

A Multimodal Discourse-Analytic Approach to Incongruity in Short-Form Video

[Un Approccio Multimodale di Analisi del Discorso all'Incongruità nei Video Brevi]

Audrey C. Willoughby*¹

¹ *Università degli Studi di Milano*

*E-mail: audrey.willoughby@unimi.it

Original article

Ricevuto il 4 dicembre 2025; accettato il 4 febbraio 2026

ABSTRACT

IT L'umorismo nei media audiovisivi brevi è sempre più prodotto attraverso l'uso coordinato di molteplici risorse semiotiche che si sviluppano nel tempo, piuttosto che attraverso il solo linguaggio verbale. Sebbene le teorie pragmatiche dell'umorismo spieghino la violazione delle aspettative e la reinterpretazione inferenziale, esse offrono strumenti metodologici limitati per analizzare come questi processi siano organizzati temporalmente nel discorso audiovisivo multimodale. Questo articolo sviluppa e illustra quindi un approccio di analisi del discorso multimodale all'incongruità nei video brevi. L'approccio procede attraverso due passaggi analitici. In primo luogo, i video vengono segmentati in fasi interpretativamente stabili sulla base di cambiamenti osservabili nei segnali semiotici, quali testo, sguardo, gesti, suono e montaggio. In secondo luogo, le occorrenze di incongruità vengono analizzate distinguendo tra una cornice interpretativa iniziale (Meaning 1) e una successiva interruzione o reinterpretazione (Meaning 2), identificando le modalità coinvolte in ciascuna fase. L'articolo dimostra l'applicazione dell'approccio ai video TikTok di tipo point-of-view (POV) e propone un quadro trasparente per l'analisi dell'umorismo nel discorso digitale multimodale.

Parole chiave: Multimodal discourse analysis, incongruity, short-form video, audiovisual discourse, POV genre

EN Humour in short-form audiovisual media is increasingly produced through the coordinated use of multiple semiotic resources unfolding over time rather than through language alone. While pragmatic theories of humour account for expectation violation and inferential reinterpretation, they offer limited methodological tools for analysing how these processes are organised temporally within multimodal audiovisual discourse. This paper develops and illustrates a multimodal discourse-analytic approach to incongruity in short-form video. The approach proceeds in two analytical steps. First, videos are segmented into interpretively stable phases based on observable shifts in semiotic cues such as text, gaze, gesture, sound, and editing. Second, instances of incongruity are analysed by distinguishing between an initial interpretive frame (Meaning 1) and a subsequent disruption or reinterpretation (Meaning 2), identifying the modes involved in each phase. The paper demonstrates the approach using TikTok point-of-view (POV) videos and proposes a transparent framework for analysing humour in multimodal discourse.

Keywords: Multimodal discourse analysis, incongruity, short-form video, audiovisual discourse, POV genre

Introduction: Humour in audiovisual material

Humour research has long emphasised that humorous effects emerge through the disruption of expectations and the inferential work this disruption invites. Across linguistic and pragmatic approaches, humour is understood as a process in which recipients recognise an apparent interpretive problem as intentional and playful, and revise their understanding accordingly (Attardo, 2020; Tsakona, 2020; Yus, 2021). From this perspective, humour is not a property of linguistic material alone, but the outcome of interpretive activity carried out by recipients in context.

Despite this emphasis on interpretation, much humour research has remained centred on language-based material, such as jokes, conversational humour, scripted dialogue, and written texts (e.g. Raskin, 1985; Attardo, 1994, 2001). Even when humour is treated as inferential and context-dependent, analytical attention has tended to focus on verbal triggers and linguistic mechanisms. As a result, non-linguistic semiotic resources are often treated as secondary or illustrative rather than as integral components of humour production. More recent work has increasingly argued that such approaches risk overlooking the multimodal reality of humour in interaction and digital communication (e.g., Bernad-Mechó & Girón-García, 2023; Gironzetti, 2022).

Researchers have begun to address humour as a multimodal phenomenon (Priego-Valverde, 2024), particularly through the study of static image–text formats such as memes and image macros. Work on visual–verbal humour has shown how incongruity can arise from the juxtaposition of captions and images that frame situations in conflicting or unexpected ways (Dynel, 2016; Abdel-Raheem, 2018; Yus, 2018; Vázquez & Aslan, 2021). These studies demonstrate that humour can be distributed across semiotic resources and that tension between modes can play a central role in humorous interpretation.

