

The Effect of Narration Language Type in Tourism Short Videos on Viewers' Travel Intention

Yuhan Wu*

College of Economics and Management, Nanjing Forestry University, Nanjing 210037, China

**Corresponding author: Yuhan Wu.*

Abstract

Tourism short videos have become an important marketing tool for destination promotion in the digital era. As an integral part of the sound symbol of short videos, the narration language affects the audience's perception of the image and atmosphere of the destination. At present, the existing research on the difference in the role of dialect and Mandarin and its impact mechanism are still insufficient. This article takes social presence as an intermediary variable, and compares the two types of travel short video language types: Nanjing dialect commentary and Mandarin commentary. 241 valid samples were obtained through offline experiments and questionnaires, and the structural equation model was used to test the research hypothesis. The results show that the type of narration language has a significant impact on social presence. Compared with Mandarin commentary, dialect commentary is easier to enhance the audience's perception of the narrator's presence, interactive atmosphere and proximity to the destination; Social presence has a significant positive impact on travel intention. The stronger social presence, the easier it is for the audience to have a tendency to further understand and visit the destination on the spot; The direct impact of the type of explanation language on travel intention has not reached a significant level. However, the indirect impact of social presence is significant, which shows that social presence plays an intermediary role between the type of language commentary and travel intention. Localized language can affect the audience's travel intention by improving the sense of realism and closeness. This article can provide reference for cultural tourism short video explanation language selection and destination marketing content optimization.

Keywords

travel short videos, social presence, travel intention, dialect, Mandarin Chinese

1. Introduction

Short videos have become an important channel for audience information acquisition and travel decision-making [1]. Language carries the functions of emotional communication and social interaction. In short travel videos, it not only conveys information about scenic spots, but also shapes the audience's perception of the destination atmosphere and local culture. Tourism short videos employ multiple symbolic elements to construct destination images, among which the difference in language interpretation in the sound symbol is one of them [2]. In recent years, many travel bloggers have used dialects to create personal characteristics and improve the recognition of communication. The distinct regional accent and colloquial style make the audience feel the local atmosphere of the destination from the hearing. Compared with dialects, standard Mandarin has

stronger universality and understanding convenience in travel short videos, which can reduce the cost of information reception for audiences in different regions and clearly convey destination information among a wider audience.

At present, dialect and Mandarin, two types of narration languages, exist in parallel in travel short videos, but the difference in the impact of the audience's travel intention and the path of action still lack direct testing. The existing tourism short video commentary research has been carried out from multiple angles. Hu Mingming and other scholars emphasized that commentary is an important part of information transmission, and different intonation and commentary themes will affect the audience's communication identification [3]. The study of heritage tourism commentary distinguishes between knowledge-based commentary and emotional commentary, and the difference in narrative style also causes differences in tourists' cognition and emotional response [4]. In the research related to red tourism, language expression, video communication and narrative can reproduce tourism scenes and stimulate emotional resonance [5]. However, under specific circumstances, these studies have not revealed the specific path of dialect and Mandarin commentary to affect travel intention, and whether the impact is transmitted through a certain mechanism, which lacks direct verification.

Therefore, this article takes Nanjing dialect as the main reference, compares the two types of tourism short video language types: dialect commentary and Mandarin commentary, and conducts empirical analysis through questionnaire survey and structural equation model. The study introduces a social presence as an intermediary path to examine whether the type of explanation language will first affect the audience's perception of presence, interactive atmosphere and proximity to the destination, and then affect travel intention. The research expands the theory of social presence to the linguistic dimension of travel short videos, and tests whether dialects as a local clue can affect travel decision-making by enhancing the on-site experience, thus providing an empirical basis for the cross-sectional research of language psychology and tourism marketing.

2. Literature Review and Research Hypotheses

2.1 Narration Language and Communication Effectiveness

Language is the carrier of conveying information. Depending on the type of expression, the audience's cognitive response and emotional response will also be different. Cognitive language is biased towards information description, while emotional language is easier to mobilize the audience's imagination and emotional investment. Different language expressions affect the effect of destination communication [6]. Language style will also change people's understanding of the image of tourist destinations, which in turn affects travel intention [7]. Language not only introduces the landscape itself, but also participates in the construction of its historical value, cultural significance and experience atmosphere [8].

In the travel short video, the narration language not only explains the basic information of the destination, but also carries the local culture, lifestyle and identity information here. As a language form formed by the long-term accumulation of regional culture, dialects are often associated with local memory, daily life and cultural identity. Tourists' perception of regional dialects includes local characteristics, cultural attractiveness, functionality and understanding. Dialects are an important factor in identifying local culture [9]. This kind of recognition not only stays at the perception level, but also extends to specific consumption scenarios. Often dialect symbols can affect tourists' perceived value and consumption attitude [10].

Therefore, when the audience hears the dialect, it is easy to associate the sound with the specific place. Social presence includes emotional expression, sense of interaction and group cohesion [11]. There are also studies that understand social presence as co-presence, psychological involvement and behavioral interaction [12]. Dialect accents may make the person behind the voice more specific, and the everyday tone and title will also make short videos closer to daily communication [13]. Different accents can also trigger the audience's evaluation of reputation, attractiveness and closeness [14]. Based on this, this article believes that dialect interpretation may enhance social presence through the levels of presence, co-presence, interaction and emotional relationships. In contrast, Mandarin commentary emphasizes the clarity of information transmission and the convenience of audience understanding. Both types of languages have their own advantages in communication.

