

# Research on the Construction of an Electronic Speech Dictionary of Hangzhou Dialect

Yao Tan, Ziyu Li, Jianlin Wang

*Communications University of Zhejiang, Hangzhou, Zhejiang, China*

**Abstract:** With the acceleration of urbanization and the popularization of Mandarin, Hangzhou dialect is facing the severe challenge of rapid decline and extinction. To explore new paths for the digital protection and dynamic inheritance of dialects, this study proposes constructing an electronic phonetic dictionary of Hangzhou dialect integrating "query, learning, and cultural inheritance" functions. This paper systematically elaborates on the core positioning, functional design, and auxiliary functions of the dictionary that incorporate cultural annotations and personalized learning, and explores the principles of lexical selection and compilation for the corpus. At the technical implementation level, a complete roadmap is planned, ranging from the selection of representative speakers, standardized phonetic data collection, to the design of structured databases and the architecture of front-end query systems. Meanwhile, the challenges and current limitations encountered during the construction process are also analyzed. This study aims to provide an operable theoretical framework and practical scheme for the digital preservation and contemporary dissemination of Hangzhou dialect, and also offer a reference for the protection of other endangered dialects.

**Keywords:** Hangzhou Dialect; Corpus; Phonetic Dictionary; Digital Protection

## 1. Introduction

Hangzhou dialect is a relatively distinctive branch of the Neo-Wu Yue language family. Unlike the native dialects spoken in the surrounding areas of Hangzhou, its formation and evolution are inextricably linked to the city's social history and cultural background [1]. As a renowned historical and cultural city, Hangzhou served as the capital of the Wuyue Kingdom under the Qian clan during the Five Dynasties

period and later became the capital of the Southern Song Dynasty for over a century. Influenced by this historical legacy, modern Hangzhou dialect bears a strong northern linguistic imprint compared with the dialects of neighboring regions, leading some scholars to simply refer to it as a semi-mandarin dialect [2]. In recent years, with the accelerated pace of urbanization and the globalization of economic development, the number of people using Hangzhou dialect for daily communication has been declining sharply, placing the dialect at risk of rapid atrophy and even extinction. As we know, language is the carrier of culture, and the dialects of a nation represent the diversity of its distinctive regional cultures, embodying the roots of ethnic culture in the most natural and intrinsic way [3]. At present, the protection of dialects has become an imperative task. The construction of an electronic phonetic dictionary of Hangzhou dialect holds profound practical significance for the inheritance and development of Hangzhou's local culture, and even for the development of Zhejiang as a culturally prominent province.

Compared with existing research on dialect protection and digitization, our research team proposed in 2024 a systematic integration of textual and phonetic resources of Hangzhou dialect, aiming to construct a multi-dimensional and highly practical electronic phonetic dictionary. This initiative is designed to develop a comprehensive digital query system, thereby providing a practical tool for the dynamic inheritance of Hangzhou dialect.

## 2. Positioning and Functional Design of the Hangzhou Dialect Electronic Pronunciation Dictionary

The design of the Hangzhou Dialect Electronic Pronunciation Dictionary takes alleviating the "inheritance gap" of the local dialect as its core objective. It caters to the dual needs of popularization and academic research, thus achieving an organic integration of popular

accessibility and scholarly rigor.

## 2.1 Positioning: Dialect Preservation and Cultural Inheritance

The electronic pronunciation dictionary is not merely a query tool, but a comprehensive carrier integrating the digital preservation of Hangzhou dialect and the dissemination of popular local culture. At the popularization level, it targets three core user groups—dialect beginners (young Hangzhou natives and external learners), daily users (citizens who need to communicate with elders in the dialect), and culture enthusiasts (tourists interested in Hangzhou folk customs). Adhering to the principle of *simplified operation and scenario-oriented functions*, it lowers the threshold for use.

The core functions of word and example sentence query are simplified into a "input-and-click" process (e.g., users can listen to the pronunciation by entering the dialect word *tongdian* [meaning "money"] and clicking the "play" button). Academic terminology is avoided by hiding the International Phonetic Alphabet (IPA) in the default interface, which only displays dialect pinyin and definitions, facilitating easy use for middle-aged and elderly users. In terms of content compilation, vocabulary and example sentences prioritize high-frequency daily scenarios such as shopping, greetings and asking for directions. Special emphasis is placed on collecting unique Hangzhou dialect words, commonly-used proverbs, and restoring authentic and natural spoken intonation, helping users quickly master idiomatic expressions of Hangzhou dialect. Meanwhile, the dictionary links dialect with regional culture through annotations and folk knowledge snippets. For instance, when querying the dialect term *congbaogui* (a local snack), a pop-up window will present the background information: "A traditional snack with roots in the Southern Song Dynasty during the resistance against the Jin army; the character 'gui' refers to Qin Hui, a historical figure reviled for betraying his country". This design enables dialect learning and regional cultural cognition to complement each other, enhancing user experience and communication value.

