

A Multimodal Discourse Analysis of Cross-Border E-Commerce Livestreaming: A Case Study of Textile Livestreams on the TikTok Platform

Xiangchen Zhang, Hanyu Xia, Han Zheng

School of Foreign Languages, Shaoxing University, Shaoxing, China
Email: 363470161@qq.com

How to cite this paper: Zhang, X. C., Xia, H. Y., & Zheng, H. (2025). A Multimodal Discourse Analysis of Cross-Border E-Commerce Livestreaming: A Case Study of Textile Livestreams on the TikTok Platform. *Open Journal of Modern Linguistics*, 15, 707-719. <https://doi.org/10.4236/ojml.2025.154041>

Received: July 25, 2025

Accepted: August 25, 2025

Published: August 28, 2025

Copyright © 2025 by author(s) and Scientific Research Publishing Inc. This work is licensed under the Creative Commons Attribution International License (CC BY 4.0). <http://creativecommons.org/licenses/by/4.0/>



Open Access

Abstract

As a burgeoning retail format, cross-border e-commerce has become increasingly influential in shaping international trade, particularly within the textile sector. This study adopts a multimodal discourse analysis framework to investigate how live-streaming is utilized as a marketing tool in cross-border textile e-commerce. Drawing on corpus data from TikTok live sessions, the research examines how various semiotic modes—such as language, visuals, and sound—interact under limited communication conditions to optimize information delivery, reinforce meaning through complementarity, and compensate for missing sensory channels. The analysis reveals that TikTok live-stream discourse in this context is marked by both high functionality and strategic intent. Through coordinated multimodal integration, it facilitates effective message transmission and enhances the immersive experience. Furthermore, hosts deploy audience-centered discourse techniques to promote engagement, foster trust, and strengthen customer loyalty and brand attachment.

Keywords

Multimodality, Cross-Border E-Commerce Live Streaming, Discourse Analysis

1. Introduction

The accelerated growth of the global digital economy has positioned cross-border e-commerce as a transformative force reshaping global consumption behaviors (Lin, 2015). In the wake of the COVID-19 pandemic, live-streaming e-commerce—an innovative model that merges product display, real-time interaction, and direct sales—has gained global prominence. Leveraging its short video and live-streaming functionalities, TikTok has emerged as a vital platform enabling interna-

tional brands and sellers to penetrate overseas markets. Unlike conventional e-commerce, cross-border live-streaming integrates diverse semiotic resources—verbal, visual, auditory, and gestural—forming a real-time, interactive marketing discourse that significantly elevates user engagement and purchase intention. This is particularly evident in the live promotion of textile goods, where fabric texture, style, and appearance demand rich multimodal representation. The synergy between linguistic narration and visual demonstration becomes essential for effective product communication. Moreover, the complexities of cross-cultural interaction necessitate flexible discourse strategies capable of adapting to diverse linguistic and cultural backgrounds (Guan & Qin, 2025). Therefore, examining the discourse characteristics and communicative tactics of TikTok-based textile live-streaming through a multimodal lens not only addresses the practical demands of international e-commerce but also contributes to the theoretical advancement of multimodal pragmatics and digital discourse studies.

2. Literature Review

Since the 1990s, multimodal discourse has drawn considerable scholarly attention, with early studies focusing on genres such as newspapers, posters, cinema, and public speeches. As digital technologies evolved, research expanded to encompass dynamic formats like websites, digital advertisements, and other multimodal artifacts laden with layered meanings. These developments have broadened the analytical scope of discourse studies by emphasizing the interplay between verbal and visual elements to convey meaning beyond what language alone can achieve.

Multimodal discourse analysis represents a significant departure from language-centric paradigms, highlighting the interdependent roles of images, sound, gesture, and text in constructing meaning. Pioneering work by Kress and van Leeuwen (1996), drawing on Halliday's (1994) metafunctional framework, introduced “visual grammar” to explain representational, interactive, and compositional functions within non-verbal communication. Barthes (1977) further conceptualized multimodal integration, laying foundational groundwork for this evolving field.

