

Safety extending beyond the cabin: A social semiotic study of the multimodal engagement strategies in Chinese preflight safety videos

Yixin Chen¹, Enhua Guo² & Mei Li¹

Shanghai International Studies University (China)¹, Ocean University of China (China)²

yixin.chen@shisu.edu.cn, guoenhua@ouc.edu.cn, meinonlee@shisu.edu.cn

Abstract

Flight safety remains a top priority in the aviation industry, with preflight safety videos (PSVs) playing a critical role in preparing passengers for potential emergencies. In recent years, airlines worldwide have increasingly incorporated multimodal resources into PSV design. However, limited research has examined how these resources are configured to engage passengers and enhance safety communication. Drawing on visual grammar (Kress & van Leeuwen, 2006), the study adopts a social semiotic approach to explore how the concept of safety is multimodally constructed in PSVs and how passengers' safety information retention is enhanced. The dataset comprises 15 PSVs released by 10 major Chinese airlines between 2019 and 2025, totaling approximately 80 minutes. The analysis identifies four primary multimodal engagement strategies: 1) deployment of pop-up visual aids; 2) variation in camera shots during safety demonstrations; 3) coordination of video character's embodied modalities; and 4) sociocultural recontextualization of the cabin environment. The findings suggest that PSVs engage viewers by presenting a concept of safety that extends beyond the physical cabin, enriched with culturally meaningful semiotic resources.

Keywords: Visual grammar, social semiotics, multimodality, PSVs, engagement.

Resumen

La seguridad más allá de la cabina: un estudio socio-semiótico de las estrategias multimodales de implicación en los vídeos chinos de seguridad previos al vuelo

La seguridad en los vuelos sigue siendo una prioridad fundamental en la industria de la aviación y los vídeos de seguridad previos al vuelo (abreviado como PSV en inglés) desempeñan un papel crucial en la preparación de los pasajeros ante posibles emergencias. En los últimos años, las aerolíneas de todo el mundo han incorporado recursos multimodales en el diseño de estos vídeos. Sin embargo, la investigación sobre cómo se configuran estos recursos para involucrar a los pasajeros y mejorar la comunicación en materia de seguridad es limitada. Con base en la gramática visual (Kress & van Leeuwen, 2006), este estudio adopta un enfoque semiótico social para explorar cómo se construye multimodalmente el concepto de seguridad en los PSV y de qué manera se mejora la retención de información. El corpus se compone de 15 vídeos publicados por 10 grandes aerolíneas chinas entre 2019 y 2025, con una duración total de aproximadamente 80 minutos. El análisis identifica cuatro estrategias principales de implicación multimodal: 1) el uso de apoyos visuales emergentes (*pop-up*); 2) la variación de los planos de cámara durante las demostraciones de seguridad; 3) la coordinación de las modalidades corporizadas de los personajes del vídeo; y 4) la recontextualización sociocultural del entorno de la cabina. Los resultados sugieren que los PSV logran captar la atención de los espectadores al presentar un concepto de seguridad que va más allá del espacio físico de la cabina, enriquecido con recursos semióticos culturalmente significativos.

Palabras clave: Gramática visual, semiótica social, multimodalidad, PSV, implicación.

1. Introduction

Over the past decades, air travel has experienced substantial growth, driven by rapid advancements in transportation technologies and the expanding forces of globalization. The aviation industry relies heavily on the effective implementation of safety measures (Winter et al., 2021), with preflight education representing a central concern. Research in aerospace psychology has demonstrated that innovatively designed preflight safety videos (PSVs) can enhance both passenger engagement and retention of safety information (Seneviratne & Molesworth, 2015). For instance, the use of humor and celebrity appearances in PSVs has been shown to improve the effectiveness of cabin safety messaging (Molesworth et al., 2018). However, passenger engagement with safety briefings and videos may often be undermined by passengers' overconfidence in personal safety knowledge, flight anxiety (Winter et al., 2021), as well as cabin noise and language barriers (Molesworth et al., 2018). In recent years, airlines worldwide have

increasingly incorporated multimodal resources into PSV design, including visual, verbal, and embodied modes, which capture passengers' attention and facilitate the delivery of safety education across diverse sociocultural contexts (Bernad-Mechó & Valeiras-Jurado, 2023; Wong, 2016; Xia & Hafner, 2021). The strategic configuration of multimodal resources in safety communication offers a lens to examine how passengers are engaged in safety education and how the very concept of safety is multimodally constructed. Although interest in multimodality has grown within visual discourse studies, its application to the airline industry, particularly in the context of preflight safety instruction, remains underexplored.

To address this gap, the present study investigates the multimodal engagement strategies employed in PSVs of Chinese airlines, which occupy a significant position in the global aviation industry. Adopting a visual grammar framework (Kress & van Leeuwen, 2006) informed by systemic functional linguistics (Halliday & Matthiessen, 2014), the analysis focuses on how diverse semiotic resources are orchestrated to engage passengers while conveying essential safety information. This investigation not only advances understanding of the communicative design of PSVs but also contributes to broader discussions on the role of multimodality in regulated public discourse.

2. Literature review

2.1. Safety communication in aviation discourse

In the high-risk context of aviation, communication plays a key role in minimizing human error and maintaining safety (Kines et al., 2010). Effective communication depends on the use of clear and concise language, sensitivity to human factors and many other aspects (Chionis et al., 2021, 2022; Yıldız, 2024). Therefore, cultivating a culture of clear, standardized, and context-sensitive communication is crucial in sustaining and advancing aviation safety (Tajima, 2004).

With technological developments, new affordances have emerged and transformed how safety information is communicated in the cabin. In recent years, airlines increasingly invest in PSVs that balance regulatory clarity with emotional and cognitive engagement, marking a broader transformation in safety discourse. Narrative techniques, humor, visual metaphors, and various strategies have been incorporated to enhance message delivery (Chang &

Liao, 2009; Seneviratne & Molesworth, 2015). These innovations reflect a growing awareness of the need for affective, attention-sensitive design in safety communication, in response to challenges including anxiety, overconfidence, and linguistic diversity (Molesworth et al., 2018; Winter et al., 2021).

