

Article

Not peer-reviewed version

Digital Gastrodiplomacy: A Multimodal Semiotic Analysis of How YouTube Food Travel Vlogs Construct Destination Image in Uzbekistan

[Iroda Mukhammadieva](#)*

Posted Date: 18 March 2026

doi: 10.20944/preprints202603.1412.v1

Keywords: gastrodiplomacy; food tourism; multimodal discourse analysis; destination image; YouTube food vlogs; parasocial interaction; social semiotics; Uzbekistan; cultural representation; nation branding



Preprints.org is a free multidisciplinary platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This open access article is published under a [Creative Commons CC BY 4.0 license](#), which permit the free download, distribution, and reuse, provided that the author and preprint are cited in any reuse.

Disclaimer/Publisher's Note: The statements, opinions, and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions, or products referred to in the content.

Article

Digital Gastrodiplomacy: A Multimodal Semiotic Analysis of How YouTube Food Travel Vlogs Construct Destination Image in Uzbekistan

Iroda Mukhammadieva

Department of Hospitality and Tourism Management, International Graduate School, Namseoul University, Cheonan 31020, Republic of Korea; iroda@nsu.ac.kr or irodastrong87@gmail.com

Abstract

This study investigates how YouTube food travel vloggers semiotically construct destination images and function as informal gastrodiplomacy agents, using Uzbekistan as a case study of emerging tourism markets. Although digital content creators are increasingly recognised as shaping tourism flows, systematic understanding of the multimodal semiotic mechanisms through which food travel vlogs construct destination meanings remains limited. Using multimodal discourse analysis, this study examines six YouTube food travel videos on Uzbekistan (over 28 million combined views) from two prominent creators, Mark Wiens and Best Ever Food Review Show. The analysis integrates Kress and van Leeuwen's visual grammar, Halliday's systemic functional linguistics, van Leeuwen's sound semiotics, and Norris's multimodal interaction analysis to code a 60-segment corpus. Independent samples t-tests reveal 25 statistically significant differences between the two creators, identifying two distinct semiotic pro-files. Mark Wiens primarily follows a parasocial intimacy model marked by direct gaze (89.2%), frequent second-person address (78.4%), and comparatively minimal editing. In contrast, Best Ever Food Review Show adopts a cinematic documentary model characterised by first-person narration (56.5%), constructed visuals (60.9%), and gastronomic heritage narratives (34.8%). Despite these divergences, shared conventions—centred food composition, upbeat music, positive evaluation, and sharing gestures—indicate a stable semiotic grammar of food travel vlogging. The findings provide evidence that digital content creators may function as informal culinary ambassadors through gastrodiplomacy mechanisms, constructing destination awareness and cultural meaning for international audiences. The study contributes to theory on multi-modal destination image construction and offers implications for how emerging tourism destinations can leverage multi-creator strategies to build culturally grounded destination brands.

Keywords: gastrodiplomacy; food tourism; multimodal discourse analysis; destination image; YouTube food vlogs; parasocial interaction; social semiotics; Uzbekistan; cultural representation; nation branding

1. Introduction

The tourism sector has been deeply influenced by digital media, which now plays a central role in shaping how destinations are explored, evaluated, and ultimately experienced by potential travellers [1,2]. Among today's travel-related digital media, YouTube food travel vlogs represent a distinct and influential form of destination mediation. These videos, particularly those centred on food tourism, have emerged as influential mediators of cultural experience and destination perception. Travel vlogs can shape tourists' perceptions of a destination's authenticity by presenting personal narratives that resonate with viewers [3,4]. As a form of user-generated content, these videos provide accessible and seemingly unfiltered glimpses into local life, shaping destination image in ways that conventional marketing campaigns rarely achieve [5,6].

Despite this demonstrated influence, relatively limited research has systematically examined how food travel vloggers construct authenticity and desirability through multimodal semiotic strategies. Existing studies have approached travel vlogs through parasocial interaction frameworks [7], cultural and power-dynamic analysis [8], destination image assessment [9] or multimodal city image construction [10] but few have analysed how visual, verbal, aural, and gestural resources interact simultaneously to produce meaning in audiovisual travel content. Comparative research investigating how different creators represent the same destination through distinct semiotic choices remains particularly scarce, leaving important questions about creator-specific representation strategies unexplored. Moreover, tourism scholarship increasingly recognises the need to critically examine how digital content mediates cultural representation, yet the specific semiotic mechanisms remain understudied.

The concept of gastrodiploacy—the strategic use of national cuisine to shape a country’s international image and enhance foreign perceptions [11]— provides a pro-ductive lens for examining these dynamics. Food can be understood as a communicative resource through which cultural identity and national narratives are articulated to international audiences. South Korea’s Hansik globalisation campaign[12] and Peru’s Cocina Peruana para el Mundo initiative, where the government built an entire nation brand around its culinary heritage [13]. Nations increasingly recognise food culture as a distinctive asset for nation branding [14]. Cabral et al.’s [15] systematic review of 86 studies at the intersection of food and diplomacy confirms that gastrodiploacy has matured into an established area of public diplomacy scholarship. However, this body of work has focused largely on state-led or institutional initiatives and has given comparatively limited attention to independent digital creators who reach large international audiences through platforms such as YouTube. This gap matters because highly followed food travel vloggers can function as informal gastrodiploacy actors, shaping destination images through multimodal semiotic choices that warrant systematic investigation.

Uzbekistan, a Central Asian nation with a culinary heritage shaped by Silk Road trading traditions, offers an ideal case for investigating these dynamics. Since liberalising its visa regime in 2018, the country has experienced sustained tourism growth, welcoming 6.6 million foreign visitors in 2023 [16], and has actively positioned its cuisine as a distinctive national asset—most prominently through plov (Uzbek pilaf), inscribed on UNESCO’s Intangible Cultural Heritage list in 2016[17]. In parallel, Uzbekistan has facilitated visits by prominent YouTube food travel creators as part of its broader destination promotion strategy. Mark Wiens (10.8 million subscribers) and Best Ever Food Review Show (11.2 million subscribers) have produced content generating combined viewership exceeding 28 million views. The analytical focus on two creators rather than one enables a comparative dimension rarely present in existing multimodal tourism studies, where most scholarship examines single creators and treats findings as representative of the genre without testing whether different creators construct systematically different destination images. For Western audiences with limited prior familiarity with Central Asian cuisine, such food travel vlogs may play a disproportionate role in shaping initial destination perceptions, making Uzbekistan an analytically rich case for examining how digital content constructs destination images [18,19].

This study addresses these gaps by examining how YouTube food travel vloggers construct destination images through multimodal semiotic choices and how, in doing so, they may function as informal agents of digital gastrodiploacy. With a focus on Uzbekistan’s culinary landscape, the study analyses six YouTube videos and codes 60 segments using a systematic multimodal framework across four semiotic dimensions. This approach enables detailed assessment of how authenticity and desirability are produced in vlog narratives and how such representations relate to nation branding and destination image formation.

The study makes three contributions. First, it provides a systematic multimodal semiotic analysis of food travel vlogs in the Central Asian context, which remains under-represented in tourism media research. Second, using comparative statistical tests (independent-samples t-tests), it identifies two distinct creator-specific semiotic profiles that construct different versions of the same

destination. Third, it demonstrates that shared genre conventions, together with creator-specific stylistic variation, function as informal gastrodiploacy mechanisms, with implications for how emerging destinations manage their digital representation.