## 2.2 Core Function Design: Building an Integrated "Query-Learning-Culture" Experience

In the development of the Hangzhou Dialect

Electronic Pronunciation Dictionary, the design revolves around three core dimensions—pronunciation query, assisted learning, and cultural extension—with the aim of creating a comprehensive digital platform that serves daily communication, disseminates regional culture, and supports academic research [4].

### 2.2.1 Word and character query: from "understanding" to "speaking"

As the core function, the dictionary supports input of individual characters or words (e.g., *máo* [very/quite] and *fāyè* [amusing/funny]), and quickly returns standard pronunciation, clear definitions, and usage contexts. In response to the literary-colloquial pronunciation variation (a phenomenon where a character has distinct pronunciations in formal and spoken contexts), such as for the character *jiē* (street), the system marks both its literary pronunciation (*jiē*) and colloquial pronunciation (*gā*), with corresponding audio playback provided for each. Users can toggle the display of International Phonetic Alphabet (IPA) symbols (e.g.,  $te^{i33}$  and  $ka$ ) with one click, catering to both public and professional needs.

To lower the barrier to use, the system supports fuzzy matching (e.g., entering *tóngbǎn* [copper plate] automatically links to the dialect term *tóngdiàn* [money]) and synonym mapping (e.g., *zìxíngchē* [bicycle] matches the dialect equivalent *tà'érchē*). The definition content includes core meanings, typical usage scenarios, and Mandarin equivalents. For example, the dialect word *tóngdiàn* is annotated as: "Money; used in shopping and bargaining; equivalent to Mandarin 'qián'". This helps users not only "recognize the word" but also understand how to use it appropriately.

### 2.2.2 Example sentence query: emphasizing colloquial and scenario-based expressions

The example sentence function focuses on "how to say and how to use", avoiding the problem of "being able to read but not apply". When users input keywords (e.g., *luòyǔ* [rain]) or scenario terms (e.g., *gòuwù* [shopping]), the system returns natural colloquial example sentences categorized by daily scenarios, such as "*Luòyǔ le, dài sǎn ó*" (It's raining, take an umbrella) and "*Zhège jǐduō tóngdiàn?*" (How much is this?).

All example sentences are limited to 30 characters or less. Priority is given to authentic expressions with modal particles like *ó* (oh) and *zāi* (a particle indicating emphasis or exclamation), while formal long sentences are

excluded. The audio adopts natural conversational intonation: for instance, "Fàn chīguò de ó?" (Have you had your meal?) is pronounced with a rising tone to convey a question, and "Hǎokàn shà zāi!" (It's extremely beautiful!) uses a falling tone to express exclamation.

Each example sentence is accompanied by context explanations (e.g., "Casual greeting among acquaintances, used at noon or night") and syntactic prompts (e.g., marking the "object-fronting" structure), helping users grasp the unique grammatical logic of Hangzhou dialect.

### 2.2.3 Supplementary functions: turning dialect learning into a cultural journey

Dialect is a living fossil of culture. For example, the dictionary features a dedicated cultural annotation function for vocabulary rich in historical and folk connotations, such as *cōngbāoguì* (a local snack) and *Chénghuáng Shān* (Chenghuang Hill). When querying these terms, users can tap an icon to view relevant allusions, folk cultural significance, and current usage scenarios, seamlessly extending language learning to the understanding of local customs and traditions. For instance, searching for regionally distinctive words like *Chénghuáng Shān* triggers an automatic pop-up of cultural notes that briefly introduce their historical origins.

### 2.2.4 Personalized learning: a customized dialect assistant

To facilitate memorization, users can favorite words, phrases, and example sentences. Based on the collection history, the system intelligently pushes related example sentences to strengthen associative learning. Additionally, all saved content can be exported for offline review and organization.

