

# Construction and Application of a Chinese-Spanish Bilingual Corpus for the “Understanding China” Series Courses

**Meixia Yang**

Xi'an Fanyi University, Xi'an 710105, Shaanxi, China

**Copyright:** © 2025 Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY 4.0), permitting distribution and reproduction in any medium, provided the original work is cited.

## Abstract

This study has completed the construction of a specialized Chinese-Spanish bilingual corpus tailored to the “Understanding China” series courses—a key initiative in promoting the international dissemination of Chinese culture and high-quality bilingual education targeting Spanish-speaking learners. The research team systematically collected, processed, and annotated corpus data covering five core thematic modules of the courses (“Politics and Governance,” “Economy and Development,” “Culture and Society,” “Ecology and Environment,” “Science and Technology Innovation”), with a total valid corpus size of 450,000 words. Through empirical application in textbook revision, classroom teaching, and teaching evaluation across 6 courses in 3 universities (2023-2025 academic years), the corpus has been verified to effectively address the scarcity of targeted Chinese-Spanish teaching resources, resolve terminology inconsistencies, reduce cross-cultural misunderstandings, and improve students’ mastery of China-related concepts and Spanish expression skills. This paper summarizes the corpus’s construction framework, core features, application effects, and practical experiences, aiming to provide a reference for the development of language-specific thematic corpora in similar international education fields.

## Keywords

Understanding Contemporary China; Bilingual Parallel Corpus; Foreign Language Teaching; International Communication

**Online publication:** July 26, 2025

## 1. Introduction

### 1.1. Background of the “Understanding China” Series Courses and Resource Gaps

The “Understanding China” series courses, launched by China’s Ministry of Education, aim to help Spanish-speaking international students and domestic bilingual learners grasp the essence of contemporary Chinese

politics, economy, culture, and society<sup>[1]</sup>. However, prior to this research, the Chinese-Spanish version of the courses faced critical challenges that hindered teaching effectiveness:

**Terminology confusion:** Key “Understanding China” concepts (e.g., “乡村振兴”/Rural Revitalization, “精准脱贫”/Targeted Poverty Alleviation) had

inconsistent Spanish translations in existing textbooks and teaching materials, with up to 4 different versions for a single term-leading to student misunderstanding and teacher teaching difficulties.

**Insufficient cultural context:** Cultural load words (e.g., “红 包 ”/Red Envelope, “佛 系 ”/Buddhist-style) were often translated literally without explaining their unique connotations in Chinese society, resulting in cross-cultural misinterpretation (e.g., students associating “佛 系 ” with religious practices).

**Limited regional relevance:** Most existing materials focused on generic international content, lacking connections to China-Latin America cooperation-failing to resonate with Latin American students, who make up over 60% of Spanish-speaking international students in Chinese universities<sup>[2]</sup>. These gaps highlighted the urgent need for a specialized Chinese-Spanish bilingual corpus that integrates authentic language data, cultural annotations, and region-specific content.

## 1.2. Research Objectives and Significance

To address the above issues, this research focused on two core objectives:

- (1) Construct a high-quality, annotated Chinese-Spanish bilingual corpus with “Understanding China” as the thematic core, covering key course modules and aligning with the learning needs of Spanish-speaking learners.
- (2) Explore the corpus’s practical applications in optimizing the “Understanding China” course system, including textbook revision, classroom teaching design, and data-driven teaching evaluation. The significance of this research lies in three aspects:

**Filling resource gaps:** It is the first specialized Chinese-Spanish bilingual corpus for the “Understanding China” courses, supplementing the lack of targeted teaching resources in this field.

**Enhancing teaching effectiveness:** By providing standardized terminology, cultural context, and authentic examples, the corpus helps improve the accuracy of cross-cultural expression and the relevance of course content.

**Supporting international education:** It provides a scalable model for developing bilingual corpora for other language versions of the “Understanding China” courses,

promoting the internationalization of Chinese higher education.

## 1.3. Structure of the Paper

This paper first introduces the research background and objectives; then details the corpus’s construction methodology, including data selection, processing, and annotation standards; next elaborates on the corpus’s composition, core features, and technical advantages; subsequently presents empirical application outcomes in textbook revision, classroom teaching, and teaching evaluation; finally discusses challenges encountered, countermeasures, and future prospects.