2. Literature Review

2.1. *Tourism Semiotics and Destination Image*

Tourism is often described as a semiotic practice in which destinations become meaningful through signs that are produced, circulated, and interpreted by travellers and audiences [20]. Destination image, defined as the mental representations and associations individuals hold about a place, significantly influencing travel intentions and behaviours [21]. Within this broad field, social semiotics, grounded in Halliday's systemic functional linguistics [22] and extended by Kress and van Leeuwen [23], provides an analytical apparatus for examining how semiotic resources operate within specific social and cultural contexts. Unlike purely linguistic approaches, social semiotic and multimodal perspectives treat communication as inherently multi-modal, where image, language, sound, and gesture each possess meaning potential shaped by cultural conventions and social practice [24,25].

Recent studies have expanded multimodal perspectives on tourism discourse in several ways. Chen et al. [10] examined transnational travel vlogs across 20 port cities, demonstrating that different camera techniques and editing strategies produce different thematic patterns and viewer relationships, even when representing similar destination content. Du and Cheong [26] developed an integrated framework for examining how sustainable tourism imaginaries are co-constructed through the interplay of visual, verbal, and interactive semiotic resources on TikTok. Wang and Feng [27] applied Kress and van Leeuwen's visual grammar to promotional videos from Xi'an, showing how visual compositions strategically position destinations for different audience segments. Moving from production-oriented analyses to audience reception, Xu et al. [28] demonstrated that audience involvement in travel vlogs operates through three distinct dimensions—narrative involvement, media personae identification, and parasocial interaction—each producing different effects on travel intention. Mirzamurodova Kizi [9], analysing 25 YouTube travel vloggers' representations of Seoul, found that vlogs provide richer destination image insights than written accounts because of the affective engagement generated through facial expressions, vocal tone, and visible excitement. These studies collectively confirm that multimodal analysis provides essential tools for understanding digital tourism representation, yet food-focused travel vlogs—with their distinctive combination of culinary spectacle, embodied tasting, and parasocial address—have not received equivalent systematic attention. Urry and Larsen's [2] concept of the tourist gaze remains helpful for interpreting how travel vlogs organise what viewers see and how they are invited to see it. The tourist gaze involves particular ways of seeing shaped by cultural frameworks, expectations, and power relations—it is not neutral observation but a socially organised visual practice that determines what is deemed worthy of viewing. YouTube food vlogs extend this gaze into the digital realm, where algorithmic amplification determines which gazes circulate widely and which remain marginal. The vlogger functions simultaneously as a gazing subject (experiencing the destination) and as a producer of gazes for the audience (constructing what is worth seeing), creating a layered mediation that complicates straightforward theories of destination image formation.

2.2. *Gastrodiploacy and Nation Branding*

Rockower conceptualizes gastrodiploacy as a form of public diplomacy that uses food to influence foreign publics [11]. The concept gained greater scholarly attention following high-profile initiatives such as Thailand's "Global Thai" programme, which promoted Thai restaurants abroad as instruments of cultural diplomacy and national image building [11,14]. South Korea's long-term efforts to promote Korean cuisine, often symbolised through kimchi, offer a further example of sustained culinary diplomacy, while Peru's positioning of ceviche as a flagship national dish

illustrates how food-based narratives can support nation branding and strengthen a country's visibility in global tourism [13]. Cabral et al. [15], show that gastrodiploamacy has shifted from a marginal policy concept to an established area of public diplomacy research, underpinned most frequently by frameworks of nation branding, public diplomacy, cultural diplomacy, and soft power.

For Uzbekistan, gastrodiploamacy is particularly relevant because the country's culinary heritage offers a highly visible and culturally resonant basis for international image-building. The country's culinary heritage, rooted in Silk Road trading traditions and centred on dishes such as plov, samsa, and lagman, provides a rich repository of cultural content [17,29]. Forman [30] suggests that culinary narratives should be understood as more than tourism products, as they can also communicate wider national values and cultural meanings. Building on this view, Parasecoli [31] examines how food identity politics operate within globalisation, arguing that countries often mobilise cuisine to balance claims of local authenticity with the need to make food understandable and appealing to international audiences. This tension between preserving cultural specificity and enhancing global legibility is also evident in the Uzbekistan vlogs analysed in this study.

Uzbekistan's approach to gastrodiploamacy has evolved since the country's tourism reforms beginning in 2016, when visa liberalisation opened the country to international visitors. The government's strategy of inviting prominent food and travel content creators represents an informal variant of gastrodiploamacy—one that operates through commercial facilitation rather than direct editorial control. This approach differs from the highly structured programmes of Thailand and South Korea, where government agencies developed explicit culinary branding guidelines. In the Uzbekistan case, the gastrodiploamacy function is emergent rather than planned: content creators exercise creative autonomy within the framework of government-facilitated access, producing representations that serve nation branding purposes without being formally coordinated as diplomatic instruments [11,12,14].

2.3. YouTube Travel Vlogs and Parasocial Interaction

YouTube is widely used for travel inspiration and trip planning, and travel-related videos attract substantial global viewership. Travel vlogs function as hybrid digital texts that blend amateur and professional production, shaping how destinations are imagined and evaluated through user-generated storytelling and social influence processes [32,33]. Within this ecosystem, travel vlogs occupy a hybrid position. They are neither entirely amateur productions nor fully institutionalized media outputs; rather, they combine elements of personal diary, cultural reportage, and entertainment into a distinctive communicative form. The platform's structure, including its recommendation algorithm, subscription model, and comment-enabled interactivity, shapes which creators and destinations gain sustained visibility, while others remain marginal[33]. Understanding why certain vlogs prove so effective at shaping destination perceptions requires attention to the interpersonal dynamics they create between vlogger and viewer.

Parasocial interaction theory offers a particularly robust framework for analyzing these dynamics. Horton and Wohl [34] originally described parasocial interaction as the illusion of face-to-face intimacy that audiences develop with media performers—a one-sided sense of knowing and being known by someone who is, in reality, unaware of the viewer's existence. In the context of YouTube travel vlogs, this illusion is not incidental but is actively constructed through specific semiotic choices. Atad and Cohen [35] in a controlled experiment with 255 participants, demonstrated that direct address—looking into the camera and speaking to “you”—produces significantly stronger parasocial experiences than indirect address, and that this intensified experience directly enhances the viewer's perception of the speaker's credibility. This finding has direct relevance for the present study, where Mark Wiens employs direct address in 78.4% of segments compared to 21.7% for BEFRS.

Several other recent studies deepen the picture of how parasocial dynamics operate in travel vlog contexts. Dewantara et al.[7] identified multiple attributes that make travel vlogs attractive to viewers, including what they term “destination attractiveness”—the quality of the destination as it is represented through the vlogger's choices about what to show, how to frame it, and what to say about

it. This suggests that destination image is not separable from the parasocial relationship through which it is communicated. Roy and Attri [36] found that parasocial bonding with travel vloggers significantly predicts both tourist engagement and electronic word-of-mouth dissemination.

