

# Sustainable beauty in the Digital era: A multimodal discourse study of tiktok live commerce in Indonesia

*Wirdatul Khasanah*<sup>1\*</sup>, *Suyanti Fatma Umayfa*<sup>1</sup>, *Imam Hanafi*<sup>1</sup>, *Rachmat Efendi*<sup>1</sup>, and *Hind Mulfi*<sup>2</sup>

<sup>1</sup>English Literature, PSDKU Universitas Negeri Surabaya, Indonesia

<sup>2</sup>English Literature, School of English, Communication, and Philosophy, Cardiff University, UK

**Abstract.** The rise of TikTok Live Commerce has reshaped Indonesia's beauty industry into a digital arena where sustainability becomes both a market value and a communication strategy. This study investigates how sustainability is linguistically and visually constructed in beauty live streams through Multimodal Discourse Analysis (MDA). Ten live sessions featuring high-end and mass-market brands were examined to identify patterns of tone, mood, gesture, and visual composition. Findings reveal two semiotic tendencies: high-end sellers perform symbolic capital through calm tone, scientific vocabulary, and minimalist aesthetics, while mass-market sellers display emotional capital through expressive speech, slang, and colorful visuals. These modes collectively shape a digital discourse of sustainable beauty that intertwines class, trust, and consumer engagement. The study argues that digital communication operates as a form of symbolic and cultural capital in the sustainable economy, supporting SDG 8 (Decent Work and Economic Growth) and SDG 12 (Responsible Consumption and Production).

## 1 Introduction

TikTok shop has been a game changer and a fresh air in Indonesia's digital market since its first launch in April 2021. TikTok Shops allows brands to market their products not only just through online videos and collaborations with influencers, but also through live videos streaming where sellers and buyers can interact in real time through persuasive talk, expressive gestures, and visually engaging product demonstrations [1]. In Indonesia, TikTok Live commerce environment has a clear distinct branding style between high-end and mass-market brands. High-end brands tend to adopt a soft spoken and refined linguistics style such as using polite address forms, slower speech tempo, and carefully chosen vocabulary that signals prestige and exclusivity [2]. On the contrary, the mass market brand usually market their products using energetic expressive and informal communication style; moreover, they also frequently use humour, bright visual and fast movements and crowded on screen settings

---

\* Corresponding author: [wirdatulkhasanah@unesa.ac.id](mailto:wirdatulkhasanah@unesa.ac.id)

to create a sense of excitement and accessibility [3]. These linguistic and multimodal differences reinforce class distinction between societies. As a result, the digital economy, while open for everyone, often reproduces social inequalities found offline which give more privilege to those who can perform classed identities effectively.

While research on TikTok commerce has grown rapidly in recent years, most studies primarily focused on marketing strategies on TikTok [4], consumer behaviour business models [5], and emphasising the influence on of influencers credibility with consumer trust [6]. Far fewer studies have explored how language and multimodal performance construct social meaning such as class, prestige, or accessibility within live commerce interactions. The ways in which linguistics choices, gestures, tone, and visual presentations remain underdeveloped. Moreover, research between discourse analysis (especially on multimodal analysis) and the sustainability of the digital economy is still limited.

To address the research gap, this study investigates how sellers in Indonesian TikTok Live beauty commerce mobilize linguistic and multimodal resources to build brand class identity. More precisely, it investigates how interpersonal linguistic attributes including mood, modality, appraisal, and lexical choice are used to create an impression of authority, intimacy, or prestige during live streaming interactions. In addition, our paper also explores the role that visual and performative modes (gesture, gaze, camera framing, and visual composition) play in enacting class distinctions during real-time commercial discourse. By combining these linguistic and multimodal factors, this study also seeks to address how these communicative practices express in particular questions about sustainability, trust, and social inequality in Indonesia's digital economy.

This study aims to analyse how language and multimodal aspects construct brand identity in Indonesia TikTok Live commerce by focusing on how sellers construct brand's identity through linguistic choices, gestures, tones, and visual presentation. By examining these practices, this study aims to reveal how symbolic hierarchies are reproduced or potentially transformed within the digital marketplace in Indonesia. This study directly contributes to Sustainable Development Goals 8 (Decent Work and Economic Growth) by highlighting how digital entrepreneurship becomes more equitable when communicative competence and linguistics awareness are recognized as economic assets. It also aligns with SDGs 10 (Reduced Inequalities) by looking at how classed discourse may exclude or disadvantage certain groups of sellers and by advocating for more inclusive and culturally sensitive communication strategies. In a broader sense, this research promotes inclusive communication in digital entrepreneurship which encourages development of discourse practices that empower diverse voices and sustain fair participation in Indonesia's digital economy.

