Silvia Masi*

University of Pisa, Italy
silvia.masi@unipi.it

HUMOUR IN TED TALKS: A MULTIMODAL ACCOUNT

Abstract

TED talks represent a popular digital genre for the dissemination and popularisation of knowledge of multiple domains, and humour is one of its endemic characteristics. While past research has mainly focused on the linguistic expression of humour in the talks, the present contribution explores multimodal ensembles comprising visuals (i.e., slides with images and videos), words and gestures which jointly contribute to humour generation in various ways. The data come from a small corpus of talks from the domains of Technology, Economics and Law. The domain of Technology proves to be the richest in humorous episodes, some of which are illustrated through multimodal transcription and a detailed qualitative analysis. The latter leads to the identification of the causes for humour and its varied functions within the talks, which go beyond the management of interpersonal relations and sometimes intertwine with argumentation and strategies of popularisation for the development of the main topics of the talks. As humour understanding may be a particularly challenging task for an international audience such as the one addressed by TED speakers, a more comprehensive grasp of its dynamics may hopefully shed light on its subtleties, and on its engagement or captivating potential for ESP learners and EFL extramural contact.

Key words

humour, multimodality, knowledge dissemination, popularisation, TED talks, digital genre.

* Corresponding address: Silvia Masi, Via Santa Maria 67, 56126 Pisa, Italy.

328

1. INTRODUCTION

TED talks are now a popular web genre for the dissemination and popularisation of knowledge from a large number of domains via short and effective presentations. They can be accessed by a co-present audience and web users through the TED website, [1] they address both specialists and non-specialists, and are rich in multimodal contents. These cover visuals such as images, photographs, graphs, short video clips embedded into slides (Harrison, 2021; Masi, 2020a; Meza & Trofin, 2015; Theunissen, 2014; Xia, 2023), along with prosody and other extra-linguistic elements such as, for example, facial expressions, head movements and gestures of speakers (Harrison, 2021; Masi, 2016, 2019, 2020a, 2020b; Valeiras-Jurado, 2017; Valeiras-Jurado & Ruiz-Madrid, 2019). Over time, the talks have become highly influential digital materials used in diverse educational settings and have increasingly been explored for their potential as a pedagogical resource (cf. e.g., Carney, 2014; Chang & Huang, 2015; Dummett et al., 2016; García-Pinar, 2019a, 2019b; García-Pinar & Pallejá-López, 2018; Takaesu, 2013; Wingrove, 2017). Numerous studies have been conducted to identify the typical linguistic, and more recently, multimodal features underpinning the rhetoric of promoting 'ideas worth spreading' specified in their motto. Typical traits include the use of humour as an endemic characteristic (Scotto di Carlo, 2013), an informal register that encourages participation and proximity (Scotto di Carlo, 2014), reduced technicality and the presence of personal anecdotes (Mattiello, 2017), the use of markers of engagement, together with epistemic verbs for the expression of stance (Caliendo & Compagnone, 2014; Compagnone, 2014). From a multimodal perspective, research has shown that different semiotic resources (including gestures and visuals as slides with images and videos, see Harrison, 2021; Jiang & Lim, 2022; Wu & Qu, 2020; Xia, 2023) can be significantly co-deployed with words in this genre of popularisation, thus contributing to multimodal ensembles (Kress, 2010), i.e., combinations of meaning-making resources or modes, of different degrees of complexity and with a varied distribution across distinct knowledge domains (Masi, 2019, 2020a).

Of all their features, the use of humour especially distinguishes TED talks from other forms of more traditional popularisation, as highlighted by Scotto di Carlo (2013) (also see Mattiello, 2017; Peruzzo, 2021). Although such accounts mainly focus on the linguistic expression of humour, "much of humour is not expressed or communicated linguistically" (Attardo, 2020: 95). Ruiz-Madrid and Fortanet-Gómez (2015), for example, highlight that non-verbal cues, such as gesturing, gaze, and prosody, have an important role in communicating humorous intention in oral academic discourse, while Crawford Camiciottoli (2021) suggests that linguistic and extra-linguistic features have a synergistic relationship in humorous episodes in university lectures. The role of extra-linguistic resources in the generation of

humour in TED talks, however, has been largely neglected so far, especially as far as the contribution of visuals is concerned. The present account is part of work in progress on the multimodal analysis of TED talks and intends to bridge such a gap by exploring a range of multimodal manifestations of humour in a small corpus of talks from different disciplinary domains, so as to expand on the variety of causes and functions of such a carefully orchestrated phenomenon in this genre of popularisation. Humour understanding may be a particularly challenging task for an international audience such as the one addressed by TED speakers. A more comprehensive grasp of its dynamics may then inform teaching strategies for assisting learners of English for Specific Purposes (ESP) who might use the talks as educational materials, and also shed light on its engagement or captivating potential for English as a Foreign Language (EFL) informal contact and learning (Pavesi & Ghia, 2020) via enjoyable resources freely accessible on the web.

2. PAST RESEARCH ON HUMOUR IN TED TALKS

Humour is the cognitive process or stimulus that causes mirth as an emotional effect, with laughter and smiling as possible physical manifestations or markers, although there is no one-to-one correlation between the stimulus and the manifestations (Attardo, 2020: 17-18). The identification of humour in communication generally requires the recognition of an incongruity (Attardo, 2020: 46-49) that arises from an unexpected opposition between overlapping semantic scripts or world knowledge scenarios, experienced as fictional or in a playful mode. From a pragmatic point of view, then, humour is non-cooperative, i.e., non-bona fide (Attardo, 2020: 165-166), for example, for the purpose of entertainment. Humour in discourse may perform different functions depending on context; for instance, it is often used to build common ground with the interlocutor or audience (thus fulfilling a social management function), it may help mediate sensitive topics thanks to its playful nature, and may also work as a politeness strategy (e.g., hedging criticism) (Attardo, 2020: 273-278).

In TED talks in particular, humour is extensively used with the consequence of enhancing the 'enjoyability' of the talks (Scotto di Carlo, 2013), a feature that, in turn, increases attention and memorability in the audience (Waknell, 2012). Scotto di Carlo (2013) draws from classifications of humour theories by Raskin (1985) and Attardo (1994) and identifies three main conditions (cognitive/incongruity, social/derisive, psychoanalytic/release) that are responsible for the emergence of humour in this genre of popularisation. Mattiello (2017) appears to confirm Scotto di Carlo's findings and especially identifies humour at the beginning and at the end of the talks, deployed through exaggerations and euphemisms. Peruzzo (2021) focuses on humour in talks that are related to mental health disorders, and also explores the ways in which it intertwines with speakers' storytelling.