However, as mentioned above, this research has largely focused on static or spatially fixed compositions. Because memes and image macros present their cues simultaneously within a single visual frame, the organisation of meaning remains stable over time. While viewers may shift their attention between modes, the material structure of the text does not unfold sequentially. As a result, these approaches offer limited tools for describing how expectations are progressively established, maintained, and later reconfigured as new cues appear, as is typical in audiovisual material such as video.

This limitation becomes increasingly apparent in contemporary humour practices. A substantial share of everyday humour now circulates through short-form audiovisual platforms such as TikTok and Instagram, where meaning is produced through tightly coordinated combinations of written text, moving images, sound, gesture, facial expression, performance, and editing. These videos are brief, fast-paced, and highly reliant on platform-specific conventions such as audio templates, textual overlays, and recognisable performance formats. In this environment, humour often emerges through shifts and tensions across modes over time rather than through a single verbal punchline, defined as climactic

humorous turns typically positioned at the end of a text (Attardo, 2001) or jablines, defined as humorous cues distributed earlier or recurrently across a discourse without closing it (Tsakona 2003).

Some studies have explored humour multimodally in audiovisual contexts, including work on gesture, performance, and visual design (Balirano & Corduas, 2008; Bernad-Mechó & Girón-García, 2023; Masi, 2023). While this research offers valuable insights into how different semiotic resources contribute to humour, it tends to focus on general patterns of coordination or on isolated humorous moments. Questions about how humorous interpretation develops dynamically over the course of a short video therefore remain relatively underexplored.

As mentioned above, short-form audiovisual videos pose a particular challenge for humour studies because they combine dense cueing with rapid temporal progression. Viewers are required to form an initial interpretation very quickly and to revise it just as quickly in response to later cues. Although existing humour theories, such as incongruity-based and pragmatic approaches, account well for expectation violation and reinterpretation in principle (e.g. Raskin, 1985; Attardo, 1994, 2020; Tsakona, 2020), they provide few tools for describing how these processes are guided by the unfolding organisation of audiovisual discourse across modes.

The approach discussed here addresses this gap by examining humour in short-form video through a perspective that integrates pragmatic accounts of humour with multimodal discourse analysis. Focusing on TikTok point-of-view (POV) videos, it seeks to describe how expectations are established, challenged, and reinterpreted through the temporal organisation of semiotic resources in contemporary audiovisual media.

1. Genre and expectation in TikTok POV videos

Humour in short-form video is strongly shaped by genre conventions and platform-specific formats. On TikTok, recurring video types function as recognisable templates that help viewers quickly orient to what kind of communicative event is taking place. These formats rely on repetition with variation, a logic that closely parallels earlier accounts of memetic humour, where humour emerges not from isolated texts but from the audience's familiarity with shared structures and conventions (Shifman, 2013; Zulli & Zulli, 2020). This logic is realised through short audiovisual performances, as creators reproduce familiar formats through a combination of visual, linguistic, and embodied resources, as well as timing and sequencing.

The POV genre is a clear example of how these genre conventions operate. POV videos position the viewer within a fictionalised scenario or social role, typically through explicit framing devices that establish perspective, relationship, or situation at the outset. Textual overlays such as "POV: you forgot your homework" or "POV: me to my man after a fight" act as condensed genre signals, invoking familiar

interactional scripts and social types that viewers are expected to recognise immediately (Trillò, 2024). In this way, POV videos function as a form of staged perspective-taking, where humour depends on the rapid activation of shared cultural knowledge.

These textual overlays do more than simply describe a situation. They operate as interpretive instructions that guide how the unfolding audiovisual performance should be understood. By assigning the viewer a role and invoking a recognisable scenario, the overlay establishes a relatively stable baseline for interpretation, shaping expectations about behaviour, stance, and likely developments (Willoughby, forthcoming). This early framing is crucial for how expectation and violation are organised in the POV genre, as it creates the conditions under which later deviations can be recognised as intentional and meaningful.

