buildings

MDPI

Article

Differentiated Soundscape Experiences in Public Spaces of a Tourism-Oriented Historic District: A Large Language Model-Assisted Study of Huishan Ancient Town

Bozhu Sheng [1] and Jingying Guo [2,*]

[1] School of Humanities, Faculty of Arts and Humanities, University of Southampton, Avenue Campus, Highfield Road, Southampton SO17 1BF, UK
[2] School of Arts and Creative Industries, College of the Arts and Society, Coventry University, Priory Street, Coventry CV1 5FB, UK
* Correspondence: guoj40@coventry.ac.uk

Abstract

Soundscape assessment in tourism-oriented historic districts requires linking auditory experiences to activities in historic buildings and public spaces. We developed a prompt-guided, large language model-assisted framework using social media reviews of Huishan Ancient Town, China. Task-specific prompts constrained soundscape screening, verbatim extraction, emotion and sentiment coding, and place identification, complemented by model comparison, manual review, BERTopic modeling, an attention–sentiment matrix, and spatial clustering. Analysis of 4301 fragments identified 27 topics. The mixed-sound topic comprising conversations, broadcasts, and music accounted for 45.57% of fragments; topic shares varied across time periods. The model frequently assigned calmness or enjoyment labels to fragments about spring water, rustling leaves, and some quiet conversations. Irritation labels were common in topics about phone or video audio played aloud and repeated loud conversations. The widely reported mixed-sound topic had near-neutral mean sentiment, while the less frequent playback topics contained a concentration of fragments coded as negative. The 2353 place-linked fragments covered 89 grid cells at the 25 m scale. Clustering the cells by their review fragments' topic proportions yielded six groups. The resulting narrative topic profiles summarize reports linked to halls, ancestral hall courtyards, pavilions, and waterfront locations. Reviews linked to the same place also expressed different evaluations of reported sound events. The framework compares how reviewers describe and evaluate sounds across activities and places. These comparisons identify settings for on-site assessment.

Keywords: soundscape perception; historic buildings; public spaces; large language models; prompt engineering; sentiment analysis; Huishan Ancient Town

Check for updates

Academic Editor: Jian Kang

Received: 9 October 2026
Revised: 21 November 2026
Accepted: 23 November 2026
Published: 26 November 2026

Copyright: © 2026 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

1. Introduction

The sustainable development of historic districts requires coordination of the conservation of built heritage with contemporary uses of space [1–3]. United Nations Sustainable Development Goal 11 places heritage conservation and safe, inclusive, and accessible public spaces on the urban development agenda [4]. In historic districts, conserving buildings, streets, and open spaces requires attention to sociocultural practices and the connections between physical settings and everyday use [1,5,6]. Growing interest in heritage tourism has also made visitor experience a concern for the conservation and management of these

districts [2,7,8]. The ICOMOS International Charter for Cultural Heritage Tourism prioritizes heritage conservation and emphasizes interpretation and presentation as means of enhancing visitor experience [7]. Sustaining the cultural continuity of built heritage while responding to users' spatial experiences is therefore a central challenge for the sustainable development of historic built environments.

Sound shapes how people perceive buildings and public spaces, forming the auditory dimension of spatial experience. The International Organization for Standardization (ISO) defines soundscape as an "acoustic environment as perceived or experienced and/or understood by a person or people, in context" [9]. This emphasis on context links soundscape evaluations to spatial characteristics, activities, and users' experiences and expectations [10–12]. In historic districts, sounds associated with traditional activities and everyday life can convey local memories and cultural experiences [13,14]. Soundscape evaluations can also be associated with users' place identity and attachment [13–15]. Environmental characteristics, including street greenery and spatial enclosure, together with the functions of different places, further shape the spatial context of these evaluations [11,12]. Understanding soundscapes in historic districts therefore requires situating soundscape perception within the use of specific buildings and public spaces and examining how sound-related activities fit their functions and cultural characteristics.

Research on auditory experiences in historic spaces has examined soundscape quality and the joint perception of acoustic and lighting environments [16], as well as the local memories and place authenticity associated with culturally meaningful sounds [17]. Other studies have investigated the suitability of sound sources for commercial, cultural, and residential spaces and the design of sound combinations [18]. Common methods include acoustic measurements, soundwalks, questionnaires, interviews, and audiovisual experiments, which provide evidence on environmental conditions, perceptual evaluations, and cultural experiences [16,18,19]. These methods yield detailed evidence in specific settings, but scaling up on-site soundscape assessment can be labor-intensive and costly [20]. One audiovisual study focused on matching sound sources to static spatial settings and did not systematically incorporate temporal variation, activity density, or seasonal changes [18]. Social media texts containing temporal and place information offer a complementary source of voluntarily shared auditory experiences. They can help identify the soundscape topics that attract visitors' attention and the sentiments expressed toward them, while enabling comparisons across time periods and places. Such analyses provide supplementary evidence on the relationship between soundscapes and the use of space [12,21].

The conservation and tourism use of historic districts in China provide diverse settings for soundscape research. Figures from the Ministry of Housing and Urban-Rural Development reported in June 2026 indicate that China had more than 1300 historic and cultural districts [22]. Studies have used online reviews to identify cultural ecosystem services and associated sentiment evaluations in historic districts [23] and examined the effects of commercial gentrification on residents' everyday lives [24]. Research on Wuzhen has also traced the long-term transformation of a traditional village tourism destination from a glocalisation perspective [25]. Studies of spatial experience have linked audiovisual perception to the sense of culture in everyday heritage [26] and identified differences between residents and visitors in soundscape perception and place attachment [27]. This evidence is shaped by the data and experimental settings used: the study of cultural ecosystem services in Fuzhou relied on a single review platform [23], while the audiovisual study of everyday heritage acknowledged that simulated settings could not fully reproduce multisensory and emotional experiences in real environments [26]. At the analytical level, an integrated examination of topic composition, sentiment evaluations, and temporal variation in visitors' soundscape narratives remains limited. How topic attention and sentiment

evaluations can jointly inform soundscape management in historic districts also requires further investigation.

Building on these research needs, this study develops a large language model-assisted framework for comparing soundscape narratives, expressed evaluations, and their temporal and spatial contexts in tourism-oriented historic districts. Using Huishan Ancient Town in Wuxi, China, as the study setting, it applies the framework to a corpus of reviews from multiple social media platforms. The analysis combines topic identification, sentiment analysis, and attention–sentiment matrix to examine temporal variation in soundscape topics and their evaluations, then explores their spatial distributions and the functions of associated places. Three research questions guide the analysis: (1) Which soundscape topics are expressed in the corpus, and how does their composition vary across the years represented? (2) How do soundscape topics differ in their relative frequency and expressed sentiment within the corpus? (3) How does the topic composition of place-linked review fragments vary across locations and their associated functions?

2. Literature Review

2.1. Cultural and Emotional Value of Soundscapes in Historic Districts

Research on soundscapes in historic districts assesses cultural significance, perceptual attributes, and emotional responses as distinct dimensions. Studies of acoustical heritage consider the historical significance of sounds, sound-producing activities, and listening conditions [28,29]; a cultural capital perspective also addresses local identity, belonging, and nonmonetary value [30]. Cultural significance is understood through the relationships among sounds, spatial settings, social practices, and the people and memories associated with them [28,30]. Pleasantness describes an immediate affective response, which may vary among listeners even when a sound has shared cultural significance. Empirical studies assess perceived cultural character [31], pleasantness and perceived restoration [32], and emotional responses in religious settings [33], revealing different evaluative outcomes and their associations with audiovisual conditions and listeners' experiences. Higher ratings of cultural character can coexist with weaker feelings of relaxation [31], indicating that cultural perceptions and emotional states require separate interpretation. Spatial studies also assess different dimensions: soundwalks and mapping identify local sound sources and activities [34]; church studies examine perceived grandeur and auditory envelopment [35]; and studies of interfaces between historic streets and green spaces compare quietness, comfort, and pleasantness [36]. These measures capture different aspects of soundscape experience. Given differences in measurement priorities and study settings, clarifying the relationship between cultural perceptions and affective responses requires identifying the sounds being evaluated and specifying what each evaluation means.

2.2. Soundscape Topics and Sentiment Evaluations on Social Media

Soundscape evaluations in online texts encompass perceptual content, emotional expression, and information dissemination. Research on semantic validation shows that the actual use of soundscape descriptors does not fully align with translations selected by experts [37]; analyses of garden reviews reveal that auditory descriptions co-occur with accounts of cultural experiences and other sensory content [38]. These findings highlight the need to interpret soundscape topics in context. Online review studies further extract references to sensory stimuli, cognition, and emotions separately [39]. Other studies examine how multisensory review content relates to perceived review usefulness [40,41], distinguishing reviewers' overall experience ratings from readers' judgments of review helpfulness [41]. Analyses of themes and emotions in cultural heritage interviews similarly link auditory experiences to activities, memories, and religious experiences, providing a

basis for interpretation within cultural contexts [42]. Heritage research using social media has tracked thematic and emotional variation across years and seasons [43], while another study examined changes in posting activity around major events [44]. These studies also noted that neutral classifications may reflect factual information [43] and that a lack of expressed negative feelings does not establish their absence [44]. These approaches provide more detailed accounts of experiential content. However, positive emotions expressed in an overall review may refer simultaneously to buildings, services, and activities, making evaluations of sound difficult to isolate. Identifying the sounds, perceptual attributes, and evaluations expressed in reviews shows which experiences become visible in the corpus and how they are represented. These digital narratives are culturally mediated accounts, interpreted in relation to contributors' posting choices, language and relationships to the place, as well as in relation to platform purposes and audiences.

2.3. Soundscape Perception and Spatial Use in the Built Environment

Relationships between spatial form and soundscape evaluations vary with the sound sources present and the evaluation measures used. Street canyon experiments show that traffic sounds alter perceived enclosure [45]. Experiments in residential settings indicate that the relationship between the green view index and pleasantness varies with acoustic conditions [46], while questionnaire-based assessments in virtual reality link architectural aesthetics and environmental cleanliness to soundscape comfort [47]. In historic districts, traditional buildings, streets, and sounds associated with local customs are related to perceived vitality [48]; research has also examined visit purpose and activity participation in relation to experiences of the acoustic environment [49]. Studies of soundscape appropriateness identify spatial variation in its relationship with acoustic comfort [50], while tranquility studies report different audiovisual predictors across place types [51]. These findings indicate that comfort and tranquility need to be interpreted in relation to different spatial conditions. Objective measurements link traffic and commercial land use to noise exposure [52], while field interventions show that improvements in experience at the same location vary with listener position, sound installation content, and background sound [53]. These findings reveal differences between places as well as variation in experiences within individual places. Assessments organized around fixed sound-source categories and place types can summarize spatial differences but provide only a coarse account of specific sound events and their activity contexts. Average differences between place types alone cannot fully explain the appropriateness of different sound-related activities within the same type of space or the divergent evaluations they receive.

3. Materials and Methods

3.1. Research Framework

Using Huishan Ancient Town as the study setting, we developed a natural language processing (NLP) framework based on a corpus of social media reviews from multiple online platforms. The framework examined how soundscape experiences were described and evaluated in the review corpus and how these accounts varied across the sampled periods and places (Figure 1). It comprised three interconnected stages. First, the review texts were collected and cleaned. Large language models (LLMs) assisted in screening soundscape content and extracting semantic fragments to produce an analytical corpus with temporal, source, and place information. Second, word frequency analysis and word clouds summarized common expressions in soundscape descriptions, while BERTopic identified topics and their variation across time periods. An attention–sentiment matrix compared relative attention and mean sentiment and guided the selection of contrasting topics for spatial analysis. This stage also linked fragments with identifiable locations to named places and mapped topic and

sentiment distributions onto 25-m grid cells. Grid cells were grouped by similarity in their narrative topic proportions. Each cluster was then described using sentiment summaries, POI labels of associated places, and the activities reported in the texts. Third, we used the resulting profiles to identify reported sound–activity relationships and places for on-site assessment. These questions inform the discussion of soundscape management.

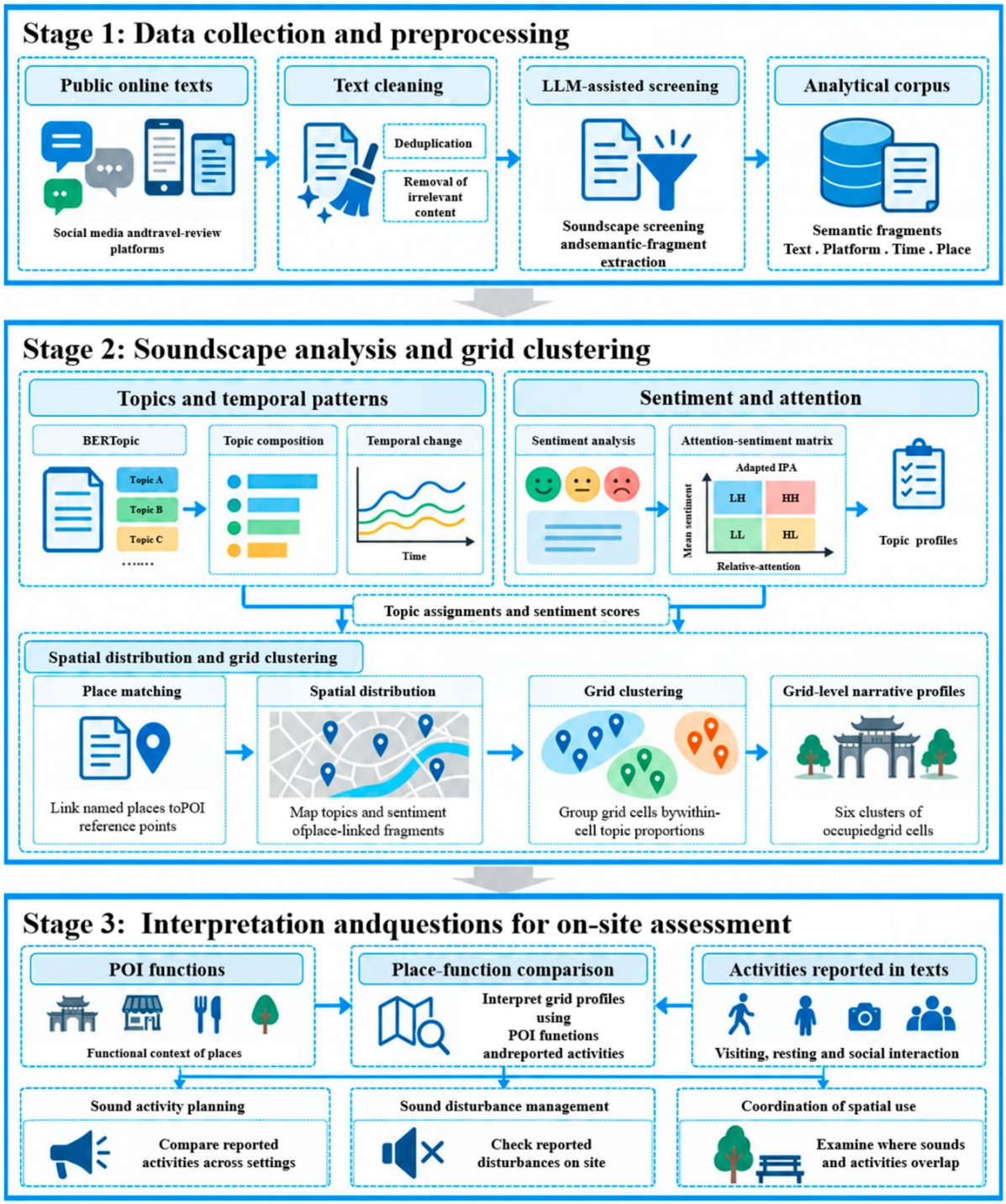

Figure 1. Research framework of the study.

3.2. Study Area

Huishan Ancient Town is located in Wuxi, Jiangsu Province, China. It is distinguished by its ancestral hall complex. Traditional streets and lanes, together with gardens incorporat-

ing hills and water features, form its historic spatial structure. The district conservation plan approved in 2025 assigned protection grades to 118 existing ancestral halls and emphasized the integrated conservation of streets and lanes, springs, and gardens [54]. The "Huishan Ancestral Halls" were included in the sixth batch of Major Historical and Cultural Sites Protected at the National Level in 2006, recognizing the conservation value of this built heritage [55]. Huishan Ancient Town Scenic Area was designated a national 5A tourist attraction in 2020 [56] and recorded more than 6.9 million visits from January to August 2025 [57]. Its gardens, ancestral halls, and historic streets have become regular venues for presentations and performances of intangible cultural heritage [57]. Heritage presentation, cultural activities, sightseeing, and recreation coexist within the same destination, creating distinct contexts of spatial use in halls, courtyards, streets and lanes, and gardens with springs. The combination of diverse building and public space types with active tourism use provided a setting for comparing soundscape topics and sentiment evaluations across places. It also supported an examination of the relationships between auditory experiences and the use and management needs of the historic built environment (Figure 2).