However, humour is first and foremost a communicative event in which language is only one of the several components. A broader, multimodal approach to its analysis becomes necessary especially when tackling its expression in media that involve different interrelated semiotic resources (Attardo, 2020: 25, 95). As a matter of fact, recent research in Natural Language Processing (NLP) has shown that cues from the modalities of language, acoustics, and vision – in the form of speakers' facial expressions – are important co-determinants of humour in the TED talks (Hasan et al., 2019). The findings appear to be in line with the rationale at the basis of the triangulation approach for humour identification (Attardo, 2020: 291-293), which advocates the use of multiple types of evidence (e.g., intonational and prosodic cues, laughter or smiling, facial expressions and gestures, *inter alia*) next to semantic incongruity. A multimodal approach can then be considered as more reliable heuristics for humour identification and analysis than a purely linguistic approach. Hasan et al.'s (2019) study was aimed at developing a computational framework for multimodal humour detection, but was mainly intended for the natural language processing community and did not take into account other extra-linguistic resources such as gestures or visuals as slides with their varied contents. In fact, gestures have been found to contribute to the humour often employed in the talks to entertain and involve the audience (Masi, 2020b), and different types of contents of slides (e.g., photos, videos, words), too, have been found to perform the function of entertaining through humour, as well as making specialised information more tangible (Theunissen, 2014). It is thus worth exploring humour in the talks through a more comprehensive, multimodal approach that takes into account the speech-gesture-visual interplay.

3. METHODOLOGY

The present study hinges on video-based analyses of a random selection of thirty TED talks from the domains of Technology, Economics and Law (ten from each domain, see Table 1 below), as preceding research had highlighted a diversified distribution of semiotic resources across such domains, with visuals being especially involved in that of Technology (Masi, 2020a). A random selection from domains of specialisation was preferred over The Funniest TED talks playlist,[2] as the latter was hybrid in terms of domain composition and the talks were often delivered by professional comedians, humourists or actors rather than domain-specific experts. A selection from domains of specialisation, instead, is more likely to provide contents that are more relevant to the needs of ESP education. The talks cover a time span from 2012 to 2022 and all speakers sound as either native or expert users of English as Lingua Franca. Each talk was given by individual speakers of different gender. Although gender differences may well have an impact on

---

[2] https://www.ted.com/playlists/235/the_funniest_ted_talks (last accessed January 8th 2023).

humour styles, this aspect was not taken into account yet, but could be further explored in future developments of the research.

Domain	Title	Speaker	Length in minutes
Technology	*The game that can give you 10 extra years of life*	Jane McGonigal	19:14
Technology	*404, the story of a page not found*	Renny Gleeson	3:51
Technology	*Everyday cybercrime -- and what you can do about it*	James Lyne	17:13
Technology	*New video technology that reveals an object's hidden properties*	Abe Davis	17:48
Technology	*How Amazon, Apple, Facebook and Google manipulate our emotions*	Scott Galloway	18:57
Technology	*3 ways to make better decisions -- by thinking like a computer*	Tom Griffiths	11:39
Technology	*How an algorithm can fight election bias so every vote counts*	Brian Olson	9:26
Technology	*The danger of AI is weirder than you think*	Janelle Shane	10:20
Technology	*What happens when a Silicon Valley technologist works for the government*	Matt Cutts	5:56
Technology	*AI isn't as smart as you think -- but it could be*	Jeff Dean	18:22
Economics	*Bitcoin. Sweat. Tide. Meet the future of branded currency*	Paul Kemp-Robertson	10:39
Economics	*When you're making a deal, what's going on in your brain?*	Colin Camerer	13:37
Economics	*What the Social Progress Index can reveal about your country*	Michael Green	14:48
Economics	*The US needs paid family leave -- for the sake of its future*	Jessica Shortall	15:36
Economics	*Why you should know how much your coworkers get paid*	David Burkus	7:21
Economics	*Know your worth, and then ask for it*	Casey Brown	8:13
Economics	*How we'll earn money in a future without jobs*	Martin Ford	14:28
Economics	*A healthy economy should be designed to thrive, not grow*	Kate Raworth	15:45
Economics	*What if all US health care costs were transparent?*	Jeanne Pinder	6:55
Economics	*The cost of work stress -- and how to reduce it*	Rob Cooke	10:29
Law	*How I defend the rule of law*	Kimberley Motley	15:17
Law	*Why gun violence can't be our new normal*	Dan Gross	13:30
Law	*How the US should use its superpower status*	Ian Bremmer	15:28
Law	*How judges can show respect*	Victoria Pratt	15:55
Law	*What if we ended the injustice of bail?*	Robin Steinberg	14:16
Law	*3 ways businesses can fight sex trafficking*	Nikki Clifton	12:21
Law	*What happens to people in solitary confinement*	Laura Rovner	11:59
Law	*What you need to know about face surveillance*	Kade Crockford	12:40
Law	*The end of Roe v. Wade -- and what comes next for reproductive freedom*	Kathryn Kolbert	16:52
Law	*What happens to people's donated eggs and sperm after they die?*	Ellen Trachman	12:43

Table 1. Selection of TED talks from the three domains of Technology, Economics, and Law

The total size of the corpus is 384.58 minutes. Its limited size does not allow quantitatively based generalisations, but it is valid for a qualitative analysis. Multimodal ensembles clustering around the audience's laughter (signalled in the transcripts available on the TED platform) were captured from the videos by making screenshots. Laughter was indeed regarded as a possible marker of humour, whose presence was confirmed by the identification of some forms of incongruence and by

the presence of other extra-linguistic evidence, for example, speakers' smiling faces and/or variation in prosodic features, gestures, and head movements. The contours of the ensembles were established on the basis of what seemed to fulfil the intention of generating humour and consequent laughter.

The humorous instances were selected manually and analysed qualitatively on the basis of an integrated method of multimodal transcription inspired by Lazaraton (2004) (originally accounting for co-speech gestures), later expanded by Masi (2020a) to cover visuals. The resulting method (accounting for the speech-gesture-visual interplay) was further developed for the present study and the ensembles were analysed on the basis of:

- semiotic types of the visuals shown during the talks (e.g., scriptural, only visual, scriptural-visual, etc., cf. Theunissen, 2014),

- what was shown (e.g., types of participants and actions) and how it was shown in visuals (e.g., with a direct gaze of depicted participants, from a frontal or oblique perspective, etc. – more on that in the section on the analysis of examples), (cf. Kress & van Leeuwen, 1996, 2021 – reference was made to the resources that contribute to different metafunctions in language and communication, the ideational-representational and interpersonal-interactive ones in particular),[3]

- the contribution of gestures (e.g., deictic, iconic, metaphoric, beats, etc., cf. McNeill, 1992) and other non-verbal cues such as facial expressions, when relevant,

- the intersemiotic relations between different resources (as far as visuals and words were concerned, reference was made to an adaptation of Royce's (2006: 63) intersemiotic complementarity, i.e., the ways different modes of communication, within a single text, complement each other and project meaning),

- the functions of the ensembles within the rhetorical structure of the talks (Chang & Huang, 2015; Xia, 2023).