Substantial research has examined verbal communication in radiotelephony interactions (Bogush & Kovtun, 2019; Ishihara & Lee, 2021), cabin instruction strategies (Arminen et al., 2014; Tuccio & Garcia, 2020; Zhao, 2022), as well as PSVs (Molesworth et al., 2018; Seneviratne & Molesworth, 2015). However, only a limited number of studies (e.g. Haas & Van Erp, 2014) has underscored the value of integrating multiple modes of communication to enhance the delivery and reception of critical safety information. Despite their growing complexity and public-facing function, what remains underexplored is PSVs' discursive structures, linguistic features, and multimodal engagement strategies. A deeper understanding of their design can yield important insights for enhancing safety communication, improving passenger engagement, and informing best practices in airline media production.

2.2. Multimodality and social semiotics

Communication involves the interaction of multiple semiotic modes in the process of meaning-making (Kress & van Leeuwen, 2001). A mode is defined as “a set of socially and culturally shaped resources for making meaning” (Jewitt et al., 2016, p. 9), and includes elements such as speech, image, gaze, and gesture, with each shaped by distinct social conventions and cultural contexts. In addition, each mode offers distinct affordances that shape how meaning is conveyed (Kress, 2010). According to Jewitt (2009), the relationships among these modes and their combined communicative effects, along with their presence, constitute the essence of multimodality.

As a social semiotic approach to multimodality, visual grammar emphasizes both individual semiotic modes and the interactions among them (Kress & van Leeuwen, 2001). From this perspective, meaning-making is a purposeful process during which sign-makers select and combine resources to achieve communicative aims (Jewitt, 2009). In the context of moving images, these resources encompass filming and editing choices, which contribute significantly to the construction of meaning (Burn, 2013). Crucially, meaning

arises not from a single privileged mode but through the integration of multiple semiotic resources in what Jewitt (2016, p. 72) terms “multimodal ensembles.” More attention should be given to the ways in which semiotic modes are configured and orchestrated in multimodal communication (Unsworth, 2008).

Kress and van Leeuwen (2006) identified three key metafunctions in the visual design of multimodal discourse. These include: 1) representational meaning, which refers to the depiction of entities, actions, and events through the interaction of modes; 2) interactive meaning, which concerns the image-viewer relationship shaped by gaze, distance, and modality; and 3) compositional meaning, which involves the arrangement of visual elements to create coherence and guide interpretation. These metafunctions provide a systematic framework for our understanding of how visual narratives in PSVs construct and convey the concept of safety, while simultaneously involving viewers in the communication.

2.3. Multimodal engagement strategies in visual discourse

Engagement in visual media discourse refers to the viewer’s active involvement in interpreting and responding to multimodal content (Jewitt, 2009; O’Halloran, 2008). Sustained viewer attention and comprehension depend on the coherent integration of modes, the use of interactive cues, and the inclusion of affective elements (Bateman et al., 2017). In PSVs, understanding passenger engagement requires a close examination of how visual resources are strategically orchestrated to direct attention, facilitate comprehension, and sustain interest. These videos are not merely vehicles for conveying technical information but meticulously designed to elicit both cognitive and emotional engagement within strict temporal limits. Analyzing PSVs through the lens of social semiotics is critical for uncovering how meaning is produced not only through the content presented but through the manner of its presentation (Kress & van Leeuwen, 2006). The underlying communicative logic embedded in the design, sequencing, and multimodal integration of the videos significantly shapes how safety messages are interpreted and internalized by viewers.

Previous research on multimodal engagement strategies in visual discourse has yielded valuable insights across diverse contexts, including education, advertising, and city branding. Coherence among semiotic modes has been

identified as a key factor in effective communication, shaping viewers' interpretation and reception of messages (Bateman et al., 2017; Zappavigna, 2016). Particularly, embodied modes such as structured gestures and long gaze are effective in fostering a sense of connection between speakers and audiences (Bernad-Mechó & Girón-García, 2023; Xia & Hafner, 2021). Visual design elements, such as color schemes, additionally help sustain interest and communicate key messages (Lirola & Chovanec, 2012; Wagner, 2015; White, 2010). In addition, linguistic strategies, including rhetorical questions, emotional references, and personal pronouns, further enhance engagement by creating a conversational tone (Tereszkiewicz, 2023; Xiao et al., 2025). In general, key engagement strategies in visual discourse include reducing social distance through embodied modes, employing emotion-related language, and maintaining visual coherence to support comprehension and retention. These insights offer a useful point of departure for analyzing Chinese PSVs, which, like other visually constructed discourses, aim to capture audience attention while communicating safety messages effectively. Building on this dual function, this study offers a social semiotic perspective on audience interaction within a highly controlled, safety-critical context.

2.4. Research gap and research questions

Existing studies have primarily examined passenger perceptions of innovative PSV designs from psychological and behavioral viewpoints (Molesworth, 2014; Molesworth et al., 2018; Seneviratne & Molesworth, 2015). However, there is a need for more detailed analysis of how multimodal resources are configured to support both viewer engagement and safety communication. The present study investigates the design of Chinese PSVs through a social semiotic lens. By identifying and interpreting multimodal engagement strategies, the study contributes to both theoretical understandings of meaning-making in safety discourse and practical insights for the production of aviation safety media. The findings can inform the design of future PSVs and other broader practices in multimodal safety communication across public-facing institutional contexts.

Accordingly, this study asks: How are multimodal strategies employed in Chinese PSVs orchestrated to engage passengers? And how is the concept of safety multimodally constructed in PSVs?

3. Methodology

3.1. Data collection

This study investigates how multimodal strategies are employed in Chinese PSVs to engage passengers and promote flight safety awareness. The dataset includes 15 PSVs released by 10 major Chinese civil airlines between 2019 and 2025, totaling approximately 80 minutes. The selected airlines reflect a broad spectrum of the Chinese aviation sector in terms of geographical coverage, passenger volume, and international operational scope.

Accessible from official airline websites and popular video platforms, these PSVs are typically shown via in-flight entertainment systems, such as seatback screens or overhead drop-down monitors. Each video ranges from 3 to 7 minutes in length, consistent with the standard duration of preflight safety presentations delivered prior to takeoff. To maintain genre and modality consistency, the analysis excludes live safety demonstrations by flight attendants and verbal briefings delivered post-boarding. The focus is exclusively on recorded video materials designed for in-cabin viewing, ensuring a uniform basis for multimodal analysis.