Expectation-setting in POV humour, however, is not confined to language alone, nor is it evenly distributed across modes. While textual overlays frequently establish the initial scenario, role, or premise, this framing is elaborated and stabilised through visual composition, setting, facial expression, gesture, and platform-specific sound templates. From a multimodal perspective, meaning in audiovisual texts emerges through the coordinated organisation of semiotic resources over time rather than through language in isolation (Bateman et al., 2017; Wildfeuer, 2014). In the 16 POV videos analysed by Willoughby (forthcoming) linguistic framing most often functions as the entry point for initial interpretation, while embodied and visual cues play a decisive role in reshaping or undermining that premise. Platform conventions such as recognisable audio templates and synchronised sound further contribute to this process by anchoring genre expectations that can later be playfully violated, positioning viewers' inferential work as central to the humorous effect.

Humour can emerge when later audiovisual cues fail to align with those initial established expectations. Rather than confirming the earlier understanding, later cues may introduce information that clashes with the expectations, prompting reinterpretation. As shown in Willoughby (forthcoming), humour in this genre relies less on discrete verbal punchlines or jablines and more on modal interplay, in which meaning established in one mode is undermined or reconfigured by another. In this specific genre, embodied modes and sound frequently carry this disruptive function, while expectation-setting tends to occur more often in linguistic and visual modes such as textual overlays, which is most often the case, or verbal speech (Willoughby, forthcoming)

Because POV videos foreground perspective and expectation so explicitly, they offer particularly analytically accessible material for observing humour as a process of expectation-building followed by reinterpretation. The genre makes visible how humorous meaning is organised across modes and over time, and how shared genre knowledge allows viewers to recognise deviations as playful rather than accidental. In this respect, POV humour exemplifies broader patterns of digital humour, where repetition,

genre literacy, and multimodal design jointly structure humorous interpretation (Shifman, 2013; Zulli & Zulli, 2020).

2. Analysing incongruity multimodally in audiovisual discourse

The analytical approach discussed here builds on the assumption outlined above: in short-form audiovisual media, humour does not appear all at once but unfolds over time through the coordinated use of multiple semiotic resources. To examine this unfolding process, this method focuses on how videos guide viewers toward particular expectations and then prompt them to revise those expectations when later cues disrupt the initial understanding.

3.1 Segmentation

The first step in the analysis is to divide each video into segments. Segmentation in this approach is not based on fixed time intervals or formal units such as scenes; instead, segments are identified at points where a noticeable shift occurs in how the viewer is likely to interpret what is happening, prompted by changes in text, gaze, facial expression, gesture, sound, editing, or camera framing (Wildfeuer, 2014).

Segmentation and annotation were carried out manually using Microsoft Excel. While this approach does not offer automated temporal alignment (such provided by tools such as ELAN), it allows transparent control over interpretive segmentation decisions and is readily accessible for students and early-career researchers. Each segment corresponds to a phase in which a relatively stable interpretation holds, until new semiotic cues invite the viewer to update or reconsider their understanding. This approach makes it possible to trace how meaning develops step by step over time, reflecting the way viewers perceive boundaries in unfolding audiovisual events rather than treating the video as a single undifferentiated whole (Wildfeuer, 2014; Zacks & Swallow, 2007).

3.2 Expectations, violation, and guided inference

Once a video has been segmented, the analysis proceeds by distinguishing between two interpretive phases for each instance of incongruity: Meaning 1 and Meaning 2. A video can have one or multiple observable incongruities.

Meaning 1 refers to the initial premise that the video constructs that sets up what will later be understood as part of an incongruous interpretation. This is the expectation established through one or more modes, such as a textual overlay that defines a scenario, a tone of voice that signals seriousness, or visual and embodied cues that establish a familiar role or situation.

Meaning 2 refers to the alternative interpretive contribution introduced by later semiotic cues, which disrupts, destabilises, or reconfigures the expectation established earlier. This occurs when one or more modes provide information that no longer aligns with the initial premise, making reinterpretation necessary.

An incongruity is identified when Meaning 2 does not align with Meaning 1 and requires the viewer to reinterpret what they have already seen. This reinterpretation is not treated as a psychological process to be measured, but as an interpretive outcome made available by the organisation of semiotic cues in the video. In this sense, inference is guided by the material itself: the text supports a particular interpretive outcome and then viewers are encouraged to revise it when new cues make the earlier interpretation untenable or incomplete.