Through the above functional design, the dictionary not only addresses the linguistic question of "how to speak" but also naturally conveys Hangzhou's urban memories and cultural genes during usage, achieving an organic integration of dialect preservation and public dissemination.

## 3. Lexicographical Principles for Vocabulary Selection and Corpus Compilation

The corpus serves as the **data core** of the Hangzhou Dialect Electronic Pronunciation Dictionary. The quality of its vocabulary selection and compilation directly determines the dictionary's accuracy, comprehensiveness, and

dynamic vitality, and it is required to fully present both the contemporary features and historical evolution of Hangzhou dialect.

### 3.1 Vocabulary Sources: Multi-channel Integration, Balancing Historical Authority and Contemporary Dynamism

Breaking free from over-reliance on a single type of literature, the vocabulary sources of the dialect corpus need to balance phonetic coverage and practical query demands. It adopts an integrated model of *authoritative literature tracing + fieldwork-based collection*, ensuring that no historical words are omitted, no active colloquial words are missing, and all practical vocabulary is fully covered [5].

#### 3.1.1 Authoritative literature tracing

Centering on dialectologist-verified authoritative works such as *Hangzhou Dialect Dictionary*, *Commonly-used Dictionary of Zhejiang Dialects*, and *Qiantang River Dialect*, this approach extracts non-controversial basic vocabulary, words with recorded historical phonetic changes, and region-specific nouns, thus guaranteeing lexical accuracy [6-8]. It also incorporates all the characters, words, and phrases specified in *Survey Manual of the China Language Resources Audio Database · Chinese Dialects*. Based on the phonological system (initials, finals, and tones) of Hangzhou dialect, representative characters and words covering all combinations of initials, finals, and tones are selected to ensure the integrity of the dialect's phonetic system.

#### 3.1.2 Fieldwork-based collection

To capture active colloquial words unrecorded in existing literature (e.g., *máogūgū* [a rough estimate] and *hāngbulāngdāng* [collectively/in total]) as well as pronunciation variations across different sub-regions, data collection subjects are stratified along three dimensions: age, region, and occupation. The age group covers the elderly, middle-aged, and young people; the geographic scope includes downtown areas, suburban districts, and exurban counties; and the occupational spectrum encompasses dialect inheritance practitioners, groups who use dialect frequently in daily life, and professionals who primarily use Mandarin. This multi-dimensional sampling comprehensively captures intergenerational and regional pronunciation differences.

### 3.2 Scope of Inclusion: Hierarchical Coverage Balancing Dialect Uniqueness and

### Evolutionary Traits

The corpus incorporates and collects four categories of vocabulary—daily colloquial terms, distinctive lexicons, intergenerational variant words, and endangered archaic words in Hangzhou dialect. This approach not only caters to contemporary query needs but also documents the historical evolution of the dialect.

#### 3.2.1 Distinctive lexicons

Priority is given to vocabulary unique to Hangzhou dialect or those with a significantly higher usage frequency than in other regions. Distinctive daily conversational terms are classified into three subcategories:

Time-related terms (e.g., *zǎojiāntóu* [early morning], *chéngguāng* [time])

Object-related terms (e.g., *yángsǎn* [umbrella], *tà'érchē* [bicycle])

State-related terms (e.g., *máogūgū* [a rough estimate], *mùlǎolǎo* [a large amount])

These words are selected based on their notable divergence from Mandarin or other Wu dialects (e.g., the Shanghai dialect term *jiǎotàchē* [bicycle] corresponds to the Hangzhou dialect *tà'érchē*). Distinctive lexicon associated with folk culture includes cuisine-related terms (e.g., *piàn'érchuān* [a local noodle dish]), geographical terms (e.g., *Chénghuáng Shān* [Chenghuang Hill], *Héfāng Jiē* [Hefang Street]), and folk custom terms (e.g., *jìzào* [Kitchen God Worship], *duānwǔ* [Dragon Boat Festival], *làbāzhōu* [Laba Congee]). All such entries are accompanied by cultural background notes to highlight their heritage value.