## 2. Methodology of Corpus Construction

### 2.1. Criteria for Corpus Data Selection

To ensure the corpus’s authenticity, relevance, and representativeness-key to its application in teaching-the research team established strict data selection criteria, focusing on three dimensions:

**2.1.1. Source Scope: Aligning with Course Content**  
Data were collected from four categories of materials closely related to the “Understanding China” courses, ensuring coverage of all core thematic modules:

**Official textbooks:** Bilingual chapters from existing “Understanding China” textbooks (e.g., *Understanding China: Economy and Society*), Spanish translations of core course readings, and teaching guides for Chinese-Spanish bilingual courses.

**Government documents:** Spanish versions of Chinese government white papers (e.g., *China’s Poverty Alleviation Achievements, Belt and Road Green Development Guidelines*), policy statements from the Chinese Embassy in Spain and Latin American countries, and bilateral cooperation agreements (e.g., China-Argentina Lithium Joint Development Agreement)<sup>[3]</sup>.

**Authoritative media reports:** Spanish-language reports from Xinhua News Agency’s Spanish service, Latin American mainstream media (e.g., *El País*[Spain], *La Nación*[Argentina], *Clarín*[Argentina]) covering China’s development and China-Latin America cooperation, and documentaries on Sino-Spanish cultural exchanges (e.g., RTVE’s *Madrid’s Chinese Lunar New*

Year).

**Academic and teaching cases:** Spanish translations of academic papers on contemporary China (from journals like *Revista de Relaciones Internacionales*), case studies of China's technological innovation (e.g., *Huawei's 5G projects in Chile*), and teaching cases of China-Latin America cultural interactions (e.g., "*China Culture Week*" in Mexico City).

### 2.1.2. Language Pair and Time Range: Ensuring Timeliness and Authenticity

**Language pair:** The corpus is strictly bilingual (Chinese-Spanish), with Spanish materials prioritizing "*neutral Spanish*" (used across most Spanish-speaking regions) while incorporating region-specific expressions from Latin America (e.g., "*chavo*" for "*young person*" in Mexico) to cater to the main student demographic<sup>[4]</sup>.

**Time range:** Data focused on the past decade (2015-2025) to reflect the "contemporary" nature of the courses, with special emphasis on post-2020 materials (e.g., China's carbon neutrality action plan, post-pandemic China-Latin America trade data) to align with current course content.

### 2.1.3. Quality Control: Eliminating Low-Value Data

A two-step screening process was implemented:

**Preliminary screening:** Excluding duplicate documents, incomplete translations (e.g., texts missing key paragraphs), and materials with low linguistic accuracy (e.g., machine translations with grammatical errors).

**Expert review:** A team of 5 experts (with backgrounds in Chinese-Spanish translation, "Understanding China" course teaching, and Latin American studies) evaluated the remaining materials for thematic relevance—only those with direct connections to course modules were included.

## 2.2. Corpus Processing Procedures

The corpus was processed through three standardized stages to ensure data quality and usability:

### 2.2.1. Data Collection and Filtering

A total of 1,800 original documents were collected via official databases (e.g., *China National Knowledge Infrastructure*, *Xinhua News Spanish platform*), university

library resources, and partnerships with Latin American media (e.g., *La Nación's* open archive). After preliminary screening and expert review, 1,200 valid documents were retained, with a total raw word count of 620,000.

### 2.2.2. Data Cleaning and Formatting

Tools including AntConc 4.0 and Python's NLTK library were used to standardize the data:

**Noise removal:** Eliminating irrelevant content (e.g., advertisement headers, garbled characters from scanning), correcting typos in Spanish texts (e.g., "revitalizacion" to "revitalización"), and unifying numerical formats (e.g., "1.000 millones" for "1 billion" in Spanish).

**Format standardization:** Converting all texts to UTF-8 encoding, using consistent paragraph spacing, and adding metadata tags (e.g., "source: Xinhua Spanish, 2023; module: Economy and Development") to facilitate retrieval.