3.2. Data analysis procedure

Drawing on the visual grammar framework, our analysis is grounded in a social semiotic approach, which offers a systematic means of examining how meaning is constructed through the interaction of image, text, sound, and embodied modes. All 15 videos were viewed multiple times by all authors to develop a comprehensive understanding of their content, stylistic features, and communicative intentions (Jewitt, 2016).

Initial analytical notes focused on recurring themes, visual patterns, and narrative strategies related to passenger engagement and safety instruction. To ensure a rigorous and systematic coding procedure, the study adopts the transcription framework proposed by Bezemer and Mavers (2011, pp. 194–195), which emphasizes how meaning is constructed through framing, selecting, and highlighting. This approach ensures a principled method for coding the data and facilitates a deeper understanding of the multimodal engagement strategies.

More specifically, in order to facilitate the analysis and demonstrate the application of multimodal engagement strategies, simplified transcripts of the data will be provided. First, the coding process is framed by the research

question of this study: rather than providing a comprehensive rendering of every semiotic detail in the videos, the transcripts were produced as analytical artefacts oriented to our research questions about passenger engagement. Second, video segments were selected based on their functional relevance to the communicative goals of the safety videos. The selection process prioritizes segments that are key to conveying the intended messages about safety and passenger engagement, ensuring that only the most representative and critical moments are included in the analysis. Finally, we made screenshots by giving visual and textual prominence in the transcript to those semiotic resources most integral to the strategy under discussion. Therefore, screenshots of the preflight safety videos are taken at significant moments when the multimodal resources are saliently interacting to engage the potential audience, and the voice-over in each screenshot is also translated into English for understanding.

In the subsequent analysis, we explicate key multimodal engagement strategies identified in the dataset and illustrate how various semiotic resources are orchestrated to involve passengers. These strategies are presented in Sections 4.1-4.4 through four recurrent multimodal configurations.

4. Findings

This study identifies four key multimodal strategies in Chinese PSVs as follows, and each strategy is analyzed for its role in engaging passengers and constructing the concept of safety.

4.1. Deployment of pop-up visual aids

The first multimodal engagement strategy involves the use of pop-up visual aids to support and enhance the delivery of safety messages. These animated visuals interact dynamically with speech and gesture, not only reinforcing meaning but also directing viewer attention, particularly beneficial for audiences with varying levels of language proficiency or limited auditory access. As a technologically enhanced modality, pop-up visuals extend the communicative scope of safety discourse beyond conventional embodied and verbal modes, enriching the multimodal repertoire of instructional design.

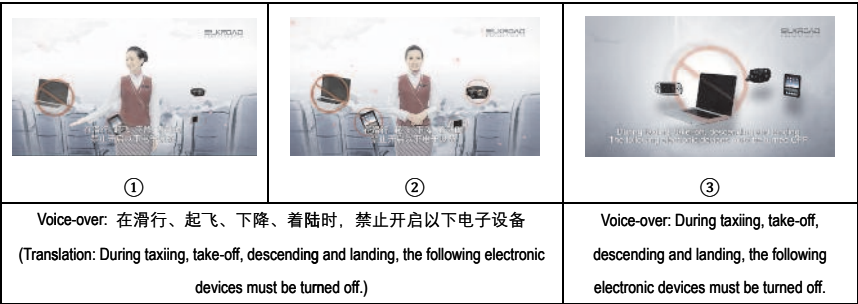


Figure 1. Prohibiting the use of electronic devices.

Figure 1 from China Southern Airlines illustrates how the prohibition of the use of electronic devices during critical flight phases is communicated through the coordinated use of speech, gesture, and animated pop-up visual aids. The sequence begins with a flight attendant performing a sweeping gesture from right to left (①), synchronized with animated pop-ups of devices such as a laptop, tablet, gamepad, and DV camera (②, ③). Each device icon is overlaid with a red prohibition symbol, with variations in image size assigning different degrees of salience to each category, guiding viewer focus. As the spoken language transitions from Mandarin to English (③), the icons rotate forward and are re-covered by a prominent red symbol, reinforcing the rule visually (Kress & van Leeuwen, 2006, p. 177). This multimodal synchronization, i.e., linking gesture, verbal cues, and animated visuals, effectively conveys representational meaning by clearly identifying prohibited items while reinforcing the regulation. The layered design also compensates for potential auditory limitations due to cabin noise, allowing visual elements to function independently of sound. Moreover, contrasting the high-modality (i.e., presenting the world “as it is”) and realistic images of electronic devices, with a low-modality (i.e., stylized, simplified interpretation of reality) and illustrated background serves to visually foreground the key informational content and reinforce the classificational structure of the message (Kress & van Leeuwen, 2006, pp. 154-174). Overall, this multimodal configuration enhances the salience and accessibility of the safety instruction by coordinating visual emphasis, representational clarity, and cross-modal reinforcement to ensure the regulation is communicated effectively.