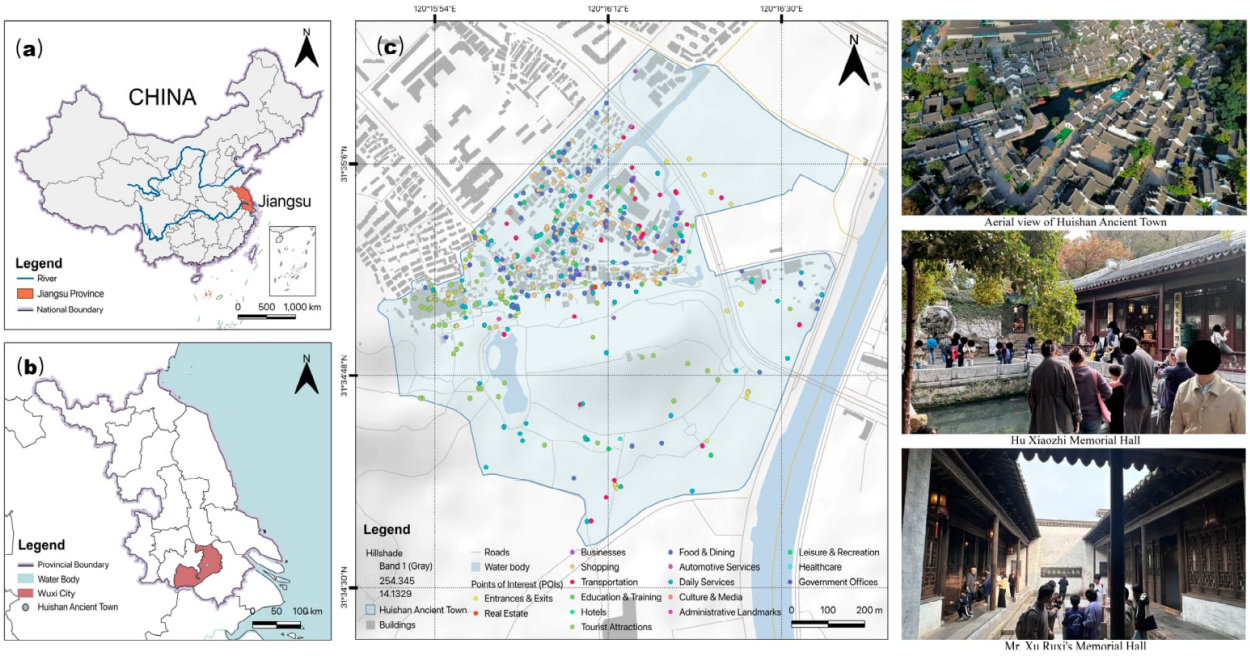

Figure 2. Study Area. (Sources: Copernicus DEM GLO-30 (terrain); Tianditu (administrative boundaries); © OpenStreetMap contributors (roads and hydrography; Available online: https://www.openstreetmap.org/copyright (accessed on 31 September 2026)); and Baidu Maps (points of interest). All photographs were taken by the authors during fieldwork).

3.3. Research Data and Preprocessing

Relevant reviews were retrieved using "惠山古镇" (Huishan Ancient Town) as the search keyword from travel-service platforms (Tongcheng Travel and Ctrip), local consumer-review platform Dianping, and social-media platform Weibo. The initial corpus contained 42,871 records. Fields for review text, date, source, and original review identifier were retained to support subsequent topic, sentiment, and spatial analyses. Contributors' residence and community affiliation were not established in this analysis. First, duplicate imports were identified using the source and original review identifier, and 299 duplicate records were removed. Two authors then prescreened the content using a common set of rules. They excluded 2426 reviews for reasons including advertising or institutional promotion, irrelevance to the study area, and the absence of substantive text, leaving 40,146 reviews. Soundscape screening used prompt engineering [58] to specify inclusion criteria

for auditory cues, soundscape experiences, and evaluative expressions. The prompts required the models to extract contiguous, semantically complete passages verbatim from the original reviews. The gpt-5.6-luna and qwen3.8-flash models were compared using the same soundscape screening prompt. Manual review assessed soundscape relevance, excerpt completeness, and consistency with the original text. qwen3.8-flash was used for batch screening. This stage excluded 35,532 reviews that did not meet the soundscape inclusion criteria and retained 4614 relevant reviews.

Because a single review could describe multiple soundscape experiences, soundscape text fragments were used as the unit of analysis, and 5248 fragments were extracted from the relevant reviews. Extracted passages were checked against their source reviews, and recoverable extraction errors were corrected. A total of 171 fragments were excluded because they lacked clear auditory evidence, contained incomplete context, or could not be matched to a contiguous passage of the original text and could not be repaired. A further 202 duplicate outputs were removed using the original review identifier and excerpt position, leaving 4875 valid soundscape fragments. Identical expressions in different reviews were retained as separate records. Text preprocessing included character and whitespace normalization, Chinese word segmentation, matching against a custom dictionary, synonym consolidation, stopword processing, and low-frequency term removal. We excluded 574 fragments with fewer than three valid word tokens after preprocessing, leaving the same 4301 fragments for topic and sentiment analysis. Topic analysis used the segmented texts. Sentiment assessment used the original wording of the retained fragments to preserve negation, degree modifiers, and necessary context (Supplementary Material B). Each fragment was linked to its original review identifier, source, date, and available place information. Fragments without explicit place information remained in the overall analysis; place-linked records were selected separately for spatial analysis. Counts at each processing stage are reported in Table 1, and the screening rules and text preprocessing resources are provided in Supplementary Materials A and B.

Table 1. The number of records for each processing stage of the data corpus.

Processing Stage	Input Count	Excluded Count	Retained or Output Count	Counting Unit
Initial corpus construction	-	-	42,871	Reviews
Removal of duplicate imports	42,871	299	42,572	Reviews
Manual content prescreening	42,572	2426	40,146	Reviews
AI-assisted screening for soundscape relevance	40,146	35,532	4614	Reviews
Extraction of soundscape text fragments	4614	-	5248	Reviews → fragments
Excerpt quality review	5248	171	5077	Fragments
Removal of duplicate outputs from the same review	5077	202	4875	Fragments
Input screening for topic analysis	4875	574	4301	Fragments

3.4. Research Methods

The analysis operationalized three properties of the review corpus. Topics summarized the sounds, activities, and experiences described in the fragments. Attention measured the proportion of fragments assigned to each topic, and sentiment represented the expressed evaluative direction through polarity labels and scores. These variables support comparisons of reported experiences.

3.4.1. Word Frequency Analysis

Word frequency analysis was used to describe common expressions in the soundscape texts [59]. The analysis covered the base corpus of 4875 soundscape text fragments. After

word segmentation, synonym consolidation, and stopword processing, all occurrences of each retained term were counted across the corpus. Repeated occurrences within the same fragment were counted individually. Terms were ranked in descending order of frequency, and the 60 most frequent terms and their counts were reported. A word cloud visualized the overall distribution, with font size indicating relative frequency [60]. The results summarized references to sounds, activities, and place names, as well as expressions of experience, and provided a reference for interpreting the topics subsequently identified with BERTopic.

3.4.2. BERTopic Topic Identification and Comparisons Across Time Periods

Topic identification grouped semantically similar soundscape narratives. Within the BERTopic framework [61], the preprocessed texts of the 4301 fragments used for joint topic and sentiment analysis were encoded as 384-dimensional sentence embeddings [62] using sentence-transformers/paraphrase-multilingual-MiniLM-L12-v2 [63]. The input texts retained descriptions of sounds, activities, places, and experiences. Evaluative wording within these descriptions could contribute to the semantic similarities used to form the topics. Sentiment labels and scores were not supplied to the embedding or clustering stages. Dimensionality reduction used UMAP [64] with n_neighbors = 15, n_components = 2, min_dist = 0, metric = "cosine", random_state = 42, and init = "pca". HDBSCAN [65] used min_cluster_size = 30, metric = "Euclidean", and cluster_selection_method = "eom", with prediction_data = True and gen_min_span_tree = True. BERTopic was configured with language = "multilingual", nr_topics = 51, and calculate_probabilities = True. Keywords were extracted using class-based term frequency-inverse document frequency (c-TF-IDF) [61]. Topic names were assigned after clustering using these keywords and representative fragments. To examine possible overlap with sentiment analysis, we checked the retained inputs, keywords, and stored representative token sequences across all 27 topics (Supplementary Material C.9). Parameter settings were compared over a grid. The selected settings and resulting topic granularity are reported in Supplementary Material C.2. Four strategies for reassigning outlier fragments were compared [66]. The embeddings strategy was selected because it yielded a higher topic coherence score, and the topic representations were updated. The resulting fragment-to-topic assignments were linked to temporal labels. Topic counts and shares were calculated within each time period to compare changes in the soundscape content discussed. Sentiment outputs were then linked to the topic assignments at the fragment level to summarize the evaluations expressed within each narrative group. Detailed parameters are provided in Supplementary Material C.1, and outlier handling and temporal aggregation are described in Supplementary Material C.3.

3.4.3. Sentiment Analysis Using LLMs

Sentiment analysis summarized evaluative direction and emotion within the narrative groups identified by topic modeling. LLM-based text annotation [67] followed predefined rules and used gpt-5.6-luna and qwen3.8-flash, with model names recorded as their API identifiers. Both models used identical prompts with temperature = 0 to annotate the same set of 4301 fragments. The input consisted of the original text, preserving negation, degree modifiers, and context. Outputs included a dominant emotion, a three-class sentiment polarity label, a score ranging from 0 to 1, keywords from the original text, and an explanation. Three annotators independently annotated the sentiment polarity of the 4196 eligible fragments. The final manually reviewed labels were used as the reference for the paired model comparison. gpt-5.6-luna achieved a polarity agreement rate of 81.89% and a macro-averaged F1 score of 0.8056, compared with 65.18% and 0.6462, respectively, for qwen3.8-flash. The gpt-5.6-luna outputs were therefore adopted. Polarity and emotion

were evaluated separately, as detailed in Supplementary Materials C.4 and C.7. Scores closer to 1 indicated more positive evaluations, scores closer to 0 indicated more negative evaluations, and scores around 0.5 indicated neutrality. Emotion labels were consolidated into eight descriptive reporting categories. Fragment-level codes were linked to topics and time periods to summarize emotion distributions and mean sentiment scores. Keyword strings were split using "&" as the delimiter. Terms of one character or fewer were removed, and keyword occurrences were totaled within each emotion category. API parameters are provided in Supplementary Material C.1; the human reference labels, comparison sample, and polarity evaluation results in Supplementary Material C.4; and the complete prompts and rules for counting keywords in Supplementary Material C.5.

The 128 distinct model-generated labels were organized by the saved mapping dictionary into eight reporting categories specific to this corpus. The scheme distinguishes calmness and comfort, enjoyment and appreciation, and interest and engagement. It also separates irritation and activity interference from low mood and fatigue, and from anxiety and tension. Mixed and Reflective Feelings retains ambivalent, uncertain, and reflective expressions as a heterogeneous residual group. No Explicit Emotion Expressed records descriptive or atmospheric labels, rather than a distinct emotion. These categories describe expression types and can span more than one polarity. They were used for descriptive distributions and keyword profiles across topics, periods, and spatial clusters; spatial clustering itself used topic proportions. For the separate emotion comparison, the final human-reviewed fine-grained labels and the model labels were mapped with the same dictionary. Of 4301 fragments, 4297 had interpretable emotion references, including the 101 mixed-polarity cases excluded from the polarity comparison. Agreement was 55.81%, with a macro-averaged F1 score of 0.4763. This comparison measures label agreement under the shared mapping. The conceptual boundaries of the categories require separate validation, while cultural interpretation draws on evidence about listening practices and local meanings (Section 5.1). The continuous sentiment scores require separate calibration. Category definitions, the emotion–polarity cross-tabulation, and class-level performance are provided in Supplementary Material C.7; Supplementary Material C.10 provides the complete label mapping.

3.4.4. Comparison of Relative Attention and Mean Sentiment

The attention–sentiment matrix adapts the two-axis display of Martilla and James [68] to describe soundscape topics in the review corpus. Relative attention represents digital visibility as the percentage of the 4301 fragments assigned to each topic. Mean sentiment averages the fragment-level sentiment scores within a topic to summarize expressed evaluation. Equations (1) and (2) define the indicators. The 75th percentile of each indicator across topics was used as the cutoff on its respective axis. The resulting groups describe combinations of higher or lower relative attention and mean sentiment relative to these cutoffs (Supplementary Material C.6). The matrix was used to compare topic profiles and select contrasting cases for spatial examination. Interpretation considered the reported activities and emotion distributions alongside topic positions. Threshold sensitivity compared mean-based cutoffs and the 50th, 65th, 70th, 75th, 80th, and 85th percentiles across the 27 equally weighted topics, applying the same criterion to both axes. Quadrant assignments were compared with the 75th-percentile baseline. The influence of T25 was examined by excluding it from cutoff estimation while holding all 27 topic values fixed (Supplementary Material C.8).

$$A_k = \frac{n_k}{N} \tag{1}$$

$$E_k = \frac{\sum_{i \in k} s_i}{n_k} \tag{2}$$

Here, n_k denotes the number of fragments assigned to topic $\boldsymbol{k}$, N is the total number of fragments in the joint analysis, and s_i us the sentiment score of fragment i. $\boldsymbol{A_k}$ and $\boldsymbol{E_k}$ denote relative attention and mean sentiment, respectively. Groups are reported as HH, HL, LH, and LL. The first letter refers to relative attention and the second to mean sentiment. H denotes a value at or above its cutoff, and L a value below it.

3.4.5. Spatial Mapping of Place-Linked Soundscape Narratives

Spatial mapping used the experience location field extracted from each fragment to associate reported experiences with named buildings and public spaces. After manual checking and standardization of place names, each identified place was represented by its POI reference point. The study boundary and place coordinates were transformed into a common projected coordinate system in meters (EPSG:32651). A base grid of 25 m × 25 m cells was generated, with boundary cells clipped to the study boundary. The 25 m grid was used as a working unit for local spatial aggregation, allowing narratives linked to named buildings and public spaces to be summarized within a common spatial framework. The POI reference point anchors the place named in the text; it does not establish where the sound originated or where the reviewer listened. Grid cells are aggregation units and do not delineate rooms, courtyard enclosures, or paths of sound transmission through walls and openings. Sound propagation and attenuation between source and listener were not modelled. We examined the sensitivity of this specification to finer and coarser grids rather than treating 25 m as a uniquely optimal scale. The 2353 place-linked fragments were assigned to grid cells using their reference points, with each fragment counted in one cell. For each grid cell, the total number of fragments, topic richness (number of distinct topics), and mean sentiment score were calculated. Topic-specific fragment counts and mean sentiment scores were then calculated for each cell and displayed in paired maps of place-linked narrative frequency and expressed evaluation. Mean scores for cells without fragments were retained as missing values; the original maps used 0.50 as a plotting fill value. Each occupied cell's narrative topic profile provided the input for clustering. Sentiment summaries and POI labels described the resulting clusters.

Grid-scale sensitivity was assessed at 10, 20, 25, 30, and 50 m using the same 2353 place-linked fragments and 109 reference locations. The study boundary, EPSG:32651 projection, and grid origin at the minimum projected coordinates of the boundary were held fixed. Topic proportions and mean sentiment scores were recalculated at each scale. For cross-scale comparison, each location was assigned the statistics and cluster label of its containing cell, providing the same set of locations at every scale. Relative to 25 m, we compared site-level sentiment rankings and absolute differences, Jensen–Shannon distances between topic-proportion vectors, and the adjusted Rand index (ARI) for cluster assignments. Clustering settings were held constant as described in Section 3.4.6. Aggregation characteristics and comparison results are reported in Supplementary Material H.