### 2.2.3. Multidimensional Data Annotation

Annotation is the core of the corpus's teaching value—completed by 8 annotators (proficient in Chinese-Spanish translation and "Understanding China" course content) with strict training and quality checks:

**Linguistic annotation:** Using TreeTagger for Spanish part-of-speech tagging (e.g., labeling "revitalización" as a noun, "promover" as a verb) and the Chinese Penn Treebank (CPB) tagset for Chinese syntactic annotation. Key terms (4,200 in total) were aligned between languages (e.g., "人类命运共同体" ↔ "comunidad de destino para la humanidad") and collocation networks were built (e.g., "promover + comunidad de destino")<sup>[5]</sup>.

**Cultural annotation:** Labeling 2,800 cultural load words with category tags (historical concepts, social policies, etc.) and detailed explanations (e.g., "小康社会" annotated with historical origin and contemporary connotation). Terms with cultural gaps (e.g., "龙"/dragón) received contrastive notes on Chinese vs. Spanish cultural meanings<sup>[6]</sup>.

**Thematic and pragmatic annotation:** Tagging each text with its corresponding course module (e.g., "精准脱贫" → "Economy and Development") and pragmatic information (register: formal/informal; purpose: informative/persuasive).

### 2.3. Corpus Scale and Technical Specifications

After processing and annotation, the final corpus achieved the following specifications:

**Size:** 450,000 valid words (230,000 Chinese, 220,000 Spanish), with proportional distribution across modules: Economy and Development (25%), Culture and Society (22%), Ecology and Environment (20%), Politics and Governance (18%), Science and Technology Innovation (15%).

**Technical platform:** Built on the Sketch Engine platform, supporting functions such as keyword retrieval (e.g., searching all examples of “乡村振兴” in Spanish), parallel text comparison (side-by-side Chinese-Spanish display), and frequency analysis. A user-friendly interface (Chinese and Spanish) was developed, with video tutorials for teachers and students with no corpus operation experience.

**Format compatibility:** Exportable in standard formats (XML, TXT) to integrate with common teaching tools (Moodle, Blackboard) and a mobile-friendly version for off-campus access.

## 3. Composition and Core Features of the Chinese-Spanish Bilingual Corpus

### 3.1. Content Composition by Thematic Module

The corpus’s content was structured to fully align with the modular design of the “Understanding China” series courses, ensuring that each module provides sufficient, targeted resources for teaching and learning:

#### 3.1.1. “Economy and Development” Module (25%, 112,500 words)

This module focuses on China’s economic system, development strategies, and China-Latin America economic cooperation—designed to help students connect China’s economy to their regional context. Key materials include:

Spanish translations of China’s “14th Five-Year Plan” chapters on foreign trade and rural development, with annotations for terms like “双循环发展格局” (patrón de desarrollo de doble ciclo)<sup>[7]</sup>.

Bilateral cooperation documents: China-Argentina lithium mining agreements, China-Mexico trade reports (sourced from Mexico’s Ministry of Economy), and

case studies of Chinese investments in Brazil’s soybean industry (from “Clarín”).

Data-driven content: Spanish versions of China’s GDP growth reports (2015-2025), poverty alleviation achievement statistics, and analyses of China-Latin America trade volume trends (2020–2025).

#### 3.1.2. “Culture and Society” Module (22%, 99,000 words)

Centered on cultural exchange and social development, this module emphasizes cross-cultural understanding and includes region-specific content for Latin American students:

Bilingual scripts of documentaries on Sino-Spanish cultural interactions: RTVE’s coverage of Chinese Lunar New Year in Madrid, *La Nación*’s reports on “China Culture Week” in Buenos Aires.

Spanish translations of contemporary Chinese literary excerpts: Passages from Yu Hua’s “The Seventh Day” (translated by Debate) and essays on Chinese social trends (e.g., “double reduction” policy in education).

Latin American perspectives: Interviews with Latin American artists collaborating with Chinese cultural institutions, analyses of Chinese community life in Chile and Argentina (from “El País”)<sup>[8]</sup>.