In the digital commerce context, cross-border e-commerce platforms serve as prototypical multimodal environments. For example, Lu et al. (2024) investigated Amazon's interface design, illustrating how text, imagery, color schemes, and animation work synergistically to enhance communicative effectiveness and improve user experience through multimodal translation strategies. Their findings underscore how coordinated modality use fosters more efficient cross-cultural interaction.

Compared to static webpages, live-streaming e-commerce offers heightened interactivity and real-time responsiveness, making it an even more dynamic and complex object of multimodal analysis.

3. Methods

3.1. Research Questions

- 1) What are the key expressions and high-frequency lexical items employed by

hosts during TikTok-based cross-border e-commerce live broadcasts of textile products? How do these linguistic patterns serve commercial intentions and communicative purposes?

2) How is multimodal theory operationalized in the context of cross-border live streaming? In what ways do different semiotic modes—such as language, imagery, and sound—interact to enhance product presentation and elevate the consumer experience?

3.2. Research Methods

3.2.1. Overview of Research Methods

In cross-border e-commerce, corpus analysis has evolved far beyond its original use as a translation aid. It now plays a critical role in interpreting consumer data, guiding the creation of promotional content that aligns with the linguistic habits and search behaviors of international audiences. Corpus technology contributes not only to content generation but also to data-driven decision-making across multiple dimensions of e-commerce operations. Through access to authentic, abundant, and easily searchable datasets, corpus tools have helped shift e-commerce practices away from intuition-led approaches toward evidence-based analytics. Functions such as frequency counts, collocation extraction, and text retrospection offer insights into consumer preferences, competitive benchmarking, and category-specific market trends. These capabilities support strategic decisions in product selection, marketing campaigns, and inventory management. Additionally, the COVID-19 pandemic has accelerated the digital transformation of China's foreign trade sector, exemplified by innovations like the Cloud Canton Fair. This transformation has imposed greater demands on live-streaming production—ranging from visual background design to host performance and linguistic delivery. Given the inherent constraints in remote, screen-based interaction, applying multimodal discourse analysis allows for a deeper understanding of how various communicative modes can complement one another to enhance overall messaging impact.

3.2.2. Corpus Research

Corpus linguistics represents a research paradigm that systematically analyzes large-scale, real-world language data using computational tools. Since the 1960s, it has gained prominence for its empirically grounded approach, which prioritizes authentic language use over speculative theorizing. Scholars like Sinclair have emphasized the value of corpus-based investigation, arguing for a departure from intuition in favor of data-derived insights.

In live-streaming e-commerce discourse, corpus-based methods enable a structured and objective analysis of linguistic features, including the identification of frequent lexical items, recurring discourse patterns, and their communicative functions. By facilitating pattern recognition at scale, corpus analysis offers a robust foundation for understanding the structural and functional dynamics of cross-border marketing discourse.

3.2.3. Multimodal Discourse Analysis

Kress and van Leeuwen (2006) conceptualized “multimodal discourse” and developed a framework of visual grammar that includes representational, interactive, and compositional meanings. Building on this, Duncum (2004) observed that digital communication involves enriched semantic expressions through diverse modes. Among these, visual literacy forms the cornerstone for multimodal comprehension, while media literacy enables more effective decoding of messages in live-streaming contexts.

Zhang (2009) defines multimodal discourse as a communicative phenomenon involving multiple symbolic systems—language, image, sound, gesture—delivered via different sensory channels. In cross-border live-streaming, where tactile and olfactory channels are absent, the interaction between visual and auditory modes becomes especially vital for constructing meaning. He (2025) further argues that this framework emphasizes a multi-level analytical approach—from cultural context (such as the symbiosis of commercial logic and emotional narrative), to modal coordination (the affective empowerment of semiotic systems), and ultimately to meaning-making (the transformation from symbolic expression to social practice). This layered perspective provides a foundational basis for interpreting the sociocultural implications of multimodal signs in cross-border livestreaming.