3.3 Cross-modal incongruity

As discussed above, incongruity in this framework is treated as a relation between an initial interpretation made available by the video and a later disruption of that interpretation. In many cases, the semiotic resources that establish the initial interpretation are not the same resources that destabilise it. To capture how different modes contribute at different points in the interpretive process, the analysis distinguishes cross-modal incongruity from incongruity that is realised primarily within a single mode.

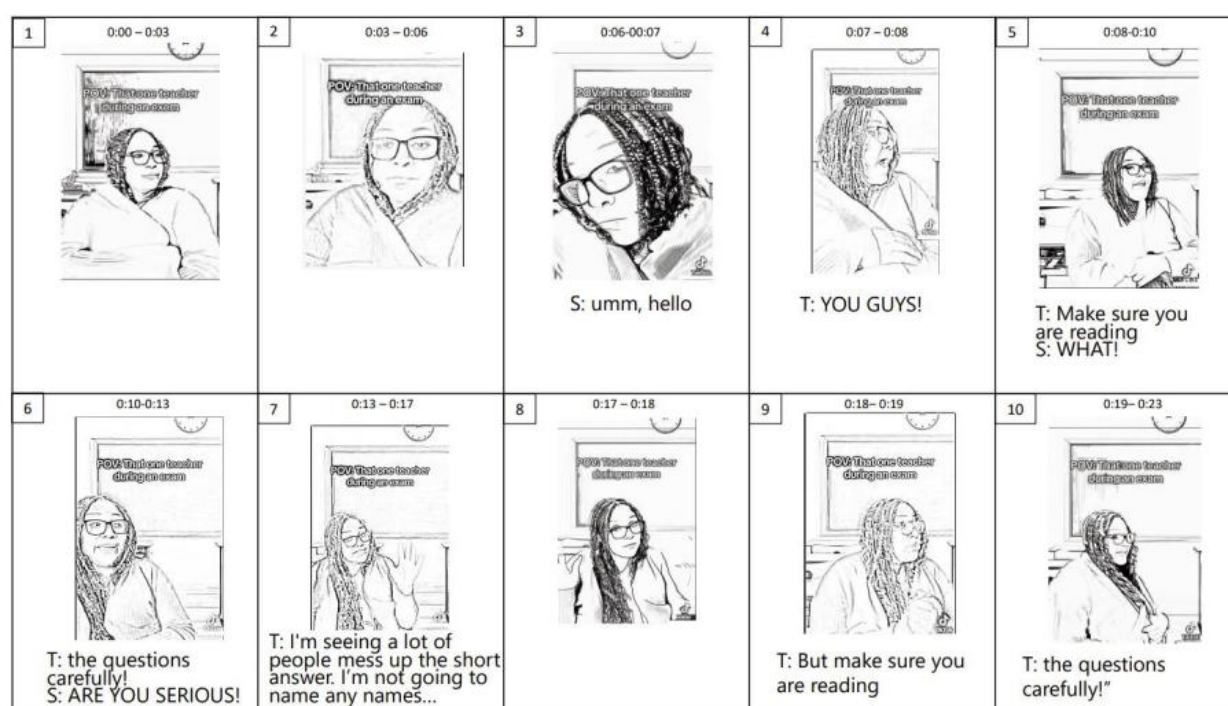
Cross-modal incongruity occurs when Meaning 1 is established predominantly through one set of semiotic resources, while Meaning 2 is introduced predominantly through another. This distinction is not simply descriptive. It allows the analyst to specify where the video places the work of expectation-setting and where it places the work of violation. For instance, a textual overlay may instruct the viewer to adopt a serious scenario, while facial expression or gaze renders that scenario implausible; alternatively, a calm verbal delivery may be undermined by an audio template that signals an ironic interpretation. In such cases, the tension is not located in any single element in isolation, but in the temporal interaction between cues that initially support different interpretations and invite reinterpretation.