2.2 The Influence of Narration Language on Social Presence

Language is an important social clue that affects the audience's perception. According to the study of social presence, language expression will affect the psychological distance of the audience from the object of communication, which in turn affects the perception of social presence [15]. Dialects often have stronger geographical attributes and emotional colors. The phonetics, vocabulary and expression habits in the dialect can make the audience feel the atmosphere of local life, and the narrator is more like a familiar local person making a live introduction [16]. Dialect services can allow tourists to obtain local uniqueness and cultural learning clues, and enhance the recognition of local experiences [17]. Local accents can affect the positive reaction and willingness of tourists to travel in AI travel explanations, indicating that localized voice expression will change tourists' perception of the explanation content [18].

The perception of destination dialect in the host-guest interaction will also affect the emotional and behavioral feedback of tourists [19]. Regional dialects can be understood by tourists as part of local characteristics and cultural attractions [20]. These contents all point to the same problem: dialect is not just a sound form, and it also participates in the construction of a sense of locality, closeness and realism.

The dialect commentary short video can be regarded as an expression innovation under the background of the homogenization of tourism content. The sound expression with unique regional characteristics may make the audience feel that the narrator is more specific, the scene is more realistic, and the local atmosphere is stronger. Therefore, in the short video situation, dialect commentary may be easier than Mandarin to enhance the audience's perception of presence, interactive atmosphere and proximity to the destination.

Based on this, this paper puts forward the hypothesis:

H1: In travel short videos, compared with Mandarin, the dialect narration language type can trigger a social presence.

2.3 Social Presence and Travel Intention

Social presence refers to individuals perceiving the psychological effects and subjective feelings of the presence of others when communicating with the help of communication media. The theory was first formally proposed by Short, Williams and Christie in 1976 to explain the impact of telecommunications media on interpersonal interaction [21]. The short video communication scene expands the explanatory space of social presence. When receiving tourism content, viewers will form a judgment of the existence of others and the authenticity of the scene through the narrator's voice, interactive behavior and scene presentation.

According to the theory of planned behavior, travel intention refers to the influence of individual behavioral intentions under the influence of behavioral attitude, subjective norms and conscious behavior control (Ajzen, 1991) [22]. The smooth experience, perceived pleasure and perceived usefulness in short videos will affect users' travel intention [23], and the technical characteristics and social characteristics of the platform will also affect the attitude and willingness of tourists to travel [24].

The relationship between on-site experience and travel intention has been discussed in scenarios such as travel short videos, live broadcasts and virtual reality. The spatial sense of presence and social presence of travel short videos can predict the audience's destination support and social presence, among which the influence of social presence is stronger [25]. Social presence in the live broadcast will affect travel intention with the help of the sense of interaction and trust [26]. At the same time, virtual reality-related research has also confirmed that on-site experience can change the cognition, emotion and willingness of tourists to visit [27].

The stronger social presence, the easier it is for the audience to feel the presence of the narrator, the more accessible the destination is, and travel intention is more likely to be stimulated.

Based on this, this paper puts forward the hypothesis:

H2: Social presence positively affects travel intention.

2.4 The Direct Effect of Narration Language on Travel Intention

Sound is an important sensory factor in watching short videos, which will affect the audience's freshness, credibility and emotional investment in the content. Mandarin commentary has the advantages of being clear, standardized and easy to understand, which is more suitable for large-scale information dissemination; Dialect commentary may break the audience's usual expectations for short travel videos, bringing a sense of intimacy and fun with accents, tone and local vocabulary. The language attractiveness in travel short videos will affect the audience's travel intention, and the interpretation language is involved in the formation of the attraction of the destination [28]. The commentary in the travel film will also change the audience's judgment of the image of the destination and affect the intention of subsequent travel [29]. The story expression and sound stimulation in tourist destination advertisements will affect the emotional arousal and memory effect of the audience, indicating that the sound content can help the destination information be remembered [30].

Dialects and Mandarin bring different feelings to the audience. Dialect dubbing is easier to bring immersion than Mandarin dubbing, and can also enhance social presence of potential tourists [31]. Scholars have confirmed that Mandarin and Cantonese will have different persuasive effects in advertising, and language choice will affect the audience's attitude and behavioral intention [32]. Regional dialects may also improve the audience's sense of closeness and satisfaction in service communication, and promote consumption intentions when the dialect is matched [33].

Therefore, dialect commentary may not only play a role through a social presence, but also directly affect the audience's cognition and emotions. After hearing the dialect, the audience will associate the sound with the specific place, local life and local people's stories. The sense of freshness, intimacy and life-oriented expression brought by dialect can make the audience more interested in the destination, and then form travel intention.

Based on this, this paper puts forward the hypothesis:

H3: In the travel short video, the dialect narration language type directly affects travel intention.

2.5 The Mediating Role of Social Presence

Social presence reflects the individual's perception of the existence of others, the state of interaction and the authenticity of the scene in the media environment. Travel short videos, travel live broadcasts and virtual tourism are all forms of digital tourism communication. In the process of watching, the audience gets not only destination information, but also the psychological experience brought by sound, picture, narrative and interaction.

Studies have found that the influence of media content on travel intention often requires the transformation of psychological experience. Interactive content such as tourism bullet comments can influence social presence through social presence, which shows that the sense of presence and interaction formed by viewers in the media environment can become an important psychological path for tourism information to act on behavioral intentions [34]. The sense of participation, interactive experience and presence in interactive videos will also affect travel intention. Digital tourism content needs to change the audience's presence experience first before it is easier to transform into travel tendency [35]. Travel short video marketing research proves that psychological variables such as authenticity perception, celebrity attachment and remote presence will participate in the formation process of travel intention, and the impact of short video content on travel intention also has a psychological transformation link [36].

The sense of presence is also involved in the psychological link in the research of virtual reality and live tourism. VR destination marketing research found that the sense of presence can affect cognitive and emotional evaluation, which in turn affects the willingness to visit [37]. In the study of live broadcast of tourism, social presence is an important psychological variable that affects travel intention in the live situation [38].

