular science web video. *Journal of Science Communication*, *15*(04), A02. https://doi.org/10.22323/2.15040202

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

Olausson, U., & Berglez, P. (2023). Communicating climate change in the era of populism and misinformation. *Journalism Studies, 24*(1), 45-62. https://doi.org/10.1080/1461670X.2022.2100917

Perez, S. (2023, August 25). *YouTube demystifies the Shorts algorithm, views and answers other creator questions*. TechCrunch. Retrieved from: https://techcrunch.com/2023/08/25/youtube-demystifies-the-shorts-algorithm-views-and-answers-other-creator-questions/

Querol-Julián, M. (2011). *Evaluation in discussion sessions of conference paper presentations: A multimodal approach.* London: LAP Lambert Academic Publishing. ISBN: 978-3-8443-2402-0

Rajendran, P.T., Creusy, K., & Garnes, V. (2024). Shorts on the rise: Assessing the effects of YouTube Shorts on long-form video content. *arXiv:2402.18208.* https://doi.org/10.48550/arXiv.2402.18208

Ruiz-Madrid, N., & Fortanet-Gómez, I. (2019). A multimodal discourse analysis approach for EMI teacher training: The case of pedagogical affordances in the mechanical engineering field. In A.C. Lahuerta-Martínez & A. Jiménez-Muñoz (Eds.), *Empirical studies in multilingualism: Analysing contexts and outcomes.* (pp. 203-229). Bern: Peter Lang. ISBN: 9783034335225, 9783034335201.

Ruiz-Madrid, N., Fortanet-Gómez, I., Bernad-Mechó, E. (2023) GRAPE-MARS (multimodal analysis research software). Castelló de la Plana: Universitat Jaume I.

Schäfer, M., & Yan, X. (2023). News and social media imagery of climate change. In *The Routledge handbook of climate change communication*. London: Routledge. https://doi.org/10.4324/9781003409748-11

Scotto di Carlo, G. (2015). Stance in TED talks: Strategic use of subjective adjectives in online popularisation. *Ibérica*, *29*, 201-222.

Shapiro, M.A., & Park, H.W. (2015). More than entertainment: YouTube and public responses to the science of global warming and climate change. *Social Science Information*, *54*(1), 115-145. https://doi.org/10.1177/0539018414554730

United Nations. (2024). *Communicating on climate change.* Retrieved from: https://www.un.org/en/climatechange/communicating-climate-change

Valeiras-Jurado, J. (2019). Modal coherence in specialised discourse: A case study of persuasive oral presentations in business and academia. *Ibérica*, *37*, 87-114.

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*, 101-118. https://doi.org/10.1016/j.jeap.2022.101118

Van Dijck, J. (2013). *The culture of connectivity: A critical history of social media.* Oxford: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199970773.001.0001

Van Leeuwen, T. (2005). *Introducing social semiotics.* London: Routledge. https://doi.org/10.4324/9780203647028

Welbourne, D.J., & Grant, W.J. (2016). Science communication on YouTube: Factors that affect channel and video popularity. *Public Understanding of Science*, *25*(6), 706-718. https://doi.org/10.1177/0963662515572068