## **2 Literature review**

### **2.1 Multimodal Critical Discourse Analysis (MCDA)**

This study uses Multimodal Critical Discourse Analysis (MCDA) as its primary analytical framework. MCDA has been widely used to analyse multimodal meaning-making in digital discourse, yet many existing studies are focused on stable and text-like artefacts. This focus risks missing out on the ways power, class, and ideology are actively negotiated in live, performative commercial contexts. Drawing on TikTok Live commerce in Indonesia, this research expands MCDA to explore how classed identities and symbolic hierarchies emerge dynamically through real-time linguistic and multimodal interaction.

## **2.2 Systemic functional linguistics**

To analyse the verbal dimension of branding discourse, this study uses Systemic Functional linguistics (SFL) developed by Halliday in 1994 and later expanded by Eggins in 2004 [7]. In accordance with the structural logics of Systemic Functional Linguistics (SFL), this study primarily foregrounds the interpersonal metafunction, which is most apt for studying live commercial interaction in TikTok Live commerce. The ideational metafunction pertains to representational experiences; whereas the textual metafunction relates to message organisation. The interpersonal metafunction, however, encompasses ways that speakers negotiate social relations, authority, intimacy, and evaluation through language use like mood, modality, and appraisal. These interpersonal resources are also important in live-streamed selling contexts, where trust, credibility and audience engagement are significantly influenced by these resources. Accordingly, this study emphasizes interpersonal meanings to uncover how sellers linguistically project both brand identity and class distinction in the context of contemporary digital commerce.

## **2.3 Multimodality**

In order to analyse the visual dimensions, the theory of Visual grammars is used [5]. According to Kress & Leeuwen, images is just like a language where it conveys meaning through structured form which realises representational, interactive, and compositional meanings. These meanings can be seen from gaze, colour, lighting, gestures, and any other visual aspects which may present. In the context of TikTok live commerce, these multimodal features play a central role in how sellers perform brand identity and class. For example, gestures and spatial arrangement contribute to the performance of intimacy or distance between seller and audience. By analysing this performative cues, this study investigates how brand class is visually presented and how multimodal design contributes to the perception of prestige or accessibility in Indonesia's digital market.

## **2.4 Critical and social dimension**

To examine how power and ideology shape communicative practices in digital branding, this study uses Fairclough's [8] model of Critical Discourse Analysis (CDA). Fairclough views discourse as both a form of social practices and other semiotic modes contribute to the production and maintenance of power relations. CDA sees communicative choices such as gestures, tones, and visual elements not only as rhetorical strategies but as the reflection of broader social structures. CDA critically analyses how language, ideology, and power relates one another in order to uncover unequal society. In the context of TikTok Live Commerce, the way the sellers talk, interact and represent themselves could be seen as a symbolic sign which positions the way they put their brand in the digital market. By combining both critical perspectives of Fairclough and Class theory of Bourdieu [8], this study defines multimodal practices as not just communicative strategy but also performative practice where social inequality is negotiated and maintained within the digital economy.

## **3 Methodology**

This study adopts qualitative interpretive design using Multimodal Critical Discourse Analysis (MCDA) as the basis of analysis. This is chosen because this study focuses on meaning-making within discourse analysis. MCDA allows the researchers to analyse deeply on how linguistics elements, visuals and performative practice are used in digital communication [9]. TikTok Live Commerce fits to the current research as it includes

language, gestures, intonation, face expression, and visuals which interact one another simultaneously to create message and branding identity.

### **3.1 Data source**

The data on this study was collected from 10 TikTok Live commercials focusing on beauty products such as skincare, haircare, and cosmetics products. Those live videos were collected by using purposive sampling to represent high-end beauty brands and mass market beauty brands. The product was classified based on three main categories: (1) product category and price range, (2) linguistics and visual tone displayed during livestream, and (3) branding aesthetic such as gesture, camera framing, and colour scheme.

The beauty industry was chosen because it represents the most dynamic and competitive sector in the Indonesian digital market where the sellers depend a lot on visual appearance to build trust and emotional closeness to the audience. Every TikTok Video was recorded 2 to 5 minutes long, screen recorded then transcribed to be analysed using linguistics aspects and multimodal analysis. All data were live streamed publicly from Early September to October 2025.