#### 3.2.2 Intergenerational variant words

Hangzhou dialect has been profoundly influenced by Mandarin, resulting in distinct differences in pronunciation and vocabulary usage between the older and younger generations. Including these variant words provides a tangible reflection of the dialect's contemporary evolution. The corpus comprehensively documents intergenerational variations in three aspects:

(1) Pronunciation (e.g., the term for "grandmother" is pronounced *nga bu* by the older generation vs. *wài pó* by the younger generation)

(2) Lexical choice (e.g., "bicycle" is referred to as *tà'érchē* by the older generation vs. *zìxíngchē* by the younger generation)

(3) Grammar (e.g., the older generation says *fàn chīguò de ó* [Have you had your meal?] vs. the younger generation's *chīguò fàn le ó*)

Brief explanations of the causes behind these variations are provided to support research on dialect evolution.

#### 3.2.3 Endangered archaic words

Endangered archaic words are identified through three methods: literature comparison, oral accounts from elderly residents, and association with traditional scenarios. For instance, interviews are conducted with Hangzhou residents aged 80 and above to record vocabulary that is "only heard from the elderly but no longer used by young people", such as *yánghuǒ* [matches], *yángyóu* [kerosene], and *tàshuǐchē* [pedal waterwheel]. For words that exist only in written records with undetermined pronunciations (e.g., *Hāiningcháo* [Haining Tide]), the corpus specifies their documentary sources and usage contexts, achieving the dual goal of preserving both vocabulary and cultural memory.

### 3.3 Example Sentence Selection: Anchored in "Grammar + Context" for Authenticity, Contextualization and Practicality

Example sentence selection closely aligns with the grammatical features and daily contexts of Hangzhou dialect, ensuring that users can directly imitate and apply them. The following criteria are strictly followed:

#### 3.3.1 Grammar-structure orientation

Although the grammar of Hangzhou dialect is similar to that of Mandarin, it has distinctive features such as object fronting and abundant modal particles. The dictionary enables users to naturally grasp these grammatical differences when querying example sentences. For instance, sentences like "*Fàn chīguò de ó?*" (Have you had your meal?) and "*Shū kànwán de zāi*" (I've finished reading the book) are used to illustrate the object-fronting structure; utterances including "*Shēntǐ hǎo ó?*" (How are you feeling?) and "*Hǎokàn shà zāi!*" (It's extremely beautiful!) highlight the usage of characteristic modal particles; and expressions such as "*Chī dé bǎo de*" (Can eat one's fill) and "*Zǒu dé dòng de*" (Can walk properly) demonstrate the uniqueness of the complement marker *de*. 3–5 example sentences are designed for each core grammatical point, covering different vocabulary items.

#### 3.3.2 Context-of-use orientation

The selection targets high-frequency scenarios including shopping (e.g., "*Néngbu néng piányi diǎn?*" [Can you make it cheaper?]), greetings

(e.g., "Fàn chīguò de ó?" [Have you had your meal?]), travel (e.g., "Chéngguāng Shān zěnme zǒu ó?" [How do I get to Chengguang Hill?]) and catering (e.g., "Piàn'érchuān yào chóng tāng" [Extra soup for the Pian'erchuan noodles]). Abstract sentences without specific contexts are excluded. All example sentences are verified by both older and younger generations of Hangzhou residents to ensure their authenticity. In addition, the dictionary includes 30–50 Hangzhou dialect proverbs (e.g., the dialect version of "Hángzhōu luóbo Shào xīng zhǒng" [Hangzhou radishes are of Shaoxing stock] and "Shàng yǒu tiāntáng, xià yǒu Sū-Háng" [Above there is paradise; below there are Suzhou and Hangzhou]) as well as 20–30 nursery rhymes (e.g., the Hangzhou-adapted version of "Yáo a yáo, yáo dào wàipó qiáo" [Rock-a-bye, rock to Grandma's bridge]), preserving the cultural attributes of the dialect.

### 3.3.3 Practicality and contextualization orientation

The selection adheres to the principle of "conciseness, simplicity and naturalness". Each sentence is limited to 5–15 characters, with lengthy expressions split into concise ones. Centered on contextualization, the example sentences cover more than 10 scenarios such as small talk (e.g., "Nǐ dào nǎlǐ qù lǎ?" [Where are you off to?]), "Jīnzhāo tiānqì mǎn hǎo de ma!" [Nice weather today, isn't it!]) and daily services (e.g., "Qǐngwèn gōngjiāo zhàn zài nǎlǐ?" [Excuse me, where is the bus stop?]), "Wǒ yào qù Xīhú jǐngqū" [I want to go to the West Lake Scenic Area]). Colloquial expressions are prioritized while formal written language is avoided. Consistency among vocabulary, example sentences and pronunciation is ensured, laying a solid foundation for the sentence query function.