Research shows that the influence of media content on travel intention is often not completed in one step, but first changes the viewer's sense of reality, immersion, emotional connection or psychological distance, and then affects the behavioral intention. The type of explanatory language that this article focuses on may also affect the audience's perception of the presence of the narrator, the interactive atmosphere and the sense of closeness of the destination, and then affect travel intention. Therefore, this article introduces a social presence as an intermediary path.

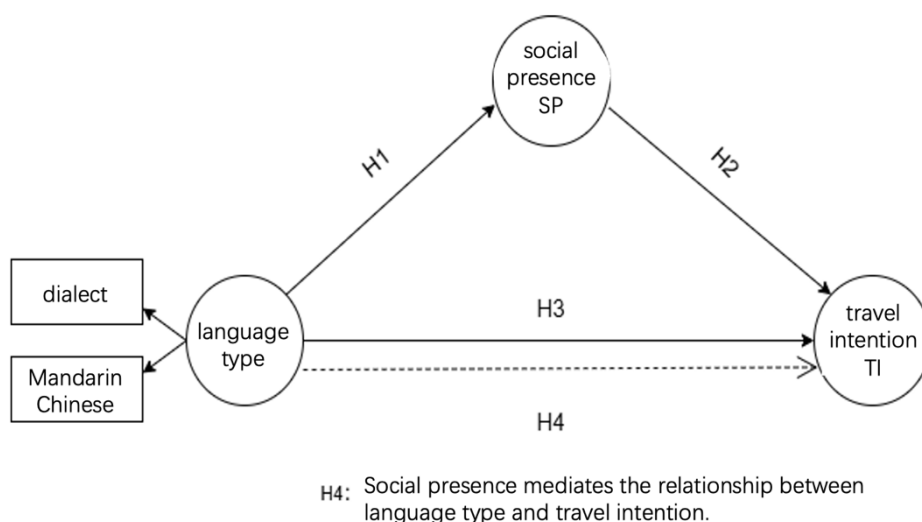
Based on this, this paper puts forward the hypothesis:

H4: Social presence plays an intermediary role in explaining the influence of language types on travel intention. Specifically, dialect commentary (compared with Mandarin commentary) enhances the audience's social presence, thereby enhancing their travel intention.

2.6 Research Model

The specific conceptual model of the article is shown in Figure 1.

Figure 1: Conceptual Model



3. Research Methodology

3.1 Experimental Design

3.1.1 Experimental Materials

This article first adopts the experimental method. The independent variable is the explanatory language type, including Mandarin and dialect. In order to ensure that the experimental materials are in line with the research theme, this article selects a short video of Nanjing flower-viewing guide video released by the blogger “Nanjing Wuyuege” on the WeChat Channels in March as the basic material. The video has received more than 55,000 likes and retweets, which is very influential. Its narration language is Nanjing dialect, and the theme is related to the seasonal tourist hotspot “flower viewing”. The content describes all kinds of scenic spots and areas where you can enjoy flowers in Nanjing, which can better present the cityscape and memory of Nanjing.

The experiment intercepted a fragment of 1 minute and 30 seconds from the original video and used it as an experimental stimulation material. The selection of this clip mainly considered that the video clip is highly related to the theme, and the front and back connection is natural, which can reduce the obstacles to understanding caused by the unfamiliarity of the content of the subject. Secondly, the length of the clip is moderate to avoid the decline in attention caused by the subject's excessive viewing time.

In order to control the interference of non-verbal factors on the experimental results, the experimental materials made a Mandarin commentary version on the basis of maintaining the consistency of the picture content, subtitle information, background music, video duration and editing rhythm. The Mandarin version was re-dubbed by radio and television company personnel with similar tones in standard Mandarin and similar intones. Except for the different explanatory language of the two materials, the overall content, intonation and speed of speech are almost consistent to reduce the impact of unrelated variables on the experimental results.

3.1.2 Experimental Procedure

The experiment is an offline experiment, and the scenic spots (Confucius Temple, Zhanyuan, Yuyuan) with a high correlation with the Nanjing urban tourism situation in the video are selected as the experimental sites. The experimental subjects are tourists or potential tourists who are randomly contacted at the scene and participate in the experiment after obtaining their consent. Before the beginning of the experiment, the researchers explained to the subjects that the theme of this study was “travel short video viewing experience and audience feedback”, and did not directly inform them that there were two different versions of dialect and Mandarin, so as to reduce the subject's subjective pandering caused by the exposure of the research purpose. Subjects were randomly divided into two groups to watch the commentary of Nanjing flower viewing strategy (group 1: Nanjing dialect; group 2: Mandarin), and asked the subjects to watch it completely in a relatively quiet environment to ensure that they could clearly receive the picture and commentary information. Fill in the questionnaire immediately after watching. The content of the questionnaire includes the language selection of watching the video, social presence, travel intention, and demographic information. In order to ensure the quality of the data, the questionnaire has set the necessary screening rules, including the short answer time (less than 60 seconds), continuous answering of the same option (choose the same option for 4 consecutive questions), missing key questions, choosing the wrong language type, etc. Samples that do not meet the requirements will be eliminated in the subsequent data analysis.

3.2 Sample and Data Collection

This experiment and questionnaire survey were conducted from April 2 to April 10, 2026, and a total of 285 questionnaires were collected. After data screening, 241 valid questionnaires were obtained, and the recovery was efficient. Among them, 123 were in the dialect group and 118 in the Mandarin group. The sample size of the two groups was balanced, which met the basic requirements of subsequent data analysis.