### **3.2 Data collection**

Data were collected using video recording and transcribed and noting some multimodal analysis to catch complexity in communication. The live streaming was recorded to preserve the verbal component and visual interaction. The recording then transcribed verbatim to represent the linguistics aspect from the sellers including intonation, pauses, and expressions. In addition, screen captures were taken at the key moments that document relevant gestures, facial expressions, camera framing, product placement, and visual layout. A set of notes accompanied each video to record the non verbal expression such gaze direction, body posture, and interaction with the audience through chat or comment responses. This multi-layered data collection procedure ensures that both linguistics and visual performances are represented accurately to enable comprehensive analysis of how brand class is constructed through multimodal communication in beauty-related live commerce contexts.

### **3.3 Data analysis**

The data analysis was conducted in three interrelated stages to capture both linguistics and multimodal dimensions of meaning making. Step 1, linguistics analysis focused on examining lexical choice, tone, mood, and evaluative language using SFL. This stage aimed to identify how sellers constructed authority, friendliness, or prestige through specific word choices and interpersonal meanings. Step 2, multimodal analysis. This multimodal analysis uses the visual grammar [5] to analyse gestures, gaze, camera framing, visual composition, and colour grading to uncover how visual elements and body strengthen the brand identity as well as class identity. Step 3 focused on interpretative synthesis. This was done by integrating the findings from both linguistics and visual analyses to interpret how class identity was constructed across semiotic modes. In this way, in order to raise the analytical rigour, data coding and interpreting process involved more than one researcher. Linguistic and multimodal features were analyzed independently, which were reviewed and discussed with the researchers. Interpretation differences were resolved by iterating through discussion until it reached an analytical consensus. This technique allowed to reduce individual bias and thus increased credibility of qualitative findings.

## 4 Results and analysis

### 4.1 Linguistics findings

From the linguistics perspective, interpersonal metafunction highlights how speakers, in this context in the sellers, build relations with the audience through mood, modality, and appraisal. The result of this study shows that there is clear distinction between the lexical choice used by the high-end brand sellers and mass-market sellers.

#### 4.1.1 High-end brand

High-end sellers display a consultative and professional tone with declarative sentence and moderate modality. They prefer technical vocabulary and scientific explanation to project authority and trustworthiness.

**Excerpt from the video:**

*“Kalau kulit kalian cenderung sensitif, hindari ingredients yang bisa memicu atopic reaction, ya.”*

**Translation:**

“If your skin tends to be sensitive, avoid ingredients that might trigger an atopic reaction, okay?”

This utterance uses declarative mood and low-intensity appraisal which focuses on objective information rather than emotion. The polite particle “ya” softens the tone to show approachableness while still represent to be an expert.

Moreover, premium-level sellers regularly couch their product-promoting messages within a problem–solution sequence structure and act as consultants not as aggressors when promoting the product. This interactional pattern is realised linguistically in the form of interrogative clauses that prompt viewers to express skin concerns, followed by declarative explanations that propose product use as a rational response. The controlled speech rate, cautious evaluative language, and targeted incorporation of scientific terminology (for instance, biome, atopic, and fungal) act as interpersonal resources that create the conditions for epistemic authority and professional distance. These lexical choices together contribute to a discourse of expertise and exclusivity, associating high-end branding with symbolic capital and trust-based engagement, rather than emotional urgency.

In addition to the previous example, high-end brand sellers consistently employ cautious evaluative language and controlled modality to construct professional authority.

**Excerpt 2 (High-end brand):**

*“Produk ini diformulasikan khusus untuk menjaga skin barrier, jadi relatif aman digunakan untuk pemakaian jangka panjang.”*

**Translation:**

“This product is specifically formulated to maintain the skin barrier, making it relatively safe for long-term use.”

This utterance is realized through a declarative mood combined with low to moderate modality. In this case, instead of stating a statement of certainty, the seller mitigates the assertion of scientific responsibility and credibility. The incorporation of technical lexical items like formulated and skin barrier places the seller in a context of expert knowledge and professional competence. According to the appraisal approach, the evaluation is constructed via graduated appreciation, avoiding exaggerated affect and instead stressing a balanced and measured reliability. This linguistic approach helps create a perception of expert confidence

while also supporting the symbolic status of high-end brands as trustworthy, cautious, and science-oriented through its use of expert-specific language.

#### 4.1.2 Mass-market

In contrast, mass market sellers adopt an energetic and intimate speech which is dominated by imperatives, repetition, and emotional markers.

**Excerpt (Original):**

*“Bestiee! Ini serum bikin glowing banget! Buruan check-out sekarang!”*

**Translation:** “Bestie! This serum makes you super glowing! Hurry and check out now!”

This statement is governed by short imperative clauses (hurry up and check out now) and exclamatory expressions, which serve to create immediacy and promotional urgency. The use of informal address (bestie) reduces social distance and establishes familiarity between the seller and the audience. From an appraisal perspective, the expression glowing so much conveys high-intensity positive appreciation, amplifying enthusiasm and emotional involvement.