## 4. Key Technical Implementation Paths for Dictionary Construction

The development of the pronunciation dictionary follows a technical workflow of data collection → processing and annotation → database construction → function design and implementation, where each step is closely interconnected. The details are as follows:

### 4.1 Collection and Processing of Phonetic Corpus

Phonetic recordings of dialect corpus serve as the core of the electronic pronunciation dictionary. A high-quality corpus must be

constructed through scientific collection standards and professional processing technologies, laying the foundation for the subsequent pronunciation query function.

#### 4.1.1 Selection of speakers: establishing a representative speaker system

The selection of speakers aims to cover the mainstream variants of Hangzhou dialect, with multi-dimensional screening criteria formulated to prevent individual differences of a single speaker from affecting the representativeness of the corpus.

##### (1) Geographic Screening

Priority is given to long-term residents (residential duration  $\geq 20$  years) in the old urban areas of Hangzhou—such as the core districts of Shangcheng and Gongshu, which fall within the scope of the former "Hangzhou Prefectural City". The dialect in this region retains the typical features of the integration of "Mandarinization" and "Wu dialect substratum" in Hangzhou dialect, avoiding interference from dialects of surrounding districts and counties (e.g., Xiaoshan dialect and Yuhang dialect). Meanwhile, 1–2 speakers from suburban areas (e.g., the outskirts of Qiantang and Xihu districts) are added to record the differences in dialect variants.

##### (2) Age Stratification

A three-tier age structure of elderly, middle-aged, and young groups is adopted. The elderly group is familiar with traditional Hangzhou dialect vocabulary and proverbs, and can provide endangered dialect expressions; the middle-aged group represents the mainstream usage status of current Hangzhou dialect; the young group reflects the evolution of the dialect among the youth (e.g., the use of loanwords from Mandarin). Each group includes no fewer than 3 speakers, ensuring comprehensive coverage of the phonetic features of all age groups.

##### (3) Language Proficiency and Gender Balance

Speakers must be native speakers of Hangzhou dialect, free from stuttering or vocal cord disorders, and capable of pronouncing long sentences clearly and stably. The gender ratio is maintained at approximately 1:1 to avoid the impact of gender-induced tone differences (e.g., slightly higher tones in female speech) on the consistency of the corpus.

##### (4) Screening Process

A three-step process is implemented: dialect oral test (e.g., reading Hangzhou dialect short sentences and narrating daily experiences) →

phonetic sample analysis (evaluation of pronunciation standardization by dialectologists) → final confirmation. This ensures that all selected speakers meet the criteria, ultimately forming a multi-speaker system of 10–15 individuals.

#### 4.1.2 Collection standards: establishing a standardized collection system

The collection standards cover three aspects—corpus selection, equipment and environment, and collection procedures—to ensure the integrity and consistency of recordings. To prevent environmental noise and equipment differences from affecting corpus quality, unified standards are formulated as follows:

##### (1) Environmental Requirements

Recordings are conducted in professional recording studios or soundproofed rooms (with background noise  $\leq 30$  decibels, monitored in real time by a sound level meter). Echo and electromagnetic interference are avoided to minimize external noise.

##### (2) Equipment Parameters

Professional condenser microphones (e.g., AKG 414B) are used, with the sampling rate set at 44.1 kHz, bit depth at 16 bit, and mono channel selected (to reduce redundant data and facilitate subsequent segmentation). Pop filters and monitoring headphones are used as auxiliary equipment to ensure the consistency of recording devices [9].

#### 4.1.3 Recording collection and preliminary processing procedures

The collection process follows standardized procedures to guarantee the quality of corpus acquisition:

##### (1) Preparatory Work

Recording scripts are distributed to speakers in advance, with 1–2 days reserved for familiarization to avoid pronunciation hesitations caused by unfamiliarity with the texts. Prior to recording, brief training is provided to speakers, clarifying pronunciation requirements (e.g., moderate speaking speed, full tones, and avoiding deliberate slowdown or acceleration).

##### (2) Formal recording

A segmented recording mode is adopted, with each segment containing 5–10 sentences. After each recording session, both speakers and collectors conduct joint listening checks. Segments with unclear pronunciation or missing readings are re-recorded immediately. Speakers are given a 10-minute break after every 30 minutes of recording to prevent sound quality

degradation due to fatigue.