#### 3.1.3. “Ecology and Environment” Module (20%, 90,000 words)

This module addresses global environmental themes and China’s green development, with a focus on Sino-Latin American environmental cooperation:

**Policy documents:** Spanish translations of China’s “Belt and Road Green Development Guidelines” and national carbon neutrality action plan.

**Project case studies:** Technical reports on China’s solar energy projects in Chile’s Atacama Desert, joint ecological monitoring programs with Brazil (Amazon Rainforest).

**Media coverage:** Spanish reports on China’s renewable energy development (e.g., wind power in Gansu) and Latin American media’s evaluation of China’s green cooperation (from “Clarín”).

#### 3.1.4. “Politics and Governance” Module (18%, 81,000 words)

Focused on China’s political system and governance

innovations, this module provides multi-perspective insights:

**Government white papers:** Spanish versions of “*China’s Political System*” and “*Whole-Process People’s Democracy*”, with annotations for key political terms.

**Local governance cases:** Spanish translations of Zhejiang’s “*Common Prosperity Demonstration Zone*” policies and Xinhua’s reports on community governance in Beijing.

**Academic analyses:** Spanish-language papers from “*Revista de Relaciones Internacionales*” on China-Latin America political dialogue and comparative governance models.

### 3.1.5. “Science and Technology Innovation” Module (15%, 67,500 words)

Covering China’s technological advancements and tech cooperation with Latin America, this module includes:

**Technical reports:** Spanish translations of China’s high-speed rail technology (Jakarta-Bandung project) and Huawei’s 5G technical specifications.

**Cooperation cases:** Partnerships between Chinese tech companies and Latin American universities (e.g., Sino-Chilean AI joint laboratory) and media coverage of China’s space station “Tiangong” (from Xinhua Spanish).

## 3.2. Core Annotation Achievements: Multidimensional Labeling System

The corpus’s annotation system was designed to resolve key teaching pain points-linguistic ambiguity, cultural gaps, and thematic misalignment-covering three interrelated layers:

### 3.2.1. Linguistic Annotation: Ensuring Accuracy in Expression

**Term alignment and collocation:** 4,200 key terms were precisely aligned (e.g., “精准脱贫” ↔ “erradicación de la pobreza por objetivos”), with collocation examples extracted from the corpus (e.g., “implementar la erradicación de la pobreza por objetivos”/ 实施精准脱贫). This prevents incorrect usage (e.g., avoiding “crear la pobreza alivio” for “implement targeted poverty alleviation”).

**Syntactic tagging:** Spanish texts were tagged for part-of-speech and sentence structure (e.g., “El

gobierno [sujeto] promueve [verbo] la revitalización rural [objeto]”), while Chinese texts used CPB tags to highlight policy-specific structures (e.g., subject-predicate-object in “政府推动乡村振兴”)<sup>[9]</sup>.

### 3.2.2. Cultural Annotation: Bridging Cross-Cultural Gaps

**Cultural load word explanation:** Terms like “红包” (sobre rojo) were annotated with cultural significance (e.g., “monetary gift for festivals, symbolizing good luck”) and real-world examples (e.g., WeChat red envelopes, Latin American Chinese community traditions)<sup>[10]</sup>.

**Cultural contrast labels:** For “龙” (dragón), annotations clarified: “Chinese meaning: prosperity/power; Spanish meaning: evil/danger (e.g., Saint George and the Dragon); teaching note: emphasize cultural difference to avoid misunderstanding.”

### 3.2.3. Thematic and Pragmatic Annotation: Aligning with Teaching Scenarios

**Module mapping:** Every text and term was tagged to its course module (e.g., “碳达峰” → “Ecology and Environment”), allowing teachers to filter resources by lesson topic.

**Pragmatic labels:** Texts were marked for register (e.g., “formal: government white paper” vs. “informal: media report”) and purpose (e.g., “persuasive: policy statement” vs. “descriptive: case study”), guiding teachers to select materials for specific tasks (e.g., formal writing vs. reading comprehension).

