

Article

Teaching Translation Verification with AI Corpora: A Case Study of Chinese Culture-Loaded Terms

Sijing Xin^{1,*}

¹ School of Foreign Languages, Chengdu Jincheng College, Chengdu, China

* Correspondence: Sijing Xin, School of Foreign Languages, Chengdu Jincheng College, Chengdu, China

Abstract: Students of Chinese–English translation frequently render culture-loaded terms by pinyin transliteration or word-for-word literal translation, rarely verifying their choices against usage evidence. This paper reports a case study of integrating an AI corpus platform (AI-Corpora) into verification training for culture-loaded terms in an undergraduate translation course at an applied university in China. A five-step verification model was designed and implemented—independent drafting, corpus search, evidence evaluation, guided revision, and term archiving—across a teaching unit containing three illustrative tasks: a cured-meat term in a Sichuan food text, teahouse-culture terms, and solar-term names. Analysis of before-and-after student outputs shows that corpus evidence helps learners move from pinyin-only or literal renderings toward accepted and audience-appropriate translations, and fosters a data-driven habit of translation decision-making. The paper further discusses typical misreadings of corpus evidence, the scaffolding role of the teacher, and implications for corpus-assisted translation pedagogy in similar institutional contexts.

Keywords: AI corpus; translation verification; culture-loaded terms; translation teaching; corpus-assisted learning

1. Introduction

Culture-loaded terms—lexical items that carry meanings specific to a particular cultural community—are widely recognized as one of the most persistent sources of error in Chinese–English (C–E) translation [1]. In the context of China’s ongoing effort to “tell China’s stories well” in international communication, the accurate and acceptable translation of such terms has acquired additional significance: terms relating to food, festivals, opera, tea culture, and everyday customs appear with high frequency in the text types that Chinese EFL learners are most likely to translate, including promotional materials, news features, and literary essays.

Classroom observation suggests that student translators tend to handle culture-loaded terms in one of two habitual ways: pinyin transliteration without explanation, which transfers the burden of interpretation entirely to the target reader, or literal translation, which frequently produces misleading or unintentionally comic renderings. Behind both habits lies a common deficit—students rarely verify their lexical choices against usage evidence. When asked “how do you know this translation is acceptable?”, the typical answer is intuition or memory of a textbook gloss rather than any checkable source.

Corpus tools offer a well-established remedy. Since the relevance of corpus linguistics to translation studies was argued in [2], a substantial literature has demonstrated the value of corpus consultation in translator education [3, 4]. More recently, AI-enhanced corpus platforms have lowered the technical threshold of corpus use by combining large-scale bilingual evidence with natural-language search and machine-translation assistance. However, research on how such AI corpus platforms can be pedagogically integrated into routine translation teaching—particularly for the micro-

Received: 25 August 2026

Revised: 10 October 2026

Accepted: 23 October 2026

Published: 26 October 2026



Copyright: © 2026 by the authors. Submitted for possible open access publication under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

skill of verifying culture-loaded terms—remains limited, especially in applied university contexts where students have modest proficiency and limited research experience.

This paper addresses the gap through a teaching case study conducted in the course Basics of E–C/C–E Translation at an applied university in Chengdu, China. It pursues three questions: (1) How can an AI corpus platform be embedded in a teachable verification procedure for culture-loaded terms? (2) What changes can be observed in student translations of such terms after verification training? (3) What difficulties arise, and what do they imply for the teacher’s role? The remainder of the paper reviews relevant literature, presents the teaching design, analyzes three classroom cases, and discusses outcomes and limitations.

2. Literature Review

2.1. Culture-Loaded Terms in C–E Translation

Culture-loaded terms, also termed culture-specific items, pose a dual challenge: the translator must both identify the source-culture meaning accurately and choose a target-language strategy that preserves accessibility for the target reader. Strategy taxonomies commonly range from conservation-oriented options (transliteration, literal translation) to substitution-oriented ones (naturalization, deletion, functional equivalence) [1]. In C–E practice, a widely used three-tier heuristic has emerged in Chinese translator training: transliteration plus explanation for terms without established equivalents, adoption of official or conventional translations where they exist (e.g., the Dragon Boat Festival for DuanWuJie), and free or sense translation where formal correspondence is impossible (e.g., DongZhi rendered as the winter solstice).

Choosing among these tiers is an evidence problem rather than a purely linguistic one. Whether an English rendering is “accepted” depends on its currency in target-language usage—precisely the kind of information that corpora can supply. Yet the three-tier heuristic is typically taught as a static checklist, leaving students without a procedure for testing candidate translations against real usage.