In today’s digitally driven marketplace, incorporating multisensory channels not only strengthens the dissemination of information but also elevates the perceived value of the message. Enhancing consumers’ visual decoding abilities is key to improving their engagement with product representations in a live-streamed environment.

Typically, cross-border live-streaming relies heavily on two primary modes: visual and auditory. Visual input includes the on-screen presentation—product visuals, textual overlays, and the host’s appearance (attire, expressions, posture, and gestures). Auditory elements encompass spoken content, voice modulation, pace, and even ambient product sounds. Together, these modalities co-construct meaning, foster immersion, and amplify persuasion.

Zhang (2024) points out that differentiated strategies of multimodal coordination emerge across distinct stages of livestreaming, including the opening, product display, and promotional segments. For instance, during the product presentation stage, verbal descriptions are complemented by close-up imagery, while in the promotional phase, urgency-inducing language is synchronized with dynamic visual cues to enhance conversion effectiveness.

3.3. Research Process

This study employs AntConc, a widely-used corpus analysis tool developed by Laurence Anthony, to conduct textual analysis. Known for its functionality, flexibility, and user-friendly interface, AntConc is ideal for examining language patterns and uncovering underlying discourse structures. The dataset consists of two-

hour live-streaming sessions selected from ten representative TikTok textile sellers recorded between January and March 2025 with totally 127,493 tokens. Six live-stream recordings were selected for in-depth analysis, covering five product subcategories: coats, jumpsuits, underwear, stockings, and women's handbags. Among them, one category—overcoats—was covered in two livestreams to capture internal variation within the same product type.

The selection criteria followed a stratified purposive sampling strategy aimed at reflecting diverse marketing discourses across product functionalities (e.g., outerwear vs. intimate wear) and consumer appeal (e.g., practicality vs. aesthetic). All livestreams were conducted by top-performing influencers, ensuring the corpus captures mainstream persuasive strategies and stylistic norms within TikTok's fashion commerce ecosystem.

By applying AntConc to extract high-frequency vocabulary, the study explores the discourse strategies underlying these linguistic patterns.

4. Data Analysis and Discussion

4.1. Corpus Analysis

Latent Dirichlet Allocation (LDA) is a generative probabilistic model for collections of discrete data such as text corpora. It assumes that each document is a mixture of a small number of latent topics, and each topic is characterized by a distribution over words. LDA enables the automatic discovery of thematic structures in large corpora, making it highly useful for topic modeling in text mining and corpus linguistics.

Originally proposed by [Blei, Ng, and Jordan](#), LDA has since become a foundational method for unsupervised topic discovery, applied in diverse domains ranging from social media analysis to academic research classification ([Blei et al., 2003](#)).

This section presents the thematic classification of linguistic data based on keyword frequency and the communicative characteristics of textile-focused live broadcasts. To investigate the thematic structure of livestream discourse, this study employed Latent Dirichlet Allocation (LDA) topic modeling using the `gensim` library in Python. The number of topics was set to three, a decision informed by both the optimization of the model's topic coherence score and the empirical observation of three distinct discourse categories in the corpus: product attributes and presentation, promotional and purchasing guidance, and user experience and feedback. The topic coherence metric used was the C_v score, which reached 0.54, indicating a moderate level of semantic interpretability. Additionally, the model's perplexity was -7.2 , suggesting a satisfactory fit given the size of the corpus and the chosen topic granularity. By comparing the model's output with the categories of high-frequency terms listed in the word frequency table (see [Table 1](#)), and validating through manual annotation, the resulting topic clusters were found to be highly consistent with the actual discourse types in semantic terms. Using a combination of LDA topic modeling and manual coding, the study identifies three

primary discourse themes. Each theme incorporates a range of high-frequency nouns, verbs, adjectives, adverbs, and pronouns, reflecting the strategic and functional nature of live-stream discourse.