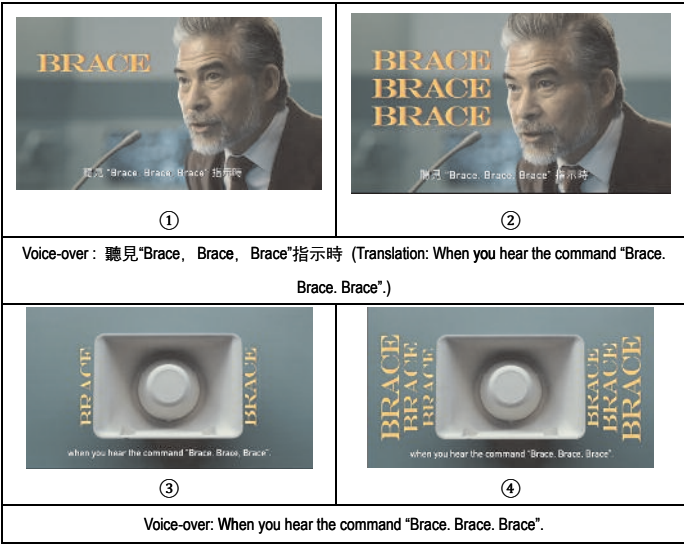


Figure 2. Brace position command

As the instruction “Brace, Brace, Brace” is given, the capitalized word “BRACE” appears on screen in sequence, its size progressively increasing in rhythm with the narration. This rhythmic animation adds visual dynamism and creates a multisensory crescendo that heightens affective intensity and elicits an instinctive, embodied response. The use of bold, vintage-style fonts in high-contrast red and yellow not only ensures legibility, even for passengers with limited auditory access, but also signals urgency and enacts the emotional escalation of the auditory cue (Kress & van Leeuwen, 2001; Yao & Zhuo, 2018). The animations of texts that move top-down and center-outward further sustain attention and emphasize the salience of the directive through visual novelty (White, 2010). These kinetic and visual properties produced by motion and contrast are known to trigger heightened alertness, ensuring that the message remains cognitively and emotionally indelible in high-stress situations (Franconeri & Simons, 2003; Vavourakis et al., 2022).

Importantly, text in these sequences does not merely supplement the narration but functions as a sayer in a verbal process, visually echoing and amplifying the spoken command. This modal concurrence (Unsworth, 2008) transforms the pop-up into a co-articulated, multimodal speech act. It constructs safety as an urgent, actionable command, not just an

informational message. The repetition and growth of “BRACE” mirror the physical act of bracing, embedding the instruction into viewers’ sensorimotor memory (Vavourakis et al., 2022). The increasing scale, combined with the stark color contrast, frames the directive as a critical survival mechanism which is immediate, non-negotiable, and universally interpretable (Kress, 2010). In high-stakes environments where cognitive load is high and attention may be fragmented, such synchronized multimodal strategies ensure clarity, urgency, and memorability, constructing safety not just as knowledge, but as embodied readiness.

In sum, pop-up textual elements in Chinese PSVs exemplify a technologically mediated semiotic strategy that transforms verbal instruction into a multisensory, embodied experience. By aligning typography, motion, and timing with auditory cues, these sequences optimize attentional capture, reinforce message salience, and facilitate sensorimotor encoding of safety procedures (Wagner, 2015). Rather than functioning as passive captions, the pop-up visual aids act as active components in a multimodal speech act, visually echoing the spoken command and amplifying its urgency. This strategy reflects a broader shift in safety communication toward an audience-centered approach, effectively engaging formats that leverage digital affordances for greater clarity, retention, and behavioral readiness.

4.2. Variation in camera shots during safety demonstrations

A range of camera shots —long, medium, and close-up— are strategically utilized to shape the interactive meaning of safety demonstrations and modulate the perceived social distance between viewers and the depicted participants or objects. These framing choices create a dynamic viewing experience that enhances viewer involvement and attentional focus (Kress & van Leeuwen, 2006).



Figure 3. Instructing how to fasten the seatbelt.

In Figure 3 from China Eastern Airlines, a safety narrative opens with a focus on the seatbelt-fastening sequence. Medium shots (①) are used to establish an appropriate social distance that aligns with instructional conventions. The participant's upper body is framed frontally in alignment with the audio prompt or voice-over "Please fasten your seatbelt after seated". A clear vector is thus formed to direct viewer attention toward the seatbelt-fastening action. As an entry point, the medium shot evokes a classroom-like setting, offering a balance between accessibility and authority (Forey & Feng, 2016). It invites identification with the modeled behavior while maintaining the formality typical of safety demonstrations, effectively setting the stage for viewer engagement.

As the sequence unfolds, the camera zooms into close-up shots (② and ③) of the participant's hands fastening (②) and unfastening (③) the seatbelt. This transition reduces social distance, creating a more intimate and immersive perspective that draws viewers into a first-person mode of engagement (Jewitt et al., 2016). The hands are centrally framed, amplifying the "informational value" of the gesture (Kress & van Leeuwen, 2006, p. 177) and foregrounding the action as a key instructional moment. By narrowing the focus from whole-upper-body framing to detailed hand movement, the video shifts from passive instruction to active cognitive simulation, encouraging viewers to mentally rehearse the procedure themselves (Xia & Hafner, 2021). This heightened involvement can enhance memory retention, as the viewers mentally mirror the activity on screen.

The progression from medium to close-up shots constructs safety as both a collective norm and an individual responsibility. The medium shot situates the action within the collective flight travel environment (Kress & van Leeuwen, 2006, p. 126), signaling that safety is a communal expectation. In contrast, the close-up (Kress & van Leeuwen, 2006, p. 177) emphasizes personal accountability by isolating the specific, replicable mechanics of seatbelt use. The detailed hand gestures function as a visual contract, underscoring the need for precision and attentiveness in fulfilling this safety requirement.

In sum, Chinese PSVs strategically manipulate visual framing to manage viewer engagement and enhance instructional clarity. The interplay between medium shots and close-up shots is what makes the viewing experience dynamic (Lim, 2017). This progression mirrors how we naturally shift our attention from the general to the specific, creating a narrative flow that feels intuitive and engaging. The constant change in perspective prevents the

video from becoming monotonous, keeping passengers visually stimulated (Forey & Feng, 2016). At the same time, it strategically guides their focus to essential safety elements, ensuring that key messages, i.e., how to secure a seatbelt or locate an exit, are both noticed and remembered.

4.3. Coordination of video characters’ embodied modalities

In Chinese PSVs, gaze, gesture, and body posture are strategically coordinated in guiding viewer attention and constructing engagement. These multimodal strategies facilitated by advances in video production can enable precise control over viewer positioning and attentional focus (Hood, 2011; Kress & van Leeuwen, 2006).