2.2. Corpus-Assisted Translation Teaching

Corpus use in translator education has been advocated under the banner of “discovery learning” [3]: learners consult concordances to observe patterns, test hypotheses, and develop autonomous strategies. Empirical work has reported gains in lexical choice, collocation awareness, and translator self-efficacy [4]. Two obstacles have nevertheless constrained classroom adoption: the steep learning curve of traditional concordancers, and the time cost of query formulation—both particularly acute for undergraduates in applied programs.

AI-enhanced corpus platforms mitigate these obstacles by allowing natural-language queries and by integrating search results with machine translation and polishing functions. Such affordances shift the pedagogical design question from “how to operate the tool” to “how to reason with the evidence the tool returns.” This reasoning skill—here termed translation verification competence—is the target of the teaching model reported below.

2.3. The Present Study

Existing studies of corpus-assisted translation teaching focus largely on general lexis and collocations; work specifically on culture-loaded terms is scarce, and classroom-embedded models designed for AI corpus platforms are scarcer still. The present case study therefore contributes (a) a five-step verification model operationalized on an AI corpus platform, and (b) illustrative evidence of its effects on student renderings of Chinese culture-loaded terms.

3. Teaching Design

3.1. The Platform: AI-Corpora

AI-Corpora is a web-based corpus platform developed for language learners and translators. Three of its functions are pedagogically central to this study. The Search function returns bilingual concordance lines drawn from large-scale Chinese and English text collections, allowing users to check how a term or phrase is actually rendered in authentic contexts. The Translate and Polish functions provide machine-translation drafts and stylistic refinement, which students can use to revise their own renderings after consulting the evidence [5]. A Memory function allows users to save verified term pairs to a personal term base, supporting cumulative terminology management across tasks.

3.2. A Five-Step Verification Model

To convert these platform functions into a teachable procedure, a five-step verification model was designed and posted as a one-page student guide (Table 1).

Table 1. The Five-Step Verification Model for Culture-Loaded Terms

Step	Student action	Platform function	Verification question
1. Independent drafting	Translate the term without any tool	—	What is my intuitive rendering?
2. Corpus search	Query the term and its candidate translations	Search	How is it rendered in authentic texts?
3. Evidence evaluation	Compare frequency, context and register of candidates	Search results	Which rendering is accepted, and in what contexts?
4. Guided revision	Revise the translation; refine the sentence	Translate / Polish	Does the revised version read naturally?
5. Term archiving	Save the verified pair with a usage note	Memory	What have I learned for future tasks?

Three disciplines accompany the model: (a) students must draft independently before consulting the platform, so that corpus use tests rather than replaces their own hypotheses; (b) every final decision must be justified by citing at least one concordance line as evidence; and (c) verified terms must be archived with a one-line usage note, so that verification produces a reusable personal term base.

3.3. Task Design

The model was implemented in a teaching unit within a one-semester translation course for third-year English majors ($n = 32$) at an applied university in Chengdu. Three task clusters of increasing difficulty were selected from course materials: (1) a cured-meat term (LaRou) in a Sichuan food text, representing terms with deceptive “obvious” equivalents; (2) teahouse-culture terms (GaiWanCha , PingShu) in a Chengdu local-culture text, representing terms requiring tier selection among transliteration-plus-explanation, conventional translation, and free translation; and (3) solar-term and festival names (LiDong , QingMing) in a seasonal customs text, representing terms with both official conventions and folk registers. For each cluster, students submitted a first-draft translation and a verified revision accompanied by evidence citations.

4. Classroom Cases and Analysis

4.1. LaRou — Escaping the “Obvious” Equivalent

In the first-draft submissions, 26 of 32 students rendered LaRou as bacon, presumably retrieving the nearest English meat category from memory; five used the literal dried meat; one used smoked pork. A corpus search for bacon in comparable food texts returned concordance lines in which bacon consistently co-occurred with breakfast, frying, and slices—collocational evidence that bacon denotes a specific cured pork product sliced thin and cooked before eating, typically at breakfast. Searching the Chinese term and English food-media texts yielded frequent occurrences of cured pork and lap

cheong/lap yuk for related Cantonese products, while preserved meat appeared in more general descriptive contexts.

Guided by the evidence, most students revised to Sichuan-style cured pork or cured pork (larou), with several adding a one-phrase gloss. Post-revision, 29 of 32 submissions used an evidence-supported rendering, and students' cited evidence showed correct reasoning about why bacon misleads: it activates an inappropriate frame (breakfast, frying pan) for a product that is air-cured and used as an ingredient. The case demonstrates the core value of verification: the "obvious" dictionary equivalent is precisely the candidate most in need of testing.