Li et al. [37] developing a storytelling scale specifically for travel vlogs, established that verbal narrative techniques including direct audience address enhance travel intention through the perceived personal connection they create between vlogger and viewer. Xu et al. [28] demonstrate that audience involvement in travel vlogs is multidimensional rather than unitary, operating through three distinct pathways: narrative involvement (immersion in the unfolding story), media personae identification (affective alignment with the vlogger as a persona), and parasocial interaction (the perceived experience of direct address). Each pathway exerts differential effects on travel intention. This framework provides an analytical lens for interpreting the present study's comparative findings. Mark Wiens's semiotic profile appears oriented toward maximizing parasocial interaction through sustained direct gaze and second-person address, whereas BEFRS privileges narrative immersion and cultural contextualization through first-person narration and documentary-style editing. This contrast raises a central question: do distinct engagement pathways systematically construct different destination images? The present study addresses this question through comparative multimodal analysis.

2.4. Multimodal Semiotic Analysis in Tourism Video Research

The analytical toolkit for studying tourism video has developed unevenly across semiotic modes. Visual analysis is the most mature. Kress and van Leeuwen's [23] grammar of visual design, originally developed for still images, has proven widely adaptable to the analysis of audiovisual and digital media contexts. Their framework distinguishes between "demand" images -where a depicted participant gazes directly at the viewer the scene from a more detached perspective. It also differentiates between intimate close-up framing and more distant long shots. These distinctions map closely onto the visual strategies frequently employed in travel vlogs, where framing and gaze are used to construct varying degrees of intimacy between the vlogger, the viewer, and the destination. In the verbal mode, Halliday and Matthiessen's systemic functional linguistics [22,38] analytical tools for examining not only what vloggers say about food and place, but how meaning is constructed through language choices. Linguistic features such as evaluative expressions, hedging or assertive statements, forms of direct address (e.g., "you"), and first-person narrative positioning shape how viewers interpret the authenticity and desirability of a destination. These linguistic strategies function as semiotic resources that position the viewer in a particular relationship with both the vlogger and the place being represented. The aural and gestural modes have received less systematic attention in tourism research [25,39], despite being central to how food vlogs actually work. Van Leeuwen [40] developed what remains the most thorough semiotic account of sound, distinguishing between music's affective functions, ambient sound's capacity to authenticate a depicted environment, and the role of sound effects in constructing sensory presence. In a food vlog, the amplified crunch of bread being torn or the sizzle of meat hitting a grill is not incidental—it is deliberately designed to draw the viewer into a bodily experience they can only imagine. Norris [41] provided the equivalent framework for gesture and embodied action, treating facial expressions, body orientation, and physical interaction with objects as communicative modes in their own right rather than as mere accompaniments to speech.

Recent research has begun applying multimodal approaches specifically to food and lifestyle vlog content. Lacsina [42] analysing food review vlogs, found that creators employ a recognisable repertoire of communicative strategies including enthusiastic evaluations, rhetorical questions, and direct invitations for viewers to "try this," alongside visual techniques such as extreme close-ups of food. Together, these multimodal elements construct both credibility and a sense of shared experience. Similarly, Torjesen [43] studying YouTube lifestyle influencers, identifies the phenomenon of "professional amateurism," whereby creators deliberately maintain an apparently informal or unpolished style. Features such as handheld camera movement, spontaneous laughter,

or visible uncertainty in pronouncing local dish names operate as semiotic resources that signal authenticity to audiences who may perceive highly polished content as less trustworthy.

Despite these advances, few studies have integrated visual, verbal, aural, and gestural modes within a single systematic framework when analysing tourism video content. The multimodal tourism studies published to date have concentrated heavily on East Asian destination representations in both travel vlog and promotional video contexts [10,27]. Central Asia remains largely absent from this body of literature. Uzbekistan's increasing visibility on international YouTube food channels, combined with its distinctive culinary traditions, provides a valuable case for examining how multimodal storytelling shapes destination image when the cultural distance between vlogger and destination is substantial.

By integrating visual, verbal, aural, and gestural analysis within a unified analytical framework, the present study responds to recent calls for more comprehensive multimodal approaches in tourism media research and contributes empirical insight into how food-focused travel vlogs mediate cultural representation.

3. Materials and Methods

3.1. Research Design and Corpus

This study applies multimodal discourse analysis [23,38,40,41] to YouTube food travel vlogs that feature Uzbekistan. Using purposive sampling [44], we selected six videos from two high-profile creators: Mark Wiens (10.8M subscribers; three videos) and Best Ever Food Review Show (BEFRS; 11.2M subscribers; three videos). Videos were included if they (1) focused primarily on Uzbekistan, (2) treated food and culinary experiences as the central theme, (3) came from channels with more than five million subscribers, (4) were at least 10 minutes long to allow detailed multimodal analysis, and (5) were available in full, unedited form at the time of analysis.

The Mark Wiens corpus includes of three videos: 'Street Food in Uzbekistan—1,500 KG. of RICE PLOV (Pilau) + Market Tour in Tashkent!' (31:39), 'Central Asian Food—TEARDROP SOMSA (SAMOSA) and HUGE UZBEK DINNER in Tashkent, Uzbekistan!' (25:21), and 'Ultra-Tender TANDOORI LAMB!! | Food Tour + Attractions in Bukhara—Silk Road Uzbekistan!' (21:05), collectively totalling approximately 78 minutes and generating over 15 million views. The BEFRS corpus comprises 'Asia's Biggest Frying Pan! Over 3,000 POUNDS of Rice and Meat Cooked Each Day!' (14:44), 'WAGYU LAMB!!! Uzbekistan's UNKNOWN Nomad Mountain Meat' (16:28), and 'Death by Meat! Street Food in Tashkent, Uzbekistan!' (14:38), totalling approximately 45 minutes with over 13 million views. Together, the two corpora offer substantial reach while reflecting clearly different production styles.

Both creators operate in the top tier of YouTube food travel content. Mark Wiens is known for an enthusiastic, highly personal on-camera style, whereas BEFRS (hosted by Sonny Side) relies more heavily on narrative structuring, professional cinematography, and post-production editing. Notably, both channels disclosed facilitation by Uzbekistan's Ministry of Tourism in their Uzbekistan content—an important contextual factor for interpreting how authenticity is constructed. Across the six videos, total viewership exceeds 28.2 million views, providing a strong empirical basis for examining how food travel vlogs shape destination images for large international audiences.

3.2. Analytical Framework

The analytical framework structures the coding process around four semiotic modes, each grounded in established theoretical resources. The visual mode draws on Kress and van Leeuwen's [23] grammar of visual design, codes for contact (direct gaze/demand vs. indirect gaze/offer), social distance (intimate close-up vs. public long shot), involvement (frontal vs. oblique angle), composition (centred vs. polarised), and visual naturalness (handheld vs. highly edited). The verbal mode applies systemic functional linguistics Halliday's [38], coding for evaluative language (positive evaluation, superlative expressions), modality (high certainty vs. hedged), pronominal address (direct 'you',

first-person 'I', inclusive 'we'), and topical focus (food description, local people references, place-specific naming). The aural mode follows van Leeuwen's [40] sound semiotics framework, coding for music (upbeat, traditional, none), sound effects (amplified eating sounds, environmental ambient), and tonal quality. The gestural mode uses Norris's [41] multimodal interaction analysis, coding for facial expressions (surprise, satisfaction, smile), body movements (pointing, leaning), and food-related gestures (sharing, savoring). A fifth analytical dimension—thematic analysis—operates across all modes to identify overarching narrative patterns including spectacle, authenticity, heritage, personal discovery, and Orientalist framing.