In the next section, I will present the main findings and three examples that were selected based on their representativeness, i.e., their potential to illustrate (1) the synergistic interplay of different semiotic resources in the generation of humour, and (2) different humour functions in the talks.

333

4. DATA ANALYSIS: FINDINGS AND EXAMPLES

Most of visuals in the humorous multimodal ensembles from all the three domains had a prominent role in humour generation, as they displayed a complementary relation with words for the expression of humour, while, in other cases, visuals

---

[3] The metafunctions are drawn from Halliday's (1985) work on linguistic analysis (ideational, interpersonal, textual metafunctions). Kress and van Leeuwen (2021) applied them to the analysis of images and changed their names into representational, interactive, and compositional. The first represents the participants, relations, processes and attributes of the entities depicted, the second mediates the relations between images and their viewers, the third concerns the way in which the preceding elements are integrated into a meaningful whole (Kress & van Leeuwen, 2021: 179).

appeared to have a less crucial, supportive function.[4] Overall, the highest number of instances of laughter signalling humour were found in the talks from the domain of Technology, which also contained the highest number of multimodal ensembles involving visuals (slides with images/photos or videos)[5] (see Table 2 below).

Domains	Technology	Economics	Law
Instances of laughter	51	33	36
Visuals with complementary relationship with words for the expression of humour	20 (in all the selected 10 talks for this domain)	7 (in 5 talks)	4 (in 2 talks)
Visuals supporting the expression of humour	7	3	1
Total of visuals in humour-related multimodal ensembles	27	10	5

Table 2. Distribution of laughter and visuals in humour-related multimodal ensembles across domains

The multimodal ensembles clustering around the laughter instances in the Technology talks also highlighted the most varied range of communicative functions, e.g., social management and mitigation of criticism – more traditionally associated with humour, but also its contribution to strategies of popularisation via illustration of past experiments or tests done during the talk, analogy and exemplification (even with videos).[6] The examples that follow are some cases in point from this domain, which were chosen on account of the rich multimodal interplay causing the humour they contain, where humour also appears to contribute to different functions.

The analysis of each example is preceded by its transcription summarising the main aspects of the interplay and of its interpretation in terms of humour causes and functions. Each transcription comprises:

334

---

[4] As for the distinction between the two categories of visuals, while those supporting humour could be removed and the humourous effect would still be present, the removal of the visuals with a complementary relation with words would disrupt the humorous effect altogether.

[5] Such findings appear to be in line with the higher use of visuals in the domain of Technology (than in that of Business and Economics and that of Law) as emerged from Masi (2020a).

[6] Scientific popularisation involves the reformulation and recontextualisation of specialised knowledge for a lay audience. Indeed, popularisation strategies cover the following (see Calsamiglia & Van Dijk, 2004): denomination (for the introduction of specific concepts), definition/description (to explain unknown terms and ideas), reformulation or paraphrase (through the use of parentheses, dashes, metalinguistic expressions, etc.), generalization (based on conclusions derived from specific cases), exemplification (based on specific examples), and analogies or associations (e.g., via similes and metaphors which establish links between different conceptual domains).

    - a screenshot[7] of one or more visuals (on the left hand side) with a caption at the bottom that specifies the semiotic types they belong to;

    - the verbal part (on the right hand side), where the words occurring simultaneously with the visuals have been marked in bold type, and where gesture description has been integrated and given in italic style and in between brackets, while the words co-occurring with the gesture have been underlined. The name of the speaker is specified below the text, followed by the time of occurrence of the ensemble in the talk, the label for the domain of specialisation and year when the talk was recorded;

    - the description of what is provided, and how, both visually and verbally (on the basis of the resources that contribute to the previously mentioned metafunctions);

    - the interpretation of the intersemiotic relations;

    - the identification of the causes and functions of humour.

4.1. Example 1 – Humour for social management

The first case below comes from the talk entitled 'AI isn't as smart as you think – but it could be'[8] (length: mins. 18:22), in which Jeff Dean, the head of Google's AI, explains the technology at the basis of artificial intelligence and also suggests how to build systems that have a deeper understanding of the world.

VISUAL	VERBAL TEXT AND OTHER NON-VERBAL CUES
Circa 1990 **Visual type**: Scriptural-visual	I first got interested in neural networks when I took a class on them as an undergraduate in 1990. At that time, neural networks showed impressive results on <u>tiny problems</u> (*left hand half raised at front, with little distance between thumb and index*), but they really couldn't scale to do real-world important tasks. **But I was super excited**. (simultaneous display with visual) (Dean, 03:36, TECH, 2021) (*Speaker's and audience's audible laughter*) (after and during the display of the visual)
What is represented: Face of younger speaker	**What is expressed**: Focus on emotional attribute of participant

---

[7] Using TED screenshots for research is permitted by CC BY – NC – ND 4.0 International.
[8] https://www.ted.com/talks/jeff_dean_ai_isn_t_as_smart_as_you_think_but_it_could_be (last accessed March 25th 2023).

How it is represented: Direct gaze; personal/social distance; frontal perspective for involvement; sensory-affective validity	**How it is expressed**: Declarative structure with subjective adjective + metaphoric gesture
Intersemiotic relations: Repetition on ideational-representational level Interpersonal-interactive level - Visual mode: reinforcement of address, stronger involvement and sensory-affective validity; overall attitudinal congruence between modes	
Humour causes and function	
Incongruity rests on clash between causes for excitement referred to verbally and the unexpected, exaggerated and involving way in which the emotion is conveyed visually Communicative function: representation of emotion within narrative for social management	

Transcription and interpretation of example (1)

The visual of this ensemble belongs to the scriptural-visual type, as it consists of both a photo and a verbal component, i.e., the indication, at the bottom of the slide, of the year when the photo was taken. On the right hand side, the transcription covers the description of a metaphoric gesture that emphasises the idea of smallness (verbally expressed by the adjective 'tiny' in the co-occurring "tiny problems") through the little distance between thumb and index of the half raised left hand of the speaker during the talk. In so doing, the gesture also appears to contribute to the contrast (verbally expressed through the conjunction 'but') between the relatively few results obtained at the time (cf. "they really couldn't scale to do real-world important tasks") and the speaker's intense (i.e., "super") excitement.