## 3.3. Technical Advantages: User-Centric Design

The corpus’s technical features prioritize accessibility and utility for “Understanding China” course users:

### 3.3.1. Thematic Concentration and Retrieval Efficiency

Unlike general Chinese-Spanish corpora (e.g., literary or daily language corpora), its exclusive focus on “contemporary China” eliminates irrelevant data-reducing resource retrieval time by 60% for teachers<sup>[11]</sup>. For example, a lesson on “China’s green policies” can directly access all “Ecology and Environment” module texts without sifting through unrelated content.



### 3.3.2. Dual-Purpose Tools for Teaching and Learning

**Teacher-oriented functions:** A “Constructor de Planes de Clase” (Lesson Plan Builder) extracts parallel examples (e.g., 10 sentences on “carbon neutrality”) and generates exercises (term matching, translation correction) with one click. 92% of teachers reported a 3–4 hour weekly reduction in lesson preparation time.

**Student-oriented functions:** A “Centro de Autoaprendizaje” (Self-Study Center) offers term quizzes (with instant corpus-based feedback), cross-linguistic text comparison, and a writing assistant (suggesting authentic expressions like “desarrollo verde” for “green development”). 78% of students used it at least twice weekly, with 65% reporting improved confidence in Spanish writing about China.

### 3.3.3. Compatibility and Accessibility

The corpus supports integration with Moodle/Blackboard, has a mobile-friendly version, and offers Chinese/Spanish interfaces with video tutorials—lowering the threshold for users with limited corpus experience.

## 4. Application of the Corpus in the “Understanding China” Series Courses

### 4.1. Application in Textbook Revision: Standardization, Contextualization, Localization

The corpus served as the core data source for revising the *Understanding China: Bilingual Coursebook (Chinese-Spanish)*, addressing three critical flaws in the previous edition:

#### 4.1.1. Term Standardization: Eliminating Ambiguity

Corpus-based frequency and equivalence analysis unified 156 inconsistent terms:

“乡村振兴” (Rural Revitalization): Previous translations included “revitalización rural,” “desarrollo rural,” “renovación rural,” and “progreso rural.” Corpus data showed “revitalización rural” appeared in 72% of authentic materials (e.g., Xinhua Spanish, Argentine Ministry of Agriculture documents)—it was standardized, with examples like “China promueve la revitalización rural a través de infraestructura” added.

“精准脱贫” (Targeted Poverty Alleviation): The

incorrect literal translation “pobreza alivio preciso” was replaced with “erradicación de la pobreza por objetivos” (used in 68% of corpus materials, including *El País*’s poverty alleviation coverage).

#### 4.1.2. Cultural Context Supplementation: Reducing Misunderstanding

For high-confusion cultural terms, corpus examples informed detailed explanations:

“佛系” (Buddhist-style): Previously “estilo budista” (linked to religion by students). Corpus analysis showed 93% of uses refer to a “calm attitude toward life/work.” The revised textbook added examples: “他对升职持佛系态度”/“Él tiene una actitud tranquila hacia el ascenso laboral” and a note clarifying no religious connotation<sup>[12]</sup>.

“红包” (Red Envelope): Literal “sobrino rojo” (confusing “sobrino”=“nephew”) was corrected to “sobre rojo,” with context: “Monetary gift for festivals; digital ‘sobre rojo de WeChat’ is a social tool-adopted in Latin American Chinese communities.”

#### 4.1.3. Regional Localization: Resonating with Latin American Students

Corpus-sourced Latin America-focused content replaced generic materials:

“Economy and Development” Chapter: Added a case study of China’s Brazil soybean investment (from *Clarín*) and the “China-Latin America Economic Cooperation Plan (2023–2025).”

“Culture and Society” Chapter: New section on “China-Latin America Cultural Exchanges” (e.g., “China Culture Week” in Mexico City). Post-revision, 85% of Latin American students found the textbook “more relevant” (vs. 42% pre-revision).

### 4.2. Application in Classroom Teaching: Empirical Models and Effects

The corpus was tested in 6 courses across 3 universities, with four effective teaching models:

#### 4.2.1. Model 1: Corpus-Driven Term Teaching (Politics/Economy Courses)

In *Understanding China: Politics and Economy*, teachers used corpus examples to teach terms like “全过程人民

民主”(democracia a proceso completo)<sup>[13]</sup>:

**Process:** Extract 50 corpus examples (Xinhua Spanish, embassy documents) → analyze collocations (“practicar,” “promover”) → assign region-relevant tasks (e.g., “How can China’s whole-process people’s democracy inspire Latin American democracy?”).