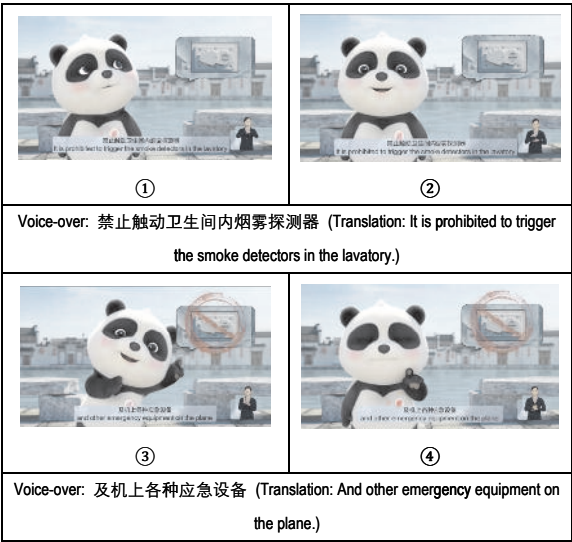


Figure 4. Prohibiting triggering of smoke detectors.

Figure 4 from Air China features a cartoon panda bearing the airline’s logo, demonstrating the no-smoking policy. The panda begins by gazing indirectly toward the top-right corner of the screen, where a smoke detector appears, before shifting to direct eye contact with the viewer. This gaze sequence functions as a narrative guide, with the initial indirect gaze (①) introducing “new” information (Kress & van Leeuwen, 2006) and the direct gaze (② and ③) forming an interactive demand that prompts viewer acknowledgment of the no-smoking rule. This transition transforms its interactivity with viewers

from passive observation into a dialogic exchange. The indirect gaze builds anticipation, drawing viewers into the flight safety context; the direct gaze, by engaging the human instinct to respond to eye contact, enhances emotional receptivity and sustained attention (Xia & Hafner, 2021). Through this coordinated gaze shift, safety is reframed as both collective and personal responsibility. It is not merely as an external, imposed regulation (via indirect gaze) but as an internalized and personalized commitment.

Concurrently, the panda's gestures of pointing toward the smoke detector (③), followed by a head shake and finger wag (④), serve as embodied reinforcements of the verbal message. These gestures act as visual anchors: the pointing gesture directs attention precisely to the detector, while the head shake and finger wag visually enact the prohibition (Wong, 2016). This gesture sequence not only communicates the rule but also embodies its enforcement. The shift from pointing to wagging creates a rhythmic tempo that maintains viewer engagement. The animated, playful exaggeration of these movements leverages animation's flexibility, making the instruction memorable and entertaining, especially for younger or less attentive audiences. The gesture trajectory constructs safety as a proactive and enforceable norm: while the pointing gesture establishes the detector as a protective tool, the head shake and finger wag assert a clear boundary, framing safety as a non-negotiable rule that passengers must respect. Additionally, both body posture and facial expression balance authority with approachability. These expressions amplify emotional signals beyond what a human actor might convey. Drawing on Cuddy et al. (2011), this blend of warmth through playfulness and competence through sternness fosters viewer trust.

In sum, Figure 4 employs a carefully choreographed interplay of gaze, gesture, posture, and facial expression to construct safety as an accessible, authoritative, and culturally resonant message. Chinese PSVs transform the no-smoking policy into an engaging, culturally rich safety lesson through a multimodal ensemble, i.e., gaze, gesture, posture, and expression, in a tightly coordinated manner (Jewitt, 2016). The panda's indirect gaze sets the narrative frame: its pointing gesture anchors attention; the direct gaze and wag deliver the behavioral directive; and its expressive posture modulates tone. Together, these resources reduce social distance, blend authority with warmth, and construct safety as a shared but individually fulfilled responsibility. This multimodal strategy not only sustains viewer attention but also fosters a deeper, more relatable connection to safety protocols.

4.4. Sociocultural recontextualization of the cabin environment

Chinese PSVs also employ a sophisticated multimodal strategy through sociocultural recontextualization (Bezemer & Kress, 2008), relocating safety instructions from the sterile, high-stress environment of an aircraft cabin to culturally resonant everyday or symbolic settings. This recontextualization not only localizes the meaning of safety but also reduces the emotional intensity associated with emergency scenarios, thereby promoting viewer receptiveness and engagement through affective and aesthetic familiarity. Figure 5 below from Air China is a compelling example of this strategy.

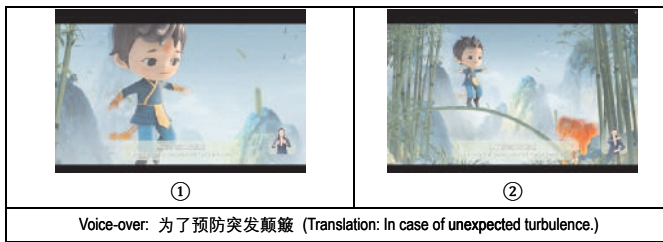


Figure 5. Representation of unexpected turbulence.

In Figure 5, the experience of in-flight turbulence is visually reimaged within a traditional bamboo forest. A boy dressed in classical Chinese attire attempts to balance on a bamboo stalk, a familiar Kung Fu motif, while watercolor-style mountains form the backdrop. The rhythmic flow of Chinese-style instrumental music complements the scene, creating a multisensory composition that draws from shared cultural memory and artistic conventions.

This multimodal orchestration reframes turbulence, a typically distressing experience, as a culturally meaningful metaphor of balance and resilience. The scene compares the instability of in-flight turbulence to the sensation of standing on one foot atop a bamboo stalk, a familiar Kung Fu motif. The boy's precarious pose visually echoes the physical instability of turbulence, but within a narrative of self-mastery, adaptability, and poise. The bamboo forest in watercolor style, paired with graceful movement and soothing Chinese music, turns the safety instruction into a calm, almost meditative experience (Unsworth, 2008). By leveraging the symbolic power of Chinese philosophical values, particularly harmony and equilibrium, the video shifts the emotional tone from panic to serenity. This affective shift not only

reduces tension but also enhances cultural resonance and encourages viewers to internalize safety as both a personal commitment and a culturally grounded practice.

Figure 6 from Xiamen Airlines presents a sequence of images in which an in-flight emergency is recontextualized, however, through a culturally vibrant urban narrative set in Shaweipo of Xiamen city in southeast China.



Figure 6. Adopting brace positions for emergency.