336

     The portion of verbal text in bold (co-occurring with the visual) is the 'core' of the ensemble, as it generates the subsequent instances of laughter (both from the audience and the speaker). What is represented visually – ideational-representational metafunction – is an image of the speaker at a younger age, in which his facial expression (i.e., widely opened eyes and slightly opened mouth) particularly attracts the viewer's attention. On the verbal level, the part in bold defines the emotional attitude of the participant in the photo, preceded by the narration of the circumstances which led to that emotional state.

     From the point of view of how things are represented, i.e., the interpersonal-interactive metafunction, the participant's gaze (i.e., which in this case identifies with the speaker's) in the visual represents an instance of direct address to viewers, thus demanding for their reaction and contributing an effect of interaction. The participant is captured through a close/middle-shot, thus embodying personal/social distance from the viewers, rather than impersonal detachment through a long shot. The relevance of a frontal perspective contributes to a strong involvement effect. As for the degree of visual validity of what is shown,[9] relevant

---

[9] Validity indicates the degree to which what is represented is considered real. Kress and van Leeuwen (2021) introduce this notion to replace that of modality (Kress & van Leeuwen, 1996). While linguistic truth is based on modality judgements in terms of possibility, probability or

markers score high within a naturalistic coding orientation, yet emphasis does not seem to be placed on the authenticity of the facial expression (as normally encountered in real life), but on a somehow exaggerated representation mimicking (and drawing attention to) emotional meaning. Moreover, a lack of greater visual contextualisation appears to shed doubts on the circumstances in which the photo was actually taken (hence on the actual spontaneity at the origin of its content). Rather, the 'loaded' quality of the representation appears to enhance the affective appeal of the visual (cf. its sensory-affective validity), as it portrays an intense emotion and enables to experience the ensemble in a playful mode, allowing for humour detection. On the verbal level, a declarative structure (as a statement of information) with a subjective adjective is used, the latter expressing a high degree of intensity in a colloquial register (cf. the combining form 'super-') – which brings the speaker closer to the audience.

From the point of view of intersemiotic relations, both words and visuals represent the same entity (emotion) on the ideational-representational level. However, on an interpersonal-interactive level, the visual, in particular, complements words by reinforcing the effect of direct address to viewers[10] (and correlated demand for their reaction and involvement), as resulting from the resources of direct gaze, personal distance and frontal perspective mentioned above. The exaggerated representation of the facial expression is at the basis of the affective appeal of the visual (cf. its sensory-affective validity); overall, attitudinal congruence between the visual and the verbal text appears to be relevant, on account of the fact that a high degree of intensity of emotion is expressed through both modes.[11]

337

As for the causes of humour in the ensemble, incongruity appears to rest on a clash between the (little) reason for excitement of the speaker, referred to verbally, and the surprising/unexpected, almost exaggerated and involving way in which the emotion is represented visually. The preceding gesture also paves the way to this

---

frequency, visual truth "is based on the idea of *realism* as assessed and felt by the evidence of what can be seen" (Kress & van Leeuwen, 2021: 154). Validity indeed accounts for different types of representations of what is real, depending on the affordances of different semiotic modes. Different validity criteria, that is, may operate in different contexts, as captured via distinct coding orientations, e.g., naturalistic (dominant in Western society, based on a resemblance to reality as usually portrayed in naturalistic photography), abstract (used in science), sensory (in contexts that tend to invoke affective-emotional reactions), etc. On a visual level, validity markers are, e.g., colour saturation and differentiation, absence/presence and degree of articulation of the background, degree of representation of detail, etc., and they acquire different values depending on the coding orientation in question.

[10] In the present study, 'reinforcement of address' is not (necessarily) regarded as describing an identical form of address used in the different modes (as in Royce, 2006: 69), but it can account for a strong(er) manifestation in either mode.

[11] Attitudinal congruence (Royce, 2006: 69) is said to depend on modality features (replaced here by the more comprehensive notion of validity) but also on other kinds of attitudes (as expressed, for instance, by attitudinal epithets in the form of subjective adjectives on the linguistic level).

incongruity by emphasising the contrast between what is expressed verbally and visually.

As for the communicative function of humour in the ensemble, which occurs at a rather early stage of the talk, it supports the representation of emotion within the speaker's narrative by setting the scene for the following development of the topic. This instance of humour can indeed be viewed as contributing social management that establishes common ground with the audience by making the speaker look funny (possibly awkward).

4.2. Example 2 – Humour hedging criticism within counter-argumentation

Example 2 comes from the talk entitled 'The game that can give you 10 extra years of life'[12] (length: mins. 19:14), in which game designer Jane McGonigal explains how games can boost resilience in life.

VISUAL	VERBAL TEXT AND OTHER NON-VERBAL CUES
"On your deathbed, are you really going to wish you spent more time playing Angry Birds?" **Visual type**: Scriptural-visual	Now, 21 billion hours, it's a lot of time. It's so much time, in fact, that the number one unsolicited comment that I have heard from people <u>all over the world</u> (*hands part at front*) <u>since I gave that talk</u> (*thumb pointing back at previous slide*), <u>is this</u> (*index pointing towards the audience at front*): Jane, games are great and all, but on your deathbed, are you really going to wish **you spent more time playing Angry Birds?** (simultaneous display with visual) (McGonigal, 01:44, TECH, 2012) (*Audience's audible laughter*) (after the display of the visual)
What is represented: Old-time picture of person on deathbed surrounded by other people, and superimposed animated image of Angry Bird **How it is represented**: No direct gaze; impersonal distance (esp. old-time picture); no frontal perspective (lack of involvement); abstract validity	**What is expressed**: Focus on reported comment, addressed to speaker as addressee – and to the audience of the talk, too **How it is expressed**: Interrogative structure (introduced by deictic gesture)

[12]https://www.ted.com/talks/jane_mcgonigal_the_game_that_can_give_you_10_extra_years_of_life/ (last accessed March 25th 2023).