Table 1. Frequency of keywords under different thematic classifications.

Subject Classification	High-frequency words (frequency in brackets)
Product Attributes and Display	color, (434) size (638), material (101), quality (157), weight (70), discount (15), sale (94), coupon (44), price (92), review (13), love (390), satisfaction (2)
Promotion and Purchase Guidance	beautiful (462), comfortable (110), soft (50), durable (23), stretchy (11), limited (10), special (40), affordable (16), unbeatable (14), perfect (162), happy (82), amazing (181), gorgeous (181)
User Experience and Feedback	show (292), check (197), see (596), look (278), try (49), order (49), purchase (301), grab (119), buy (170), click (81), enjoy (162), feel (205), recommend (7), trust (62), now (492), today (226), quickly (17), we (1474), you (6344), your (1148)

The distribution of word types and usage frequency of keywords under different themes reveals the language characteristics of both functionality and strategy in live-streaming discourse:

First, product attributes and presentation. In this theme, the host uses nouns (such as color, size, material, quality, weight) (see **Table 2**) frequently to precisely describe the basic physical attributes of the product, constructing a product information framework. At the same time, the frequent appearance of adjectives (such as beautiful, comfortable, soft, durable, stretchy) strengthens the subjective evaluation of the product's experiential value, highlighting selling points such as comfort, durability and aesthetics.

Take “color” as an example:

Table 2. Example of product attributes and presentation.

appal color. For those people ones interested in the applical	color,	you can also go ahead and purchase it right
There's a different size to the size to the	color.	You can will check directly show you the price
purchase your so many different colors. You can choose any	color	you want to purchase. Rainfield. Okay, we have so

Furthermore, the relevant verbs (such as order, show, check, see, look, try) (see **Table 3**) reflect the dynamic demonstration features during the live broadcast process, guiding users to focus on specific product characteristics through visual actions. This discourse practice combining “physical description + sensory rhetoric + dynamic demonstration” helps enhance consumers' intuitive perception and purchase interest.

Second, promotion and purchase guidance. Promotion and purchase guidance type of discourse occupies an important position in the live broadcast discourse system. High-frequency nouns (such as discount, sale, coupon, price, free) prominently present key information such as promotional activities and price advantages. The frequent use of adjectives (such as limited, special, affordable, unbeatable) further enhances the attractiveness and uniqueness of the products through descriptive expressions, strengthening the sense of urgency for consumption. Additionally, verbs such as order, purchase, grab, buy, click directly point to the purchasing behavior, accompanied by adverbs like now, today, quickly, which impose an immediate pressure in terms of time, prompting users to place orders quickly. The characteristics of this type of discourse demonstrate the deep utilization of emotional mobilization and behavior guidance mechanisms by live broadcast discourse in sales conversion.

Take “order” as an example:

Table 3. Example of promotion and purchase guidance.

this one. But if our sister want to place an	order	for this one, feel free to let me know.
they are gone for good. Uh, little fairy has placed an	order	for those productlic Holly up and pay. Oh, otherwise
right now as a freebie time, because it plays an	order	for too bad. You can get our crucial real

Third, User Experience and Feedback. The theme of user experience and feedback mainly builds brand trust and user identification by presenting positive evaluations and emotional expressions. High-frequency nouns (such as feedback, review, love, satisfaction) directly reflect consumers’ experiences and emotional attitudes, while adjectives (such as perfect, happy, amazing, gorgeous) further enhance positive feelings and increase emotional resonance.

In this theme, verbs (such as enjoy, feel, recommend, trust) frequently appear, strengthening the experience-centered discourse orientation. At the same time, the extensive use of high-frequency pronouns (such as you, your, we) reflects the discourse strategy in live streaming that focuses on users and emphasizes interaction and a sense of belonging. This discourse style, which builds on “emotional evaluation + direct substitution”, helps reduce the perception of transaction risks and enhance consumers’ psychological security and loyalty.