3.3 Measurement of Variables

3.3.1 Social Presence

Based on the theoretical framework of this study, social presence and travel intention are measured. The measurement indicators of the two variables come from the existing mature scales. The topic of social presence is adapted from the maturity scale of social presence such as Gunawardena and Zittle (1997), Gefen and Straub (2004) and Biocca et al. (2001) (see Table 1), Make appropriate adjustments in combination with the short video situation in terms of interaction, sense of existence, co-presence and emotional connection. The final scale measures whether the subjects feel the presence of the narrator, the natural atmosphere of communication and the proximity of the destination scene when watching travel short videos. Among them, there are a total of 6 questions set up for social presence, which are measured by the 7-point Likert scale. 1 means “very disagree” and 7 means “very much agree”.

Table 1: Measurement Items of Variables

Dimension	Item Code	Measurement Item	Reference Source
Social Presence	SP1	I feel that the vlogger in the video is communicating with me.	Gunawardena & Zittle (1997)
Social Presence	SP2	I feel that the people appearing in the video are real.	Biocca et al. (2001)
Social Presence	SP3	I do not feel alone while watching the video.	Biocca et al. (2001)
Social Presence	SP4	I experience a sense of companionship while watching the video.	Gefen & Straub (2004)
Travel Intention	TI1	I am willing to visit this tourist destination in the future.	Ajzen (1991)
Travel Intention	TI2	I am willing to learn more about this tourist destination.	Chi & Qu (2008)
Travel Intention	TI3	I am willing to recommend this tourist destination to others.	Chi & Qu (2008)
Travel Intention	TI4	If circumstances permit, I would give priority to this tourist destination when making travel plans.	Crompton (1979)

3.3.2 Travel Intention

Travel intention is used to reflect the attitude and tendency of the subjects towards the destination after watching short videos. This dimension has a total of 4 questions to measure, including: behavioral intention, information search, recommendation intention and priority selection. All questions have been specifically adjusted according to the travel short video scene to ensure the rigor and scientificity of the overall measurement. All questions use the 7-point Likert scale, and the subjects need to be evaluated between 1 point (very disagree) and 7 points (very agree).

3.4 Data Analysis Methods

This article uses descriptive statistics to analyze the basic characteristics of the sample to understand the distribution of subjects in terms of gender, age, education, income, short video viewing frequency and travel experience. Descriptive statistics can show whether the sample structure is relatively balanced, and also helps to judge whether the sample meets the basic characteristics of the travel short video audience and potential tourists.

The structural equation model (SEM) can simultaneously analyze the measurement relationship between latent variables and dominant indicators, as well as the path relationship between latent variables. This paper conducts a confidence test and a fit test on the measurement data, and verifies the research hypothesis according to the path coefficient and significance level. The Bootstrap method is used to test the intermediary effect. Bootstrap repeats sampling 5,000 times, and calculates the 95% confidence interval to determine whether the indirect effect is significant. If the indirect effect confidence interval does not include 0, it means that social presence plays an intermediary role between the type of explanation language and travel intention.

4. Empirical Results

4.1 Descriptive Analysis

The statistical analysis results of the sample are shown in Table 2. The sample includes groups of different genders, ages, education and short video viewing frequency, as follows:

Among the subjects participating in the survey, women accounted for 56.8%; men accounted for 43.2%, and the gender structure of the sample was balanced. The age is mainly 18-30 years old, accounting for 67.6%, and the characteristics of the youth group are obvious; Academic qualifications are concentrated in undergraduate and master's degrees, accounting for 86.3%, and the overall education level is relatively high. Most of the monthly disposable income is less than 3,000 yuan, accounting for 41.5%, which is consistent with the relatively high proportion of students. The place of residence and place of origin are mainly non-Nanjing areas, and the sample source is wide. Short videos account for 69.7% of those who are interested in travel content and above account for 72.6%. 95.0% of those who have travel experience in the past year, indicating that the subjects have certain travel experience.

In general, the sample is mainly young people, bachelor's degree or above and high-frequency viewers of short videos, with a wide range of geographical sources. Most of the subjects are interested in tourism content and have travel experience, which can provide a rich data basis for subsequent analysis.

4.2 Model Fit Assessment

This study uses SmartPLS to construct a SEM structural equation model. The model matching results show (Table 3), $\chi^2/df=2.365$, the gap between the model and the sample data is not large, and the relationship structure is reasonable; GFI (0.928), TLI (0.950), CFI (0.962) are all greater than 0.9. It shows that the model can better present the relationship structure of variables in the sample data; The RMR is 0.075 and the SRMR value is 0.043, which is less than 0.08. The above data all meet the fitting standard, pass the suitability test, and have a good fit.

Table 2: Descriptive Statistics of the Sample

Basic Information	Group	Frequency	Percentage
Gender	Male	104	43.2%
	Female	137	56.8%
Age	18 years and below	4	1.7%
	18—25 years	89	36.9%
	26—30 years	74	30.7%
	31—40 years	45	18.7%
	41—50 years	22	9.1%
	51—60 years	5	2.1%
	61 years and above	2	0.8%
Educational Attainment	Junior high school and below	1	0.4%
	Senior high school or vocational school	15	6.2%
	Associate degree	16	6.6%
	Bachelor's degree	174	72.2%
	Master's degree	34	14.1%
Monthly Disposable Income	Doctoral degree	1	0.4%
	RMB 3,000 and below	100	41.5%
	RMB 3,001–6,000	53	22.0%
	RMB 6,001–10,000	54	22.4%
	RMB 10,001–20,000	29	12.0%
Place of Residence	Above RMB 20,000	5	2.1%
	Nanjing	76	31.5%
	Outside Nanjing	164	68.0%
Place of Origin	Other	1	0.4%
	Native of Nanjing	50	20.7%
	Non-native of Nanjing	191	79.3%
Frequency of Watching Short Videos	Frequently	168	69.7%
	Occasionally	50	20.7%
	Rarely	17	7.1%
	Almost never	6	2.5%
Interest in Tourism-Related Content	Not interested	12	5.0%
	Neutral	54	22.4%
	Moderately interested	133	55.2%
	Very interested	42	17.4%
Number of Trips Taken in the Past Year	0 trips	12	5.0%
	1–2 trips	125	51.9%
	3–4 trips	80	33.2%
	More than 4 trips	24	10.0%