3.4.6. Grid-Cell Clustering and Associated Place Functions

Clustering grouped grid cells with similar narrative topic profiles. Input features comprised the proportions of fragments assigned to each of the 27 topics within the 89 grid cells with text records (Equation (3)). Each topic column was standardized using StandardScaler [69], following Equation (4). K-means clustering [70] was performed for K = 2–10, and the within-cluster sum of squares and mean silhouette coefficient [71] were compared across solutions. A six-cluster solution was adopted for subsequent analysis. The random seed was set to 42, with 50 initializations for each value of K and a maximum of 500 iterations per run. In the grid-scale sensitivity analysis, topic proportions were standardized separately at each scale, while K = 6, random_state = 42, n_init = 50, and max_iter = 500 were held constant to isolate changes associated with spatial aggregation. Cluster topic

centers were calculated as the arithmetic means of the original topic proportions across member grid cells. After cluster assignment, emotion category proportions and POI label proportions were calculated separately to describe the review fragments in each cluster. Both used the total number of fragments in each cluster as the denominator. The arithmetic mean of fragment sentiment scores was also calculated for each cluster. The place contexts of the grid clusters were described using the associated POI labels as indicators of place function [72], together with the named buildings and public spaces.

$$p_{gk} = \frac{n_{gk}}{n_g} \tag{3}$$

$$z_{gk} = \frac{p_{gk} - \mu_k}{\sigma_k} \tag{4}$$

Here, n_{gk} denotes the number of fragments assigned to topic k in grid cell g, n_g is the total number of fragments in that cell, and p_{gk} is the corresponding topic proportion. The μ_k and σ_k parameters denote the mean and population standard deviation, respectively, of the proportions for topic k across the 89 grid cells. The standardized input is z_{gk}.

4. Results

4.1. Characteristics of High-Frequency Terms in Soundscape Descriptions

Word frequency analysis covered 4875 fragments in the base soundscape corpus, yielding 1090 distinct retained terms with 34,236 total occurrences. Among expressions related to sounds and activities, “Chatting” (570 occurrences) and “Talking” (394) were prominent. “Footsteps” (322) and “Human voices” (268) also ranked highly, reflecting the prominence of everyday social interaction and walking in soundscape narratives. “Conversation” (193), “Calling to others” (183), and “Guided explanations” (134) added further descriptions of interpersonal communication. Terms related to the natural environment included “Birdsong” (226), “Wind blowing” (155), “Tree leaves” (148), and “Water sounds” (48). Playback devices and audio content were represented by “Music” (222), “Phone speaker playback” (157), and “Public address announcements” (101). Together, these terms described sounds and their associated activities. Table 2 lists the 60 most frequent terms and their counts, and Figure 3 presents the overall term distribution as a word cloud. Soundscape narratives also described locations, the course of activities, and experiences. Among place-related terms, “Huishan” occurred 242 times, while “Square”, “China Azalea Garden”, “Ancestral halls”, and “Jichang Garden” occurred 184, 154, 97, and 71 times, respectively. Thus, the soundscape vocabulary included a range of place references. “Loudly” (247) and “In a quiet voice” (202) provided linguistic descriptions of loudness. The corpus contained “Quiet” (194), “Relaxation” (165), and “Quite comfortable” (161), alongside “Slightly annoyed” (126), reflecting different descriptions of the acoustic environment and feelings at the time. “Interruption” (160) and “Distraction” (132) referred to changes in attention during these experiences. The two most frequent terms, “For a while” (634) and “Nearby” (624), together with “A spell” (538) and “Slowly” (420), primarily conveyed time, location, or manner of activity. Overall, the high-frequency vocabulary encompassed identifiable sounds and sound-producing activities, as well as place references, temporal and spatial expressions, and experiential terms.

Table 2. The 60 most frequent terms in the corpus.

Rank	Term	Frequency	Rank	Term	Frequency	Rank	Term	Frequency
1	For a while	634	21	In a quiet voice	202	41	China Azalea Garden	154
2	Nearby	624	22	Following along	195	42	Walking past	151
3	Chatting	570	23	Footfall sounds	194	43	Reminders	149
4	A spell	538	24	Quiet	194	44	Softly	149
5	Slowly	420	25	Conversation	193	45	Tree leaves	148
6	Companions	394	26	At the site	191	46	In a raised voice	147
7	Talking	394	27	Details	189	47	Moving about	141
8	A while	362	28	Square	184	48	Position	135
9	Footsteps	322	29	Calling to others	183	49	Buildings	135
10	Sounds and activity	290	30	Taking photographs	178	50	Guided explanations	134
11	Immediate surroundings	269	31	Visitors	170	51	Mood	133
12	Human voices	268	32	Relaxation	165	52	Reducing pace	132
13	A little	260	33	Coming to a stop	163	53	Distraction	132
14	A few utterances	257	34	Quite comfortable	161	54	Coming from elsewhere	128
15	Loudly	247	35	Interruption	160	55	A few sounds	126
16	Huishan	242	36	A small	160	56	Slightly annoyed	126
17	Resting one's feet	239	37	Phone speaker playback	157	57	Shouting to others	125
18	Birdsong	226	38	Explanatory information	155	58	Noticing	124
19	Music	222	39	Wind blowing	155	59	Speech sounds	123
20	Within view	216	40	Paying attention	155	60	Family members	123



Figure 3. Word cloud diagram of soundscapes description.

4.2. Soundscape Topic Composition and Variation Across Time Periods

4.2.1. Topic Composition and Semantic Content

Following outlier reassignment, the 4301 fragments used for joint analysis were distributed across 27 narrative topics. Table 3 reports the full topic labels, fragment counts, and shares. The topic distribution was markedly uneven. "Mixed Sounds of Conversations, Broadcasts, and Music" (T25) contained 1960 fragments (45.57%), making it the largest topic. The remaining 26 topics accounted for 54.43% of the fragments, with individual counts ranging from 40 to 175. "Footsteps during Visits to Pavilions" (T18) contained 175 fragments (4.07%), "Wind and Rustling Leaves during Visits to the Azalea Garden" (T1) contained 161 (3.74%), and "Distant Conversations along the Waterfront and Distraction" (T26) contained 153 (3.56%). The smallest topic, "Verbal Greetings and Photography near Pavilions" (T8), contained 40 fragments (0.93%). Twelve topics contained at least 100 fragments, while the remaining 15 contained fewer than 100. The corpus thus comprised one broad mixed-sound topic and multiple topics tied to specific contexts.

Table 3. Narrative topic labels, fragment counts, and shares.

Topic	English Label	Number of Segments	Percentage (%)
T1	Wind and Rustling Leaves during Visits to the Azalea Garden	161	3.74
T2	Conversations and Phone Calls while Walking through Halls	59	1.37
T3	Car Door Sounds and Conversations in Parking Areas	61	1.42
T4	Phone Speaker Playback and Disrupted Attention	109	2.53
T5	Loud Phone Calls Disrupting Architectural Appreciation	61	1.42
T6	Bridge Crossing and Shouted Exchanges	70	1.63
T7	Conversations from Adjacent Seats in Shops	48	1.12
T8	Verbal Greetings and Photography near Pavilions	40	0.93
T9	Video Audio Playback and Interrupted Rest	66	1.53
T10	Conversations in Ancestral Hall Courtyards	55	1.28
T11	Conversations and Explanations during Ancestral Hall Visits	113	2.63
T12	Occasional Calls and Perceived Quiet at Viewing Platforms	114	2.65
T13	Cup Sounds and Music in Tea Seating Areas	80	1.86
T14	Rustling of Wind-Blown Leaves	152	3.53
T15	Occasional Birdsong among Trees	42	0.98
T16	Repeated Sounds from Trees and Auditory Annoyance	43	1.00
T17	Footsteps and Perceived Quiet during Courtyard Visits	152	3.53
T18	Footsteps during Visits to Pavilions	175	4.07
T19	Quiet Parent–Child Conversations about Musicians	46	1.07
T20	Repeated Loud Conversations and Auditory Fatigue	74	1.72
T21	Raindrops beneath Eaves and Splashing Footsteps	53	1.23
T22	Recurrent Bird Calls	117	2.72
T23	Sounds of Flowing Spring Water	101	2.35
T24	Quiet Family Conversations and a Sense of Warmth	85	1.98
T25	Mixed Sounds of Conversations, Broadcasts, and Music	1960	45.57
T26	Distant Conversations along the Waterfront and Distraction	153	3.56
T27	Disruptive Loud Conversations at Waterside Resting Spots	111	2.58

The keywords and representative fragments combined references to sounds, sound-producing activities, place contexts, and experiences. Topics related to natural sounds included wind-rustled leaves, birdsong with different patterns of occurrence, flowing spring water, raindrops beneath eaves, and splashing footsteps. The Azalea Garden topic also included the names of pavilions and halls within the garden alongside descriptions

of wind and rustling leaves. Interpersonal communication appeared in descriptions of halls, shops, ancestral halls, bridge approaches, and waterfront areas. The texts referred to quiet exchanges, explanations and storytelling, calls across distances, and repeated loud conversations. Walking appeared in topics concerning visits along courtyard edges and to pavilions. Fragments describing tea seating areas mentioned clinking cups and music. Topics concerning phone speaker playback, video playback, and loud phone calls included expressions of disrupted attention, interrupted rest, or interference with viewing architecture. Fragments describing parent–child storytelling and family conversations included experiential expressions of feeling moved and a sense of warmth. Figure 4 presents the ten highest-ranked keywords and their c-TF-IDF weights for each of the 27 topics. T4's keywords included "helplessness" and "interruption", while T17 included "quiet" and "mood" (Supplementary Material C.9). These representations contain perceptual or evaluative language alongside sound-related content.



Figure 4. Top ten keywords and their c-TF-IDF weights for each of the 27 soundscape topics.

The topic similarity heatmap shows the semantic proximity among the 27 topics (Figure 5). Cosine similarity across the 351 unique topic pairs ranged from 0.37 to 0.97, with a median of 0.68. Similarity was highest between "Wind and Rustling Leaves during

Visits to the Azalea Garden" (T1) and "Distant Conversations along the Waterfront and Distraction" (T26), at 0.97. It was lowest between "Verbal Greetings and Photography near Pavilions" (T8) and "Quiet Family Conversations and a Sense of Warmth" (T24), at 0.37. Topics with related content also showed varying degrees of similarity. Similarity was 0.79 between "Occasional Birdsong among Trees" (T15) and "Recurrent Bird Calls" (T22), and 0.75 between "Phone Speaker Playback and Disrupted Attention" (T4) and "Video Audio Playback and Interrupted Rest" (T9). The footstep topics for courtyard and pavilion visits (T17 and T18) had a similarity of 0.76, compared with 0.80 for the two waterfront conversation topics (T26 and T27). Similarity between the mixed sound topic (T25), which contained the most fragments, and the other topics ranged from 0.49 to 0.80. It was highest for T1 and lowest for T24, indicating varying semantic proximity to topics describing different sounds and place contexts. Supplementary Material D shows the relationships among these 27 topics using Ward linkage and Euclidean distances between L2-normalized c-TF-IDF topic vectors.

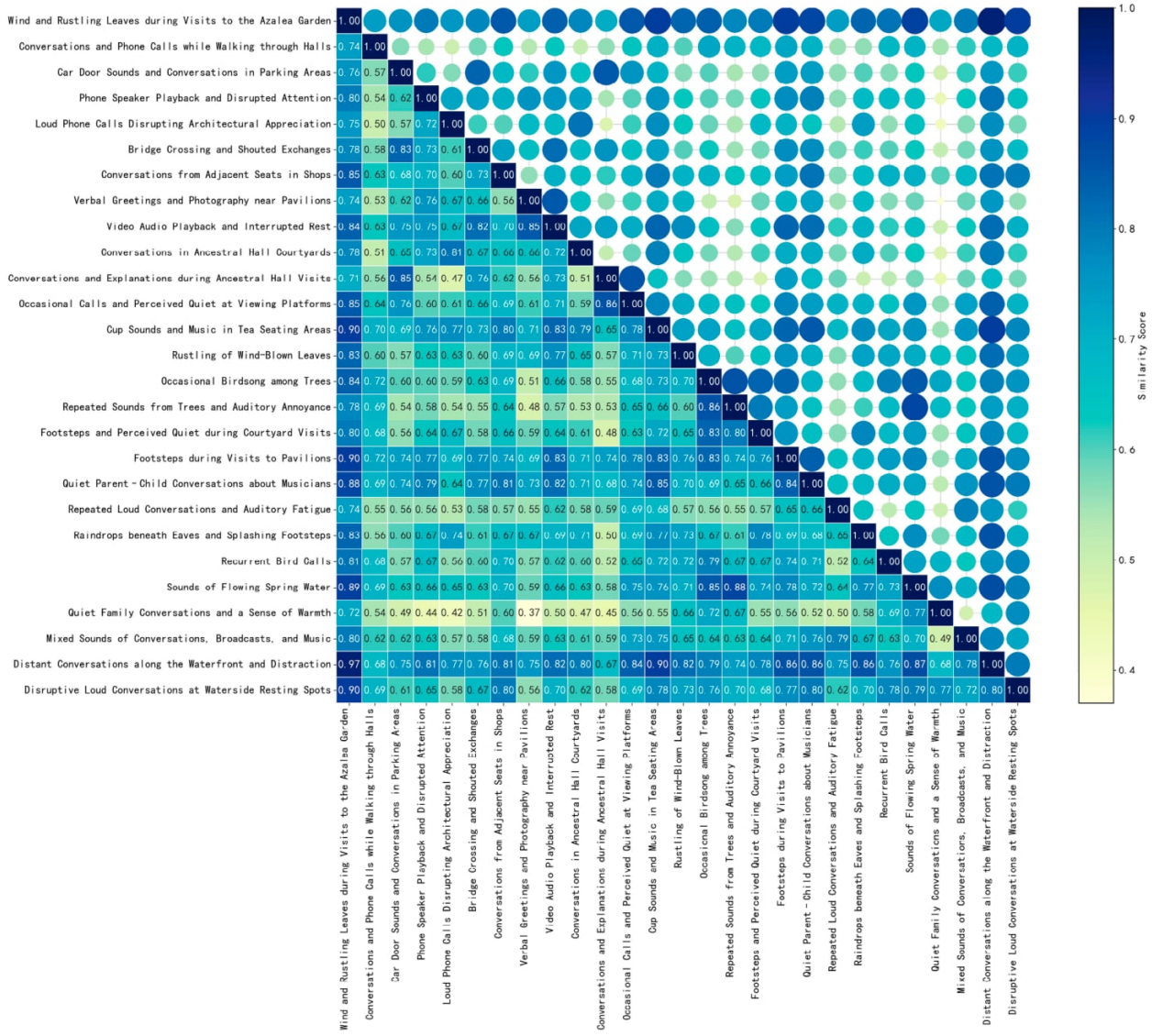

Figure 5. Cosine similarities between the 27 topic embeddings. *The lower triangle reports similarity scores; color and bubble size in the upper triangle encode the same similarities.*

4.2.2. Variation in Topic Composition Across Time Periods

Across the 27 topics, the periods of 2022 and earlier, 2023, 2024, 2025, and 2026 contained 43, 1253, 781, 864, and 1360 fragments, respectively. The numbers of observed topics were 8, 26, 27, 25, and 27, respectively (Figure 6). The figure label "2022" includes records from that year and earlier, whereas "2026" covers records through September. Topic shares were calculated using the total fragment count in each period as the denominator. Because sample sizes were uneven and proportions in the earliest period were sensitive to small changes in counts, the main comparison focuses on the two largest samples, 2023 and 2026. Figure 6 retains all five periods as descriptive context.

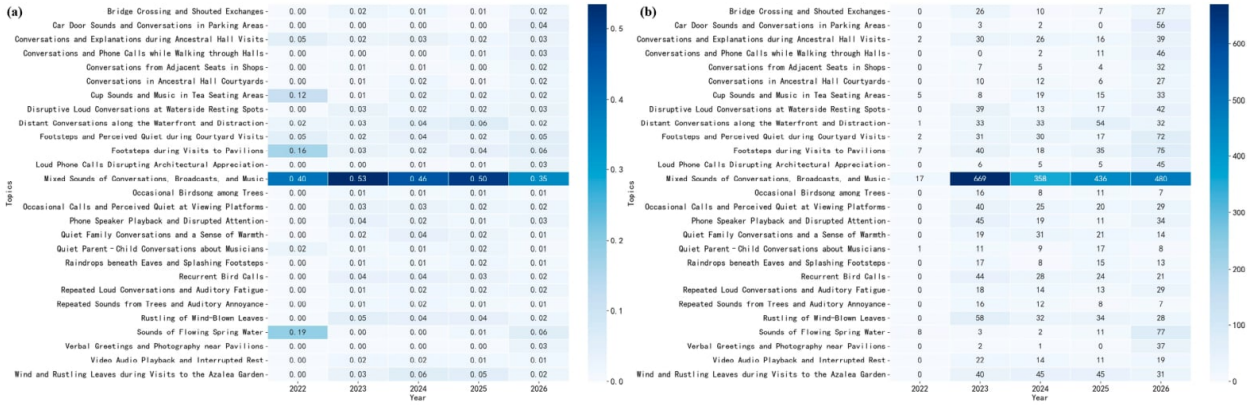

Figure 6. (**a**) Topic proportions within each temporal stage. (**b**) Numbers of soundscape text fragments assigned to each topic in each temporal stage.