##### (3) Preliminary screening

Upon completion of collection, the first round of quality inspection is carried out on the recording files. Segments with obvious background noise or pronunciation errors are eliminated, with a screening pass rate of over 90% to ensure the basic quality of the corpus entering the subsequent processing stage.

##### (4) Phonetic processing

Standardization and validation of recording data are achieved. Sound quality is optimized through technical means, dynamic range is controlled, and the WAV lossless format is adopted, laying the foundation for the subsequent docking of the pronunciation query function.

##### (5) Phonetic segmentation

Query units are accurately segmented, with words and sentences as the core query units to achieve precise segmentation. Each segmented unit is fully aligned with the corresponding textual content.

## 4.2 Database Design and Construction

The database serves as the data hub of the electronic pronunciation dictionary. It is required to implement core functions such as corpus storage, associative query, and dynamic update through structured design, while ensuring data security and retrieval efficiency, so as to provide stable support for front-end query and user interaction.

#### 4.2.1 Data table structure design: building a structured system with multi-table association

Data table design follows the Entity-Relationship (E-R) model, clarifying the attributes and association logic of core data entities including vocabulary, example sentences, audio files, and annotations. This avoids data redundancy and inefficient query issues, forming a complete data chain through "primary key-foreign key" associations.

##### (1) Field design and functional orientation of core data tables

Combined with the query requirements and data characteristics of the Hangzhou dialect dictionary, the following core data tables are designed:

**Lexeme Table:** Includes lexeme ID, Chinese character headword, International Phonetic Alphabet (IPA), Hangzhou dialect pinyin, basic definition, cultural annotation, and associated lexeme ID;

**Pronunciation File Table:** Includes audio ID,

associated lexeme ID, associated example sentence ID, audio file path, speaker ID, and audio duration;

Example Sentence Table: Includes example sentence ID, example sentence text, scenario description, and associated lexeme ID;

Speaker Information Table: Includes speaker ID, speaker code, age group, and regional code [10].

(2) Data table association logic

A five-level association system of "lexeme-example sentence-audio-annotation-speaker" is constructed through foreign keys:

Core Chain: The Lexeme Table is associated with the Example Sentence Table to realize integrated "word-sentence" query;

Phonetic Chain: The Lexeme Table and Example Sentence Table are respectively associated with the Pronunciation File Table to enable one-click retrieval of "text-audio" resources;

Traceability Chain: The Pronunciation File Table is associated with the Speaker Information Table to achieve traceability of pronunciation sources.

4.2.2 Data annotation: endowing the corpus with "searchable and analyzable" attributes

Data annotation is the key link connecting raw corpus and database query. It endows the corpus with structured attributes through standardized annotation, ensuring query accuracy and analytical feasibility:

(1) Phonetic Annotation

Text transcription is based on recording scripts, with audio segments matched segment by segment (e.g., the phrase *mùlǎolǎo hǎochī* [extremely delicious] shall not be abbreviated). Dialect-specific pronunciations are labeled as "colloquial variants" (e.g., *nǎlī* [colloquial equivalent of "where"]).

Phonetic transcription prioritizes the International Phonetic Alphabet, following the phonological specifications in *Zhejiang Dialect Resource Collection · Hangzhou* (e.g., the term *zǎojiāntóu* [early morning] is transcribed as "tsɔ<sup>33</sup> teie<sup>51</sup> t<sup>h</sup>ou<sup>21</sup>"). Hangzhou dialect pinyin (e.g., "zǎo jiān tou") is supplemented simultaneously to accommodate general users.

A mechanism of "dual annotator labeling + cross review" is adopted, with the annotation discrepancy rate controlled at  $\leq 3\%$  [11].

(2) Attribute Annotation

Each audio segment is linked to a speaker ID, with labels for age group, regional code, and gender added.

Vocabulary items are annotated with parts of speech (e.g., *qièyì* [comfortable] labeled as "adjective") and usage scenarios (e.g., *jìdào* [Kitchen God Worship] labeled as "folk custom"). Scenario annotations are aligned with the design of example sentence texts, reserving expansion space for the front-end function of "scenario-based classified navigation".

### 4.3 Frontend Query System Architecture

The frontend query system takes user experience as its core. Through a closed-loop design of *frontend interaction – backend processing – data feedback*, it achieves the functional goals of convenient query, efficient learning, and open contribution, while adapting to multi-terminal usage scenarios.