4.2. *GaiWanCha and PingShu* — *Choosing the Right Tier*

First drafts of GaiWanCha showed pinyin-only renderings (gaiwan tea) or awkward literalisms (covered-bowl tea); PingShu was commonly rendered as storytelling without qualification, losing its status as a named performing art. Corpus consultation showed gaiwan tea occurring in China-facing English media with explanatory appositions (e.g., tea served in a lidded bowl), supporting a transliteration-plus-explanation strategy for this non-conventionalized term. For PingShu, evidence supported pingshu storytelling or simply storytelling when the context made the art form explicit, illustrating how register and audience mediate tier choice. After revision, most students adopted gaiwan tea (tea in a lidded bowl) and pingshu storytelling, and—crucially—could articulate why a conventional translation such as Peking Opera was unavailable for these terms: the corpus returned no established English convention.

4.3. *LiDong and QingMing* — *Convention and Register*

The third cluster targeted terms with partially conventionalized renderings. Students easily retrieved the official convention for QingMing (the Qingming Festival / Tomb-Sweeping Day) but initially rendered LiDong variously as winter begins, start of winter, or lidong. Corpus evidence showed Lidong (the Beginning of Winter) in China Daily-style reporting and the winter solstice reserved for DongZhi—a contrast that clarified the boundary between a solar term with an established descriptive gloss and one requiring transliteration plus explanation. Students also learned to mark folk registers: the saying "ChiLeJiaoZiBuDongErDuo" was revised from a flat it is said that to folk wisdom has it that..., a formulation several students located verbatim in corpus lines and correctly judged as register-appropriate.

Across the three clusters, a consistent progression was visible in student work: first drafts relied on memory and produced pinyin-only or literal renderings; verified revisions adopted evidence-supported forms; and evidence citations became more selective, moving from copying the first concordance line to choosing lines that matched the target register.

5. Discussion

Three outcomes merit emphasis. First, verification training changed the locus of authority in the classroom: students began to ask "what does the evidence say?" rather than "what does the teacher say?", a shift that aligns with the discovery-learning rationale of corpus pedagogy [3]. Second, the model externalized the three-tier heuristic for culture-loaded terms: instead of memorizing strategy labels, students experienced tier selection as a consequence of evidence—transliteration-plus-explanation when no convention exists, conventional translation when the corpus attests one, and free translation when register demands it. Third, the archiving step converted single-task verification into cumulative terminology management; by semester's end, student term bases contained verified pairs with usage notes that were reused in later assignments.

Two difficulties were equally instructive. Some students over-trusted frequency, treating the most frequent rendering as automatically correct and ignoring register mismatches; others misread concordance lines, citing evidence from irrelevant genres (e.g., justifying a literary rendering with a technical-manual line). Both difficulties required teacher scaffolding in the form of mini-lessons on reading concordance lines: checking

source genre, co-text, and date. These observations suggest that AI corpus platforms do not remove the teacher's role but relocate it—from providing answers to training evidence literacy.

For similar teaching contexts, three recommendations follow: begin with deceptive “obvious” equivalents (such as LaRou /bacon), because the visible failure of intuition is the strongest motivation for verification; constrain early tasks to one term per sentence so that evidence lines are easy to interpret; and require evidence citation in every submission, since the citation habit is what distinguishes verification from mere tool use.

6. Conclusion

This case study has presented a five-step verification model for teaching the translation of Chinese culture-loaded terms with an AI corpus platform, and illustrated its classroom operation through three task clusters. Student outputs indicate that corpus-supported verification moves learners from intuition-based renderings toward accepted, register-appropriate translations, while fostering a reusable personal term base and a habit of evidence-based decision-making.

The study is limited by its single-institution design and its reliance on qualitative analysis of student work without a control group or pre-/post-test instrumentation. Future research will adopt a quasi-experimental design with parallel classes, incorporating quantitative measures of translation accuracy and retention of verified terms, and will extend the model from culture-loaded terms to sentence-level problems.

References

1. P. Newmark, *A Textbook of Translation*. New York: Prentice Hall, 1988.
2. M. Baker, “Corpus linguistics and translation studies: Implications and applications,” in *Text and Technology: In Honour of John Sinclair*, M. Baker, G. Francis, and E. Tognini-Bonelli, Eds. Amsterdam: John Benjamins, 1993, pp. 233–250.
3. S. Bernardini, “Corpora in the classroom: An overview and some reflections on future developments,” in *How to Use Corpora in Language Teaching*, J. M. Sinclair, Ed. Amsterdam: John Benjamins, 2004, pp. 15–36.
4. F. Zanettin, S. Bernardini, and D. Stewart, Eds., *Corpora in Translator Education*. Manchester: St. Jerome, 2003.
5. L. Venuti, *The Translator's Invisibility: A History of Translation*. London: Routledge, 1995.

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 Publisher and/or the editor(s). Publisher 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.