"Mixed Sounds of Conversations, Broadcasts, and Music" (T25) remained the most frequent topic in both periods, accounting for 669 fragments (53.39%) in 2023 and 480 (35.29%) in 2026 (Figure 6a). Its share was 18.10 percentage points lower in 2026. "Sounds of Flowing Spring Water" (T23) accounted for 3 fragments (0.24%) and 77 (5.66%), respectively. The corresponding counts for "Rustling of Wind-Blown Leaves" (T14) were 58 (4.63%) and 28 (2.06%), while "Recurrent Bird Calls" (T22) accounted for 44 (3.51%) and 21 (1.54%) (Figure 6a,b). These topic shares describe the composition of the retained soundscape fragments. Supplementary Material I summarizes the corresponding tourism activities and reported visitor volumes to contextualize the comparison.

4.3. Emotion and Sentiment in Soundscape Narratives

4.3.1. Emotion Category Distribution and Keyword Characteristics

Among the 4301 soundscape text fragments, the 128 distinct fine-grained emotion labels produced by the model were consolidated into eight descriptive reporting categories. The model assigned the most fragments to Calmness and Comfort, with 1872 fragments (43.52%), followed by Irritation and Dissatisfaction, with 1221 (28.39%). There were 436 fragments (10.14%) in Positive Affect and Enjoyment and 395 (9.18%) in Low Mood and Fatigue. No Explicit Emotion Expressed accounted for 144 fragments (3.35%); Anxiety and Tension, 100 (2.33%); Interest and Anticipation, 70 (1.63%); and Mixed and Reflective Feelings, 63 (1.46%). Among the fine-grained labels, "Calmness," "Irritation," and "Annoyance" occurred 1322, 573, and 206 times, respectively. Three-class sentiment polarity coding classified 1706 fragments as negative (39.67%), 1332 as neutral (30.97%), and 1263 as positive (29.37%). The overall mean sentiment score was 0.510, the median was 0.500, and scores ranged from 0.05 to 0.95. The Calmness and Comfort category included 747 positive and 1125 neutral fragments, reflecting the distinction between emotion category and sentiment polarity coding (Table 4).

Table 4. Distribution of emotion categories and sentiment polarity in soundscape fragments.

(**a**) Emotion categories		
Emotion category	**Fragments (n)**	**Share (%)**
Calmness and Comfort	1872	43.52
Irritation and Dissatisfaction	1221	28.39
Positive Affect and Enjoyment	436	10.14
Low Mood and Fatigue	395	9.18
No Explicit Emotion Expressed	144	3.35
Anxiety and Tension	100	2.33
Interest and Anticipation	70	1.63
Mixed and Reflective Feelings	63	1.46
Total	4301	100.00
(**b**) Sentiment polarity		
Sentiment polarity	**Fragments (n)**	**Share (%)**
Negative	1706	39.67
Neutral	1332	30.97
Positive	1263	29.37
Total	4301	100.00

Note: Emotion categories and sentiment polarity are separate model-output fields. Percentages use 4301 fragments as the denominator. Table C8 cross-tabulates these fields; Table C9 separately compares emotion categories against the harmonized human-reviewed reference labels. Polarity comparisons are reported in Supplementary Material C.4 and the original prompt in Supplementary Material C.5.

The eight-category comparison against the harmonized human-reviewed reference labels yielded 55.81% agreement and a macro-averaged F1 score of 0.4763 across 4297 fragments. The largest confusion involved 840 fragments with no explicit emotion in the reference being assigned to Calmness and Comfort by the model. Mixed and Reflective Feelings, which groups ambivalent, uncertain, and reflective expressions, had an F1 score of 0.0894. A sensitivity comparison pooling Interest and Anticipation with Positive Affect and Enjoyment yielded 56.27% agreement, leaving the principal confusion unchanged. The distributions summarize the labels assigned by the model. The assignment of reference No Explicit Emotion Expressed fragments to Calmness and Comfort contributes to its high share in the model output. The low F1 for Mixed and Reflective Feelings limits interpretation of its frequency. Supplementary Material C.7 presents the definitions, confusion matrix, class-level metrics, and pooling comparison.

Frequent keywords described different soundscape experiences across the emotion categories. In Calmness and Comfort, prominent terms included “Birdsong” (110 occurrences), “Quite comfortable” (78), and “Relaxation” (56), alongside “Walked slowly while viewing the scenery” (47) and “Quiet” (44). Irritation and Dissatisfaction featured “Slightly annoyed” (113), “Phone speaker playback” (77), “Was interrupted” (42), and “A loud voice” (36). These expressions concerned sound content, shifts in attention, and interruptions to activities. Positive Affect and Enjoyment was represented by “Laughter” (24), “Friendly” (19), and “Clapping” (14). Low Mood and Fatigue frequently included “A little tired” (36), “Somewhat tired” (30), and “Talking loudly” (27). Anxiety and Tension featured “Anxious” (27) and calls to assemble, while Interest and Anticipation included “Interesting” (11) and “Wanted to learn more” (9). Mixed and Reflective Feelings included both “Difficult to hear” and “Friendly” (6 occurrences each). No Explicit Emotion Expressed also included terms such as “Paying attention” (7) and “Footfall sounds” (5) (Figure 7).



Figure 7. Keyword clouds by emotion category.

4.3.2. Emotion and Sentiment Profiles by Narrative Topic and Time Period

Figure 8a shows the distribution of model-coded emotion categories among the fragments assigned to each narrative topic. The model assigned Calmness and Comfort to 98.11% of the fragments in "Raindrops beneath Eaves and Splashing Footsteps". The corresponding shares for "Rustling of Wind-Blown Leaves" and "Cup Sounds and Music in Tea Seating Areas" were 85.53% and 86.25%, respectively. Within "Sounds of Flowing Spring Water," Calmness and Comfort and Positive Affect and Enjoyment accounted for 52.48% and 38.61%, respectively. The model assigned Positive Affect and Enjoyment to 56.52% of the fragments in

"Quiet Parent–Child Conversations about Musicians". The mixed-sound topic (T25), which contained the most fragments, showed a varied distribution of emotion labels. Calmness and Comfort accounted for 39.23%, Irritation and Dissatisfaction for 28.16%, and Positive Affect and Enjoyment for 12.50%. Among topics involving audio played aloud, Irritation and Dissatisfaction accounted for 96.97% of fragments in "Video Audio Playback and Interrupted Rest" and 87.16% in "Phone Speaker Playback and Disrupted Attention". "Repeated Loud Conversations and Auditory Fatigue" comprised Low Mood and Fatigue (56.76%) and Irritation and Dissatisfaction (43.24%). All fragments in this topic and the video audio playback topic were classified as negative in the separate sentiment polarity coding.

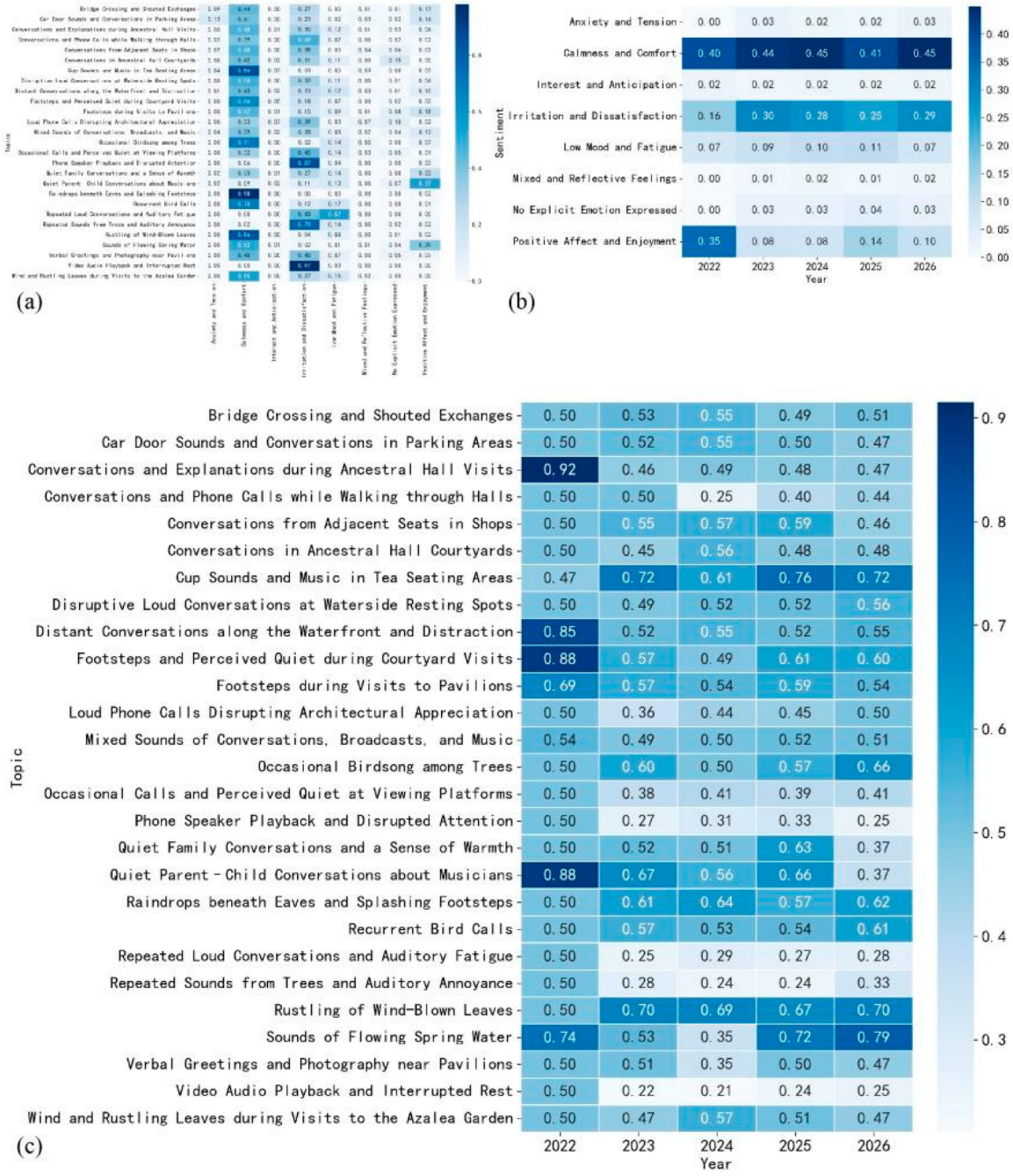

Figure 8. Coded emotion and sentiment in soundscape narratives. Panel (**a**) shows emotion category proportions by topic. Panel (**b**) shows their proportions by time period. Panel (**c**) shows mean sentiment scores by topic and time period.

Calmness and Comfort had the largest share in all five periods, ranging from 39.53% to 44.93%. Irritation and Dissatisfaction accounted for 25.46–30.25% of fragments in 2023-2026 (Figure 8b). Positive Affect and Enjoyment accounted for 34.88% in 2022 and earlier, a pooled period containing 43 fragments; the corresponding shares in 2023-2026 ranged from 7.81% to 14.35%. Consistent with Section 4.2.2, the following topic-level comparisons focus on 2023 and 2026 (Figure 8c). The wind-blown leaves topic had a mean sentiment score of approximately 0.70 in both periods, based on 58 and 28 fragments. Mean scores for cup sounds and music in tea seating areas were 0.72 (n = 8) and 0.72 (n = 33), respectively. Mean sentiment scores for fragments in the phone speaker playback, video audio playback, and repeated loud conversation topics ranged from 0.22 to 0.28 across the two periods. Fragments grouped under flowing spring water had mean sentiment scores of 0.53 (n = 3) in 2023 and 0.79 (n = 77) in 2026. The corresponding scores for quiet family conversations were 0.52 (n = 19) and 0.37 (n = 14). For parent–child conversations, they were 0.67 (n = 11) and 0.37 (n = 8), respectively.

4.4. Relative Attention and Mean Sentiment Across Soundscape Topics

Figure 9 compares digital visibility and expressed evaluation across 27 soundscape topics using relative attention and mean sentiment, respectively. The 75th-percentile cutoffs were 2.6854% for relative attention and 0.5697 for mean sentiment (Supplementary Material C.6). The HL group comprised five topics with 2566 fragments (59.66%). Within this group, "Mixed Sounds of Conversations, Broadcasts, and Music" (T25) contained the most fragments, accounting for 45.57% of the joint analysis corpus, with a mean score of 0.506. The other 4 topics were "Footsteps during Visits to Pavilions" (T18), "Wind and Rustling Leaves during Visits to the Azalea Garden" (T1), "Distant Conversations along the Waterfront and Distraction" (T26), and "Recurrent Bird Calls" (T22). Their relative attention ranged from 2.72% to 4.07%, and their mean scores ranged from 0.508 to 0.562. The HH group contained "Rustling of Wind-Blown Leaves" (T14) and "Footsteps and Perceived Quiet during Courtyard Visits" (T17). Each comprised 152 fragments and had a relative attention of 3.53%, with mean scores of 0.691 and 0.578, respectively. Both indicators exceeded their respective cutoffs.

The five topics in LH comprised 322 fragments, accounting for 7.49% of the corpus (Table 5). Fragments assigned to "Sounds of Flowing Spring Water" (T23) had the highest mean sentiment score across topics (0.762), with a relative attention of 2.35%. "Cup Sounds and Music in Tea Seating Areas" (T13) had a mean score of 0.684 and a relative attention of 1.86%. Mean scores for "Raindrops beneath Eaves and Splashing Footsteps" (T21), "Quiet Parent–Child Conversations about Musicians" (T19), and "Occasional Birdsong among Trees" (T15) ranged from 0.584 to 0.605; their relative attention was at most 1.24%. These topics combined above-cutoff mean sentiment with below-cutoff relative attention. The LL group contained 15 topics comprising 1109 fragments (25.78%), with mean scores ranging from 0.230 to 0.525. Within this group, "Video Audio Playback and Interrupted Rest" (T9), "Repeated Sounds from Trees and Auditory Annoyance" (T16), "Repeated Loud Conversations and Auditory Fatigue" (T20), and "Phone Speaker Playback and Disrupted Attention" (T4) had mean scores of 0.230, 0.268, 0.273, and 0.275, respectively. These were the four lowest topic means, with relative attention ranging from 1.00% to 2.53%.