3.3. Segmentation and Coding Procedure

Each video was divided into analytically discrete segments following established procedures in multimodal video analysis [39,45]. A segment boundary was marked only when two criteria occurred simultaneously: a visible scene transition (such as a change in location, camera setup, or editing cut) and a clear topical shift in the verbal narration.

This procedure resulted in 60 segments across the six-video corpus (Mark Wiens: $n = 37$; BEFRS: $n = 23$), with individual segments typically ranging between two and four minutes. The uneven distribution reflects substantive differences in narrative structuring. Mark Wiens produces longer videos characterized by frequent scene transitions, whereas BEFRS tends to sustain extended sequences focused on a single cooking process or cultural encounter.

The 60 segments were coded using a binary presence/absence protocol across 28 codes distributed over five analytical dimensions. The visual mode comprised seven codes: contact type (demand/direct gaze vs. offer/indirect gaze), social distance (intimate close-up vs. public long shot), camera angle (frontal vs. oblique), composition (centred vs. polarised), visual naturalness (handheld vs. stabilised), camera movement style, and editing intensity.

The verbal mode also included seven codes. Pronominal address was disaggregated into three distinct variables (direct "you," first-person "I," and inclusive "we"), alongside evaluative language, modality level, topical focus, and place-specific naming.

The aural mode comprised four codes: background music type, amplified eating sounds, ambient environmental sound, and added sound effects.

The gestural mode included five codes covering facial expressions, body movement, pointing or indicating behaviours, food-related gestures (e.g., tearing, presenting, savoring), and sharing actions.

In addition, five cross-cutting thematic codes—spectacle and grandeur, cultural authenticity, gastronomic heritage, personal discovery, and Orientalist framing—operated across all modes to capture overarching narrative patterns.

A binary coding protocol inevitably simplifies the density of multimodal meaning within any single segment. For example, a moment in which Mark Wiens locks eyes with the camera while tearing into lamb samsa may simultaneously involve visual demand, positive verbal evaluation, amplified eating sound, and a satisfaction expression. The coding framework records the presence of each semiotic resource but does not attempt to model their dynamic interaction within that moment. This trade-off is intentional: while nuance at the micro-level is reduced, the binary structure enables systematic comparison between creators, which is central to the study's comparative design. The codebook was developed iteratively rather than imposed rigidly from the outset. An initial coding scheme was derived from the four theoretical frameworks [23,38,40,41].

The final codebook specifies operational definitions and boundary conditions for each code. "Direct eye contact," for example, required the vlogger's gaze to be directed at the camera lens for at least three consecutive seconds within a segment. "Amplified eating sounds" required audible enhancement of chewing, crunching, or cooking sounds above ambient levels.

Because the primary coding was conducted by a single researcher, reliability was strengthened through iterative refinement, systematic boundary testing, and repeated review of ambiguous cases rather than through independent dual coding. Two full rounds of pilot coding were completed, each followed by clarification of code definitions and resolution of uncertainties. Although this approach

does not yield a formal intercoder reliability coefficient, it aligns with qualitative multimodal research traditions [45], where interpretive consistency is established through transparent methodological documentation and reflexive rigor rather than statistical agreement alone.

3.4. Statistical Analysis

To assess whether the two creators employ systematically different semiotic strategies, the binary-coded data were divided into two groups by creator, and independent samples t-tests (Welch's, with unequal variances assumed) were conducted on each code. Welch's t-test was selected because the group sizes are unequal (37 vs. 23), and the procedure is robust when equal variances cannot be assumed [46]. This approach enables identification of statistically significant differences in representational strategies between the two creators while accounting for the binary nature of the coded data.

4. Results

4.1. Aggregate Frequency Distribution Across Semiotic Modes

Analysis of the 60-segment corpus revealed systematic multimodal patterns across the six videos. In the visual mode, direct eye contact with the camera appeared in 70.0% of segments, while centred composition (63.3%) and intimate close-up framing (61.7%) were also highly frequent. In addition, scale emphasis through salient sizes such as large cooking vessels or expansive food displays—appeared in 61.7% of segments. These visual strategies frequently co-occurred to produce a dual orientation toward interpersonal intimacy and visual spectacle. Close-up shots reduced spatial distance and created a sense of interpersonal proximity, whereas wide shots of massive plov cauldrons and bustling bazaars constructed scenes of abundance and scale.

In the verbal mode, direct audience address using second person 'you' appeared in 56.7% of segments, making it the dominant interpersonal strategy. Food-related discussion dominated topical content (51.7%), with vloggers consistently naming dishes, describing flavours, and commenting on preparation methods. High-certainty modality (40.0%) and positive evaluative language (36.7%) together constructed Uzbek cuisine through categorical, emphatic endorsement rather than hedged or tentative assessment.

In the aural mode, upbeat background music (71.7%) functioned as a pervasive affective design element. Amplified eating sounds appeared in 46.7% of segments—a high frequency reflecting deliberate sonic design. The amplification of crunching, sizzling, and slurping transforms eating from a visual event into a multisensory experience, inviting viewers to imaginatively participate in the tasting. Ambient environmental sounds (market chatter, kitchen clatter, street noise) appeared in 30.0% of segments, contributing what van Leeuwen [40] describes as 'provenance'—the sonic trace of a real, specific location that authenticates the depicted environment.

In gestural mode, sharing gestures (66.7%) and surprise expressions (45.0%) were most prominent. Physical offers of food toward the camera, inclusive hand movements, and demonstrations of communal eating construct the vlog as a participatory encounter rather than a detached observation. Pointing and indicating gestures appeared in 35.0% of segments, typically accompanying references to specific dishes or cultural features. These deictic gestures direct the viewer's attention and position the vlogger as a knowledgeable guide.

Thematic analysis revealed spectacle and grandeur (65.0%) as the dominant narrative frame, followed by memorable food moments (53.3%), sensory richness (50.0%), personal discovery (41.7%), and cultural authenticity (36.7%).

Cross-modal coordination analysis identified three primary clusters: an Intimacy Cluster combining direct gaze, close-up framing, and direct verbal address; an Authenticity Cluster combining naturalistic visuals, ambient sound, and local people references; and a Spectacle Cluster combining edited visuals, upbeat music, and superlative language.

The identification of these clusters is analytically significant because it demonstrates that semiotic resources do not operate in isolation within food travel vlogs. A segment coded for direct gaze was significantly more likely to also feature close-up framing and second-person verbal address, suggesting that creators deploy coordinated multimodal ensembles rather than making independent choices within each mode. This coordination produces coherent communicative effects: the Intimacy Cluster creates a sense of personal connection with the viewer, the Authenticity Cluster provides evidence of genuine cultural engagement, and the Spectacle Cluster generates excitement and visual memorability. Most segments in the corpus combined elements from multiple clusters, reflecting the hybrid nature of the food travel vlog as a genre that must simultaneously entertain, inform, and build parasocial relationships.

4.2. Comparative Creator Analysis

Of the 107 codes tested across all four semiotic dimensions, 25 yielded statistically significant differences at $p < .05$. Table 1 presents the most analytically relevant significant differences, organised by semiotic mode.

Table 1. Statistically Significant Differences Between Content Creators.