Additionally, mass-market sellers frequently use lexical repetition as a persuasive tool. Promotional phrases like hurry up or check out now are often repeated in a short interactional span to emphasize urgency and prompt a quick response from consumers. Linguistically, this type of high-affect/low modality conversation values emotional connection over informational caution. These interactions position mass-market sellers as accessible and community-based compared to the restrained and expertise-driven discourse used in high-end branding.

## 4.2 Multimodal findings

The multimodal analyses also show significant visual and performative differences between the two brand classes.



**Fig. 1.** Visual and performative differences between the two brand classes

#### 4.2.1 High-end brands

High-end brand sellers appear in formal and professional attire such as blazers or clean, minimalist clothing with soft neutral colours such as black dominant or white dominant. The background is organized with white or beige tones. The body movement of the host is composed, minimal, and the facial expressions are calm. The overall visual composition reflects elegance, order, and expertise. This shows that they are knowledgeable consultants.

### **4.2.2 Mass-market**

Meanwhile, mass market sellers perform expressive and energetic gestures. They often wear flashy outfits such as pink headbands along with colourful excessive digital effects like fire animation or glowing product visuals. The background is crowded with multiple product displays. The voices are loud with some excitement. The performative style mirrors popular entertainment formats which align with the mass market goal of immediacy and inclusiveness.

### **4.3 Interpretive synthesis**

By integrating linguistics and multimodal findings, this study shows that class distinction in terms of digital branding is constructed through a combination of expertise based discourse and emotional performance. high-end brands rely on professionalism and also scientific knowledge especially to the products to represent prestige. Furthermore, the calm gestures and visual simplicity reflect a discourse of authority and exclusivity. Conversely, mass market sellers use emotional and participatory discourse to build community through shared language and informal interactions. This aligns with what Fairclough [7] calls as the conversationalization of discourse in digital capitalism. The finding shows that digital entrepreneurship opens access to more mass market but at the same time also still maintains social stratifications. Therefore, linguistics and multimodal literacy is the key to actively encourage equitable participation in Indonesia sustainable digital economy.

## **5 Discussion**

The finding reveals that class identity in digital spaces is performed through the interaction of linguistics, visual, and performative modes. In the context of TikTok Live Commerce of beauty products, the class is not only represented through prices or exclusivity of the products but also from the way the host live communicates. The high-end brand uses linguistic which feature controlled tone, scientific vocabulary, and structured explanations which conveys professionalism and expertise, these elements convey the symbolic authority and exclusivity [7]. On the other hand, mass market products rely on linguistic features that emphasises inclusivity, speed, and emotional connection which align with the conversationalization trend of modern capitalism [7]. These findings both align with the previous studies [2, 3].

This finding extends Bourdieu's notion of symbolic capital by demonstrating that communicative performance in live digital commerce operates as a form of capital in its own right. In TikTok Live interactions, authority and credibility are not derived solely from economic resources or brand status, but are actively produced through linguistic restraint, scientific vocabulary, and controlled multimodal performance. These findings suggest that symbolic capital in digital marketplaces is increasingly contingent upon sellers' ability to perform professionalism and expertise through language and multimodality.

### **5.1 Social inequality and inclusion**

From a critical viewpoint, it highlights that the linguistic and multimodal distinction between high-end and mass-market brands perform inequality in terms of hierarchy in digital platforms. Being calm, controlled gestures, and minimalist appearance and arrangement show a symbolic mode of education, professionalism, and prestige [10]. Conversely, being expressive and friendly show emotion which highlights inclusivity and closeness [11]. Simultaneously, this conversationalisation does not eliminate power relations but rather reconfigures them. While mass-market discourse appears inclusive and accessible, it

simultaneously normalises urgency-driven consumption, aligning affective engagement with the logics of digital capitalism. However, these differences also show there is still structural inequality where sellers who lack formal linguistics or aesthetic literacy are often excluded from professional credibility. Thus, class differences are sustained through discourse where certain ways of speaking, looking, and behaving are perceived as more legitimate than others.

## **5.2 Theoretical implications**

This study extends the use of Multimodal Critical Discourse Analysis (MCDA) by integrating with Systemic Functional Linguistics (SFL) into the study of live, performative digital context. This study shows that live commerce creates class identity dynamically through linguistic tone, body movement, and interaction. The findings highlight that discourse in digital entrepreneurship is both representational (showing brands' values) and performative (creating class identity). This expands current understanding of how multimodal communication operates in commercial and social spaces. This also shows that online commercial performances represent a new genre of digital identity work where class, emotion, and authority intersect through communicative design [12]. Collectively, this study shows that live, multimodal digital commerce is a critical site where discourse, power, and class identity are dynamically negotiated.