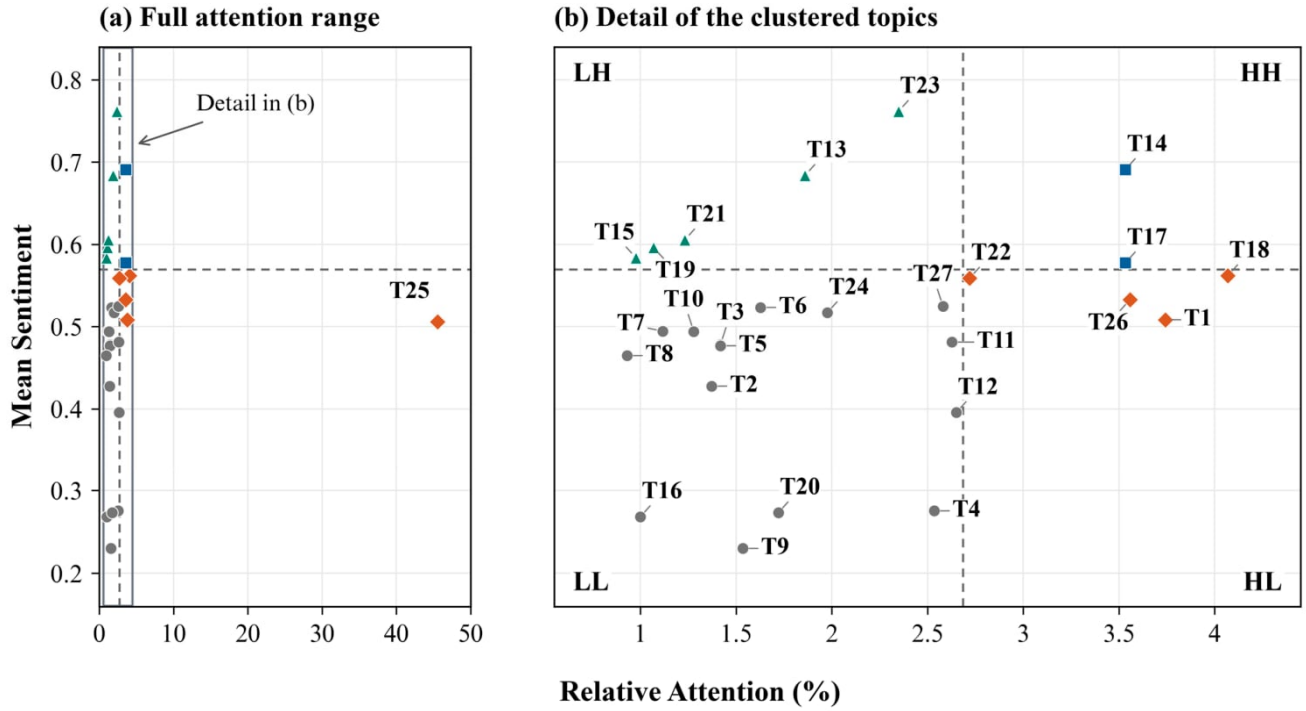

Figure 9. Relative attention and mean sentiment across 27 soundscape topics. (**a**) Full attention range. (**b**) Enlarged view of the 26 topics other than T25. Relative attention is each topic's share of the 4301 fragments. Mean sentiment is the average sentiment score within each topic. Dashed lines mark the 75th-percentile cutoffs of 2.6854% and 0.5697, respectively. In the group codes, the first letter refers to relative attention and the second to mean sentiment. H denotes a value at or above the corresponding cutoff and L a value below it.

Table 5. Topic groups by relative attention and mean sentiment. The first letter denotes relative attention, and the second denotes mean sentiment. H indicates a value at or above the corresponding cutoff. L indicates a value below it.

Group	Topic IDs	Topics (n)	Fragments (n)	Share (%)
HH	T14, T17	2	304	7.07
HL	T25, T18, T1, T26, T22	5	2566	59.66
LH	T23, T13, T21, T19, T15	5	322	7.49
LL	T12, T11, T27, T4, T24, T20, T6, T9, T3, T5, T2, T10, T7, T16, T8	15	1109	25.78
Total	T1–T27	27	4301	100.00

Note. Share (%) is the percentage of the 4301 fragments assigned to each group.

Using the 65th, 70th, 80th, and 85th percentiles on both axes retained the baseline quadrant assignment for 21/27 (77.78%), 25/27 (92.59%), 25/27 (92.59%), and 22/27 (81.48%) of topics, respectively. T1 and T25 remained in HL across the 65th–85th percentile settings. Mean-based and 50th-percentile cutoffs retained 17/27 (62.96%) and 15/27 (55.56%) baseline assignments. The mean-based rule placed no topics in HL and placed T25 in HH; T25 remained in HL under all tested percentile rules. Excluding T25 only from cutoff estimation changed nine of the other 26 assignments under the mean rule and one under each of the 50th- and 75th-percentile rules. T25 remained above the attention cutoff, while its position on the sentiment axis depended on the cutoff criterion (Supplementary Material C.8).

4.5. Distribution of Place-Linked Soundscape Narratives Across Buildings and Public Spaces

A grid-scale sensitivity analysis using 10, 20, 25, 30, and 50 m grids showed broadly similar place-level sentiment rankings, while cluster assignments varied with grid size. The spatial results are reported at the 25 m scale, with detailed comparisons provided in Supplementary Material H. The mapped values describe the distribution of narratives attributed to named places and their expressed evaluations. They do not constitute acoustic exposure maps or independent assessments of soundscape quality.

4.5.1. Spatial Variation in Topic Richness and Overall Sentiment

Soundscape fragments were linked to specific places to examine how the differences identified in the attention–sentiment matrix were distributed across buildings and public spaces in Huishan Ancient Town. Of the 4301 fragments used for joint analysis, 2353 could be linked to 109 standardized locations. These fragments represented 54.71% of the corpus and covered 89 25-m grid cells (Figure 10). All 27 topics were represented in this place-linked corpus. Cells with text records contained 1–18 topics, with a median topic richness of 4. Specifically, 27 cells contained 1–2 topics, 21 contained 3–4, 23 contained 5–7, and 18 contained 8 or more. Cells with records were concentrated in the western and central parts of the study area, with additional scattered locations. The cell associated with Jichang Garden had the highest topic richness, with 18 topics and 118 fragments. The shared cell for the Second Spring under Heaven and Luzi Shrine contained 17 topics and 120 fragments, the largest fragment count. The cell associated with Bishan Poetry Society contained 14 topics and 57 fragments, representing multiple soundscape experiences at this site.

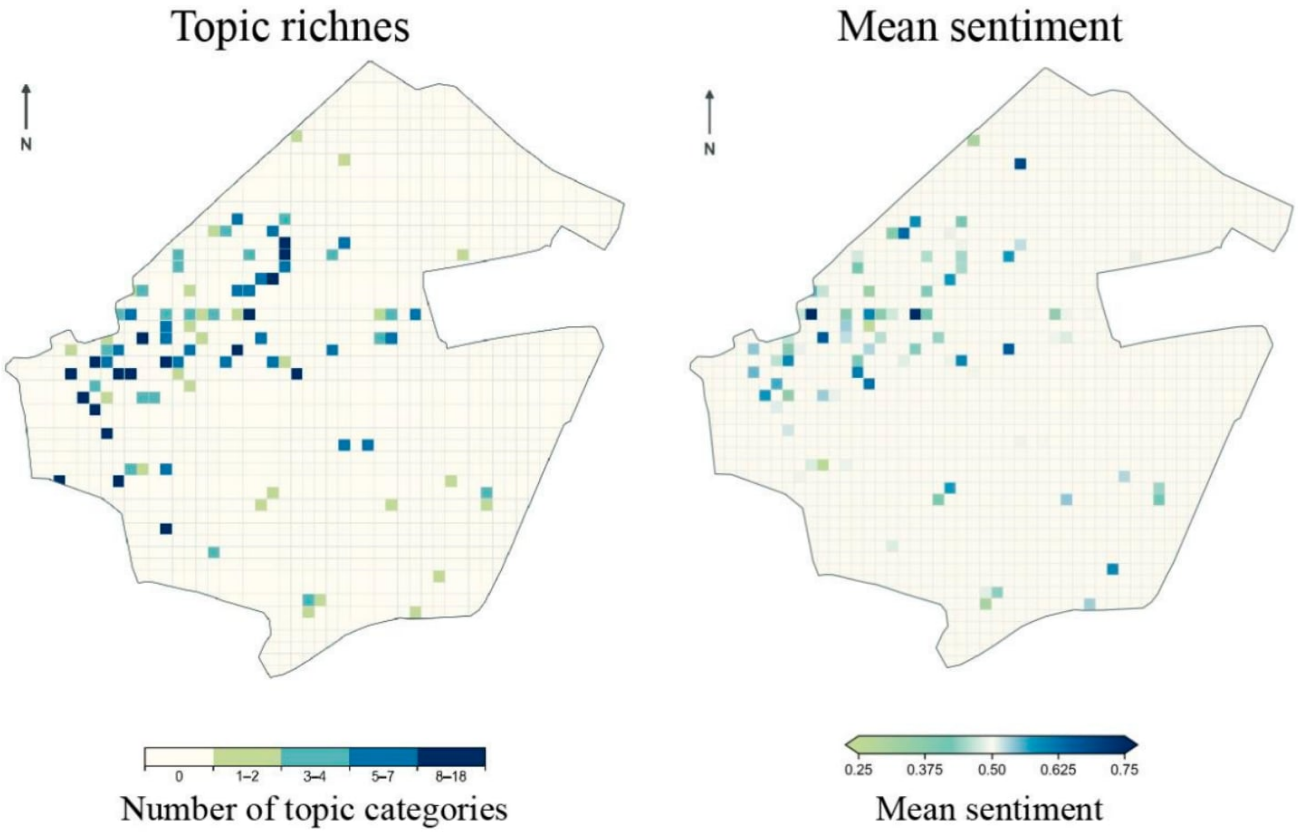

Figure 10. Grid-level distribution of narrative topic richness and mean text sentiment in the place-linked corpus.

Mean sentiment scores also varied across the cells associated with named places (Figure 10). The cell associated with Bayin Stream contained 64 fragments with a mean score of 0.789; the corresponding figures were 118 and 0.661 for Jichang Garden, and 65 and 0.552 for Huishan Temple. By comparison, the cell associated with Xiuzhang Square contained 71 fragments with a mean score of 0.430. The shared cell for the Imperial Stele Pavilion, Tingsong Stone Bed, and Baishi Pavilion contained 76 fragments with a mean score of 0.464. Cells with higher topic richness also differed in their evaluations: mean scores were 0.661 for the cell associated with Jichang Garden, 0.580 for the shared cell of the Second Spring under Heaven and Luzi Shrine, and 0.516 for the Bishan Poetry Society cell. Topic richness describes the variety of narrative topics associated with each cell, while mean sentiment summarizes the evaluations expressed in those texts. The accompanying spatial statistics workbook lists grid cells, associated place names, fragment counts, and observed mean scores; Supplementary Material E, Table E1 summarizes coverage by topic. Topic-specific maps further distinguish the soundscape experiences included in these aggregate evaluations.

4.5.2. Spatial Distribution of Individual Soundscape Topics

Representative topics from the four attention–sentiment groups were selected to compare the spatial distributions of fragment counts and sentiment scores (Figure 11). "Mixed Sounds of Conversations, Broadcasts, and Music" (T25), in HL, had 1166 place-linked fragments across 77 grid cells. These fragments accounted for 49.55% of the spatial corpus and had a mean sentiment score of 0.509. The cells associated with Xishen Square and Xiuzhang Square contained 69 and 60 fragments of this topic, with mean scores of 0.493 and 0.451, respectively. The shared cell for the Imperial Stele Pavilion, Tingsong Stone Bed, and Baishi Pavilion contained 52 fragments, with a mean score of 0.473. In the overall matrix, "Phone Speaker Playback and Disrupted Attention" (T4) belonged to LL. Its 63 place-linked fragments were distributed across 31 cells and had a mean score of 0.270. Each cell with records of this topic contained 1–6 fragments. The shared cell for the Imperial Stele Pavilion, Tingsong Stone Bed, and Baishi Pavilion contained 6 fragments with a mean score of 0.325. The cells associated with Xu Ruzi Shrine and 308 Auditorium contained 5 and 4 fragments, with mean scores of 0.360 and 0.238, respectively. Mixed-sound fragments were distributed widely. Fragments about phone speaker playback had low mean sentiment scores across several buildings and visitor locations. Topics involving natural sounds showed different spatial patterns. "Rustling of Wind-Blown Leaves" (T14), in HH, had 73 place-linked fragments across 21 cells, with a mean score of 0.703. The cells associated with Longguang Pagoda and Bishan Poetry Society contained 12 and 11 fragments, respectively. The Jichang Garden cell contained 10 fragments, with a mean score of 0.746. "Sounds of Flowing Spring Water" (T23), in LH, had 88 place-linked fragments across 6 cells, with a mean score of 0.781. The cells associated with Bayin Stream and Jichang Garden contained 58 and 23 fragments, respectively. Together, these cells accounted for 92.05% of this topic's place-linked fragments; their mean scores were 0.796 and 0.793, respectively. The paired count and sentiment maps showed reports of wind and leaf sounds across multiple places, with spring water reports concentrated at a few garden locations. The Jichang Garden cell also contained fragments on the mixed sound and phone speaker playback topics, indicating contrasting evaluations in texts linked to the same place. Complete paired maps for the remaining 23 topics are provided in Supplementary Material F.

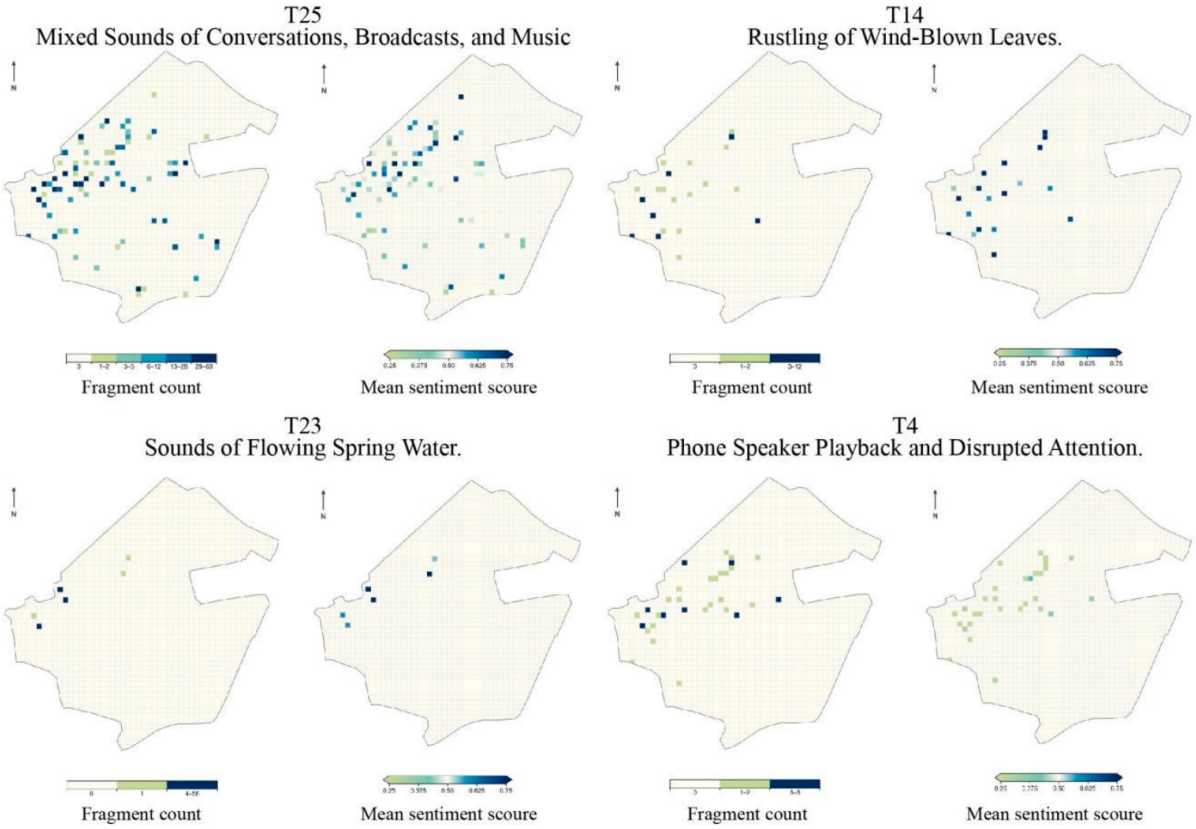

Figure 11. Fragment counts and mean text sentiment for representative topics from the four attention-sentiment groups, aggregated by POI-linked grid cell.