| **Intersemiotic relations:** |
| Repetition on ideational level |
| Interpersonal level - Direct address via verbal mode; abstract validity (visual mode) and low modality/validity (verbal mode), but some degree of attitudinal dissonance (anachronistic quality of background image adds element of criticism) |

| **Humour causes and function** |
| Incongruity rests on the unexpected juxtaposition of the ideas of death and of playing video games, both referred to verbally, complemented by the improbable visual clash (visually emphasised by anachronistic quality of background image) |
| Communicative function: Hedging the impact of the speaker's criticism – expressed by the anachronistic quality of the background image – of the opponents' viewpoints (i.e., counterargument presentation for the development of the topic) |

Transcription and interpretation of example (2)

As well as the verbal text, the ensemble includes:

- a scriptural-visual type of visual consisting, once again, of images and a verbal component, i.e., a caption, on top of the slide, repeating part of what is said by the speaker;

- three gestures, the last two being deictic ones, as they point to different components of the talk, with a parsing or organising function. Indeed, the different directions of pointing seem to act as stage directions guiding understanding. The last gesture, in particular (*index pointing to the audience at front*), could be viewed as also having a social function, while performing the move of presenting 'an opposing argument' (cf. the content of the unsolicited comment verbally referred to earlier in the ensemble, and then announced by means of a demonstrative pronoun with a cataphoric reference) (on the multifunctional analysis of gestures in the talks, see Masi, 2016).

What is represented visually is more complex than the visual in Example 1. First, we see an old-time image of a person on their deathbed surrounded by other people; a second image is then superimposed on the first, i.e., an animated Angry Bird[13] flapping wings and angrily looking at the characters represented in the first picture in the background. On the verbal level, the words in bold type are from the reported comment, originally addressed to the speaker and then projected onto the audience of the talk, too (cf. the gesture pointing to them), on hypothetical ("wish you spent") and improbable ("are you really") changes to their past lives (seen from a future perspective).

339

---

[13] From Rovio Entertainment's popular mobile video game series (which has spawned movies and merchandise since 2009).

From the point of view of how things are represented, the participants' gaze is not directly addressed to the viewers; in the old-time picture, impersonal distance is conveyed through a long shot, in line with the lack of a frontal perspective and correlated lack of involvement. Low naturalistic validity is at stake, because of the non-realistic cartoon drawing of the bird looming over the characters in the old-time picture and clashing with the more detailed style of the latter as improbable background. Some compositional features add to this non-naturalistic effect, namely the high salience of the full-colour saturation of the Angry Bird (where the red colour ties well with the anger of the character) against the low-saturation, almost monochromatic 'palette' (for antique effect) of the deathbed image in the background. Furthermore, the Angry Bird profile breaks out of the edges of the framed picture in the background and invades its left margin, which seems to hint at the fact that it does not fully belong there. Such features appear to enhance the abstract validity of the images as essential representations of other concepts (i.e., the role of game in life). On the verbal level, the speaker enacts past audience addressing her with a comment on her proposal, expressed through an interrogative structure (a rhetorical question). The speaker and the audience, that is, are led to question the plausibility of an imagined scenario in which they are at the end of their lives and regret not having spent more time playing with video games.

From the point of view of intersemiotic relations, the verbal text is once again crucial to make sense of what is depicted in this complex visual. On the ideational-representational level, the association of 'death' and 'game' is repeated in both modes. On the interpersonal-interactive level, a direct address to the audience (and correlated demand for reaction) especially emerges from the verbal mode (cf. the caption on the slide), while the pictures emphasise the low naturalistic validity (and correlated lack of plausibility) of the coexistence of what they show and convey (e.g., a serious and detached attitude in the old-time image vs. a more jocular attitude in the Angry Bird one), a coexistence also put into doubt verbally (cf. "are you really" in the rhetorical question of the speaker). The verbal and visual modes, however, do not appear to be fully congruent, as the visual mode adds the feature of 'oldness' inherent in the style of the first image. This could be motivated either by the association of old age with the end of life, or by the speaker's intention of expressing criticism of an old-fashioned attitude to the end of life and to life, and its relation with game, generally. Based on this latter interpretation, an element of attitudinal dissonance between modes seems at stake.[14]

As for the causes of humour in the ensemble, incongruity rests on the unexpected juxtaposition of the ideas of death and of playing video games, both referred to verbally, and complemented by the improbable visual clash between what stands for a contemporary video game and an anachronistic portrayal of death,

340

---

[14] In Royce's (2006: 69) framework, attitudinal dissonance consists of an opposite or ironic attitude between the verbal and visual modes. In the example in question, a critical attitude seems to be added by the visual on the basis of the second interpretation of the trait of 'oldness', which does not meet our expectations for a more contemporary type of image portraying the-end-of-life situation.

also suggesting an old-fashioned conception of the way of facing the end of life and life in general. This element of attitudinal dissonance between modes may be viewed as revealing the critical attitude of the speaker towards what she is reporting, which is indeed in line with the fact that the ensemble is part of the move of developing the topic by presenting a counterargument (whose performance is supported by gestures) at the beginning of the talk.

From a functional point of view, then, humour assists in the performance of a potentially face-threatening move in argumentation via hedging the impact of criticism of past and future or potential opponents' views.

4.3. Example 3 – Humour as part of exemplification

The following case (example 3) comes from the talk entitled '3 ways to make better decisions – by thinking like a computer'[15] (length: mins. 11:39), in which cognitive scientist Tom Griffiths shows how the logic of computers can be applied to solve everyday human problems.

VISUAL	VERBAL TEXT AND OTHER NON-VERBAL CUES
The Noguchi filing system **Visual type**: Scriptural-visual	This idea of organizing things so that the things you are most likely to need are most accessible can also be applied in your office. The Japanese economist Yukio Noguchi actually invented a filing system that has exactly this property. **He started with a cardboard box, and he put his documents into the box from the left-hand side. Each time he'd add a document, he'd move what was in there along and he'd add that document to the left-hand side of the box. And each time he accessed a document, he'd take it out, consult it and put it back in on the left-hand side**. As a result, the documents would be ordered from left to right by how recently they had been used. And he found he could quickly find what he was looking for by starting at the left-hand side of the box <u>and working his way to the right</u> (*both hands half raised at front, facing one another, right hand moves rightwards*). (*Pause*) <u>Before you dash home and implement this filing system</u> (*both hands open and apart at front, facing the audience*)-- (*Laughter – smiling face of speaker*) It's worth recognizing that you probably already have. (*Laughter*) (simultaneous with the display of second visual) **That pile of papers on your desk** (*speaker is visible, smiling*) ... typically maligned as messy and

---

[15]https://www.ted.com/talks/tom_griffiths_3_ways_to_make_better_decisions_by_thinking_like_a_computer (last accessed March 25th 2023).