Table 3: Model Fit Assessment Results

Index	X ²	χ^2/df	GFI	TLI	CFI	RMR	SRMR
Recommended Criterion	Not Significant	2.365	0.928	0.950	0.962	0.075	0.043

4.3 Reliability and Validity Analysis

This study uses SmartPLS to analyze the credibility of the scale. First of all, this paper uses Cronbach's α coefficient for reliability analysis. As shown in Table 4, the Cronbach's α coefficient of each latent variable of the questionnaire is greater than 0.8, and the confidence is very high; Secondly, the polymerization validity is tested with standard load, average variance extraction (AVE) and composite reliability (CR), and the AVE value is higher than 0.5, and the CR value is higher than 0.7, indicating that the scale has good convergent validity and internal consistency, and the measurement results are relatively stable.

Finally, use the HTMT value to judge the distinguishing validity. As shown in Table 5, the HTMT value between social presence and travel intention, language and social presence, and travel intention do not exceed 0.85, indicating that there is a certain degree of distinction between the variables. In general, the reliability and

validity of this study scale are relatively good, which is conducive to the subsequent analysis of the structural equation model.

Table 4: Results of Reliability and Convergent Validity Analysis

Variable	Cronbach's Alpha	CR	AVE
Social Presence (SP)	0.874	0.881	0.558
Travel Intention (TI)	0.888	0.887	0.667

Table 5: Results of Discriminant Validity Analysis (HTMT)

Variable	Social Presence (SP)	Travel Intention (TI)	Language Type
Social Presence (SP)	—	—	—
Travel Intention (TI)	0.847	—	—
Language Type	0.384	0.335	—

4.4 Results of Path Analysis

As shown in the results of the model path test (Table 6), the language type has a significant impact on social presence. The path coefficient is $\beta = -0.639$, and the language type has a more obvious impact on social presence; The T value is 5.499, the path result is relatively stable and $p < 0.001$, and the impact of language type on social presence is highly significant. This article assigns the dialect language type to 1 and the Mandarin language type to 2. It can be seen from the negative path coefficient that the dialect group has a higher social presence. Social presence emphasizes the perception of “the presence of others” in media communication, and language is an important clue that triggers this perception. Mandarin has strong universality and normativeness, and can clearly convey scenic spot information. Compared with Mandarin commentary, dialects have more distinct local sound characteristics. The accent, tone and expression habits carried can reduce the psychological distance between the audience and the destination, and it is also easier to convey the atmosphere of life and local identity. In the eyes of the audience, the commentator is no longer just an information publisher, but more like a familiar local friend sharing on the spot. The audience is easy to feel that “the introduction of the local people is very down-to-earth” and “someone takes me to see the destination”. The scenic spots in the video are no longer just displayed, but closer to a perceptible real scene. Therefore, dialect commentary can enhance the audience's perception of interpersonal interaction, realism and closeness, and H1 was supported.

Social presence has a significant positive impact on travel intention. The path coefficient is $\beta = 0.916$, and the coefficient is positive, indicating that there is a positive relationship between the two; The T value is 10.680, which is significantly higher than the standard critical value, and $p < 0.001$, reaching a highly significant level, indicating that there is a strong connection between social presence and travel intention, which is not an accidental result caused by sample fluctuations. In the short video of travel, social presence reflects the audience's subjective feelings about “the presence of others” and “real communication”. Short videos are not simply to present scenic spots. They can drive engage viewers at multiple perceptual levels. The narrator's voice, expression and narrative atmosphere will shape the viewer's psychological distance from the destination. When the audience feels that the narrator in the video is more realistic, the destination is more accessible, and even has a similar experience of “being taken to visit”, the scenic spots on the screen are transformed into a realistic and accessible tourist space. The audience's sense of closeness and trust in the destination gradually arises in the viewing, and the possibility of searching for relevant information, planning the itinerary or visiting the field also increases. A strong social presence can turn the image of a destination like Nanjing in the short video into a more attractive and experiential travel object. Different scenic spots and diverse flower types promote the desire of viewers to travel. Thus, H2 was supported.

The direct impact of language type on travel intention has not reached a significant level. The path coefficient is $\beta = -0.087$, and the coefficient value is close to 0, indicating that the difference between Nanjing dialect commentary and Mandarin commentary is not obvious in directly affecting travel intention; The T value is 0.852, which is lower than the significance judgment standard, $p = 0.395$, higher than 0.05, and the result lacks support. This shows that the language type itself is not enough to explain the result variable of travel intention. Travel intention is usually affected by many factors such as destination attraction, individual travel interests, time cost and emotional preferences. It is difficult to directly translate the difference between dialects or Mandarin into clear travel tendencies. What changes more about language is the audience's feelings about the content atmosphere, the narrator's image and the viewing experience. Its role needs to be further transmitted through other latent variables, so H3 was not supported.

The indirect path of language type that affects travel intention through social presence has reached a significant level. The path coefficient is $\beta = -0.585$, and the P value is 0.000. The direct effect has not reached a significant level, but the indirect effect is significant, and social presence thus constitutes a complete intermediary. This result shows that the type of explanation language does not directly change the audience's travel intention. What really plays a role in transmission is social presence evoked by language. The audience will not immediately form a travel plan just because they hear dialect or Mandarin. The real role of language is whether it can change the viewer's perception of the narrator, destination and communication situation. The narration language in the travel short video not only supplements the picture information, but also shapes the viewing atmosphere through tone, rhythm, emotion and accent, so that the audience has different degrees of realism and closeness to the content. In the experiment, the flower viewing attractions in the short video screen are gradually three-dimensional in the interaction and explanation, and social presence plays a role in this process: It transforms the local atmosphere, interpersonal contact and real communication brought by language into the audience's psychological proximity to the destination, thus promoting the formation of travel intention. Therefore, the impact of language type on travel intention is not directly completed, but is gradually realized through the intermediary mechanism of social presence, and H4 is established.