Take “enjoy” as an example (see **Table 4**). The words that are commonly used in combination with it are mostly “deal”.

4.2. Multimodal Discourse Analysis

4.2.1. Auditory Modality Analysis

In cross-border e-commerce live broadcasts, Ji & Li (2021) point out that the auditory modality mainly refers to the language expression of the host, which is the core channel for information transmission. The host needs to maintain clear

Table 4. Example of user experience and feedback.

the last four minutes right now, so if you to	enjoy	those good deal, let's go ahead and purchase
ahead and get for yourself within the flas of a	enjoy	those good deals is definitely going to be where
been so so long so as if you want to	enjoy	those good deals, let's go ahead and purchase
purchase you what is as fast as you can and	enjoy	those good deals. Trust me, they are going to
order let's. Go ahead and purchase right there and	enjoy	those good do and'm going to show you
you go ahead and get those coupons. Right there and	enjoy	those good don't have hesitate and especially if
a great deal. So let's go ahead and enjoy.	Enjoy	those good job. Definitely the best policy material to
go ahead and purchase your orders. Okay, go ahead and	enjoy	those good tools. Definitely the best quality to materials
anymore. They are sold out. So if you want to	enjoy	those good you have the favorite color ent your
go ahead and purchase your own orderss right there and	enjoy	those good zones. Okay, this is what we got

pronunciation, moderate speech rate, and appropriate stress to enhance the comprehensibility and communication effectiveness of the language. A steady speech rate helps to enhance the comfort of the audience, and in certain situations, by appropriately increasing or decreasing the speech rate, it can stimulate auditory attention and highlight key information, especially when conveying core content such as price, performance, and advantages of the product, the application of stress is particularly crucial. Taking the live broadcast of women's stockings as an example, an analysis of the auditory modality in cross-border e-commerce live broadcasts is conducted. The auditory modality has strong interactivity and is very direct in the live broadcast, it is the most important modality in cross-border e-commerce live broadcasts, and this modality is used throughout the entire live broadcast process. However, in textile live broadcasts, the host will focus on introducing the subjective perspective of consumers through the description of the product's usage experience, such as "A seamless design and anti-static technology ensure a flawless appearance and the comfortable fit will keep you feeling great all night at the party." This emphasizes the visual and tactile experience of the product in real life, guiding consumers to make purchases. During the process of forcing orders in the live broadcast, in order to create a tense and intense purchasing atmosphere, the host usually speeds up the speech rate, switches to background music with a faster rhythm, emphasizes the core advantages, discounted prices, discounts, and discount periods through stress, and adds the sound of the live

broadcast props bell to continuously stimulate the audience's auditory sense, guiding customers to place orders and rush to purchase as soon as possible.

To further demonstrate the integration of linguistic themes with multimodal resources, each of the three LDA-derived discourse themes is exemplified through corresponding auditory or visual instances drawn from the livestream data. The theme of product attributes and presentation is often paired with dynamic visual modalities, such as close-up camera shots and product rotation movements, allowing viewers to observe color, texture, and design details, while the host synchronously provides descriptive language (e.g., “this color really pops on camera” or “look at this silky finish”). The theme of promotion and purchasing guidance typically corresponds with intensified auditory signals: hosts raise their voice pitch, accelerate their speech rate, and introduce rhythmic background music to emphasize time-limited offers, e.g., “Only today! Grab it now!”, often supported visually with flashing banners or pop-up discount icons. Finally, the theme of user experience and feedback is supported by both auditory narratives (e.g., read customer reviews aloud) and static visual displays, such as rotating screenshots of 5-star feedback and on-screen overlays of emotive phrases like “99% satisfaction.” These multimodal alignments not only enhance information delivery but also reinforce persuasion, urgency, and consumer trust.