Analytically, identifying cross-modal incongruity involves three moves. First, the analyst specifies which mode or combination of modes carries the primary expectation-setting function in the relevant segment(s), that is, which resources make Meaning 1 available and stabilise it as a plausible reading. Second, the analyst identifies the moment and the mode or modes that introduce Meaning 2, focusing on the cue that renders the earlier interpretation untenable, incomplete, or in need of revision. Third, the analyst describes the relationship between these contributions, indicating whether the later cue contradicts the earlier interpretation, recontextualises it, or exposes it as performative, exaggerated, or misaligned with the implied social script.

By organising analysis around these relations, cross-modal incongruity treats humour as a temporally organised process rather than as something located in individual semiotic elements taken out of context. This allows the analyst to specify which semiotic resources establish an initial interpretation and which introduce cues that require it to be revised. In this way, cross-modal incongruity offers a precise account of how short-form videos guide inferential revision through modal interplay over time.

3. Example of application to a TikTok video

Figure 1. *Video 001*



Note. Screenshots from TikTok POV video “POV: That one teacher during an exam”, illustrating segmentation and cross-modal incongruity. Reproduced from Willoughby (forthcoming).

Figure 1 presents screenshots corresponding to each analytically identified segment of the TikTok POV video 001 from Willoughby (forthcoming). The purpose of the figure is not to illustrate isolated moments of humour, but to provide visual support for the segmentation process and the step-by-step interpretation that follows from it.

Figure 1 and the following discussion provide an illustration of how the proposed analytical approach can be applied to short-form video to trace interpretive development and identify a violation of expectation. The video, titled “POV: That one teacher during an exam,” features a single content creator

enacting a teacher through coordinated use of textual overlay, embodied conduct, gaze, and speech. The analysis proceeds by segmenting the video and examining how expectations are established, elaborated, and later violated through modal interplay.

Segments 1 and 2 are treated as separate analytical units because each introduces a perceptible change in semiotic organisation that contributes new information to the unfolding interpretation. In Segment 1, the textual overlay and the teacher's silent monitoring behaviour establish a recognisable institutional scenario and social role. Segment 2 elaborates upon this initial interpretation through intensified embodied conduct, including increased surveillance, controlled closed-off movement, and facial cues that heighten the teacher's evaluative stance. Although the core scenario remains unchanged, the shift in bodily behaviour justifies a new segment boundary because it expands what kind of authority figure is being enacted. Note that this elaboration remains fully compatible with the initial interpretation and does not introduce incongruity.

Segment 3 introduces a further development through a change in gaze orientation. The teacher's direct gaze toward the camera, which the viewers understand to be the student's point-of-view, intensifies the enactment of individualised scrutiny without altering the underlying scenario. This segment continues to contribute to the initial interpretive development through narration by refining the teacher–student dynamic.

Segments 4 through 6 mark additional interpretive shifts prompted by the introduction of spoken language and changes in address. The teacher moves from silent observation, in which she takes a moment to review the POV student's exam, to then publicly addressing the class with a warning to double check their work through verbal instructions, lecturing gestures and a mildly scolding tone. These shifts justify further segmentation because they reorganise the interactional structure of the scene, but they continue to reinforce a consistent understanding of the teacher as a strict and authoritative figure. At this stage, segmentation allows the analyst to capture how expectations are progressively built and stabilised across multiple modes and segments.

It isn't until Segment 7 that the video introduces Meaning 1 for the purpose of subsequent violation. Here, during the ongoing verbal warning to the class to double check their work, the teacher states "I'm not gonna name names," accompanied by open-palmed gestures that conventionally signal restraint and impartiality. This verbal–gestural combination establishes a specific, short-term expectation about the teacher's conduct, separate from the more general portrayal of her as a strict authority figure developed earlier in the video.

Meaning 2 is then realised in Segment 8 through an embodied cue that conflicts with Meaning 1. While maintaining the verbal posture of neutrality, the teacher makes prolonged and deliberate eye contact with the camera, or the implied first person in this scenario, thereby implicitly identifying the student's work who initiated the public scolding as our POV student. In other words, Meaning 1 assured

confidentiality through verbal means while Meaning 2 identifies the culprit through embodied cues. This moment qualifies as cross-modal incongruity because the expectation and its violation are realised in different modes and, in this case, across adjacent segments.