Table 6: Results of Path Analysis

Hypothesis	Path	Path Coefficient (β)	t-value	p-value	Conclusion
H1	Language Type $\rightarrow$ Social Presenc	-0.639	5.499	0.000	Supported
H2	Social Presence $\rightarrow$ Travel Intention	0.916	10.680	0.000	Supported
H3	Language Type $\rightarrow$ Travel Intention	-0.087	0.852	0.395	Not Supported
H4	Language Type $\rightarrow$ Social Presence $\rightarrow$ Travel Intention	-0.585	/	0.000	Supported

Note: $p < 0.05$ indicates statistical significance; $p < 0.01$ indicates a higher level of significance; and $p < 0.001$ indicates a very high level of significance.

4.5 Results of the Structural Model

4.5.1 Differences in Social Presence and Travel Intention Across Groups

In order to understand the perception differences of the audience under different types of narration languages, this article takes the language type as the group variable, and conducts an independent sample t test on the average of social presence and travel intention (see Table 7). The results show that the average social sense of presence in the dialect group is 5.50, and the standard deviation is 0.73; The average value of the Mandarin group is 4.73, and the standard deviation is 1.23. The difference reaches a significant level (significance < 0.01) and the answer focuses on the social sense of presence of the dialect group, $t = 5.890$, $p < 0.001$. Subjects who watch the Nanjing dialect commentary video are more likely to perceive the presence of the narrator, the strong interactive atmosphere and the proximity of the destination than the subjects who watch the Mandarin commentary video. The average value of the dialect group of travel intention is 5.56, and the standard deviation is 0.74; The average value of the Mandarin group is 4.86, and the standard deviation is 1.31. The difference between the two is equally significant (significance < 0.01), but the standard deviation of the Mandarin group is larger, and the difference in perception of the Mandarin explanation of different subjects is relatively more obvious. Secondly, $t = 5.121$, $p < 0.001$, it can be seen that dialect commentary can not only enhance social presence, but also relatively improve the audience's travel intention to the destination.

Table 7: Results of Independent-Samples t-Tests for Social Presence and Travel Intention Across Different Narration Language Types

Dependent Variable	Narration Language Type	Sample Size	Mean	Standard Deviation	Significance	t	p
Social Presence	Dialect	123	5.50	0.73	< 0.01	5.890	< 0.01
	Mandarin Chinese	118	4.73	1.23			
Travel Intention	Dialect	123	5.56	0.74	< 0.01	5.121	< 0.01
	Mandarin Chinese	118	4.86	1.31			

Individual differences will also have an impact on the research variables. This article takes the frequency of watching travel short videos, the degree of interest in travel videos and the number of trips in the past year as group variables, and conducts ANOVA single-factor analysis of social presence and travel intention

respectively. The results show that the frequency of watching short videos has a significant impact on the social sense of presence and travel intention, with F values of 2.748 and 3.929 respectively; The p values are 0.044 and 0.009 respectively, both of which are less than 0.05. Subjects with high frequency of watching short videos have higher scores in terms of social presence and travel intention than those in other groups with low viewing frequency, indicating that the higher the frequency of contact with short videos, the easier it is for the audience to enter the viewing situation created by travel short videos.

The influence of interest in tourism content is more obvious. There are significant differences in social presence and travel intention in different levels of interest subjects, with F values of 20.698 and 26.476 respectively; The p value is less than 0.001. Subjects who are very interested in tourism content have an average social sense of presence of 5.68, and the average travel intention has reached 5.99, both of which are higher than other groups. It shows that the higher the interest in tourism content, the more likely the audience is to form a positive perception of the destination in the video and have a stronger tendency to travel.

The number of trips in the past year also has a significant impact on the social sense of presence and travel intention, with F values of 4.837 and 6.377 respectively; The p value is 0.003 and less than 0.001 respectively. The more times the subjects travel, the higher the scores of their social presence and travel intention, especially the subjects who have traveled more than 5 times in the past year, the average value of their travel intention reached 5.73. It shows that the audience with more travel experience is more sensitive to the destination information in travel short videos, and it is also easier to turn the viewing experience into the actual travel intention.

In summary, the type of narration language, the frequency of watching short videos, the interest in travel and the number of trips will affect the audience's perception and behavioral tendencies of travel short videos. Among them, dialect commentary has obvious advantages in terms of social presence and travel intention.

5. Conclusions and Implications

5.1 Conclusions

This article compares dialect and Mandarin, introduces a social presence as an intermediary variable, and uses the SEM model to analyze the impact of travel short video narration language on the audience's travel intention. Research results show that language types have a significant impact on social presence, and dialect commentary is more likely than Mandarin commentary to trigger the audience's perception of the presence of the narrator, the interactive atmosphere and the sense of proximity of destination. Social presence has a direct positive impact on travel intention. When the audience has an emotional link with the narrator when watching a short video, as if they were in the destination space, it will promote the formation of the desire to travel.

The path test results show that the direct impact of the type of explanatory language on travel intention is not significant, but the indirect impact through social presence has reached a significant level. The type of explanation language will not directly contribute to the decision to travel, but will first affect the audience's viewing experience and then stimulate travel intention.

In terms of differences between groups, the score of the dialect group in terms of social presence and travel intention is higher than that of the Mandarin group, and localized language can enhance the sense of closeness and bringing in to short videos. The audience's own characteristics will also affect this process. Groups with high frequency of watching short videos, who are interested in tourism and have travel experience are more likely to have a positive perception of the destination in the video.