4.2.2. Visual Modality Analysis

In the specific application of visual modality, [Chen \(2020\)](#) points out that its core function lies in achieving the display and transmission of information through visual means. Visual modality encompasses static images, textual information, and dynamic images (such as videos). Although simple pictures and text have certain limitations in conveying a “realistic sense” and are difficult to create a strong sense of presence, they possess high generalization and refinement capabilities. Compared to other modalities, the combination of text and graphics has a more systematic and logical expression, can effectively present the core information, and helps the audience understand and master it within a short period of time. In addition, carefully designed images often have strong visual appeal. For example, embedding promotional posters in the live broadcast page can not only enhance the visual tension but also quickly attract users' attention visually, guiding them to pay attention to the brand or product, thereby promoting the occurrence of purchase behavior.

Regarding the static visual elements in cross-border e-commerce live broadcasts, this modality is usually regarded as a secondary modality. Although it has the characteristic of being persistent and always visible throughout the entire live broadcast, its interactivity is low, the sense of participation is not strong, and its ability to actively convey information is limited. Therefore, static visual content exists more as an auxiliary form, usually combined with dynamic modalities (such as language, actions, etc.) to enhance the overall information transmission effect and audience acceptance. For example, in the cross-border live broadcast of women's underwear, the operator will switch positive review screenshots of the

products every 5 seconds to achieve a continuous photo playback effect creating a psychological feeling that the product has received numerous positive reviews for the audience (see **Figure 1**). Through the overload input of positive review information, the trust of consumers can be increased, thereby guiding customers to place orders.

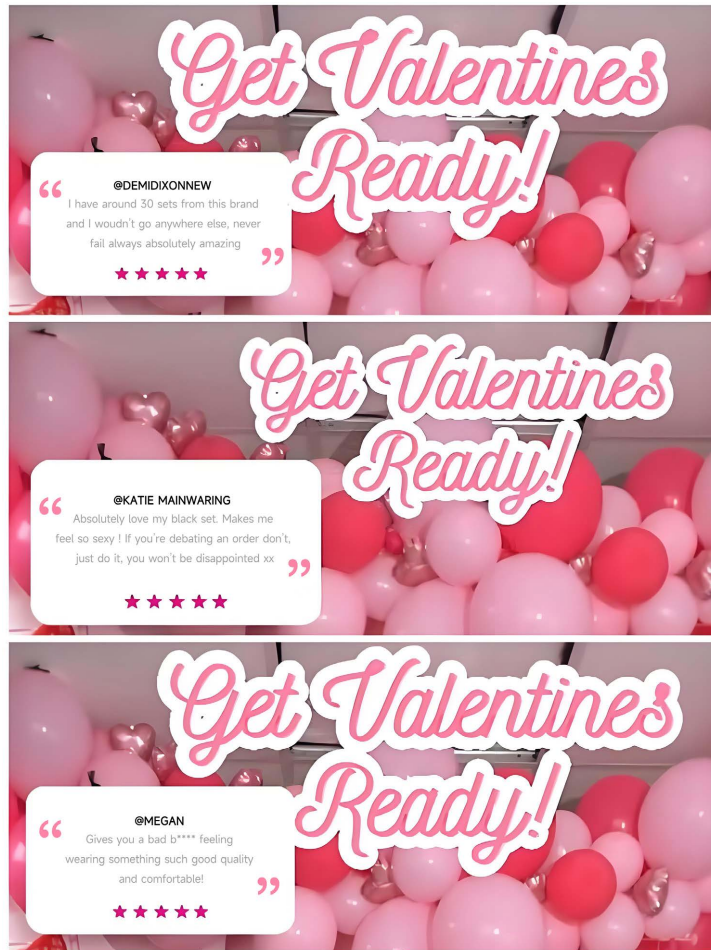


Figure 1. Example of positive review screenshots.