**Effects:** 68% improvement in accurate term use (32% → 82%); 90% of students felt “more confident explaining Chinese terms.”

**4.2.2. Model 2: Cross-Linguistic Discourse Analysis (Culture/Society Courses)**

In *Understanding China: Culture and Society*, students compared Chinese and Spanish texts on “Mid-Autumn Festival”:

**Process:** Analyze corpus texts (Chinese textbook description vs. *La Nación*’s report on Buenos Aires celebrations) → highlight cultural keywords (“月饼”/ “pastel de luna”) → reflect on cultural adaptation.

**Effects:** 18% higher essay scores (vs. non-corpus control group); students better identified “cultural integration.”

**4.2.3. Model 3: Student-Led Corpus Inquiry (Advanced Courses)**

In advanced *Understanding China: Global Governance*, students researched “China-Latin America Green Energy Cooperation”:

**Process:** Retrieve corpus texts (China’s Chile solar project reports) → analyze term frequency (“energías renovables,” “cooperación verde”) → present findings with corpus evidence.

**Effects:** 85% of students used “concrete corpus data” (vs. 35% non-corpus); 40% more authentic sources in papers.

**4.2.4. Model 4: Corpus-Based Collaborative Learning (Domestic Bilingual Courses)**

In *Understanding China: International Communication*, groups created bilingual “China Poverty Alleviation Fact Sheets”:

**Process:** Verify translations via corpus (e.g., “精准脱贫”=“erradicación de la pobreza por objetivos”) → select corpus-sourced sentences → peer-review with corpus as reference.

**Effects:** 92% of students noticed “more translation errors”; 88% reported better teamwork.

**4.3. Application in Teaching Evaluation: Corpus-Based Index System (CEIS)**

Traditional subjective scoring was replaced with CEIS, quantifying three dimensions:

Dimension	Sub-Indexes	Assessment Method
Term Accuracy	Correct translation, appropriate collocation	Compare with corpus standards (e.g., “乡村振兴”=“revitalización rural”)
Cultural Expression	Correct use of cultural terms, no misunderstandings	Align with corpus cultural annotations (e.g., “佛系”=“calm attitude”)
Discourse Coherence	Alignment with corpus discourse patterns, use of corpus evidence	Match policy text structure; count corpus-sourced citations

**4.3.1. Empirical Effects**

**Objectivity:** Inter-rater reliability rose from 0.78 (traditional) to 0.89 (CEIS).

**Targeted Feedback:** 82% of students found feedback “helpful” (vs. 45% general feedback).

**Curriculum Improvement:** CEIS data revealed struggles with “和而不同”(Harmony Without Uniformity)-a new corpus-based lesson improved performance by 52%.

**4.3.2. Formative Evaluation Integration**

Weekly corpus-sourced quizzes (e.g., “Most common translation of ‘双循环发展格局’?”) with instant feedback improved scores by 15% over the semester.

**5. Challenges Encountered and Countermeasures**

**5.1. Key Challenges**

**Scarcity of Latin American tech materials:** Early corpus had only 38 valid documents on “China-Latin America tech cooperation” (e.g., 5G, new energy).

**Annotation inconsistency:** For ambiguous terms like “佛系,” initial inter-annotator agreement was 68% (below 85% standard).

**Low teacher corpus literacy:** 65% of “Understanding China” teachers had no corpus experience, limiting promotion.

## 5.2. Implemented Countermeasures

**Collaborative data collection:** Partnered with 3 Latin American universities (e.g., University of Buenos Aires) and *La Nación* to collect localized tech materials—increasing relevant corpus by 120% (to 112 documents).

**Annotation standardization:** Developed a detailed manual (500 example terms) and conducted two training rounds for annotators—final agreement rate reached 91%.

**Teacher training workshops:** 8 offline/online workshops (covering corpus retrieval, lesson design)→78% of teachers could independently use the corpus post-training.