Visual type: Visual

disorganized, a pile of papers is, in fact, <u>perfectly organized</u> (*both hands move forward facing the audience*)— (*Laughter*) (Griffiths, 09:20, TECH, 2018)

What is represented: Visual 1 - Orderly filing method Visual 2 - Untidy pile of papers	**What is expressed**: Description of actions illustrating a process (co-speech simultaneous with Visual 1) Description of Visual 2 follows its display (and correlated laughter): Description of object with diverging attributes
	How it is expressed: Co-speech of Visual 1: Declarative structures in the past, third person subject pronouns (statement of information) + iconic gesture Speech preceding Visual 2: Declarative structure, in the present, which expresses judgements and directly addresses the audience + deictic gesture Speech following Visual 2: Declarative structure in the present with direct address to audience; contrast between negative and positive adjectives + gesture as beat
How it is represented: Visual 1 - Personal/social distance through close/middle shot; objective-scientific perspective; abstract validity Visual 2 - Personal/social distance through close/middle shot; frontal and high angle perspectives (involvement and viewer power); naturalistic validity	

Intersemiotic relations:
Ideational level - repetition (Visual 1) and antonymy (Visual 2)

Interpersonal level - Verbal and visual 'objectivity' and abstract validity (Visual 1 and its co-speech) vs. involvement and naturalistic validity of Visual 2; verbal-visual reinforcement of address (Visual 2 and its co-speech); attitudinal dissonance between Visual 1 (plus its co-speech) and Visual 2

Humour causes and function
Incongruity (esp. in the second and third instances of laughter) rests on the repeated unexpected contradiction between what is shown and what is said First instance of laughter – mainly caused by sudden shift of verbal address (supported by gesture) Second instance of laughter – hinging on the unexpected contradictory sequence of Visual 1 (plus its co-speech) and Visual 2 Third instance of laughter – based on contradiction between negative verbal description of Visual 2 (and Visual 2 itself) and its unexpected verbally expressed positive quality (supported by gesture) Communicative function: ironic use of everyday example increasing audience's familiarity for better understanding, as part of popularising strategy (for the development of the topic)

Transcription and interpretation of example (3)

The ensemble is more extended this time, covering a sequence of three distinct but interrelated instances of laughter and humour. The verbal text is accompanied by two visuals, several gestures, and the speaker's smiling face. Below is a chronological account of the most relevant aspects of the sequence.

Instance of laughter n. 1: The first visual belongs to the scriptural-visual type as it proposes an animated drawing of a box with a verbal caption on top. It shows a well-organised filing method while the speaker verbally illustrates the process of filing. From the point of view of how things are represented, it exhibits a frontal-isometric perspective "reminiscent of the impersonal style of scientific language" (Kress & van Leeuwen, 2021: 142) and abstract validity; the co-speech consists of declarative structures in the past, with third person subject pronouns (as objective statements of information). The description is then supported by the first gesture signalled in the transcription (*both hands half raised at front, facing one another, right hand moves rightwards*), which is iconic, in that it imitates part of the process that is being illustrated verbally. After a brief pause, a second gesture follows, with a more deictic quality and social function (*both hands open and apart at front, facing the audience*), as it deflects the attention from the referential content (what is being illustrated) to the interaction of the speaker with the audience. This unexpected, sudden shift in address (verbally paralleled by an indirect directive act not to "dash home and implement" the filing method just illustrated) indeed generates the first instance of laughter.

Instance of laughter n. 2: Visual n. 2 comes into play, i.e., a display of a photo (which belongs to the only visual semiotic type), simultaneously with a second instance of laughter and with part of its verbal description. The display is preceded by verbal text (a declarative structure, in the present, which expresses judgements and directly addresses the audience – "it's worth", "you probably already have") which sets expectations for an orderly process, also crucially hinging on what represented on visual n. 1 and its co-speech. Yet, such expectations are not met by what is shown visually this time. Opposite experiential meaning is sequentially conveyed by different modes on an ideational-representational level (tidiness vs. untidiness) and this unexpected 'antonymy'[16] is what triggers the second instance of laughter. The photo in fact provides a close/middle shot of an untidy pile of paper from frontal and high angle perspectives, scoring high in involvement and viewer power,[17] respectively. Involvement is also promoted verbally, by the preceding shift in address and by the concomitant "That pile of papers on your desk". The realistic scenario embodies a high degree of naturalistic validity. Such interpersonal-interactive features emphasise the 'closeness' of the audience to an example taken from their everyday life. The contradiction arising from the sequence of the first visual (and its verbal description), with its abstract scientific objectivity depicting

343

---

[16] For the intersemiotic sense relation of antonymy, see Royce (2006: 68).
[17] If what is represented is seen from a high angle, the relation between the viewer and what is represented (cf. the represented participant) is depicted as one in which the viewer "has power over the represented participant" (Kress & van Leeuwen, 2021: 138).

order, followed by the second visual, with a naturalistic and familiar scenario portraying 'imperfection', indeed conveys attitudinal dissonance in terms of an ironic opposition.

Instance of laughter n. 3: The words following visual n. 2 spell out the evident negative qualities ("typically maligned as messy and disorganized") associated with the pile of paper taken as an example. Then, a third gesture (*both hands move forward facing the audience*) appears to have an emphatic function, a sort of beat that enhances the salience of the presumable "perfectly organised" quality verbally added to the description of the pile. In so doing, the gesture places emphasis on the contradictory nature (also cf. "in fact") of what is shown on the second visual, hence the laughing reaction (third instance) of the audience.

Overall, the three steps in the ensemble appear to be largely interdependent in the generation of humour, and incongruity (esp. in the second and third instances of laughter) rests on the repeated unexpected contradiction between what is shown and what is said.

From a functional point of view, the humour associated with the first instance of laughter seems to reinforce common ground, also setting the stage for the following instances. The latter, instead, contribute to the illustrative (i.e., popularising) function of the passage by ironically making reference to an example that is likely to be familiar to the audience,[18] within the context of the description of a process (for the development of the main topic of the talk).

5. CONCLUDING REMARKS

In a recent TED talk on 'The three magic ingredients of amazing presentations',[19] Waknell (2020) said that we forget most of what we hear very quickly, and suggested that visual elements, too, should be used in effective presentations. Despite the limited size of the corpus, the present study has indeed highlighted the crucial role of visuals in the multimodal generation of humour in a selection of TED talks, especially from the domain of Technology, while fewer cases were found in the talks from the domains of Economics and Law. The study has then shown varied subtle synergies among words, visuals, and gestures in some examples of ensembles. The last mentioned, all from the domain of Technology, were chosen on account of their representativeness in terms of richness in the semiotic resources used for humour generation, and variety of functions of such humorous instances in the talks.