4.3.1 Architecture selection: b/s architecture as the mainstay, wechat mini program as a supplement

A hybrid model of Browser/Server (B/S) architecture + WeChat Mini Program is adopted. The B/S architecture supports cross-platform use on PCs and mobile phones, featuring no installation required and convenient maintenance. The WeChat Mini Program is tailored for fragmented mobile scenarios (e.g., instant query during daily conversations), and leverages voice input to enhance query convenience, satisfying the diverse needs of different users.

4.3.2 Voice query module: bidirectional adaptation for dialect and mandarin

In terms of technical selection, an input method system supporting speech recognition is adopted. The recognized text is directly matched with the *Chinese character headword* and *Hangzhou dialect pinyin* fields in the database lexeme table, returning corresponding entries and audio files quickly. An error-tolerant mechanism is implemented: if no direct match is found for the recognition result, similar entry recommendations are automatically pushed (e.g., when the user inputs *mùlǎolǎo*, the system recognizes it and recommends content related to the dialect term *mùlǎolǎo*). This reduces the impact of recognition errors on user experience.

4.3.3 Frontend interface design: balancing simplicity and functionality

The interface design follows the principle of minimalist interaction + direct access to core functions. It draws on the interface logic of mainstream dictionary products such as the *Modern Chinese Dictionary* APP, and incorporates Hangzhou regional cultural

elements (e.g., adopting color schemes inspired by the West Lake imagery). While ensuring operational convenience, it strengthens the attribute of dialect cultural dissemination.

### 5. Conclusion

In the practice of developing the Hangzhou Dialect Electronic Pronunciation Dictionary, affected by the inherent complexity of the dialect itself, the project team was confronted with three-dimensional challenges in linguistics, technical implementation, and resource sustainability. The primary difficulties in dictionary construction lie in three aspects: pronunciation confirmation, character usage standardization, and generational dialect differences. Furthermore, the project requires multidisciplinary technical support, relies heavily on sufficient resource guarantees, and needs to ensure long-term project sustainability. Restricted by the research cycle, resource investment, and technical conditions, this study is currently in the vocabulary expansion stage. It still has limitations such as limited corpus coverage, incomplete International Phonetic Alphabet (IPA) annotation, and room for further improvement in dictionary functions. With the gradual refinement of the pronunciation dictionary in the later stage, it will theoretically contribute to deepening linguistic research and sustaining cultural continuity. In practice, it will empower applications in education, technology, and social sectors. This work bridges history and future, and bears far-reaching significance for preserving the unique local accent and cultural essence of Hangzhou, a renowned historical and cultural city.

### References

[1] Xiao Nengping. A Study on the Contact between Wu Dialect and Mandarin from the

Perspective of Sandhi Tones in Hangzhou Dialect. *Nankai Journal of Linguistics*, 2024(1), 22-32.

- [2] Luo Xiaogang. Historical Development, Current Situation and Protection Strategies of Hangzhou Dialect. *Hangzhou Journal of Social Sciences*, 2016(1), 191-200.
- [3] Lv Baoli. The Influence of the Popularization of Mandarin on the Inheritance of Intangible Cultural Heritage. *Daqing Social Sciences*, 2008(6), 115-116.
- [4] Cao Zhiyun. On the protection and inheritance of Zhejiang dialect culture. *Zhejiang Social Sciences*, 2021(2), 118-124.
- [5] Wang Jianlin. Discuss the protection strategy of Chinese dialect heritage in the age of artificial intelligence. *Journal of Management and Social Development*, 2024(4), 275-283.
- [6] Li Rong & Bao Shijie. *Hangzhou Dialect Dictionary*. Jiangsu Education Press, 1998.
- [7] Cao Zhiyun. *Commonly-used Dictionary of Zhejiang Dialects*. The Commercial Press, 2025.
- [8] Xu Yue. *Qiantang River Dialect*. Hangzhou Publishing House, 2015.
- [9] Department of Language and Character Information Management, Ministry of Education & China Language Resources Protection Research Center. *Survey Manual of the China Language Resources Audio Database • Chinese Dialects*. The Commercial Press, 2016.
- [10] Hong Tuoyi. Conceptions of Constructing a Phonetic Database of Chinese Dialects. *Library Work*, 2009(3), 83-86.
- [11] Wang Wensheng & Chen Yu. *Compendium of Zhejiang Dialect Resources • Hangzhou*. Zhejiang University Press, 2025.