4.6. Spatial Patterns in Grid-Level Topic Composition

4.6.1. Grid Clusters and Associated Places

The 89 occupied grid cells formed six clusters based on the within-cell topic proportions of the 2353 place-linked review fragments (Figure 12). In numerical cluster order, they contained 4, 62, 9, 4, 2, and 8 cells, with 27, 1838, 73, 51, 93, and 271 fragments, respectively. Cluster 2 contained 69.66% of the cells and 78.11% of the place-linked fragments. It was associated with 78 places, including Jichang Garden, Xishen Square, the Second Spring under Heaven, and commercial sites. The mean grid-level share of "Mixed Sounds of Conversations, Broadcasts, and Music" was 57.25%. The other clusters described more localized combinations of review topics. Cluster 1 was associated with the Chaofang at Jichang Garden, Junzi Hall, Xiang Hall, and Chaoran Hall. "Loud Phone Calls Disrupting Architectural Appreciation" and "Conversations and Phone Calls while Walking through Halls" had mean shares of 50.76% and 36.36%, respectively. Cluster 3 was associated with 10 ancestral halls, including Fan Wenzheng Shrine and the Wu Clan Ancestral Hall. Mean shares were 28.78% for "Footsteps and Perceived Quiet during Courtyard Visits," 28.63% for "Conversations and Explanations during Ancestral Hall Visits," and 19.33% for "Conversations in Ancestral Hall Courtyards". These two profiles emphasized narratives about architectural appreciation inside halls and visits to ancestral hall courtyards, respectively. Cluster 4 included cells associated with Chunhua Pavilion, Lingxu Pavilion, Mei Pavilion, Jing Pond, and Dizang Hall; Lingxu Pavilion and Jing Pond shared a cell. "Footsteps during Visits to Pavilions" and "Verbal Greetings and Photography near Pavilions" had mean shares of 50.82% and 20.19%, respectively. Cluster 5 comprised the two cells associated with Yingshan Lake and Baoshan Bridge Sculling Boat Pier. "Distant Conversations along the Waterfront and Distraction" and "Disruptive Loud Conversations at Waterside Resting

Spots" had mean shares of 45.10% and 34.50%, respectively, highlighting reports of conversations during waterfront stays. Cluster 6 was associated with 10 places, including Chunshen Stream, Bishan Poetry Society, Longguang Pagoda, and China Azalea Garden. Mean shares were 33.65% for mixed sounds, 17.96% for wind-blown leaves, 9.82% for wind and rustling leaves during visits to the Azalea Garden, and 5.93% for recurrent bird calls. This cluster included a relatively large share of natural-sound topics alongside mixed sounds. The groups summarize similarities in the composition of mapped reviews, and cells in the same cluster can be associated with different places. Clusters 1 and 4 each covered four cells; the 93 fragments in Cluster 5 covered two cells at two places. These smaller groups summarize reviews linked to the included cells and provide cases for focused field comparison. All topic shares were calculated as unweighted means across grid cells within each cluster. Table 6 summarizes cluster sizes and associated places, and the accompanying statistical workbook lists grid cell assignments. The comparison of candidate numbers of clusters and the PCA projection are presented in Supplementary Material G.

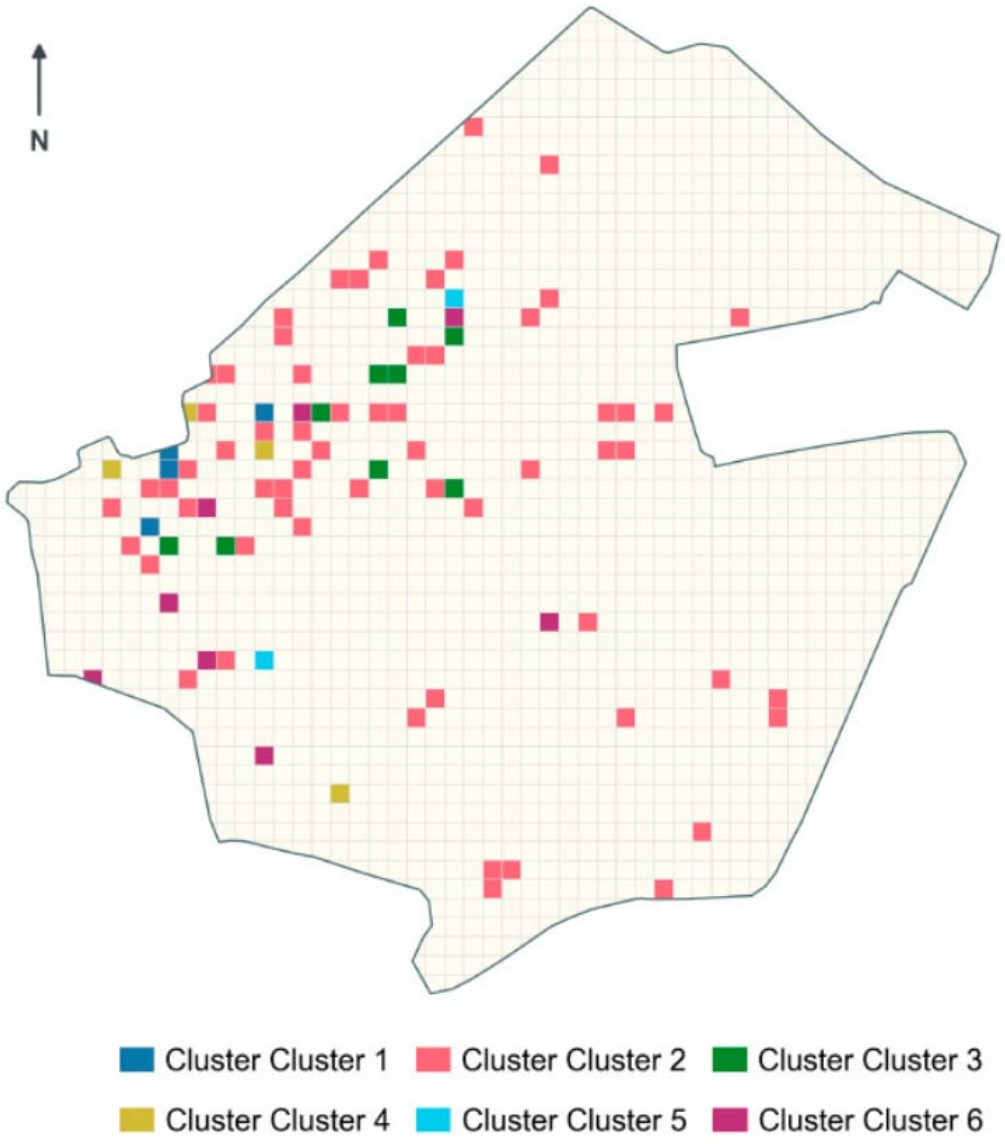

Figure 12. Spatial distribution of grid clusters based on narrative topic proportions across the 89 occupied cells.

Table 6. Grid clusters defined by topic composition: sizes, mean fragment sentiment scores, and linked buildings and public spaces.

Cluster	Grid Cells	Fragments	Mean Sentiment Score	Linked Places
Cluster 1	4	27	0.4333	Chaofang at Jichang Garden, Junzi Hall, Xiang Hall, and Chaoran Hall (4 places)
Cluster 2	62	1838	0.5279	Jichang Garden, Xishen Square, the Second Spring under Heaven, Xiuzhang Square, Hua Xiaozi Shrine, Huishan Temple, Bayin Stream, and commercial sites, among others (78 places in total)
Cluster 3	9	73	0.4641	Fan Wenzheng Shrine, Huang Dounan Shrine, Zhao Zongbai Shrine, Yuci Guangji Shrine, Wu Clan Ancestral Hall, Sima Wengong Shrine, Pan Jingxiao Shrine, Yang Zhenjie Shrine, Wang Wumin Shrine, and Hu Wenzhao Shrine (10 places)
Cluster 4	4	51	0.4988	Chunhua Pavilion, Lingxu Pavilion, Mei Pavilion, Jing Pond, and Dizang Hall (5 places; Lingxu Pavilion and Jing Pond share the same grid cell)
Cluster 5	2	93	0.5011	Yingshan Lake and Baoshan Bridge Sculling Boat Pier (2 places)
Cluster 6	8	271	0.4965	Chunshen Stream, Bishan Poetry Society, Longguang Pagoda, Yugong Valley, Huishan Garden, China Azalea Garden, Qianlu, Huishan Mountain Trail Entrance, Azalea Garden Ticket Office, and Huanggong Stream (10 places)

4.6.2. Place Function Labels and Differences in Emotion and Sentiment

Figure 13a shows variation in topic profiles across the grid clusters. The distribution of POI labels associated with fragments further describes the place contexts of each cluster (Figure 13b). All fragments in Clusters 1 and 4 were linked to the "Tourist Attractions" label. The former reflected architectural appreciation and phone calls inside halls, whereas the latter focused on pavilion visits, verbal greetings, and photography. In Cluster 2, "Tourist Attractions" accounted for 59.09% of fragments, while "Food & Dining," "Leisure & Entertainment," and "Shopping" each accounted for 9.03%. "Entrances & Exits" accounted for 5.88%, and "Culture & Media" for 3.81%, corresponding to the cluster's coverage of heritage buildings, squares, and commercial sites. Cluster 3 comprised fragments linked to "Tourist Attractions" and "Life Services," accounting for 58.90% and 41.10%, respectively. The latter label corresponded to six ancestral halls, including Yuci Guangji Shrine and the Wu Clan Ancestral Hall. In Cluster 5, "Tourist Attractions" and "Transportation Facilities" accounted for 61.29% and 38.71%, corresponding to Yingshan Lake and the boat pier, respectively. Cluster 6 was dominated by "Tourist Attractions" (85.24%); "Hotels" accounted for 11.44%, and all records with this label were linked to Huishan Garden. POI label proportions were calculated at the fragment level.

Model-coded emotion distributions differed among the grid clusters (Figure 13c). Cluster 1, associated with phone calls inside halls, was dominated by Irritation and Dissatisfaction (51.85%), while Calmness and Comfort accounted for 29.63%. The mean sentiment score of its fragments was 0.4333. In Cluster 3, associated with ancestral hall courtyards, Irritation and Dissatisfaction accounted for 34.25% and Calmness and Comfort for 43.84%. The mean sentiment score was 0.4641. In Cluster 4, associated with pavilion visits, Calmness and Comfort predominated (52.94%), while Irritation and Dissatisfaction accounted for 35.29%. In Cluster 5, associated with waterfront conversations, Calmness and Comfort

accounted for 39.78%, Irritation and Dissatisfaction for 30.11%, and Positive Affect and Enjoyment for 13.98%. In Cluster 6, the shares of Calmness and Comfort, Irritation and Dissatisfaction, and Low Mood and Fatigue were 51.66%, 28.41%, and 14.39%, respectively. Cluster 2, which covered the most grid cells, had shares of 42.76% for Calmness and Comfort, 25.79% for Irritation and Dissatisfaction, and 13.55% for Positive Affect and Enjoyment. Its mean sentiment score was 0.5279. Each cluster contained multiple emotions. Soundscape topics and emotion profiles associated with specific places together characterized differences in the reported experiences.

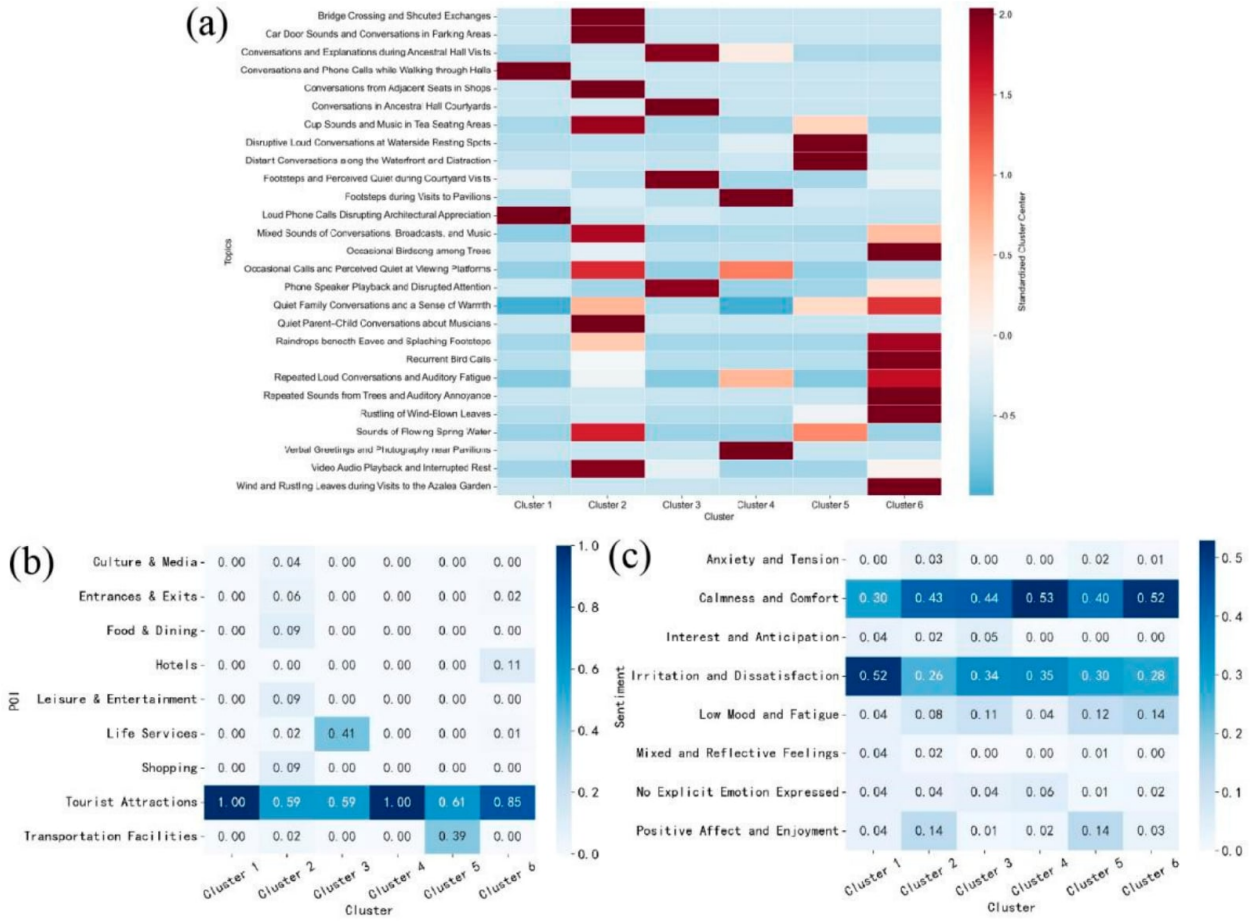

Figure 13. Grid cluster profiles: (**a**) relative differences in topic centers; (**b**) POI-label proportions; (**c**) model-coded emotion proportions. Panels (**b**,**c**) use the fragments within each cluster as their denominator.

5. Discussion

5.1. The Meaning of Sounds in Contexts of Space Use

Soundscape topics in Huishan Ancient Town commonly combined sounds, activities, and places. Within the digital accounts retained for analysis, these combinations show how reviewers situate auditory experiences in the use of space. They identify recurring descriptions and contrasting evaluations that can inform questions for field observation and community interviews. Appreciating architecture inside halls, visiting ancestral halls, taking photographs at pavilions, and resting by the water place different demands on the auditory environment. Cultural storytelling and loud phone calls may consequently receive different evaluations, although both involve human voices. Ren's field study linked perceptions of conversational speech to acoustic comfort and found that the role of multisensory factors varied across spatial types [73]. The review topics distinguished social

interaction, guided explanations, and reported interruptions. These groups show how reviewers describe and evaluate sounds alongside their accounts of using buildings and public spaces. Interviews by Hu et al. in historic districts indicated that social and mechanical sounds can also carry local memories and cultural symbolism [17]. Accounts of parent–child storytelling and quiet conversations document reported activities, affective expressions, and ways of presenting experience to an online audience. Tourism-oriented narratives can therefore be examined as part of the public representation of the district. Their possible cultural significance concerns how these activities relate to local practices, shared memories, and community relationships. Establishing those connections requires historical and community-based evidence in addition to the review texts. In experiments with virtual nature scenes, Smalley et al. found that evoked personal memories were closely associated with emotional responses [74]. This offers an explanation for such differences in terms of listeners' experiences. Soundscape assessment in historic buildings can therefore examine whether sounds support architectural appreciation, understanding, and social interaction, while distinguishing cultural significance from experiential quality.

Temporal variation should also be interpreted alongside the tourism context summarized in Supplementary Material I. Public reports document the 2023 summer night-tour programme, expanded day-to-night activities and a changing visitor mix in 2024, and Spring Festival Gala exposure and new visitor activities in 2025. The visitor-volume figures retain their reported time windows and counting scopes. Such operational and visitor changes may alter the sounds and activities encountered, while platform participation and posting preferences may affect which experiences enter the corpus. The temporal results therefore describe shifts in reported soundscape content. They do not isolate changes in perception from visitor behaviour and composition, platform use, or scenic-area operations.