---

[18] Popularising texts in general tend to refer to the knowledge background and common life experiences of the lay reader (see, e.g., Gotti, 1996; Myers, 1989). For further examples of visuals in the talks contributing to popularisation via audience-oriented recontextualisation and humour, see Masi (in press).

[19] https://www.ted.com/talks/phil_waknell_the_3_magic_ingredients_of_amazing_presentations (last accessed March 25th 2023).

In the data under analysis, humour was often the result of a cumulative effect that depended on what and how entities were represented visually and expressed verbally, also with the support of gestures (or other non-verbal cues such as facial expressions). The complexity of the intersemiotic relations at stake emerged from the multimodal transcriptions proposed – open to refinements – covering a large number of mainly ideational-representational and interpersonal-interactive features in the ensembles, and helping to identify different types of incongruities at the basis of humorous episodes, together with their functional interpretation.

The ensembles were found at various points of the talks, and appeared not only as asides – e.g., within personal narratives – managing social relations with the audience, but also intertwined with argumentation and strategies of popularisation for the development of the main topics of the talks. The contribution of humour to popularisation, in particular, was probably the most surprising aspect that emerged from the present account (once again especially prominent in the sample from the Technology domain), which surely deserves further attention. Several cases in the data in fact showed humorous multimodal ensembles as part of explanatory strategies, which appeared to have a great potential for enhancing the involving and enjoyable quality of such pivotal moments of popularisation, so that understanding of specialised notions can successfully take place.

If confirmed by future investigation, the role and functions of multimodally orchestrated humour in the talks could be profitably exploited in ESP settings. Indeed, critical multimodal analysis (García-Pinar, 2019a) of humorous ensembles from the talks could inform the design of ESP pedagogical materials and instruction for a deeper understanding of humour mechanisms, [20] and greater mastery of multimodal literacy skills and ESP knowledge. More specifically, such materials could be used not only to stimulate students' ability to notice and explore the dynamics of humorous multimodal interplays, but also to expose students to cases where multimodally generated humour is used to bridge knowledge gaps between experts and non-experts in ways that are more likely to be conducive to learning through engagement, thanks to their enjoyable nature and possible positive influence on memorability. For these same reasons, humorous ensembles in a web genre freely available online like TED talks could also have a positive effect, in terms of interest, motivation and learning outcomes, on EFL extramural contact. Further research based on more data (from the same and other domains) and further development of methodology are obviously necessary to corroborate and expand on the present findings and proposals.

[20] Including a discussion of culture-specific constraints that might be involved in humour generation and understanding.

References

Attardo, S. (1994). *Linguistic theories of humor.* Mouton de Gruyter.
https://doi.org/10.1515/9783110219029

Attardo, S. (2020). *The linguistics of humor: An introduction*. Oxford University Press.

Caliendo, G., & Compagnone, A. (2014). Expressing epistemic stance in university lectures and TED talks: A contrastive corpus-based analysis. *Lingue e Linguaggi*, *11,* 105–122. https://doi.org/10.1285/i22390359v11p105

Calsamiglia, H., & Van Dijk, T. A. (2004). Popularization discourse and knowledge about the genome. *Discourse and Society*, *15*(4), 369–389. https://doi.org/10.1177/0957926504043705

Carney, N. (2014). Is there a place for instructed gesture in EFL? In N. Sonda & A. Krause (Eds.), *JALT 2013 Conference proceedings* (pp. 413–421). JALT. https://jalt-publications.org/sites/default/files/pdf-article/jalt2013_041.pdf

Chang, Y., & Huang, H. (2015). Exploring TED talks as a pedagogical resource for oral presentations: A corpus-based move analysis. *English Teaching & Learning*, *39*(4), 29–62. https://doi.org/10.6330/ETL.2015.39.4.02

Compagnone, A. (2014). Knowledge dissemination and environmentalism: Exploring the language of TED talks. In E. Chiavetta, S. Sciarrino, & C. Williams (Eds.), *Popularisation and the media* (pp. 7–25). Edipuglia.

Crawford Camiciottoli, B. (2021). The multimodal expression of humour in university lectures: Some insights for ESP. In J. Eveno, V. Favre, V. Guignery, N. Lasorak, H. Lecomte, R. Marcon, & M. Rabecq (Eds.), *15th ESSE Conference programme and Book of abstracts* (p. 48). Université Claude Bernard Lyon 1. https://media.univ-lyon1.fr/2020/esse2020/ESSE-book-of-abstracts.pdf

Dummett, P., Stephenson, H., & Lansford, L. (2016). *Keynote*. National Geographic Learning, CENGAGE Learning. http://www.eltkeynote.com/

García-Pinar, A. (2019a). Getting closer to authenticity in the course of technical English: Task-Based instruction and TED talks. *English Language Teaching*, *12*(11), 10–22. https://doi.org/10.5539/elt.v12n11p10

García-Pinar, A. (2019b). The influence of TED talks on ESP undergraduate students' L2 motivational self system in the speaking skill: A mixed method study. *ESP Today*, *7*(2), 231–253. https://doi.org/10.18485/esptoday.2019.7.2.6

García-Pinar, A., & Pallejá-López, C. (2018). TED talks: A multimodal tool for students of technological English. *Docencia e Investigatión*, *29*, 6–24.

Gotti, M. (1996). Il linguaggio della divulgazione: problematiche di traduzione interlinguistica [The language of knowledge dissemination: Problems of interlinguistic translation]. In G. Cortese (Ed.), *Tradurre i linguaggi settoriali* (pp. 217–235). Edizioni Libreria Cortina.

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

Harrison, S. (2021). *Showing* as sense-making in oral presentations: The speech-gesture-slide interplay in TED talks by Professor Brian Cox. *Journal of English for Academic Purposes*, *53,* Article 101002. https://doi.org/10.1016/j.jeap.2021.101002

Hasan, M. K., Rahman, W., Zadeh, A. B., Zhong, J., Tanveer, M. I., Morency, L-P., & Hoque, M. E. (2019). UR-FUNNY: A multimodal language dataset for understanding humor. In K. Inui, J. Jiang, V. Ng, & X. Wan (Eds.), *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on*

Natural Language Processing (EMNLP-IJCNLP) (pp. 2046–2056). Association for Computational Linguistics. https://aclanthology.org/D19-1211.pdf

Jiang, J., & Lim, F. V. (2022). Popularizing science: Analyzing the presenter's multimodal orchestration in a TED talk. *Ibérica*, *44*, 179–206. https://doi.org/10.17398/2340-2784.44.179

Kress, G. (2010). *Multimodality: A social semiotic approach to contemporary communication* (2nd ed.). Taylor and Francis.