5.2 Practical Implications

This article uses social presence as an intermediary to study the impact of travel short video narration language on the audience's travel intention. The results show that dialect commentary can significantly improve the audience's social presence, and social presence will further enhance travel intention. In contrast, the direct impact of the type of explanatory language on travel intention is not significant, indicating that dialects do not simply directly promote tourists to travel, but need to play a role through the path of social presence. Based on this, this article puts forward the following suggestions:

(1) Reasonable use of dialect commentary in social media publicity. When making short videos, tourist destinations and cultural tourism accounts can appropriately add dialect explanations to local life scenes, street food and well-known scenic spots. The life tone and cultural background of the dialect can make the audience connect with the commentator, just like entering a real travel scene. However, dialect commentary should not appear in the content of hunting or funny, let alone affect the understanding of information. For key contents such as route, transportation, tickets, etc., it can be combined with short explanations with subtitles, so that the audience can not only feel the local characteristics, but also get travel information clearly.

(2) Optimize the explanation content around social presence, instead of just pursuing differences in language forms. The focus of short video creation should be on “let the audience feel that someone is taking them around”. Creators can reduce the expression of simply listing scenic spots, use first-person narration, bring in a live tone and describe stories. For example, when introducing the flower viewing route in Nanjing, you can not only explain “there are flowers here”, but also add content with a sense of companionship such as “enter from this alley”, “the best light in the afternoon” and “locals often walk from here”, so as to bring the audience closer to the destination.

(3) Differentiated dissemination according to the audience's interests and travel experience. The research results show that the higher the interest in tourism content, the higher the frequency of watching short videos, and the more trips in the past year, the stronger the audience's social presence and travel intention. Therefore, cultural tourism accounts can design different levels of content for different audiences. For groups with high interest and high tourism experience, personalized routes, local unpopular points or treasure attractions can be launched to meet the exploration and unique needs of these potential groups; For groups with low interest or less travel experience, the viewing threshold should be lowered to attract the audience's attention with intuitive pictures, subtitles and commentary.

6. Limitations and Future Research

Although the article verifies and understands the influence of the type of language on social presence, and the intermediary role of social presence in explaining the language type and travel intention, there is still room for expansion. This article mainly focuses on the psychological path of social presence. In fact, viewers may also be affected by other factors when watching travel short videos. In the future, more intermediary variables or adjustment variables will be added to explain the process of explaining the language affecting travel intention.

References

- [1] Wang, W. S., Zhou, Y. H., & Tang, C. C. (2025). The impact of short-video content characteristics on tourism destination image and tourists' behavioral intentions: Based on positive emotion theory. *Geography and Geo-Information Science*, 41(2), 135–144.
- [2] Zhou, H. L., Tan, Z. Y., Jiang, Y. J., & Wang, F. Y. (2024). Scale development and validation of symbolic perception of tourism short videos in ethnic villages. *Journal of Chinese Ecotourism*, 14(2), 286–303.
- [3] Hu, M. M., Hu, Q. Y., & Yang, L. M. (2025). The influence of narration audio on identification with tourism short-video communication. *Tourism Tribune*, 40(6), 68–79. <https://doi.org/10.19765/j.cnki.1002-5006.2025.06.010>
- [4] Ruan, W. Q., Wang, M. Y., Zhang, S. N., Li, Y. Q., & Su, X. (2024). Knowledge-based or affection-based? The influence mechanism of heritage tourism interpretation content on tourists' willingness to inherit culture. *Tourism Management*, 102, 104876.
- [5] Dai, B., & Ma, X. F. (2021). Reflections on promoting the high-quality development of red tourism. *Hunan Social Sciences*, (4), 77–85.
- [6] Byun, J., & Jang, S. S. (2015). Effective destination advertising: Matching effect between advertising language and destination type. *Tourism Management*, 50, 31–40.