It is important to note that this incongruity is not resolved explicitly within the narrative. From a pragmatic perspective, however, incongruity does not require full resolution in order to produce humour. As Attardo (1994, 2020) and Hempelmann and Attardo (2011) argue, humorous resolution is characteristically partial: logical mechanisms mask or shift incongruity without eliminating it entirely, leaving residual tension that sustains the humorous effect.

In subsequent segments, the teacher resumes general classroom address without acknowledging the breach. The method therefore allows the analyst to describe humour as emerging from guided inferential reinterpretation supported by the temporal organisation of semiotic cues, rather than from an explicit punchline or narrative repair. Due to space constraints, the present paper focuses on illustrating the analytical procedure rather than offering a full distributional account of incongruity across the dataset; additional analytical dimensions are discussed in more detail in Willoughby (forthcoming).

4. Potential and directions for future research

By focusing on how expectations and violations are organised across semiotic resources and over time, the approach outlined in this paper enables researchers to ask precise questions about the mechanics of meaning-making in short-form audiovisual genres. In the case of humour, it allows analysts to investigate which semiotic resources are most involved in setting up expectations within a given genre, how violations are timed in relation to earlier framing cues, and whether different formats rely on early or delayed incongruity. The method also supports the analysis of more complex humorous structures, including videos in which expectations are challenged across multiple segments, requiring viewers to repeatedly adjust their interpretation as the video unfolds, as well as the ways in which remix and reuse practices shift the location and communicative function of incongruity in platformed contexts.

Although the present paper centres on humour, the analytical logic is not restricted to comic effects and irony. The same segmentation-based procedure could be applied to, for example, persuasion or emotional framing in short-form video. In such cases, what changes is not the method itself, but the interpretive question guiding the analysis. Holding the analytic procedure constant while shifting the interpretive focus makes it possible to compare how different communicative effects are organised through similar multimodal resources and temporal structures.

A central methodological strength of this approach is that it remains closely tied to the material while still engaging with broader theoretical concerns. Rather than beginning with abstract categories or predefined humour types, it starts from observable shifts in the audiovisual text and works outward

toward interpretation. This makes the approach both analytically rigorous and pedagogically accessible. For students and early-career researchers in particular, it offers a transparent sequence of analytic steps: segmenting the video, identifying framing cues, tracing expectations and violations, and relating these patterns to genre conventions, all while still asking for reflection on inference and context.

This method could be taken into several different directions of future research. Larger datasets could be used to test whether the patterns observed specific genres extend to other TikTok genres or to short-form video trends on different platforms. Cross-cultural studies could examine how expectations and violations draw on distinct social scripts and shared knowledge. Audience-oriented methods, such as eye-tracking or response-time measures, could be combined with segmentation to investigate how viewers process multimodal cues as videos unfold in real time.

In a broader sense, the approach proposed here contributes to multimodal discourse analysis by expanding a way of systematically tracking interpretive change in unfolding audiovisual texts. By combining segmentation with a relational account of expectation and violation, it makes visible how humorous meaning is incrementally built, adjusted, and recalibrated across time and modes. This shifts analytical attention away from static descriptions of multimodal composition toward the dynamics of interpretation as organised by the text itself. In doing so the approach responds to a growing need within multimodal research for methods that can account for rapid, temporally dense forms of platformed communication, where meaning emerges not from stable configurations but from their systematic reorganisation over time.

References

- Abdel-Raheem, A. (2018). Multimodal humour: Integrating blending model, relevance theory, and incongruity theory. *Multimodal Communication*, 7(1), 1–25.
- Attardo, S. (1994). Linguistic theories of humor. Mouton de Gruyter.
- Attardo, S. (2001). *Humorous texts: A semantic and pragmatic analysis*. Berlin & New York: Mouton de Gruyter.
- Attardo, S. (2020). *The linguistics of humor: An introduction*. Oxford University Press.
- Balirano, G., & Corduas, M. (2008). Detecting semiotically-expressed humor in diasporic TV productions. *Humor*, 21(3), 227–251. <https://doi.org/10.1515/HUMOR.2008.012>
- Bateman, J. A., Wildfeuer, J., & Hiippala, T. (2017). *Multimodality: Foundations, research and analysis – A problem-oriented introduction*. De Gruyter Mouton.