Semiotic Code	MW %	BEFRS %	Diff.	t	p
Visual Mode					
Direct eye contact	89.2	39.1	+50.1	4.308	<0.001 ***
Offer/indirect gaze	8.1	60.9	-52.8	-4.646	<0.001 ***
Edited/constructed visuals	16.2	60.9	-44.7	-3.695	<0.001 ***
Cooking action	32.4	73.9	-41.5	-3.404	0.001 **
Walking action	64.9	34.8	+30.1	2.332	0.024 *
Handheld camera	40.5	17.4	+23.1	2.013	0.049 *
Verbal Mode					
Direct address ('you')	78.4	21.7	+56.6	5.078	<0.001 ***
First-person 'I'	2.7	56.5	-53.8	-4.933	<0.001 ***
Inclusive 'we'	13.5	39.1	-25.6	-2.159	0.038 *
Superlative language	35.1	13.0	+22.1	2.061	0.044 *
Local people references	24.3	52.2	-27.8	-2.171	0.036 *
Gestural Mode					
Satisfaction expression	32.4	4.3	+28.1	3.144	0.003 **
Pointing/indicating	51.4	26.1	+25.3	2.016	0.049 *
Thematic Mode					
Gastronomic heritage	8.1	34.8	-26.7	-2.397	0.023 *
Orientalism resistance	13.5	0.0	+13.5	2.372	0.023 *

(Welch's t-test). * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. MW = Mark Wiens ($n = 37$); BEFRS = Best Ever Food Review Show ($n = 23$).

4.3. Two Distinct Semiotic Profiles

The pattern of significant differences is not random but reveals two coherent, internally consistent representational strategies. Mark Wiens's profile is defined by direct interpersonal engagement: he maintains eye contact with the camera in 89.2% of segments (vs. 39.1% for BEFRS; $p < 0.001$) and addresses the audience using 'you' in 78.4% (vs. 21.7%; $p < 0.001$). His gestural repertoire reinforces this orientation: satisfaction expressions (32.4% vs. 4.3%; $p = 0.003$) and pointing gestures (51.4% vs. 26.1%; $p = 0.049$) function as embodied communication directed primarily toward the viewer rather than toward co-present participants. He favours superlative evaluative language (35.1% vs. 13.0%; $p = 0.044$) and employs handheld, minimally edited cinematography (40.5% vs. 17.4%; $p = 0.049$; editing: 16.2% vs. 60.9%; $p < 0.001$). Together, these features construct what may be

termed a parasocial intimacy model—an aesthetic of spontaneous, seemingly unmediated personal experience shared directly with the viewer.

BEFRS deploys a different semiotic logic. First-person narration ('I'; 56.5% vs. 2.7%; $p < 0.001$) and inclusive ('we'; 39.1% vs. 13.5%; $p = 0.038$) position the experience as the vlogger's journey that the viewer observes rather than participates in. The visual strategy relies on indirect gaze (60.9% vs. 8.1%; $p < 0.001$) and heavily edited sequences (60.9% vs. 16.2%; $p < 0.001$), producing a more cinematic quality. BEFRS foregrounds cooking processes (73.9% vs. 32.4%; $p = 0.001$) and references local people more frequently (52.2% vs. 24.3%; $p = 0.036$), positioning Uzbek culinary practitioners as the central subjects rather than the vlogger himself. Heritage narratives appear significantly more often (34.8% vs. 8.1%; $p = 0.023$). This configuration constitutes a cinematic documentary model that privileges cultural documentation over direct parasocial engagement.

4.4. Shared Genre Conventions

Not all semiotic resources differed significantly between creators. Several high-frequency features showed no statistically significant differences, including centred composition (MW 59.5% vs. BEFRS 69.6%; $p = 0.433$), intimate close-up framing (56.8% vs. 65.2%; $p = 0.521$), upbeat music (67.6% vs. 78.3%; $p = 0.367$), positive evaluation (35.1% vs. 39.1%; $p = 0.762$), sharing gestures (62.2% vs. 73.9%; $p = 0.347$), and surprise expressions (43.2% vs. 47.8%; $p = 0.735$). These features appear to constitute genre conventions—a shared semiotic core of the food travel vlog that operates regardless of individual creator style. Placing food centrally in the frame, filming it in close-up, accompanying it with upbeat music, and expressing positive surprise and enthusiasm appear to be requirements of the genre rather than individual choices. Table 2 summarises the key non-significant codes alongside their frequencies, documenting the stable semiotic core shared by both creators.

Table 2. Non-Significant Shared Semiotic Features Between Creators.

Semiotic Code	MarkW %	BEFRS %	p value
Centred composition	59.5	69.6	.433
Intimate close-up	56.8	65.2	.521
Eating action	56.8	65.2	.521
Upbeat music	67.6	78.3	.367
Positive evaluation	35.1	39.1	.762
Sharing gestures	62.2	73.9	.347
Surprise expressions	43.2	47.8	.735

Note. All $p > .05$ (non-significant). These features represent shared genre conventions.

5. Discussion

5.1. Parasocial Engagement vs. Documentary Observation as Paths to Destination Image

The comparative findings show that the same destination Uzbekistan can be made to feel like two quite different places depending on the creator's semiotic choices. Mark Wiens's parasocial intimacy model brings the viewer as a companion sharing the experience in real time. The combination of sustained eye contact, frequent second-person address, and visible emotional reactions work together to create what Horton and Wohl [34] described as the illusion of face-to-face closeness with a media performer. Atad and Cohen's [35] experimental results help explain why this matters: direct address strengthens parasocial experience and increases perceived credibility. In this light, Wiens's consistent reliance on direct gaze and "you" address (89.2% and 78.4% of segments, respectively) is not accidental, it functions as a deliberate strategy for intensifying viewer connection.

This strategy also shapes how "destination attractiveness" is built. Dewantara et al. [7] argue that travel vlog appeal depends partly on how the destination is mediated through the vlogger's performance and presentation. In Wiens's case, Uzbekistan becomes attractive through enthusiasm: viewers encounter the cuisine through his excitement, strong endorsements, and embodied

satisfaction. That kind of relationship may matter for behavioural outcomes. Roy and Attri [36] show that parasocial bonding with vloggers predicts tourist engagement and willingness to share recommendations, suggesting that Wiens's approach may be especially effective for generating word-of-mouth interest in destinations that feel unfamiliar to many global audiences.

BEFRS, however, works through a different logic. What emerges is closer to a cinematic documentary model. The viewer is positioned less as a direct conversational partner and more as an observer of a structured journey. First-person narration, indirect gaze, and heavy editing create distance from the host's immediate feelings and redirect attention toward the cooking process, local practitioners, and cultural detail. Cheng et al. [47] similarly note that travel vlogs can activate different engagement pathways depending on whether they lean toward personal immersion or informational framing. Xu et al. [28] further show that involvement is not a single experience but unfolds through distinct dimensions—parasocial interaction, narrative involvement, and media personae identification—each influencing travel intention in different ways. In this sense, BEFRS appears to privilege narrative involvement and cultural identification over direct parasocial intensity.

This difference becomes especially clear when both creators film environments, such as bustling bazaars. Wiens positions himself beside vendors, maintains eye contact with the camera, narrates in real time using second-person address (“you can see the colour”), and displays visible tasting reactions. BEFRS, by contrast, frames the vendor's skilled movements, narrates in first person or voiceover (“I watched as the baker shaped the samsa”), and compresses preparation through edited sequences. The same physical setting is thus transformed into different textual experiences through distinct multimodal choices.