In Figure 6, the aircraft cabin is replaced by a bustling Shaweipo street market. The scene depicts a skateboarder nearly colliding with a mother and daughter seated outside a café, an incident that mimics the sudden jolt of in-flight turbulence. By relocating the safety scenario to a vibrant urban environment infused with elements of youth culture and Xiamen’s recognizable coffee scene, the video draws on localized sociocultural knowledge to make the safety message feel familiar and personally relevant.

This urban recontextualization enhances engagement by transforming the brace instruction into a scene that resembles everyday life rather than an abstract emergency protocol. The skateboarder’s dynamic movement symbolizes modern urban energy, while the mother and daughter provide a

relatable familial anchor (② and ④). As the camera slowly zooms in from a wide-angle view of the market to a close-up of the brace position, the viewer's perspective shifts from passive observer to active participant (O'Halloran, 2008), fostering proximity and narrative immersion.

Importantly, the light-hearted tone of the encounter softens the perceived threat, using cognitive reframing to present turbulence as a manageable, even playful, challenge (Winter et al., 2021). This affective modulation helps regulate viewers' emotional responses, creating a psychologically safe space for processing critical safety information. The use of culturally resonant semiotic resources, i.e., skateboarders, café settings, lively streets, bridges the gap between flight and everyday life, reducing psychological distance and reinforcing cultural relevance. The design of these multimodal resources not only guides attention to the brace action but also fosters empathy, enhancing memorability and emotional engagement.

In sum, across Figures 5 and 6, safety is not simply taught but is performed within culturally grounded scripts that blend relational warmth with procedural clarity. The use of traditional aesthetics in one scene and contemporary urban semiotics in the other demonstrates a flexible but consistent strategy: grounding universal safety messages in local meanings to promote identification, comprehension, and retention. In doing so, Chinese PSVs redefine safety communication as not merely instructional but experiential, affective, and culturally intelligible.

5. Discussion

Drawing on visual grammar, this study contributes to safety communication and aviation discourse by examining the orchestration of multimodal resources in Chinese PSVs, a new safety-critical genre. Four multimodal engagement strategies are identified and their capacities to engage both cognitive and affective dimensions of viewer experience can be summarized as follows.

First, pop-up visual aids function as co-articulated elements in the multimodal ensemble, working synchronously with speech and gesture to optimize attention, comprehension, and retention (Xia & Hafner, 2021). Far from being supplementary, these visuals operate as modal co-sayers, reinforcing spoken directives through size, motion, rhythm, and color

contrast (Lirola & Chovanec, 2012; Yao & Zhuo, 2018). Moreover, the visuals enhance accessibility for passengers across linguistic and cognitive differences, positioning safety as universally interpretable.

Second, the variation in camera shots (long, medium, close-up) in Chinese PSVs directly contributes to viewer engagement by modulating social distance and directing attention from generalized context to specific actions (Xia & Hafner, 2021). This dynamic use of visual framing not only sustains visual interest but constructs “safety” as both a collective norm and an individual responsibility. Through this shift in proximity, PSVs attempt to help viewers to internalize and mentally rehearse critical safety behaviors.

Third, video characters’ embodied orchestration of gaze, gesture, posture, and expression further deepens engagement by transforming safety information into an interactive, affective performance (Bernad-Mechó & Girón-García, 2023; Bernad-Mechó & Valeiras-Jurado, 2023; Wong, 2015). From an engagement perspective, this strategy mobilizes the viewer’s affective and attentional resources. From a communicative standpoint, it reframes safety not as a static rule but as a socially enforceable and emotionally resonant commitment.

Last but not least, the sociocultural recontextualization of the cabin environment into symbolic or ordinary life may be the most innovative strategy. It enhances engagement by leveraging the aesthetic and emotional familiarity of the bamboo forest setting, which resonates with Chinese viewers’ cultural memory and philosophical traditions (Lakey, 2016). This approach also constructs safety as a harmonious and adaptable practice rooted in Chinese philosophical values, particularly those of balance, harmony, and equilibrium.

Taken together, these multimodal strategies do more than enhance communicative engagement: they reconstruct the concept of safety in Chinese sociocultural contexts. The four strategies converge on a common trajectory: transforming safety from a set of institutional instructions into embodied, affective, and culturally positioned practices. This transformation is especially salient in the Chinese context, where sociocultural symbolism, collectivist values, and narrative traditions provide rich semiotic affordances.

These findings reinforce the view that meaning in multimodal discourse emerges through the dynamic interplay of semiotic resources (Bezemer & Kress, 2008; Kress & van Leeuwen, 2006). They echo previous findings showing that direct gaze, gestures, and coordinated visuals (e.g., color, layout,

lighting) help reduce social distance and enhance audience engagement (Bernad-Mechó & Valeiras-Jurado, 2023; White, 2010; Xia & Hafner, 2021). Such multimodal alignment, common in advertising and public campaigns, is key to effective communication. Most importantly, the study reveals that the notion of safety in PSVs can be multimodally constructed in a broader context, transcending the cabin. Safety, in this sense, is no longer confined to rigorous safety protocols but rendered as attentional discipline, emotional regulation, relational care, and cultural identification. It shifts from being about what to do in a crisis to how to live through and with preparedness. Our findings show that sociocultural recontextualization enhances engagement in safety videos by embedding instructions in familiar, emotionally resonant scenarios. This reduces anxiety and aids recall, turning abstract procedures into relatable narratives (Lakey, 2016). Besides, safety as a harmonious and adaptable practice is rooted in Chinese philosophical values, particularly those of balance, harmony, and equilibrium ... (Wang, 2015). The incorporation of national cultural elements underscores the broader communicative potential of PSVs. By anchoring safety messages in culturally familiar settings, PSVs cultivate a sense of emotional ease and social proximity, helping both local and international passengers better relate to the content.

6. Conclusion

This study adopts a social semiotic approach to examine how multimodal resources in Chinese PSVs, i.e., visual aids, camera work, embodied coordination, and recontextualization of the cabin, are orchestrated to communicate safety information and enhance passenger engagement. By foregrounding the genre-specific features of PSVs, the study extends multimodal research into an underexplored domain and offers insights into how media design can simultaneously serve instructional and cultural functions. The findings have practical implications for airline communication strategies and the design of more culturally resonant safety media. However, the actual effectiveness of these strategies—particularly their impact on information retention—requires further empirical investigation. Investigating similar multimodal practices in other transportation contexts would also help evaluate the broader applicability of the design principles identified here.