- Bernad-Mechó, E., & Girón-García, C. (2023). A multimodal analysis of humour as an engagement strategy in YouTube research dissemination videos. *European Journal of Humour Research*, 11(1), 46–66.
- Dynel, M. (2016). “I has seen image macros!” Advice animal memes as visual–verbal jokes. *International Journal of Communication*, 10, 660–688.
- Gironzetti, E. (2022). *The multimodal performance of conversational humor*. John Benjamins Publishing Company. <https://doi.org/10.1075/ftl.13>
- Hempelmann, C. F., & Attardo, S. (2011). Resolutions and their incongruities: Further thoughts on logical mechanisms. *Humor: International Journal of Humor Research*, 24(2), 125–149. <https://doi.org/10.1515/HUMR.2011.008>
- Masi, S. (2023). Humour in TED Talks: A multimodal account. *ESP Today*, 11(2), 328–348.
- Priego-Valverde, B. (Ed.). (2024). *Interactional humor: Multimodal design and negotiation*. De Gruyter Mouton.
- Raskin, V. (1985). Semantic mechanisms of humor. Reidel.
- Shifman, L. (2013). *Memes in digital culture*. MIT Press.
- Trillò, T. (2024). “PoV: You are reading an academic article.” The memetic performance of affiliation in TikTok’s platform vernacular. *New Media & Society*. <https://doi.org/10.1177/14614448241290234>
- Tsakona, V. (2003). Jab lines in narrative jokes. *Humor: International Journal of Humor Research*, 16(3), 315–329. <https://doi.org/10.1515/humr.2003.017>
- Tsakona, V. (2020). *Recontextualizing humor: Rethinking the analysis and teaching of humor*. De Gruyter Mouton.
- Vásquez, C., & Aslan, E. (2021). “Cats be outside, how about me?” The linguistic construction of stance in digitally mediated humorous narratives. *Journal of Pragmatics*, 171, 240–253. _
- Wildfeuer, J. (2014). *Film discourse interpretation: Towards a new paradigm for multimodal film analysis*. Routledge.
- Willoughby, A. (forthcoming). Humour and modal interplay in TikTok’s POV genre. In A. Beville, F. Ciambella, & J. Culpeper (Eds.), *Bridging the gap between pragmatics and multimodality* (Special issue). *Lingue e Linguaggi*.
- Yus, F. (2018). Multimodality in memes: A cyberpragmatic approach. In P. Bou-Franch & P. Garcés-Conejos Blitvich (Eds.), *Analyzing digital discourse: New insights and future directions* (pp. 105–131). Palgrave Macmillan. https://doi.org/10.1007/978-3-319-92663-6_4
- Yus, F. (2021). Incongruity-resolution humorous strategies in image macro memes. *Internet Pragmatics*, 4(1), 131–149. <https://doi.org/10.1075/ip.00058.yus>
- Zacks, J. M., & Swallow, K. M. (2007). Event segmentation. *Current Directions in Psychological Science*, 16(2), 80–84. <https://doi.org/10.1111/j.1467-8721.2007.00480.x>
- Zulli D. & Zulli D. (2020). Extending the internet meme: Conceptualizing technological mimesis and imitation publics on TikTok, in *New Media & Society* 24 [3], pp. 614- 636, <https://doi.org/10.1177/1461444820983603>

Biografia

Audrey C. Willoughby

Audrey Willoughby è dottoranda in Linguistica presso l'Università degli Studi di Milano, nel Dipartimento di Lingue, Letterature, Culture e Mediazioni. Ha conseguito un Master in Applied Linguistics presso la Texas A&M University–Commerce, dove ha discusso una tesi intitolata *Humor Markers in Computer-Mediated Communication*, e ha inoltre ottenuto un Graduate Certificate in TESOL. La sua ricerca si concentra sulle modalità semiotiche attraverso cui l'umorismo si realizza nelle piattaforme di video brevi, adottando un approccio di analisi multimodale del discorso volto allo studio dei processi di significazione nei media digitali performativi. Ha inoltre ricoperto per diversi anni il ruolo di assistant editor della rivista *HUMOR: International Journal of Humor Studies* ed è stata editorial assistant per la *Encyclopedia of Humor*.