From a tourism development perspective, this distinction also has implications for sustainable destination promotion. The parasocial model positions the destination primarily as a backdrop for the vlogger's personal experience—Uzbekistan becomes attractive because Mark Wiens is excited about it. The documentary model positions the destination's cultural practices and practitioners as the central subjects. The sustainability implications diverge accordingly: the parasocial approach may generate immediate travel interest and word-of-mouth amplification, but risk reducing the destination to a consumption experience centred on the vlogger rather than the place. The documentary approach, by foregrounding local knowledge and culinary process, may attract visitors with deeper cultural motivations—the kind of culturally engaged tourism that destination managers increasingly recognise as more sustainable in the long term.

5.2. Authenticity Construction Through Different Semiotic Pathways

The two semiotic profiles construct authenticity through distinct multimodal pathways. In Mark Wiens's case, authenticity is primarily performative. His spontaneous reactions, handheld camera work, and relatively minimal editing create the impression of immediacy and unfiltered experience. Torjesen [43] describes this aesthetic as “professional amateurism” a strategic display of informality that signals realness rather than production sophistication. Mark Wiens's significantly lower editing rate (16.2% vs. 60.9%) and higher use of handheld camera work (40.5% vs. 17.4%) are consistent with this strategy. Sandel and Wang [50] similarly demonstrate that YouTubers construct a multimodal discourse of “realness” through choices related to camera positioning, setting, and personal expressiveness. In this model, authenticity emerges from the visibility of lived, embodied experience.

BEFRS constructs authenticity differently, grounding it more explicitly in cultural depth than in personal immediacy. The emphasis on cooking processes (73.9%), local people (52.2%), and gastronomic heritage (34.8%) signals depth of engagement with Uzbek culinary culture. Rauf and Pasha [8] identify two authenticity pathways in gastronomic tourism vlogs—personal embodied experience and cultural-historical documentation. The comparative findings here closely reflect that distinction: Mark Wiens's profile aligns with embodied authentication, while BEFRS corresponds more closely to documentary-style cultural authentication. Cohen and Cohen's [51] distinction between “hot” and “cool” authenticity further clarifies this contrast. “Hot” authenticity is grounded in immediate, affective, seemingly unmediated experience; “cool” authenticity rests on

contextualisation, expertise, and cultural framing. Wiens constructs authenticity through bodily performance and emotional intensity, whereas BEFRS builds a cooler, more documentary form of authenticity centred on cultural process and heritage.

5.3. *The Sponsorship Paradox*

Both channels disclosed facilitation by Uzbekistan's Ministry of Tourism, creating what this analysis frames as a sponsorship (or transparency) paradox. On the surface, government support—such as flights, accommodation, guides, and curated access—could raise doubts about independence. Yet in these vlogs, disclosure often works in the opposite direction. By naming the relationship openly, creators signal honesty and invite viewers to read their evaluations as still genuine: the logic is that if the sponsorship is not hidden, then the food judgments are unlikely to be covertly controlled. Steils et al. [52] describe as the tension between disclosure requirements and credibility maintenance in influencer marketing functions here as a transparency paradox. Both creators navigate this paradox, but they do so through different semiotic routes. Mark Wiens leans on the immediacy of visible emotional reactions, responses that appear too spontaneous to be scripted, while BEFRS leans on informational density, using extended process footage and cultural explanation that reads as too detailed to be mere promotion. In both cases, transparency becomes part of the authenticity performance, supporting trust even in the presence of institutional facilitation. This suggests that transparency in influencer-mediated tourism promotion does not necessarily weaken authenticity perceptions; rather, when integrated into multimodal storytelling practices, disclosure itself can function as a semiotic resource for reinforcing credibility and trust.

5.4. *Orientalism, Resistance, and Cultural Representation*

Said's [53] concept of Orientalism—the Western production of “the East” as an exotic, knowable Other—offers a useful critical lens for interpreting these vlog representations. At the same time, the patterns in the corpus are more mixed than a simple “Orientalist vs. non-Orientalist” divide. On one hand, there are clear continuities with Orientalist logics. Temporal othering appears most visibly in the dominant spectacle-and-grandeur theme (65.0%), which often frames Uzbekistan through scale, excess, and historical atmosphere in ways that can position it as existing outside everyday modernity. The t-test results reinforce this concern: BEFRS uses gastronomic heritage framing significantly more often than Mark Wiens (34.8% vs. 8.1%; $p = 0.023$). This suggests that the cinematic documentary model, while providing richer cultural information, may also be more vulnerable to exoticising discourse, especially when “heritage” is used to emphasise cultural difference and historical depth. On the other hand, the corpus also contains meaningful departures from classical Orientalism. The personal discovery theme (41.7%) frequently positions the vlogger as a learner rather than an authoritative cultural interpreter, which softens the asymmetry implied in Orientalist narration. Similarly, cultural authenticity framing (36.7%) often attributes knowledge and agency to local participants rather than treating them as background figures. Notably, Mark Wiens shows significantly higher “Orientalism resistance” coding (13.5% vs. 0.0%; $p = 0.023$). This pattern suggests that the parasocial intimacy model, because it foregrounds immediate personal reaction more than cultural interpretation—may sometimes avoid representational pitfalls that documentary-style framing can reproduce, even unintentionally, when describing non-Western settings for Western audiences.

5.5. *Implications for Sustainable Digital Tourism Marketing in Emerging Destinations*

The finding that different creators produce statistically distinct destination images has practical implications for tourism stakeholders in Uzbekistan and other emerging markets. Across the six videos, combined viewership exceeded 28.2 million consistent with prior research demonstrating the persuasive influence of social media influencers on destination image and travel intention [54]. The Ministry of Tourism's investment in facilitating vlogger visits likely constituted a fraction of the cost

of equivalent advertising campaigns, yet generated content with enduring reach through the platform's algorithmic recommendation system. This positions YouTube food vlogs as a cost-effective instrument for sustainable destination branding.

The multi-creator approach, inviting more than one creator is not simply a scaling strategy; it also diversifies destination image. The two creators in this study offer complementary representational strengths. Mark Wiens's parasocial model generates immediacy and personal connection, while BEFRS's documentary model supports cultural understanding and heritage appreciation. At the same time, the shared genre conventions identified in Section 4.4 ensure continuity: regardless of style, both creators repeatedly communicate a stable core message—Uzbekistan as an exciting and abundant food destination. Against this shared baseline, creator-specific differences add texture and range, resulting in a richer composite image than any single creator is likely to produce alone. Dewantara et al. [7] argue that “destination attractiveness” itself emerges through representation and parasocial dynamics; viewed this way, different semiotic strategies do not compete so much as foreground different facets of the same place, strengthening overall brand resilience.

For Uzbekistan, the gastrodiploacy dimension makes this especially relevant. Communal plov preparation and shared eating practices, prominent in both creators' content, communicate values of hospitality, generosity, and social cohesion [17,55]. When Mark Wiens films himself sharing a massive plov dish in a Tashkent, using intimate framing and direct address, or when BEFRS documents the ritual of osh preparation with cinematic editing and heritage narration, these semiotic choices help translate culturally specific practices into destination meanings that are legible to international audiences. The two strategies work differently. The parasocial mode makes Uzbek food feel approachable and personally accessible; the documentary mode frames it as culturally significant and historically grounded. Together, they position Uzbekistan not only as a place to eat well but as a destination with cultural depth—a framing that can support sustainable tourism aims by attracting visitors who value cultural experience rather than purely surface-level consumption [48,49].