Acknowledgements

This article is supported by the Youth Project of Tianjin Philosophy and Social Science Planning Program: A Study on China-Related News Discourse in German Television Based on a Multimodal Corpus, TJYY21-021 (天津市哲学社会科学规划青年项目, 基于多模态语料库的德国电视涉华新闻话语研究, TJYY21-021).

Article history:

Received 11 October 2025

Received in revised form 31 January 2026

Accepted 30 March 2026

References

- Arminen, I., Koskela, I., & Palukka, H. (2014). Multimodal production of second pair parts in air traffic control training. *Journal of Pragmatics*, 65, 46-62. <https://doi.org/10.1016/j.pragma.2014.01.004>
- Bernad-Mechó, E., & Girón-García, C. (2023). A multimodal analysis of humour as an engagement strategy in YouTube research dissemination videos. *The European Journal of Humour Research*, 11(1), 46-66. <https://doi.org/10.7592/EJHR.2023.11.1.760>
- Bernad-Mechó, E., & Valeiras-Jurado, J. (2023). Multimodal engagement strategies in science dissemination: A case study of TED talks and YouTube science videos. *Discourse Studies*, 25(6), 146144562311617. <https://doi.org/10.1177/14614456231161755>
- Bezemer, J., & Kress, G. (2008). Writing in multimodal texts. *Written Communication*, 25(2), 166-195. <https://doi.org/10.1177/0741088307313177>
- Bezemer, J., & Mavers, D. (2011). Multimodal transcription as academic practice: a social semiotic perspective. *International Journal of Social Research Methodology*, 14(3), 191-206. <https://doi.org/10.1080/13645579.2011.563616>
- Bateman, J. A., Wildfeuer, J., & Hiippala, T. (2017). *Multimodality: Foundations, research and analysis. A problem-oriented introduction*. De Gruyter. <https://doi.org/10.1515/9783110479898>
- Bogush, A., & Kovtun, O. (2019). Discourse "radiotelephony of civil aviation": psycholinguistic aspect. *Psycholinguistics*, 25(1), 11-32. <https://doi.org/10.31470/2309-1797-2019-25-1-11-32>
- Burn, A. (2013). The kineikonic mode: Towards a multimodal approach to moving image media. In C. Jewitt (Ed.), *The Routledge handbook of multimodal analysis* (2nd ed.). Routledge.
- Chang, Y. H., & Liao, M.-Y. (2009). The effect of aviation safety education on passenger cabin safety awareness. *Safety Science*, 47(10), 1337-1345. <https://doi.org/10.1016/j.ssci.2009.02.001>
- Chionis, D., Karanikas, N., Iordan, A., & Svensson-Dianellou, A. (2021). Contribution of risk perception and communication in aviation safety events. *Transportation Research Record*, 2676, 405-416. <https://doi.org/10.1177/03611981211051617>
- Chionis, D., Karanikas, N., Iordan, A., & Svensson-Dianellou, A. (2022). Risk perception and communication factors in aviation: Insights from safety investigators. *Journal of Risk Research*, 25, 844-859. <https://doi.org/10.1080/13669877.2022.2038246>
- Cuddy, A. J., Glick, P., & Beninger, A. (2011). The dynamics of warmth and competence judgments, and their outcomes in organizations. *Research in Organizational Behavior*, 31, 73-98. <https://doi.org/10.1016/j.riob.2011.10.004>
- Forey, G., & Feng, D. (2016). Interpersonal meaning and audience engagement in academic presentations a multimodal discourse analysis perspective. In K. Hyland (Ed.), *The Routledge handbook of English for academic purposes*. Taylor & Francis.
- Franconeri, S. L., & Simons, D. J. (2003). Moving and looming stimuli capture attention. *Perception & Psychophysics*, 65(7), 999-1014