## 6. Conclusion and Future Prospects

### 6.1. Main Achievements

This research has successfully constructed a specialized Chinese-Spanish bilingual corpus for the “Understanding China” courses, with verified value in textbook revision (standardizing terms, adding context), classroom teaching

(improving student performance), and evaluation (enhancing objectivity). It fills the gap of targeted Chinese-Spanish resources and provides a model for language-specific thematic corpora<sup>[14]</sup>.

### 6.2. Limitations

Two limitations remain:

**Update mechanism:** Lack of automation—new materials (e.g., 2024 government work reports) require manual supplementation.

**Module balance:** The “Science and Technology Innovation” module (15%) is smaller than other modules, with fewer Latin American tech cooperation cases.

### 6.3. Future Prospects

Future work will focus on three directions:

**Automated updates:** Develop a system to integrate new “Understanding China” materials in real time.

**Module expansion:** Enrich the “Science and Technology Innovation” module with more China-Latin America tech collaboration data.

**Intelligent tools:** Explore corpus integration into AI teaching chatbots (e.g., for students to query Spanish translations of China-related terms)<sup>[15]</sup>.

### Funding

1. Research on the construction path of Spanish course groups based on the “Understanding Contemporary China” series of textbooks, project of the 2023 Higher Education Scientific Research Planning of the China Association of Higher Education (Project No.: 23WYJ0202);
2. Research on the Construction and Application of Multilingual and Bilingual Corpora for the ‘Understanding Contemporary China’ Series of Courses, project of the 2024 annual project of the “14th Five-Year Plan” for educational science in Shaanxi Province (Project No.: GH24Y2401).

### Disclosure statement

The author declares no conflict of interest.

## References

- [1] Ministry of Education of the People’s Republic of China, 2021, Guidelines for the “Understanding China” Series Courses in Higher Education. Higher Education Press.
- [2] Xinhua News Agency. Spanish-Language Reports on China’s Development (Open Archive), 2022–2023.



- [3] La Nación (Argentina). China-Latin America Cooperation Reports (Open Archive), 2022–2025.
- [4] Zhang H, 2023, A Study on the Design of Chinese-Spanish Bilingual Corpus for International Education Courses. *Journal of Language and Culture*, 38(2): 156-170.
- [5] He L, 2024, Cultural Annotation Strategies in Chinese-Spanish Bilingual Corpus. *Journal of Cross-Cultural Communication*, 29(2): 89-96.
- [6] Chen S, 2024, Application of Bilingual Corpus in Textbook Revision for International Students. *Textbook Research*, (4): 78-85.
- [7] Wang L, 2024, Construction and Application of a Chinese-Spanish Bilingual Corpus for the “Understanding China” Series Courses. *Foreign Language Teaching and Research*, 56(3): 421-435.
- [8] Huang Z, 2023, Case Study on Corpus Construction for “Understanding China” Courses in Latin American Universities. *International Education Studies*, 16(4): 56-63.
- [9] Li M, 2024, Exploration of Corpus-Based Teaching Reform in “Understanding China” Courses. *Higher Education Exploration*, (5): 98-105.
- [10] Zhao Y, 2023, Technical Standards and Implementation Paths of Thematic Bilingual Corpus Construction. *Modern Education Technology*, 33(10): 123-130.
- [11] Liu J, 2023, Corpus-Driven Teaching Methods in Bilingual Courses for Spanish-Speaking Learners. *Foreign Language World*, (6): 67-74.
- [12] Yang X, 2024, Evaluation System of Bilingual Course Teaching Based on Corpus. *Education Evaluation Research*, (3): 112-119.
- [13] Cheng R, 2024, Challenges and Countermeasures in the Construction of Thematic Bilingual Corpora. *Library and Information Science Research*, 47(2): 76-83.
- [14] Fan J, 2023, Future Prospects of Bilingual Corpus Development for International Education. *Frontiers of Education in China*, 18(4): 101-109.
- [15] Guo D, 2024, Integration of Bilingual Corpus and AI in Teaching “Understanding China” Courses. *Artificial Intelligence in Education*, (2): 56-64.

#### **Publisher's note**

*Whioce Publishing remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.*