However, this study also identifies risks. The dominant spectacle-and-grandeur theme (65.0% of segments) suggests that both creators, to varying degrees, frame Uzbekistan through scale and excess in ways that could encourage spectacle-driven tourism rather than culturally engaged visitation. The documentary model's heavier reliance on heritage framing (34.8% vs. 8.1%) may inadvertently exoticise the destination, as discussed in Section 5.4. For destination managers, this finding suggests that facilitating influencer visits is not sufficient on its own—the sustainability outcomes depend on which representational strategies creators deploy, and these are shaped partly by platform incentives rather than destination priorities

A final implication concerns the role of platform algorithms in shaping what kinds of semiotic strategies become visible in the first place. YouTube tends to reward content that generates engagement signals such as watch time, likes, comments, and shares, and many genre conventions identified here (centred food shots, close-ups, upbeat music, positive evaluation) may persist partly because they reliably produce those signals. Creator-specific differences can also be read as different responses to platform incentives. A parasocial style may encourage engagement and interaction (e.g., more comments and relational engagement), while a documentary style may sustain attention through narrative structure and process footage (supporting longer viewing sessions). Understanding food travel vlogs, therefore, requires treating them not only as cultural texts but also as products shaped by platform economics—a key condition for analysing how digital gastrodiploacy operates in practice. These findings highlight how influencer-mediated food travel vlogs function not only as promotional media but also as semiotic arenas in which destination meanings are negotiated between creators, platforms, and audiences.

6. Conclusions

This study examined how two prominent YouTube food travel creators construct Uzbekistan as a tourism destination through multimodal semiotic strategies. By analysing six high-visibility food

travel vlogs featuring Uzbekistan, the study showed that destination image in digital tourism media is not produced through content alone, but through the coordinated interaction of visual, verbal, aural, and gestural resources.

The identification of 25 statistically significant differences between creators demonstrates that food travel vlogs should not be treated as homogeneous promotional texts. Mark Wiens's parasocial intimacy model and Best Ever Food Review Show's cinematic documentary model represent distinct semiotic approaches with different viewer positioning strategies, authenticity construction mechanisms, and representational implications, including different degrees of susceptibility to Orientalist framing. At the same time, shared genre conventions—centred food composition, upbeat music, positive evaluation, sharing gestures—indicate the presence of a stable semiotic core that defines the food travel vlog as a recognisable media form. These findings also suggest that semiotic strategies in food travel vlogs are shaped not only by creator preference but by the engagement incentives built into platform architecture.

For Uzbekistan and similar emerging tourism markets, the findings suggest that a multi-creator gastrodiplomacy strategy is more effective than reliance on a single promotional voice. Different creator styles construct different dimensions of destination image, and their combination produces a multifaceted representation that can appeal to varied audience segments through both affective and cognitive pathways.

Several limitations should be acknowledged. The corpus of six videos from two creators, while enabling detailed multimodal analysis, limits generalisability to other creators and destinations. The study analyses production without examining reception; future research should investigate whether the parasocial and documentary models produce systematically different effects on viewers' destination perceptions through controlled reception studies. Additionally, the binary coding protocol, while facilitating statistical comparison, necessarily simplifies the density of multimodal interaction within individual segments. Comparative analyses across multiple emerging destinations would further strengthen claims regarding the broader applicability of the identified semiotic profiles. Despite these limitations, the study demonstrates the analytical value of integrating multimodal discourse analysis with comparative statistical methods in examining how digital media content constructs destination images in contemporary tourism research. First, facilitating multiple creators with different content styles generates a richer and more resilient destination image than relying on a single creator, however popular. Second, the shared genre conventions identified here such as centred food composition, close-up framing, upbeat music, positive evaluation, and sharing gestures, appear to function as necessary conditions for effective food travel vlog content, and facilitators should ensure creators have access to settings and experiences that support these semiotic strategies while respecting creators' editorial independence. Third, the documentary model's tendency toward heritage narratives may serve long-term cultural tourism development goals, while the parasocial model's emphasis on personal excitement may be more effective for generating immediate travel interest. A sustainable gastrodiplomacy strategy would intentionally balance both approaches, using digital marketing not merely to maximise visitor numbers but to shape the kind of tourism that supports cultural preservation and equitable community benefit.

Author Contributions: Conceptualization, I.M., methodology, I.M.; formal analysis, I.M.; writing—original draft preparation, I.M.; writing—review and editing, I.M.

Funding: Funding for this paper was provided by Namseoul University

Institutional Review Board Statement: Not applicable. The data corpus comprises publicly available YouTube videos and their public comment sections.

Informed Consent Statement: Not applicable.

Data Availability Statement: The YouTube videos analysed in this study are publicly available. The coded dataset can be made available on reasonable request to the corresponding author.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Jansson, A. Rethinking post-tourism in the age of social media. *Ann. Tour. Res.* 2018, 69, 101–110.
2. Urry, J.; Larsen, J. *The Tourist Gaze 3.0*; Sage: London, UK, 2011.
3. Cheng, Y.; Wei, W.; Zhang, L. Seeing destinations through vlogs: Implications for leveraging customer engagement behavior to increase travel intention. *Int. J. Contemp. Hosp. Manag.* 2020, 32, 3227–3248.
4. Nguyen, P.M.B.; Pham, T.T.N.; Truong, D. Travel vloggers as influencers: How source credibility and inspiration affect travel intention. *J. Vacat. Mark.* 2023, 29, 543–55 .
5. Cox, C.; Burgess, S.; Sellitto, C.; Buultjens, J. The role of user-generated content in tourists' travel planning behavior. *J. Hosp. Mark. Manag.* 2009, 18, 743–76.
6. Marine-Roig, E. Measuring online destination image, satisfaction, and loyalty: Evidence from Barcelona districts. *Tour. Hosp.* 2021, 2, 62–78.
7. Dewantara, M.H.; Jin, X.; Gardiner, S. What makes a travel vlog attractive? Parasocial interactions between travel vloggers and viewers. *Tour. Recreat. Res.* 2025, 50, 107–123.
8. Rauf, A.A.; Pasha, F.M. Vlogging gastronomic tourism: Understanding Global North–South dynamics in YouTube videos and their audiences' feedback. *Tour. Geogr.* 2024, 26, 407–431.
9. Mirzamurodova Kizi, I. Tourist destination image projected by YouTube travel videos: The case of Seoul. *Int. J. Tour. Res.* 2025, Advance online publication. <https://doi.org/10.1002/jtr.70049>
10. Chen, S.; Zang, Y.; Yang, P. City images in transnational travel vlogs from a multimodal perspective. *Online Media Glob. Commun.* 2024, 3, 82–107.
11. Rockower, P.S. Recipes for gastrodiploacy. *Place Brand. Public Dipl.* 2012, 8, 235–246.
12. Pham, M.J. Food as communication: A case study of South Korea's gastrodiploacy. *J. Int. Commun.* 2013, 19, 1–16.
13. Wilson, R. Cocina Peruana para el mundo: Gastrodiploacy, the culinary nation brand, and the context of national cuisine in Peru. *Exch. J. Transdiscip. Writ. Res. Prax.* 2013, 2, 13–20.
14. Chapple-Sokol, S. Culinary diplomacy: Breaking bread to win hearts and minds. *Hague J. Dipl.* 2013, 8, 161–183.
15. Cabral, Ó.; Lavrador, L.; Orduna, P.; Moreira, R. Gastronomy as a diplomatic tool: A systematic literature review. *Int. J. Gastron. Food Sci.* 2024, 38, 101072. <https://doi.org/10.1016/j.ijgfs.2024.101072>
16. National Statistics Committee of the Republic of Uzbekistan. *Tourism and Recreation Statistics 2024*. Available online: <https://stat.uz>
17. UNESCO. *Palov Culture and Tradition*. 2016. Available online: <https://ich.unesco.org/en/RL/palov-culture-and-tradition-01166> (accessed on 24 February 2026).
18. Uzbekistan Tourism. *Gastro Tourism*. 2025. Available online: <https://uzbekistan.travel/en/v/food-tourism/>
19. (accessed on 24 February 2026).
20. Campelo, A.; Aitken, R.; Thyne, M.; Gnoth, J. Sense of place: The importance for destination branding. *J. Travel Res.* 2014, 53, 154–166.
21. MacCannell, D. *The Tourist: A New Theory of the Leisure Class*; Schocken Books: New York, NY, USA, 1976.
22. Baloglu, S.; McCleary, K. A model of destination image formation. *Ann. Tour. Res.* 1999, 26, 868–897.
23. Halliday, M.A.K. *Language as Social Semiotic: The Social Interpretation of Language and Meaning*; Edward Arnold: London, UK, 1978.
24. Kress, G.; van Leeuwen, T. *Reading Images: The Grammar of Visual Design*, 3rd ed.; Routledge: London, UK, 2021.
25. van Leeuwen, T. *Introducing Social Semiotics*; Routledge: London, UK, 2005.
26. Jewitt, C. *The Routledge Handbook of Multimodal Analysis*; Routledge: London, UK, 2009.
27. Du, S.; Cheong, C.Y.M. Beyond the scenic view: A multimodal discourse analysis of sustainable tourism imaginaries on TikTok in Anhui, China. *Humanit. Soc. Sci. Commun.* 2025, 12, 690. <https://doi.org/10.1057/s41599-025-05015-3>