The host can showcase dynamic visual modalities through body movements, thereby enhancing the visual expressiveness in live broadcasts. For instance, in daily presentations, the host usually maintains a certain distance from the camera equipment, but when promoting detailed products such as ladies' handbags, the host will consciously approach the camera and slightly lean forward, shortening the visual distance between the product and the audience. At the same time, through turns, angle changes, etc., the host enables consumers to intuitively see the overall form and detailed features of the product from multiple perspectives and feel the dynamic effect of the product in the dressing context. Such dynamic visual modalities not only enhance the vividness of the presentation but also play a role in reinforcing information, clearly conveying the key content being highlighted.

4.2.3. Modal Complementarity Analysis

In cross-border e-commerce live broadcasts, the collaboration between different modalities is particularly crucial, and the complementarity between visual and auditory modalities is the most typical. The combined use of dynamic visual modalities and auditory modalities can enhance the expressiveness of information; while static visual modalities play a supporting and supplementary role to auditory information to some extent. In the visual dimension, there is also a need for coordination between static and dynamic elements to enhance the overall communication effect (Yuan, 2024). Due to the limitations of the screen display in cross-border live broadcasts, the olfactory and tactile modalities are difficult to be directly presented. At this time, the host can substitute through their own actions and language: for example, the host can touch the product and express the tactile sensation with words, or actively smell the product and describe its odor, thereby indirectly transmitting information that cannot be intuitively perceived through the senses, enhancing the immersive experience of consumers. In practical operations, this modal complementarity strategy is often reflected in specific live broadcast scenarios. For example, when the host guides the audience to place an order, they will use words to explain: “Click the link below to receive a free gift.” Subsequently, the right lower corner of the live broadcast screen will simultaneously display product promotion images and gift images. This technique conveys the information through language (auditory modalities) and images (static visual modalities), making the keyword “free gift” not only be perceived in the auditory sense but also be presented vividly through the image form, thereby deepening the audience’s impression. Another example is when presenting ladies’ underwear with a holiday theme, the host takes the underwear out from the background wall cabinet, uses a dedicated hanger to present it in a three-dimensional way, and repeatedly touches and rotates the product to display its material and design, combined with real-time verbal explanations, forming a close linkage between language and dynamic visuals. In terms of language expression, the host uses emotionally charged rhetorical questions such as “How fabulous does that sound?” to enhance the impact of the tone, while creating an interactive atmosphere. The host also sets up scenarios, such as “If everyone puts a thousand likes right now, girls, you would be amazed.” This hypothetical conditional sentence effectively attracts user attention, stimulates interaction behavior, and creates an expectation atmosphere, thereby enhancing the live broadcast’s activity and audience participation.

5. Summary

This study conducted a corpus and multimodal analysis of the discourse of textile category cross-border e-commerce live broadcasts on the TikTok platform. It was found that such live broadcasts have a high degree of functionality and strategy. The discourse revolves around three main themes: “product attribute display”, “promotional guidance”, and “user feedback”, demonstrating the precision and directionality of language in marketing. Additionally, the synergy of visual

and auditory modalities enhances the effectiveness of information dissemination. The host uses multimodal means such as language, expressions, and actions to create an immersive viewing experience. The study also reveals that the complementarity of modalities such as language and images overcomes sensory limitations, reconstructs consumers' perception system, and enhances immersion and trust. Moreover, the user-centered discourse strategy, such as frequent use of second-person pronouns, emotional verbs, and various interactive sentence structures, effectively promotes user engagement and brand recognition, showcasing the unique discourse construction mechanism and communication advantages of cross-border live e-commerce.

The results refine Kress and van Leeuwen's metafunctional model by demonstrating how interactive, time-sensitive elements in live-commerce settings amplify the interpersonal function through real-time verbal cues and synchronized multimodal prompts, thus extending the model's applicability to dynamic, participatory digital commerce discourse.

This study reveals the guiding mechanism of frequent keywords and their corresponding discourse strategies on consumer decision-making, and points out how hosts utilize the three language dimensions and multimodal collaborative means of "product-promotion-experience" to stimulate user interest and purchase behavior. The research results provide systematic suggestions for cross-border e-commerce enterprises to optimize live content design, enhance user stickiness and conversion rate, and have high application and promotion value.