## **5.3 Toward a sustainable digital economy**

By showing how the way sellers communicate and perform during live stream sessions create accessibility, visibility, and trust from digital market, this study contributes directly to Sustainable Development Goal 8 which is Decent Work and Economic Growth and Goal 10 which is reduced Inequalities. These findings indicate that sustainability in the digital economy cannot be separated from communicative practices. Linguistic and multimodal literacy emerge as critical resources that shape participation, visibility, and symbolic value in platform-based entrepreneurship. Linguistics competitiveness and multimodality is the new shape of digital marketing to determine who will thrive in online digital economy. Having inclusive education in language and multimodal literacy can therefore help to reduce inequality and ensure that digital economic growth is not limited to those who already possess symbolic or cultural advantages. Sustainable development in this digital era should include communication equity where diverse voices and expressions are recognized as equally valuable within Indonesia's online economy.

## **6 Conclusion**

This study confirms that brand representation in Indonesian TikTok beauty live commerce is not merely determined by market category or product pricing but it is constructed through language and multimodality. High-end sellers perform a professional construction with declarative mood, moderate tone, and minimalist aesthetics while mass-market sellers perform inclusivity and friendliness through imperative mood, slang, and expressive visuals. These contrasting communicative styles show how class identity is performed from how sellers speak, move, and visually present during live sessions. Theoretically, the study contributes to linguistics scholarship by revealing interpersonal and multimodal resources operate together as symbolic capital in the construction of credibility and class distinction. Practically, the findings underline the need for inclusive linguistics and multimodal literacy in digital entrepreneurship because communication in digital pluriform becomes a new form of capital that determines visibility and trust. However, this study is limited by its focus on

the beauty industry and its emphasis on sellers' communicative performances without examining audience reception. Future research could incorporate audience responses, cross-platform comparisons, or longitudinal data to further explore how linguistic and multimodal practices shape inclusion and inequality in digital commerce.

## Acknowledgment

The authors thank to LPDP EQUITY for supporting this publication. This publication is funded by the Indonesian Endowment Fund for Education (LPDP) on behalf of the Indonesian Ministry of Higher Education, Science and Technology, and managed under the EQUITY Program (Contract No. 4312/B3/DT.03.08/2025).

## References

1. M. A. Afkar and Yusmaneli, *J. Law Econ.* **2**, 1 (2023)
2. M. A. A. Rahim, N. H. C. Rosdi, F. S. M. A. Ha, and S. W. S. M. Idris, *QUANTUM J. Soc. Sci. Humanit.* **6**, 4 (2025)
3. M. H. Mubarak, M. Madonna, and F. Reza, *J. Ilm. Ilmu Komun.* **204**, 90 (2024)
4. Z. Shen, *Sage Open* **13**, 4 (2023)
5. N. Nguyen-Thanh, T. H. Nguyen, N. N. K. Doan, M. Yusoh, T. H. Chau, and N. T. K. Anh, *Rev. Mark. Sci.* **23**, 1 (2025)
6. R. D. Anggara, S. F. Wibowo, and N. F. Fidhyallah, *J. Ilm. Ekon. Manaj. Bisnis dan Akunt.* **2**, 4 (2025)
7. M. A. Halliday and C. M. I. Matthiessen, *Halliday's Introduction to Functional Grammar*, 4th ed (Routledge, London, 2014)
8. N. Fairclough, *Discourse and social change* (Polity, Cambridge, 1992)
9. G. Kress and T. Van Leeuwen, *Multimodal discourse: The modes and media of contemporary communication* (Arnold, London, 2001)
10. P. Bourdieu, *Language and Symbolic Power* Edited and Introduced by (Basil Blackwell Ltd., Cambridge, 1991)
11. D. Suparno, I. Fitriana, N. Nadra, and F. Gunawan, *Cogent Arts Humanit.* **10**, 1 (2023)
12. H. Huang, J. Blommaert, and E. Van Praet, OH MY GOD! BUY IT!' a Multimodal Discourse Analysis of the Discursive Strategies Used by Chinese Ecommerce Live-Streamer Austin Li, In *HCI International 2020–Late Breaking Papers: Interaction, Knowledge and Social Media: 22nd HCI International Conference, HCII 2020, Copenhagen, Denmark, July 19–24 (2020)* **12427**