5.2. Sentiment Patterns in Soundscape Narratives

The topic–sentiment patterns describe how sounds, activities, and evaluations occur together in the review narratives. In the phone speaker playback topic (T4), for example, the retained text included descriptions of diverted attention and interruptions (Supplementary Material C.9). Such wording can contribute to semantic grouping and also inform sentiment coding. Both models used the same general Chinese prompt, and the reported agreement metrics summarize performance across the evaluated fragments. Dialect-specific classification performance remains unquantified. When an evaluation depends on regional wording or indirect cues, the original wording and the reported activity help contextualize the assigned label (Supplementary Material C.5). Stobbe et al. observed reduced negative affect following natural sound exposure in an experiment [75]. Yang et al. found that soundscape appropriateness mediated the relationship between sound sources and acoustic comfort, with spatial variation in these relationships [50]. The difference between the mean scores of spring-water and playback narratives linked to the same garden location suggests a question for field observation about how listening experiences vary with activity and setting. T25 combined broad representation in the corpus (45.57%; n = 1960) with a near-neutral mean sentiment score (0.506). Its HL position combines above-cutoff relative attention with below-cutoff mean sentiment. T4 was less frequently mentioned (2.53%; n = 109) but had a lower mean text sentiment score (0.275). These contrasting patterns support different lines of investigation: assessment of overlapping activities in mixed-sound settings and targeted checking of reported playback disturbances. Management responses depend on the affected activity, on-site confirmation of interference, and the feasibility of a targeted measure. Conservation decisions require historical and community evidence about sound-related practices, including sounds seldom represented in reviews.

5.3. Topic Profiles in Contexts of Space Use

The six clusters group 25 m grid cells with similar topic proportions in their place-linked reviews. The resulting narrative topic profiles summarize what reviewers reported in connection with halls, ancestral hall courtyards, pavilions, and waterfront locations. The dominant cluster spans many places, while the groups with two or four cells provide localized profiles. The scale comparison also showed changes in some groups, including the cluster associated with pavilion visits (Supplementary Material H). These profiles can guide the selection of places and reported activities for field comparison. In a field study of a historic district in Fuzhou, Wu et al. found that the spatial distribution of sound source harmony ratings was associated with place functions [76]. In the present corpus, the grid profiles distinguished narratives about viewing, guided explanations, lingering, and social interaction within the same "Tourist Attractions" label. Auditory experience in architectural spaces may also interact with spatial form. In street canyon experiments, Yilmaz et al. found that the effects of sounds on perceived enclosure and spaciousness varied with street width [45], demonstrating audiovisual interaction in architectural experience. POI-linked patterns provide cases for examining how activities generate sound and how architectural configuration shapes its transmission and interpretation. Future work could combine social media mapping with on-site acoustic soundwalks that compare listening positions inside halls, at doorways, in adjoining courtyards, and along waterfronts. Recording the sound-generating activity, source and listener positions, openings, intervening boundaries, and acoustic measurements would allow transmission, filtering, and attenuation to be examined along specific paths. Participants' evaluations and interviews about the use and history of these spaces could help explain how sounds acquire meaning in those settings. Hu et al. proposed different sound combinations for commercial, cultural, and residential spaces, with authenticity as the evaluation objective [18]. The grid clusters here summarize reported activities, with emotion and sentiment summarized from the review fragments after clustering. Cluster 6, with greater representation of natural-sound topics, did not have the highest overall sentiment score, suggesting that spatial function labels may not capture the full range of sound events. Similarly, Lee and Kim used social media texts to examine spatial patterns in reported soundscapes in relation to activities and meanings [77]. Together, these studies support examining the fit between sounds and activities in named buildings and public spaces and using specific sound events to identify divergent experiences within the same type of place.

5.4. Implications for Soundscape Management in the Buildings and Public Spaces of Huishan Ancient Town

Soundscape management in Huishan Ancient Town should begin with an assessment of the historical and community significance of sound-related practices and listening opportunities. Managers, residents, practitioners, and operators could use historical records and local accounts to decide which practices to sustain, which conditions of sound production and listening can be adjusted, and which new relationships between sounds and activities to facilitate [17]. Negative sentiment alone should not determine whether a sound is reduced. Conversation, for example, can interfere with quiet viewing in one setting and support everyday social exchange in another. Decisions should address its role, timing, and spatial setting.

For Huishan Ancient Town, management recommendations are organized into two implementation tiers based on the reported activity interference and the coordination required. Matrix group membership does not determine these tiers. The first tier prioritizes checking specific disturbance reports and applying targeted measures where interference is confirmed, particularly for phone and video playback during viewing and rest. Site staff could respond to visitor feedback, check the reported disturbance, and use brief, softly spoken reminders

to request headphones or an end to speaker playback. In halls and ancestral halls, guides and site staff could coordinate group sizes and gathering points so that explanations remain audible while other users can appreciate the architecture and move through the space. A simple record of the location, time, affected activity, and response would help staff identify recurring problems and assess whether further measures are needed.

The second tier addresses longer-term coordination of spatial arrangements and activities, particularly in squares and commercial spaces where broadcasts, music, and human voices overlap. Scenic-area management could work with shop and event operators to adjust amplification, equipment orientation, operating times and duration, and activity locations following on-site assessment. Group assembly, photography, and resting areas at pavilions and waterfront spaces could also be adjusted where activities interfere with one another. POI-linked review patterns can guide site selection for these assessments. For example, the concentration of spring-water reports at Bayin Stream and Jichang Garden identifies two sites for examining the audibility of water sounds and conditions for quiet rest. In a field intervention at traffic-exposed residential sites, Lam et al. found that adding adaptively selected natural sounds improved pleasantness and was associated with greater restorativeness [78]. Individual measures need to be evaluated against background sound and space use. Where historical and community evidence identifies a valued sound-related practice, adjustments should sustain its continuity and opportunities to be heard. Planning could also connect spaces for cultural storytelling, everyday interaction, and quiet listening through routes and stopping points that offer complementary experiences. Implementation could begin with pilot interventions at representative locations. Before-and-after comparisons could combine disturbance feedback and soundscape appropriateness ratings with residents' and practitioners' assessments of the continuity and audibility of valued practices. Experiences of viewing and rest would also inform the wider adoption of measures in both tiers.

6. Conclusions

This study integrated topic, sentiment, and spatial analyses of social media reviews to characterize reported soundscape experiences in Huishan Ancient Town. The selected model identified 27 narrative topics in the analysed corpus of 4301 fragments. Mixed sounds of conversations, broadcasts, and music formed the dominant topic, accounting for 1960 fragments (45.57%). Within the corpus, mean sentiment was higher in the narrative groups about flowing spring water and rustling leaves than in the group about phone speaker playback. These comparisons describe the wording and evaluations in the fragments that make up each group. Findings for less-represented topics, including occasional birdsong (n = 42) and quiet parent–child conversations (n = 46), and comparisons involving the earliest period (2022 and earlier; n = 43) remain sample-limited descriptions. Reviews about similar sound sources and reviews linked to the same place also contained differing evaluations. The attention–sentiment matrix distinguished the frequently mentioned mixed-sound topic from less frequently mentioned topics that included reports of activity interference. Place-level sentiment rankings were broadly consistent across the tested grid scales. At 25 m, the 89 occupied grid cells formed six clusters based on the topic proportions of their mapped review fragments. The largest cluster contained 62 cells. Three clusters of two or four cells summarize reviews linked to the included hall, pavilion, and waterfront locations. Existing soundscape frameworks recognize the importance of context. Assessments based on broad sound-source categories and place types provide limited detail on activity-related differences within the same setting. The methodological contribution is a workflow that integrates topic, sentiment, temporal, and spatial analyses to compare what reviewers report and how they evaluate sounds in relation to activities

and places. The resulting narrative patterns help select locations, activities, and questions for field observation and community interviews.

The principal limitations concern cultural interpretation, corpus coverage, sentiment validation, and spatial precision. Interpreting the cultural significance of the reported sound–activity relationships requires historical evidence and community accounts of associated practices and memories. Differences in sample size and date coverage across time periods limit continuous comparisons of long-term change. Platform participation, posting, and the study's screening rules select which experiences enter the corpus. The minimum-token rule may underrepresent experiences expressed in a few words, including noise complaints and descriptions of quietness (Supplementary Material B). The texts capture explicitly reported experiences and provide limited information on unmentioned sounds or activities; they do not represent the listening experiences of all residents or visitors. The three-class polarity comparison used final manually reviewed reference labels, but separate annotator-level labels were unavailable in the retained analysis file, preventing calculation of inter-annotator agreement. In the respective comparisons with human-reviewed labels, macro-F1 was 0.8056 for three-class polarity and 0.4763 for the eight emotion categories. The frequent assignment of reference No Explicit Emotion Expressed fragments to Calmness and Comfort and the low F1 for Mixed and Reflective Feelings limit the precision of the emotion summaries. The eight-category distributions describe model assignments under the shared mapping. The boundaries of the categories require separate conceptual validation. The continuous sentiment scores also require calibration against an independent reference. Place mentions and POI reference points enabled spatial aggregation but did not directly record actual listening positions. The sensitivity analysis indicated broadly similar site-level sentiment rankings but scale-dependent cluster assignments; the six clusters summarize the mapped review corpus at the 25 m scale, and the fixed-origin comparison does not assess uncertainty from alternative grid origins or imprecise place coordinates. POI labels also provide an incomplete account of building form and the details of space use.

Future research could connect digital monitoring, situated cultural interpretation, and intervention evaluation. Repeated collection of reviews across seasons and tourism periods could support an observatory of mentioned sounds, reported conflicts, linguistic patterns, and locations for follow-up. Soundwalks, contextualized recordings and acoustic measurements, activity observations, interviews with residents and visitors, and historical reconstruction of sound-related practices could relate these accounts to listening experiences and local meanings. This combined evidence could inform criteria for conservation, compatible uses, spatial arrangements, and listening conditions. Pilot interventions could then compare soundscape appropriateness, experiences of architectural appreciation and rest, and the continuity of valued practices before and after changes. Digital analysis would help identify questions and places for investigation, with historical and community interpretation guiding judgments about cultural significance.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/buildings16234751/s1, .Supplementary Material A: Soundscape Text Screening and Extraction Rules; Supplementary Material B: Text Preprocessing and Record Structure; Supplementary Material C: Analysis Parameters, Model Comparisons, and Time-Period Definitions; Supplementary Material D: Topic Hierarchy; Supplementary Material E: Spatial Statistics; Supplementary Material F: Grid Distributions of the Remaining Topics; Supplementary Material G: Grid Clustering Diagnostics; Supplementary Material H: Sensitivity to Spatial Grid Size; Supplementary Material I: Tourism Context for the Temporal Comparisons [57,79–85].

Author Contributions: Conceptualization, B.S.; methodology, B.S.; software, B.S. and J.G.; validation, B.S.; formal analysis, B.S.; investigation, B.S.; resources, J.G.; data curation, B.S. and J.G.; writing—original draft preparation, B.S.; writing—review and editing, J.G.; visualization, J.G.; supervision,

J.G.; project administration, J.G.; funding acquisition, J.G.; All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: The original contributions presented in this study are included in the Supplementary Material. Further inquiries can be directed to the corresponding author.

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

References

1. UNESCO. *Recommendation on the Historic Urban Landscape, Including a Glossary of Definitions*; UNESCO: Paris, France, 2011. Available online: https://www.unesco.org/en/legal-affairs/recommendation-historic-urban-landscape-including-glossary-definitions (accessed on 7 October 2026).
2. Song, Y.; Han, C.; Zhao, Y. A Study on Tourist Satisfaction Based on the Conservation and Reuse of Alleyway Spaces in Urban Historic Neighborhoods. *Buildings* **2024**, *14*, 1324. [CrossRef]
3. Navarrete-Hernandez, P.; Urrutia, J.P.; Mellouki, K. Preserving enough? A randomised controlled trial approach to determine relevant urban planning regulations for small touristic towns—A case of Chiloé, Chile. *Habitat Int.* **2024**, *146*, 103033. [CrossRef]
4. United Nations General Assembly. *Transforming Our World: The 2030 Agenda for Sustainable Development*; United Nations: New York, NY, USA, 2015. Available online: https://documents.un.org/doc/undoc/gen/n15/291/89/pdf/n1529189.pdf (accessed on 7 October 2026).
5. Tüzel, F.; Pirselimoğlu Batman, Z. User satisfaction-based landscape planning in historical city centers: Bursa-Turkey. *Humanit. Soc. Sci. Commun.* **2025**, *12*, 1589. [CrossRef]
6. Wang, Z.; Sheng, K.; Li, D.; Man, T.; Zhang, X.; He, Y.; Zhou, Q.; He, G. Cultural topography of publicness: Assessment of the publicness of public spaces in traditional settlements. *PLoS ONE* **2025**, *20*, e0332755. [CrossRef] [PubMed]
7. ICOMOS. ICOMOS International Charter for Cultural Heritage Tourism: Reinforcing Cultural Heritage Protection and Community Resilience Through Responsible and Sustainable Tourism Management. 2022. Available online: https://admin.icomos.org/wp-content/uploads/2023/01/International-Cultural-Heritage-Tourism-Charter_EN_IT.pdf (accessed on 7 October 2026).
8. Wang, Z.; Zhou, Q.; Man, T.; He, L.; He, Y.; Qian, Y. Delineating Landscape Features Perception in Tourism-Based Traditional Villages: A Case Study of Xijiang Thousand Households Miao Village, Guizhou. *Sustainability* **2024**, *16*, 5287. [CrossRef]
9. *ISO 12913-1:2014*; Acoustics—Soundscape—Part 1: Definition and Conceptual Framework. International Organization for Standardization: Geneva, Switzerland, 2014.
10. Zhang, R.; Ma, H.; Wang, C.; Zhang, Y.; Kang, J. Soundscape and its context: A framework based on a systematic review. *J. Acoust. Soc. Am.* **2025**, *157*, 4417–4436. [CrossRef] [PubMed]
11. Ye, J.; Li, S.; Chen, Y.; Ma, Y.; Chen, L.; He, T.; Zheng, Y. A Study of the Effects of Historical Block Context on Soundscape Perception. *Buildings* **2024**, *14*, 621. [CrossRef]
12. Xie, H.; He, Y.; Wu, X.; Lu, Y. Interplay between auditory and visual environments in historic districts: A big data approach based on social media. *Environ. Plan. B Urban Anal. City Sci.* **2022**, *49*, 1245–1265. [CrossRef]
13. Zhao, W.; Rui, Q.; Zhu, X.; Xu, H. Effect of Soundscape on Place Attachment for Historical Blocks: A Case Study of Harbin, China. *Buildings* **2023**, *13*, 607. [CrossRef]
14. Chen, Z.; Jin, M.; Zuo, Y.; Ding, P.; Shi, X. The Effect of Soundscape on Sense of Place for Residential Historical and Cultural Areas: A Case Study of Taiyuan, China. *Buildings* **2024**, *14*, 1394. [CrossRef]
15. Yang, L.; Liu, J. The moderating effect of visual landscape and soundscape on place attachment in world cultural heritage: A case study in Kulangsu, China. *npj Herit. Sci.* **2025**, *13*, 94. [CrossRef]
16. Flores-Villa, L.; Oberman, T.; Guattari, C.; Asdrubali, F.; Frascarolo, M.; Puglisi, G.E.; Astolfi, A.; Aletta, F. Exploring relationships between soundscape and lightscape perception: A case study around the Colosseum and Fori Imperiali in Rome. *Light. Res. Technol.* **2023**, *55*, 603–620. [CrossRef]
17. Hu, Y.; Meng, Q.; Li, M.; Yang, D. Enhancing authenticity in historic districts via soundscape design. *Herit. Sci.* **2024**, *12*, 396. [CrossRef]
18. Hu, Y.; Meng, Q.; Li, M.; Yang, D.; Yan, W.; Kang, J. Evaluation of soundscape authenticity in historic districts: Weighted formulae and spatial typology-based sound selection. *Appl. Acoust.* **2026**, *249*, 111301. [CrossRef]
19. Aletta, F.; Torresin, S. Adoption of ISO/TS 12913-2:2018 Protocols for Data Collection From Individuals in Soundscape Studies: An Overview of the Literature. *Curr. Pollut. Rep.* **2023**, *9*, 710–723. [CrossRef]
20. Zhao, T.; Liang, X.; Tu, W.; Huang, Z.; Biljecki, F. Sensing urban soundscapes from street view imagery. *Comput. Environ. Urban Syst.* **2023**, *99*, 101915. [CrossRef]