6. Limitations and Future Directions

Despite offering insights into the multimodal characteristics of TikTok-based cross-border textile livestreaming, this study is not without limitations. First, the sample size is relatively small, comprising only six livestream sessions, which may limit the generalizability of the findings. Second, the analysis focuses solely on one platform—TikTok—thus potentially overlooking modality-specific or discourse-style variations present on other platforms such as YouTube Live or Taobao Live. Third, the data were collected within a narrow time window (January to March 2025), which does not capture seasonal or temporal shifts in promotional strategies or linguistic choices. Future research should consider expanding the dataset to include multiple platforms, broader product categories, and a longer observation period. Additionally, integrating consumer reception data—such as comment analysis or engagement metrics—could further enrich the understanding of how multimodal strategies influence consumer behavior and trust-building in live commerce.

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

References

Barthes, R. (1977). *Image-Music-Text*. Hill and Wang.

- Blei, D. M., Ng, A. Y., & Jordan, M. I. (2003). Latent Dirichlet Allocation. *Journal of Machine Learning Research*, 3, 993-1022.
<https://www.jmlr.org/papers/volume3/blei03a/blei03a.pdf>
- Chen, Y. (2020). A Study on Cross-Border E-Commerce Live-Streaming Discourse from a Multimodal Perspective. *Digitization User*, 11, 128-130. (In Chinese)
- Duncum, P. (2004). Visual Culture Isn't Just Visual: Multiliteracy, Multimodality and Meaning. *Studies in Art Education*, 45, 252-264.
<https://doi.org/10.1080/00393541.2004.11651771>
- Guan, L. L., & Qin, L. (2025). Multimodal Discourse Analysis of Popular Chinese and American Livestreamers from Taobao and Amazon: Li Jiaqi and Brittany Vasseur. *Media and Communication Research*, 6, 93-107.
<https://dx.doi.org/10.23977/mediacr.2025.060113>
- Halliday, M. (1994). *An Introduction to Functional Grammar*. Edward Arnold.
- He, X. L. (2025). A Multimodal Discourse Analysis of E-Commerce Livestreaming from the Perspective of the Emotion Economy. *China Story*, 5, 90-92. (In Chinese)
- Ji, X. R., & Li, W. L. (2021). Analysis of Cross-Border E-Commerce Live-Streaming Discourse from a Multimodal Perspective. *Overseas English*, 20, 92-93. (In Chinese)
- Kress, G., & Van Leeuwen, T. (2006). *Reading Images: The Grammar of Visual Design* (2nd ed.). Routledge.
- Lin, Y. Z. (2015). Analysis of the Current Status and Opportunities of Cross-Border E-Commerce Development. *Modern Economic Information*, 21, 326-327. (In Chinese)
- Ye, L., Du, L. X., Zhang, J. J., & Wang, X. (2024). A Study on Multimodal Discourse Construction and Translation of Cross-Border E-Commerce English Websites: A Case of Amazon. *Overseas English*, 7, 82-85. (In Chinese)
- Yuan, X. Y. (2024). *A Multimodal Discourse Analysis of Chinese E-Commerce Live-streaming Discourse*. Master's Thesis, Shandong Normal University. (In Chinese)
<https://link.cnki.net/doi/10.27280/d.cnki.gsdsu.2024.000254doi:10.27280/d.cnki.gsdsu.2024.000254>
- Zhang, D. L. (2009). Exploring a Comprehensive Theoretical Framework for Multimodal Discourse Analysis. *Foreign Languages in China*, 6, 24-30. (In Chinese)
- Zhang, X. Y. (2024). Multimodal Discourse Coordination in Overseas E-Commerce Livestreaming. *Popular Investment Guide*, 27, 14-18. (In Chinese)