- Haas, E. C., & van Erp, J. B. F. (2014). Multimodal warnings to enhance risk communication and safety. *Safety Science*, 61, 29-35. <https://doi.org/10.1016/j.ssci.2013.07.011>
- Halliday, M. A. K., & Matthiessen, M. I. M. C. (2014). *Halliday's introduction to functional grammar*. Routledge.
- Hood, S. (2011). Body language in face-to-face teaching: A focus on textual and interpersonal meaning. In S. Dreyfus, S. Hood, & M. Stenglin (Eds.), *Semiotic margins: Meaning in multimodalities* (pp. 31-52). Continuum.
- Ishihara, N., & Lee, H.E. (2021). Face and (im)politeness in aviation English: The pragmatics of radiotelephony communications. *Journal of Pragmatics*, 180, 102-113. <https://doi.org/10.1016/j.pragma.2021.04.030>
- Jewitt, C. (2009). *The Routledge handbook of multimodal analysis*. Routledge.
- Jewitt, C. (2016). Multimodal analysis. In A. Georgakopoulou & T. Spilioti (Eds.), *Routledge handbook of language and digital communication* (pp. 69-84). Routledge.
- Jewitt, C., Bezemer, J. J., & O'Halloran, K. L. (2016). *Multimodality*. Routledge.
- Kines, P., Andersen, L.P., Spangenberg, S., Mikkelsen, K.L., Dyreborg, J., & Zohar, D. (2010). Improving construction site safety through leader-based verbal safety communication. *Journal of Safety Research*, 41(5), 399-406.
- Kress, G. R., & van Leeuwen, T. (2001). *Multimodal discourse: The modes and media of contemporary communication*. Arnold.
- Kress, G. R., & van Leeuwen, T. (2006). *Reading images: The grammar of visual design*. Routledge.
- Kress, G. (2010). *Multimodality: A social semiotic approach to contemporary communication*. Routledge.
- Lakey, H. (2016). *The grammar of fear: Morphosyntactic metaphor in fear constructions*. PhD Dissertation. University of Oregon. <http://hdl.handle.net/1794/20415>
- Lim, V. F. (2017). Analyzing the teachers' use of gestures in the classroom: A Systemic Functional multimodal discourse analysis approach. *Social Semiotics*, 29(1), 83-111. <https://doi.org/10.1080/10350330.2017.1412168>
- Lirola, M. M., & Chovanec, J. (2012). The dream of a perfect body comes true: Multimodality in cosmetic surgery advertising. *Discourse & Society*, 23(5), 487-507. <https://doi.org/10.1177/0957926512452970>
- Molesworth, B. R. C. (2014). Examining the effectiveness of preflight cabin safety announcements in commercial aviation. *International Journal of Aviation Psychology*, 24(4), 300-314. <https://doi.org/10.1080/10508414.2014.949511>
- Molesworth, B. R. C., Pagan, J., & Wilcock, C. (2018). Preflight safety briefings: Understanding the relationship between mode of delivery, recall of key safety messages, and mood. *International Journal of Aerospace Psychology*, 28(1-2), 1-14. <https://doi.org/10.1080/24721840.2018.1479639>
- Molesworth, B. R. C., Seneviratne, D., & Burgess, M. (2015). Selling safety: The use of celebrities in improving awareness of safety in commercial aviation. *Ergonomics*, 59(7), 989-994. <https://doi.org/10.1080/00140139.2015.1109712>
- O'Halloran, K. L. (2008). Systemic Functional-multimodal discourse analysis (SF-MDA): Constructing ideational meaning using language and visual imagery. *Visual Communication*, 7, 443-475. <https://doi.org/10.1177/1470357208096210>
- Seneviratne, D., & Molesworth, B. R. C. (2015). Employing humour and celebrities to manipulate passengers' attention to preflight safety briefing videos in commercial aviation. *Safety Science*, 75, 130-135. <https://doi.org/10.1016/j.ssci.2015.01.006>
- Tajima, A. (2004). Fatal miscommunication: English in aviation safety. *World Englishes*, 23, 451-470. <https://doi.org/10.1111/J.0883-2919.2004.00368.X>
- Tereszkiewicz, A. (2023). Engagement strategies on medical YouTube channels. *Studia Linguistica Universitatis Iagellonicae Cracoviensis*, 140(2), 139-164. <https://doi.org/10.4467/20834624sl.23.007.17756>
- Tuccio, W. A., & Garcia, A. C. (2020). "I'd a set that back at the chocks": The personal hypothetical "I would" in aviation flight instruction. *Journal of Pragmatics*, 157, 53-67. <https://doi.org/10.1016/j.pragma.2019.12.007>
- Unsworth, L. (2008). Multiliteracies and metalanguage: Describing image/text relations as a resource for negotiating multimodal texts. In D. Leu, J. Corio, M. Knobel, & C. Lankshear (Eds.), *Handbook of research on new literacies* (pp. 377-405). Erlbaum.
- Vavourakis, V., Chiew, M., & Theeuwes, J. (2022). The influence of visual salience on inattention blindness. *Attention, Perception, & Psychophysics*, 84(2), 405-418.
- Wagner, K. (2015). Reading packages: Social

- semiotics on the shelf. *Visual Communication*, 14(2), 193-220. <https://doi.org/10.1177/1470357214564281>
- White, P. (2010). Grabbing attention: The importance of modal density in advertising. *Visual Communication*, 9(4), 371-397. <https://doi.org/10.1177/1470357210382194>
- Winter, S. R., Pan, J. Y., Truong, D., & Lamb, T. L. (2021). Willingness to watch the preflight safety briefing: A structural model. *International Journal of Aerospace Psychology*, 31(3), 230-251. <https://doi.org/10.1080/24721840.2021.1883432>
- Wong, W. W. (2016). Synthetic personalization of Barack Obama at the 2008 US Democratic National Convention: A social semiotic multimodal analysis of a staged political context. *Visual Communication*, 15(4), 509-534. <https://doi.org/10.1177/1470357216632164>
- Xia, S. A., & Hafner, C. A. (2021). Engaging the online audience in the digital era: A multimodal analysis of engagement strategies in TED talk videos. *Ibérica, Journal of the European Association of Languages for Specific Purposes*, 42, 33-58. <https://doi.org/10.17398/2340-2784.42.33>
- Xiao, Q., Huang, W., Qu, L., & Li, X. (2025). The impact of multimodal information features of short sales videos on consumer engagement behavior: A multi-method approach. *Journal of Retailing and Consumer Services*, 82, 104136. <https://doi.org/10.1016/j.jretconser.2024.104136>
- Yao, Y., & Zhuo, Y. (2018). A multimodal discourse analysis of the promotional video of Hangzhou. *English Language Teaching*, 11(10), 121-131. <https://doi.org/10.5539/elt.v11n10p121>
- Yıldız, E. (2024). Navigating the skies: Unraveling communication challenges in diverse aviation accidents. *Antalya Bilim Üniversitesi Uluslararası Sosyal Bilimler Dergisi*, 5(1), 14-33. <https://doi.org/10.54969/abuijss.1459587>
- Zappavigna, M. (2016). Social media photography: Construing subjectivity in Instagram images. *Visual Communication*, 15(3), 271-292. <https://doi.org/10.1177/1470357216643220>
- Zhao, W. (2022). Gongneng yuyanxue shijiao xia de kancang anquan xuzhi huayu fenxi—Yi Aozhou Hangkong kancang anquan yanshi pian wei li [An analysis of cabin safety instructions discourse from the perspective of functional linguistics: A case study of Qantas Airlines safety demonstration video]. *Journal of Civil Aviation Flight University of China*, 33(3), 76-80.

Yixin Chen is a PhD student at the School of English Studies, Shanghai International Studies University, China. Her research interests include conversation analysis, interactional linguistics, and multimodal analysis.

Enhua Guo (corresponding author) is associate professor of linguistics at Ocean University of China. His research interests include ethnomethodology and conversation analysis, interactional linguistics, multimodal analysis, etc. Dr. Guo (co)authored articles in journals such as *Qualitative Health Research*, *Learning Culture and Social Interaction*, etc.

Li Mei (corresponding author) is associate professor at the School of English Studies, Shanghai International Studies University, China. Dr. Li's research field is conversation analysis and interactional linguistics and her major interest includes social interaction in Mandarin Chinese conversation and online interactions.