21. Wang, H.; Yu, Z.; Liu, X. Deciphering Urban Soundscapes: A study of sensory experiences at Hong Kong Victoria harbour waterfronts using social media. *Comput. Environ. Urban Syst.* **2025**, *120*, 102307. [CrossRef]
22. Li, X. Urban Renewal Focuses on 23 Key Tasks. *People's Daily*, 9 June 2026; pp. 3–4. Available online: https://paper.people.com.cn/rmrb/pc/content/202606/09/content_30162035.html (accessed on 7 October 2026). (In Chinese)
23. Pan, Y.; Nik Hashim, N.H.; Goh, H.C. Public perception of cultural ecosystem services in historic districts based on biterm topic model. *Sci. Rep.* **2024**, *14*, 11717. [CrossRef] [PubMed]
24. Zhu, J.; Sho, K.; Park, Y.; Yuan, Y.; Zhao, S. Exploring the interactive effects of internal and external factors on residents' daily lives: Perspective of commercial gentrification in historic districts. *Cities* **2026**, *170*, 106679. [CrossRef]
25. Xu, Y.; Lu, L. Probing the long-term evolution of traditional village tourism destinations from a glocalisation perspective: A case study of Wuzhen in Zhejiang province, China. *Habitat Int.* **2024**, *148*, 103073. [CrossRef]
26. Xie, J.; Kang, J.; Zhao, Z.; Hou, Y. Effects of audiovisual information and its perception on the sense of culture in everyday heritage. *Appl. Acoust.* **2025**, *239*, 110795. [CrossRef]
27. Yang, L.; Liu, J.; Albert, C.; Guo, X. Exploring the effects of soundscape perception on place attachment: A comparative study of residents and tourists. *Appl. Acoust.* **2024**, *222*, 110048. [CrossRef]
28. Zhu, X.; Oberman, T.; Aletta, F. Defining acoustical heritage: A qualitative approach based on expert interviews. *Appl. Acoust.* **2024**, *216*, 109754. [CrossRef]
29. Zhang, D.; Shan, Y.; Chen, X.; Wang, Z. Soundscape in religious historical buildings: A review. *Herit. Sci.* **2024**, *12*, 45. [CrossRef]
30. Jiang, L.; Bristow, A.; Kang, J.; Aletta, F.; Thomas, R.; Notley, H.; Thomas, A.; Nellthorp, J. Ten questions concerning soundscape valuation. *Build. Environ.* **2022**, *219*, 109231. [CrossRef]
31. Mao, L.; Zhang, X.; Ma, J.; Jia, Y. A comparative study on the audio-visual evaluation of the grand Song of the Dong soundscape. *Herit. Sci.* **2023**, *11*, 32. [CrossRef]
32. Zhu, X.; Liu, F.; Kang, J.; Liu, S.; Wu, Y.; Zhao, W.; Xue, Y.; Wang, C. Assessing the audio-visual environment at the Great Wall of China: Impacts on visitor satisfaction and restoration. *Appl. Acoust.* **2024**, *218*, 109905. [CrossRef]
33. Li, Y.; Li, X.; Gao, M. Soundscapes and Emotional Experiences in World Heritage Temples: Implications for Religious Architectural Design. *Buildings* **2025**, *15*, 2681. [CrossRef]
34. Jaff, A.A.M.; Erçin, Ç.; Onur, Z. Assessing the Soundscape Characteristics of Historical Urban Environments: An Analysis of the Historical Erbil Citadel and Its Environments. *Buildings* **2023**, *13*, 3091. [CrossRef]
35. Espinoza, F.; Bravo-Moncayo, L.; Garzón, L.; Poblete, V.; Arenas, J.P. Analysis of Acoustic and Perceptual Variables in Three Heritage Churches in Quito Using Structural Equation Modeling. *Buildings* **2025**, *15*, 2639. [CrossRef]
36. Lin, J.; Zhang, M.; Wang, Y.; Hong, X.-C.; Liu, J. Factors Influencing Natural and Cultural Soundscape Interactions on Perceptual Experiences in Forested–Historical Interface Areas. *Buildings* **2025**, *15*, 4103. [CrossRef]
37. de Mesquita Peixoto, N.G.; Ferreira, L.R.; Afonso, R.C.; Comissoli, V.G.; Monteiro, L.M. A methodological approach for soundscape attributes translation using Twitter data: A case study in the Portuguese language. *Appl. Acoust.* **2024**, *218*, 109880. [CrossRef]
38. Zhang, J.; Luo, Y.; Cao, H.; Su, S. Scrutinizing the cultural ecosystem services of Chinese Classical Gardens: A novel deep learning approach based on online reviews from a multisensory perspective. *Appl. Geogr.* **2024**, *172*, 103404. [CrossRef]
39. Luo, Y.; Zhang, J.; Su, S. How multisensory experiences shape landscape perception? A novel large language model based approach to online reviews mining. *Appl. Geogr.* **2026**, *192*, 104043. [CrossRef]
40. Sun, S.; Sun, H.; Xu, H.; Li, H.; Wang, S. Unlocking the power of multimodal online reviews: A multisensory perspective. *Tour. Manag.* **2025**, *111*, 105206. [CrossRef]
41. Kim, J.M.; Park, K.K.-c.; Marjerison, R.K. Usefulness of Online Reviews of Sensory Experiences: Pre- vs. Post-Pandemic. *J. Theor. Appl. Electron. Commer. Res.* **2024**, *19*, 1471–1492. [CrossRef]
42. Lindborg, P.; Lam, L.H.; Kam, Y.C.; Yue, R. Sensory Heritage Is Vital for Sustainable Cities: A Case Study of Soundscape and Smellscape at Wong Tai Sin. *Sustainability* **2025**, *17*, 7564. [CrossRef]
43. Tao, Y.; Xu, Y.; Li, C.; Cheng, Y.; Xu, X. Mapping themes of heritage tourism with BERTopic—Evidence from Nánxún Ancient Town. *J. Asian Archit. Build. Eng.* **2026**, 1–15. [CrossRef]
44. Foroughi, M.; de Andrade, B.; Pereira Roders, A. Capturing public voices: The role of social media in heritage management. *Habitat Int.* **2023**, *142*, 102934. [CrossRef]
45. Yilmaz, N.G.; Lee, P.-J.; Imran, M.; Jeong, J.-H. Role of sounds in perception of enclosure in urban street canyons. *Sustain. Cities Soc.* **2023**, *90*, 104394. [CrossRef]
46. Lu, Y.; Hasegawa, Y.; Tan, J.K.A.; Lau, S.-K. Effect of audio-visual interaction on soundscape in the urban residential context: A virtual reality experiment. *Appl. Acoust.* **2022**, *192*, 108717. [CrossRef]
47. Zhou, Z.; Ye, X.; Chen, J.; Fan, X.; Kang, J. Effect of visual landscape factors on soundscape evaluation in old residential areas. *Appl. Acoust.* **2024**, *215*, 109708. [CrossRef]
48. Chen, J.; Zhang, Q.; Li, X.; Weng, J.; Cao, Y.; Ye, J. Effects of Acoustic and Visual Environmental Factors on Perceived Street Vitality in Historic Districts: A Case Study of Shangxiahang, Fuzhou. *Buildings* **2026**, *16*, 1712. [CrossRef]

49. Hosseini, A.; Kowkabi, L. Measuring the Soundscape Quality in Urban Spaces: A Case Study of Historic Urban Area. *Sustainability* **2023**, *15*, 4255. [CrossRef]
50. Yang, X.; Zhang, G.; Lu, X.; Zhang, Y.; Kang, J. Contribution of soundscape appropriateness to soundscape quality assessment in space: A mediating variable affecting acoustic comfort. *J. Environ. Manag.* **2024**, *372*, 123321. [CrossRef] [PubMed]
51. Yan, W.; Meng, Q.; Yin, Y.; Yang, D.; Li, M.; Kang, J. Predictive models of tranquility in urban public open spaces based on audiovisual indicators analysis. *Build. Environ.* **2025**, *267*, 112260. [CrossRef]
52. Alizadeh, Z.; Chamani, A.; Lorestani, B.; Asgarian, A. Disentangling the drivers of urban noise: Why mitigation in compact historic cities should target both mobility and land-use. *Sustain. Cities Soc.* **2026**, *144*, 107443. [CrossRef]
53. Fraisse, V.; Tarlao, C.; Guastavino, C. Shaping city soundscapes: In situ comparison of four sound installations in an urban public space. *Landsc. Urban Plan.* **2024**, *251*, 105173. [CrossRef]
54. Sustaining the Vitality of a Millennium-Old Ancient Town. *Wuxi Daily*, 9 September 2025. Available online: https://www.wuxi.gov.cn/doc/2025/09/09/4644578.shtml (accessed on 8 October 2026). (In Chinese)
55. State Council of the People's Republic of China. Notice of the State Council on Approving and Announcing the Sixth Batch of Major Historical and Cultural Sites Protected at the National Level (Guo Fa [2006] No. 19). 2006. Available online: https://zwgk.mct.gov.cn/zfxxgkml/qt/202012/t20201206_918486.html (accessed on 8 October 2026). (In Chinese)
56. Ministry of Culture and Tourism of the People's Republic of China. Announcement on Designating 22 Tourist Attractions as National 5A-Level Tourist Attractions. 2020. Available online: https://zwgk.mct.gov.cn/zfxxgkml/zykf/202012/t20201213_919421.html (accessed on 8 October 2026). (In Chinese)
57. State-Owned Assets Supervision and Administration Commission of the Wuxi Municipal People's Government. Huishan Ancient Town Promotes the Deep Integration of Intangible Cultural Heritage and Cultural Tourism. 2025. Available online: https://gzw.wuxi.gov.cn/doc/2025/09/29/4659426.shtml (accessed on 8 October 2026). (In Chinese)
58. Liu, P.; Yuan, W.; Fu, J.; Jiang, Z.; Hayashi, H.; Neubig, G. Pre-train, prompt, and predict: A systematic survey of prompting methods in natural language processing. *ACM Comput. Surv.* **2023**, *55*, 195. [CrossRef]
59. Wang, Z.; Zhou, Q.; Tai, Y.; Zhu, R.; Wei, K. Public Value Perception and Conservation Strategies for Urban Industrial Heritage: Evidence from UGC. *Buildings* **2026**, *16*, 2391. [CrossRef]
60. Viégas, F.B.; Wattenberg, M.; Feinberg, J. Participatory visualization with Wordle. *IEEE Trans. Vis. Comput. Graph.* **2009**, *15*, 1137–1144. [CrossRef] [PubMed]
61. Grootendorst, M. BERTopic: Neural topic modeling with a class-based TF-IDF procedure. *arXiv* **2022**, arXiv:2203.05794. [CrossRef]
62. Reimers, N.; Gurevych, I. Sentence-BERT: Sentence embeddings using Siamese BERT-networks. In *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP), Hong Kong, China, 3–7 November 2019*; ACL: Stroudsburg, PA, USA, 2019; pp. 3982–3992. [CrossRef]
63. Sentence Transformers. Paraphrase-Multilingual-MiniLM-L12-v2. Available online: https://huggingface.co/sentence-transformers/paraphrase-multilingual-MiniLM-L12-v2 (accessed on 8 October 2026).
64. McInnes, L.; Healy, J.; Melville, J. UMAP: Uniform manifold approximation and projection for dimension reduction. *arXiv* **2018**, arXiv:1802.03426. [CrossRef]
65. McInnes, L.; Healy, J.; Astels, S. hdbscan: Hierarchical density based clustering. *J. Open Source Softw.* **2017**, *2*, 205. [CrossRef]
66. Grootendorst, M. Outlier Reduction. BERTopic Documentation. Available online: https://maartengr.github.io/BERTopic/getting_started/outlier_reduction/outlier_reduction.html (accessed on 8 October 2026).
67. Gilardi, F.; Alizadeh, M.; Kubli, M. ChatGPT outperforms crowd workers for text-annotation tasks. *Proc. Natl. Acad. Sci. USA* **2023**, *120*, e2305016120. [CrossRef] [PubMed]
68. Martilla, J.A.; James, J.C. Importance-performance analysis. *J. Mark.* **1977**, *41*, 77–79. [CrossRef]
69. scikit-learn Developers. StandardScaler. Available online: https://scikit-learn.org/stable/modules/generated/sklearn.preprocessing.StandardScaler.html (accessed on 8 October 2026).
70. Lloyd, S.P. Least squares quantization in PCM. *IEEE Trans. Inf. Theory* **1982**, *28*, 129–137. [CrossRef]
71. Rousseeuw, P.J. Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. *J. Comput. Appl. Math.* **1987**, *20*, 53–65. [CrossRef]
72. Yuan, J.; Zheng, Y.; Xie, X. Discovering regions of different functions in a city using human mobility and POIs. In *Proceedings of the 18th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Beijing, China, 12–16 August 2012*; ACM: New York, NY, USA, 2012; pp. 186–194. [CrossRef]
73. Ren, X. Combined effects of dominant sounds, conversational speech and multisensory perception on visitors' acoustic comfort in urban open spaces. *Landsc. Urban Plan.* **2023**, *232*, 104674. [CrossRef]
74. Smalley, A.J.; White, M.P.; Sandiford, R.; Desai, N.; Watson, C.; Smalley, N.; Tuppen, J.; Sakka, L.; Fleming, L.E. Soundscapes, music, and memories: Exploring the factors that influence emotional responses to virtual nature content. *J. Environ. Psychol.* **2023**, *89*, 102060. [CrossRef]

75. Stobbe, E.; Lorenz, R.C.; Kühn, S. On how natural and urban soundscapes alter brain activity during cognitive performance. *J. Environ. Psychol.* **2023**, *91*, 102141. [CrossRef]
76. Wu, L.; Zhang, Q.; Yan, Y.; Lan, T.; Hu, Y.; Zhang, Y.; He, T.; Ye, J. A Study on Spatiotemporal Dynamics and Spatial Dependence of Sound Source Perception in Fuzhou Historical and Cultural Districts. *Buildings* **2024**, *14*, 1753. [CrossRef]
77. Lee, H.; Kim, Y. Mapping urban soundscape patterns using user-generated content: A place model approach to acoustic environment perception. *PLoS ONE* **2026**, *21*, e0343393. [CrossRef] [PubMed]
78. Lam, B.; Ong, Z.-T.; Ooi, K.; Ong, W.-H.; Wong, T.; Watcharasupat, K.N.; Boey, V.; Lee, I.; Hong, J.Y.; Kang, J.; et al. Automating urban soundscape enhancements with AI: In-situ assessment of quality and restorativeness in traffic-exposed residential areas. *Build. Environ.* **2024**, *266*, 112106. [CrossRef]
79. Jiangnan Evening News, Reproduced by China Jiangsu Net. Wuxi Ranks Among the Ten Most Popular Travel Destinations. 8 October 2022. Available online: https://jsnews.jschina.com.cn/wx/a/202210/t20221008_3089777.shtml (accessed on 4 November 2026). (In Chinese)
80. Wuxi Municipal People's Government. Liangxi Comes Alive During the Spring Festival Holiday. 28 January 2023. Available online: https://www.wuxi.gov.cn/doc/2023/01/28/3880354.shtml (accessed on 4 November 2026). (In Chinese)
81. Wuxi Municipal People's Government. New Night-Time Offerings Attract Younger Visitors and Energize Cultural Tourism. 4 September 2023. Available online: https://www.wuxi.gov.cn/doc/2023/09/04/4053697.shtml (accessed on 4 November 2026). (In Chinese)
82. Wuxi Municipal People's Government. Huishan Ancient Town Adopts an All-Day Activity Model. 5 August 2024. Available online: https://www.wuxi.gov.cn/doc/2024/08/05/4365956.shtml (accessed on 4 November 2026). (In Chinese)
83. Wuxi Daily, Published on the Wuxi Municipal People's Government Website. Huishan Ancient Town: What Comes Next? 20 May 2025. Available online: https://www.wuxi.gov.cn/doc/2025/05/20/4576789.shtml (accessed on 4 November 2026). (In Chinese)
84. Wuxi Daily, Published on the Wuxi Municipal Publicity Department Website. Strong Supply and Demand in the Spring Festival Cultural Tourism Market. 24 February 2026. Available online: https://xcb.wuxi.gov.cn/doc/2026/02/24/4735758.shtml (accessed on 4 November 2026). (In Chinese)
85. Wuxi Daily, Published on the Wuxi Municipal Publicity Department website. Immersive Performances Throughout the Ancient Town Expand Cultural Tourism. 15 May 2026. Available online: https://xcb.wuxi.gov.cn/doc/2026/05/15/4776003.shtml (accessed on 4 November 2026). (In Chinese)

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.