28. Wang, Y.; Feng, D. History, modernity, and city branding in China: A multimodal critical discourse analysis of Xi'an's promotional videos on social media. *Soc. Semiotics*. 2021, 33, 402–425. <https://doi.org/10.1080/10350330.2020.1870405>
29. Xu, H.; Zeng, B.; Huang, Z.; Li, Z. How travel vlog audience members become tourists: Exploring audience involvement and travel intention. *Comput. Hum. Behav.* 2024, 151, 108007. <https://doi.org/10.1016/j.chb.2023.108007>
30. Dadabaev, T. "Silk Road" as foreign policy discourse: The construction of Chinese, Japanese and Korean engagement strategies in Central Asia. *J. Eurasian Stud.* 2018, 9, 30–41.
31. Forman, J. *Gastrodiplomacy*. In *Oxford Research Encyclopedia of Food Studies*; Oxford University Press: Oxford, UK, 2024. <https://doi.org/10.1093/acrefore/9780197762530.013.18>
32. Parasecoli, F. *Gastronativism: Food, Identity Politics, and Globalization*; Columbia University Press: New York, NY, USA, 2022.
33. Tussyadiah, I.P.; Fesenmaier, D.R. Mediating tourist experiences: Access to places via shared videos. *Ann. Tour. Res.* 2009, 36, 24–40.
34. Gretzel, U. Influencer marketing in travel and tourism. In *Advances in Social Media for Travel, Tourism and Hospitality*; Sigala, M., Gretzel, U., Eds.; Routledge: London, UK, 2018; pp. 147–156.
35. Horton, D.; Wohl, R. Mass communication and para-social interaction. *Psychiatry* 1956, 19, 215–229.
36. Atad, E.; Cohen, J. Look me in the eyes: How direct address affects viewers' experience of parasocial interaction and credibility? *Journalism* 2024, 25, 465–484.
37. Roy, S.; Attri, R. I bond, I engage, I visit: Investigating the effects of vloggers' tourist engagement. *J. Travel Res.* 2024, 63, 1872–1891
38. Li, M.W.; Kim, Y.R.; Liu, A.; Scarles, C.; Chen, J.L. Storytelling in travel vlogs: Scale development, validation, and application. *J. Travel Res.* 2025, 65, 3, 786–811. <https://doi.org/10.1177/00472875241312178>
39. Halliday, M.A.K.; Matthiessen, C. *Halliday's Introduction to Functional Grammar*, 4th ed.; Routledge: London, UK, 2013.
40. Bateman, J.; Tseng, C. *Multimodal Analysis of Video*, 2nd ed.; Cambridge University Press: Cambridge, UK, 2023.
41. van Leeuwen, T. *Speech, Music, Sound*; Macmillan: London, UK, 1999.
42. Norris, S. *Analyzing Multimodal Interaction: A Methodological Framework*; Routledge: London, UK, 2004.
43. Lacsina, N.E. Unveiling the art of food vlogging: A multimodal discourse analysis of food review vlogs. *Int. J. Linguist. Transl. Stud.* 2023, 4, 11–25.
44. Torjesen, A. Stylistic expressions of YouTube lifestyle influencers: Authenticity and professional amateurism in Norwegian YouTube content. *Soc. Semiotics* 2024, ahead of print, 1–19, <https://doi.org/10.1080/10350330.2024.2443019>
45. Patton, M.Q. *Qualitative Research & Evaluation Methods*, 4th ed.; Sage: Thousand Oaks, CA, USA, 2015.
46. Baldry, A.; Thibault, P.J. *Multimodal Transcription and Text Analysis: A Multimedia Toolkit and Coursebook*; Equinox Publishing: London, UK, 2006.
47. Neuendorf, K.A. *The Content Analysis Guidebook*, 2nd ed.; Sage: Thousand Oaks, CA, USA, 2017.
48. Cheng, W.; Tian, R.; Chiu, D.K.W. Travel vlogs influencing tourist decisions: Information preferences and gender differences. *Aslib J. Inf. Manag.* 2024, 76, 86–103.
49. Hall, C.M.; Gössling, S.; Scott, D. *The Routledge Handbook of Tourism and Sustainability*; Routledge: London, UK, 2017
50. Higgins-Desbiolles, F. Sustainable tourism: Sustaining tourism or something more? *Tour. Manag. Perspect.* 2018, 25, 157–160. <https://doi.org/10.1016/j.tmp.2017.11.017>
51. Sandel, T.L.; Wang, Y. Online content creators' self-disclosure: A multimodal discourse analysis of YouTubers. *Discourse Context Media* 2022, 50, 100653.
52. Cohen, E.; Cohen, S.A. Authentication: Hot and cool. *Ann. Tour. Res.* 2012, 39, 1295–1314.
53. Steils, N.; Martin, A.; Toti, J.-F. The transparency paradox in influencer marketing. *J. Bus. Res.* 2022, 152, 493–501.
54. Said, E.W. *Orientalism*; Pantheon Books: New York, NY, USA, 1978.

55. Bulumulla, D.S.K.; Epa, U.I.; Gamage, T.C. Effect of social media influencer involvement on tourists' travel intentions: Mediating role of traveler authenticity and destination image. *South Asian J. Tour. Hosp.* 2023, 3, 88–106. <https://doi.org/10.4038/sajth.v3i2.74>
56. Bessière, J. Local development and heritage: Traditional food and cuisine as tourist attractions in rural areas. *Sociol. Rural.* 1998, 38, 21–34.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.