Kress, G., & van Leeuwen, T. (1996). *Reading images: The grammar of visual design* (1st ed.). Routledge.

Kress, G., & van Leeuwen, T. (2021). *Reading images: The grammar of visual design* (3rd ed.). Routledge.

Lazaraton, A. (2004). Gesture and speech in the vocabulary explanations of one ESL teacher: A microanalytic inquiry. *Language Learning*, *54*(1), 79–117. https://doi.org/10.1111/j.1467-9922.2004.00249.x

Masi, S. (2016). Gestures in motion in TED talks: Towards multimodal literacy. In V. Bonsignori & B. Crawford Camiciottoli (Eds.), *Multimodality across communicative settings, discourse domains and genres* (pp. 146–165). Cambridge Scholars Publishing.

Masi, S. (2019). Complex mapping of words and gestures in TED talks. In V. Bonsignori, G. Cappelli, & E. Mattiello (Eds.), *Worlds of words: Complexity, creativity, and conventionality in English language, literature and culture* (pp. 115–126). Pisa University Press.

Masi, S. (2020a). The multimodal representation of 'ideas worth spreading' through TED talks. *Lingue e Linguaggi*, *36*, 155–171. https://doi.org/10.1285/i22390359v36p155

Masi, S. (2020b). Exploring meaning making practices via co-speech gestures in TED talks. *Journal of Visual Communication*, *39*(3-4), 201–219. https://doi.org/10.1080/1051144X.2020.1826223

Masi, S. (in press). Disseminating knowledge through TED talks for children. *Token: A Journal of English Linguistics*, *15*.

Mattiello, E. (2017). The popularisation of science via TED talks. *International Journal of Language Studies*, *11*(4), 69–98.

McNeill, D. (1992). *Hand and mind: What gestures reveal about thought*. University of Chicago Press.

Meza, R., & Trofin, C. (2015). Between science popularization and motivational infotainment: Visual production, discursive patterns and viewer perception of TED talks videos. *Studia Universitatis Babes-Bolyai: Ephemerides*, *60*(2), 41–60.

Myers, G. (1989). Science for women and children: The dialogue of popular science in the nineteenth century. In J. R. R. Christie & S. Shuttleworth (Eds.), *Nature transfigured* (pp. 171–200). Manchester University Press.

Pavesi, M., & Ghia, E. (2020). *Informal contact with English: A case study of Italian postgraduate students*. Edizioni ETS.

Peruzzo, K. (2021). Who says that talking about depression isn't fun? Exploring humour in TED talks on mental disorder. In J. Eveno, V. Favre, V. Guignery, N. Lasorak, H. Lecomte, R. Marcon, & M. Rabecq (Eds.), *15th ESSE Conference programme and Book of abstracts* (p. 51). Université Claude Bernard Lyon 1. https://media.univ-lyon1.fr/2020/esse2020/ESSE-book-of-abstracts.pdf

Raskin, V. (1985). *Semantic mechanism of humor*. Reidel.

Royce, T. D. (2006). Intersemiotic complementarity: A framework for multimodal discourse analysis. In T. D. Royce & W. Bowcher (Eds.), *New directions in the analysis of multimodal discourse* (pp. 63–109). Routledge.

Ruiz-Madrid, M. N., & Fortanet Gómez, I. (2015). A multimodal discourse analysis approach to humour in conference presentations: The case of autobiographic references. *Procedia: Social and Behavioral Sciences*, *173*, 246–251. https://doi.org/10.1016/j.sbspro.2015.02.060

Scotto di Carlo, G. (2013). Humour in popularisation: Analysis of humour-related laughter in TED talks. *European Journal of Humour Research*, *1*(4), 81–93. https://doi.org/10.7592/EJHR2013.1.4.scottodicarlo

Scotto di Carlo, G. (2014). The role of proximity in online popularizations: The case of TED talks. *Discourse Studies*, *16*(5), 591–606. https://doi.org/10.1177/1461445614538565

Takaesu, A. (2013). TED talks as an extensive listening resource for EAP students. *Language Education in Asia*, *4*(2), 150–162. https://doi.org/10.5746/LEiA/13/V4/I2/A05/Takaesu

Theunissen, G. (2014). *Analysis of the visual channel of communication in a corpus of TED talks presentations* [Unpublished master's thesis]. Ghent University. https://libstore.ugent.be/fulltxt/RUG01/002/162/185/RUG01-002162185_2014_0001_AC.pdf

Valeiras-Jurado, J. (2017). *A multimodal approach to persuasion in oral presentations: The case of conference presentations, research dissemination talks and product pitches* [Unpublished doctoral dissertation]. Universitat Jaume I/Ghent University.

Valeiras-Jurado, J., & Ruiz-Madrid, N. (2019). Multimodal enactment of characters in conference presentations. *Discourse Studies*, *20*(5), 561–583. https://doi.org/10.1177/1461445619846703

Waknell, P. (2012, March 12). How to make a great TED talk. *Phil Presents*. https://philpresents.wordpress.com/2012/03/12/how-to-make-a-great-ted-talk/

Waknell, P. (2020, January 28). *The three magic ingredients of amazing presentations* [Video]. TED Conferences. https://www.ted.com/talks/phil_waknell_the_3_magic_ingredients_of_amazing_presentations

Wingrove, P. (2017). How suitable are TED talks for academic listening? *Journal of English for Academic Purposes*, *30,* 79–95. https://doi.org/10.1016/j.jeap.2017.10.010

Wu, A., & Qu, H. (2020). Multimodal analysis of video collections: Visual exploration of presentation techniques in TED talks. *IEEE Transactions on Visualization and Computer Graphics*, *26*(7), 2429–2442. https://doi.org/10.1109/TVCG.2018.2889081

Xia, S. (2023). Explaining science to the non-specialist online audience: A multimodal genre analysis of TED talk videos. *English for Specific Purposes*, *70*, 70–85. https://doi.org/10.1016/j.esp.2022.11.007

348

SILVIA MASI is Associate Professor of English Language and Linguistics in the Department of Philology, Literature and Linguistics of the University of Pisa. Her scientific interests lie in the fields of lexical semantics, pragmatics, discourse analysis, scientific popularisation and translation. More recent research centres on the role of multimodality in the popularisation and translation of non-fiction for children and the development of multimodal literacy in and beyond ESP teaching.