- [7] Tong, Z., Yu, R., & Xiao, H. (2023). How should cities communicate? The interaction effect of city stereotypes and advertising language on travel intention. *Journal of Destination Marketing & Management*, 27, 100755.
- [8] Zhang, S. N., Huang, H., Ruan, W. Q., & Li, Y. Q. (2025). How can tourism advertising language make cultural heritage sites more popular? *Journal of Vacation Marketing*. Advance online publication. 13567667251320112.
- [9] Lu, Y., Chen, G., Huang, S. S., & Bao, J. (2019). Understanding Chinese tourists' perceptions of Cantonese as a regional dialect. *Tourism Management*, 71, 127–136.
- [10] Shi, S. X., Wen, T., & Lü, Q. S. (2025). The effects of dialect symbols on tourism souvenir consumption behavior. *Human Geography*, 40(5), 184–192. <https://doi.org/10.13959/j.issn.1003-2398.2025.05.019>
- [11] Rourke, L. (1999). Assessing social presence in asynchronous text-based computer conferencing. *Journal of Distance Education*, 14(2), 50–71.
- [12] Biocca, F., Harms, C., & Burgoon, J. K. (2003). Toward a more robust theory and measure of social presence: Review and suggested criteria. *Presence: Teleoperators and Virtual Environments*, 12(5), 456–480. <https://doi.org/10.1162/10547460322761270>
- [13] Purnell, T., Idsardi, W., & Baugh, J. (1999). Perceptual and phonetic experiments on American English dialect identification. *Journal of Language and Social Psychology*, 18(1), 10–30. <https://doi.org/10.1177/0261927X99018001002>
- [14] Coupland, N., & Bishop, H. (2007). Ideologised values for British accents. *Journal of Sociolinguistics*, 11(1), 74–93. <https://doi.org/10.1111/j.1467-9841.2007.00311.x>
- [15] Gunawardena, C. N., & Zittle, F. J. (1997). Social presence as a predictor of satisfaction within a computer-mediated conferencing environment. *American Journal of Distance Education*, 11(3), 8–26. <https://doi.org/10.1080/08923649709526970>
- [16] Su, L., Ye, C., & He, X. (2025). The memory effect of destination dialect advertising. *Annals of Tourism Research*, 113, 103973. <https://doi.org/10.1016/j.annals.2025.103973>
- [17] Xiong, X., Sun, D., Wong, I. A., & Lian, Q. L. (2026). Do tourists prefer dialectal service? The role of processing fluency, distinctiveness, and cultural learning cues. *Journal of Travel Research*, 65(5), 1539–1566.
- [18] Li, C., Qu, Z., & Yang, Y. (2025). Localizing matters: The effect of AI accent on tourist travel intention. *Journal of Destination Marketing & Management*, 36, 100992. <https://doi.org/10.1016/j.jdmm.2025.100992>
- [19] Geng, S., & Li, D. (2024). The impact of host–guest interactions among young people on cultural sustainability: A case study of dialects. *Sustainability*, 16(23), 10580. <https://doi.org/10.3390/su162310580>
- [20] Parker, E. B. (1978). *The social psychology of telecommunications*. Penguin Books.
- [21] Parker, E. B. (1978). *The social psychology of telecommunications*. Penguin Books.
- [22] Ajzen, I. (1991). The theory of planned behavior. *Organizational Behavior and Human Decision Processes*, 50(2), 179–211. [https://doi.org/10.1016/0749-5978\(91\)90020-T](https://doi.org/10.1016/0749-5978(91)90020-T)
- [23] Liu, J., Wang, Y., & Chang, L. (2023). How do short videos influence users' tourism intention? A study of key factors. *Frontiers in Psychology*, 13, 1036570. <https://doi.org/10.3389/fpsyg.2022.1036570>
- [24] Qiu, L., Li, X., & Choi, S. H. (2024). Exploring the influence of short video platforms on tourist attitudes and travel intention: A social–technical perspective. *Journal of Destination Marketing & Management*, 31, 100826. <https://doi.org/10.1016/j.jdmm.2024.100826>
- [25] Wu, W. X., Shu, B. Y., Zhang, H. M., & Zhang, Y. C. (2024). Virtual experience, real behavioral intention: The impact of tourism short-video presence experience on viewers' behavioral intentions. *Tourism Science*, 38(12), 17–31. <https://doi.org/10.16323/j.cnki.lykx.2024.12.003>

- [26] Zhang, W., Wang, Y., & Zhang, T. (2021). Can “live streaming” really drive visitors to the destination? From the aspect of social presence. *SAGE Open*, 11(1), 21582440211006691. <https://doi.org/10.1177/21582440211006691>
- [27] Ying, T., Tang, J., Ye, S., Tan, X., & Wei, W. (2022). Virtual reality in destination marketing: Telepresence, social presence, and tourists’ visit intentions. *Journal of Travel Research*, 61(8), 1738–1756. <https://doi.org/10.1177/00472875211047287>
- [28] Gan, J., Shi, S., Filieri, R., & Leung, W. K. (2023). Short video marketing and travel intentions: The interplay between visual perspective, visual content, and narration appeal. *Tourism Management*, 99, 104795. <https://doi.org/10.1016/j.tourman.2023.104795>
- [29] Liu, S., Gao, J., Zhang, G., Cheng, L., & Zhang, R. (2024). How to speak destinations more efficiently: A study on the effects of narration in tourism films on viewer perceptions and intentions. *Asia Pacific Journal of Tourism Research*, 29(12), 1509–1528.
- [30] Ma, J., Scott, N., & Wu, Y. (2024). Tourism destination advertising: Effect of storytelling and sensory stimuli on arousal and memorability. *Tourism Review*, 79(3), 671–687.
- [31] Wang, L., Xu, Y., Li, J. Y., & Wang, Y. (2026). Immersing in nostalgia and local sentiment: The positive effects of dialect voice-over in rural tourism promotional videos. *Tourism Science*, 40(2), 156–170. <https://doi.org/10.16323/j.cnki.lykx.20240624.001>
- [32] Liu, S., Wen, X., Wei, L., & Zhao, W. (2013). Advertising persuasion in China: Using Mandarin or Cantonese? *Journal of Business Research*, 66(12), 2383–2389. <https://doi.org/10.1016/j.jbusres.2013.05.019>
- [33] Mai, R., & Hoffmann, S. (2011). Four positive effects of a salesperson’s regional dialect in services selling. *Journal of Service Research*, 14(4), 460–474. <https://doi.org/10.1177/1094670511414551>
- [34] Yang, Y., Zou, Y., & Jin, D. (2025). Information processing of travel barrage reviews: Effects of social presence and opinion leaders on the relationship between barrage reviews and travel intention. *Current Issues in Tourism*, 28(11), 1746–1767.
- [35] Lin, S., Ren, Y., Ye, S., Ma, S., & Hu, X. (2025). How interactive videos shape travel intentions of Generation Z: An extension of presence theory. *Journal of Hospitality and Tourism Technology*, 1–28.
- [36] Zhu, C., Fong, L. H. N., Li, X., Buhalis, D., & Chen, H. (2024). Short video marketing in tourism: Telepresence, celebrity attachment, and travel intention. *International Journal of Tourism Research*, 26(1), e2599. <https://doi.org/10.1002/jtr.2599>

Funding

This research received no external funding.

Conflicts of Interest

The authors declare no conflict of interest.

Acknowledgment

This paper is an output of the science project.

Open Access

This chapter is licensed under the terms of the Creative Commons Attribution-NonCommercial 4.0 International License (<http://creativecommons.org/licenses/by-nc/4.0/>), which permits any noncommercial use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

