

Publication Date: 02 September 2026

Archs Sci. (2026) Volume 75, Issue 6 Pages 48-55, Paper ID 2025606.
<https://doi.org/10.68304/as/75606>

Context-Sensitive Translation of Chinese Four-Character Idioms in *To Live*: An Occurrence-Based Parallel-Corpus Study

Xinyi Cheng

College of English Language and Culture, Xi'an Fanyi University, Xi'an, Shaanxi 710105, China

Corresponding author: Xinyi Cheng (e-mail: look-you@163.com).

Abstract Chinese four-character idioms are highly condensed conventionalized expressions whose meanings may combine lexical content, figurative imagery, pragmatic force, and culturally specific associations. Their translation therefore provides a useful test case for examining how literary translators negotiate source-text imagery, contextual meaning, and target-language readability. This study investigates four-character idioms in Yu Hua's *Huozhe* and Michael Berry's English translation *To Live* through an occurrence-based parallel-corpus design combining quantitative frequency analysis with qualitative textual interpretation. A specialized corpus containing 153 distinct source-text idioms and 229 aligned translation occurrences was analyzed. To avoid overlap between broad labels such as "literal" and "free" translation, each occurrence was assigned to one of four mutually exclusive operational categories: literal translation, paraphrase, idiomatic substitution, or omission. The corpus coding yielded 120 literal translations (52.40%), 94 paraphrases (41.05%), 8 idiomatic substitutions (3.49%), and 7 omissions (3.06%). Representative, non-borderline source–target pairs were then interpreted with reference to the linguistic, cultural, and communicative dimensions of eco-translatology. The results show that the dominant pattern lies between retention of recognizable source imagery and direct semantic restatement in ordinary English. Literal translation is favored where source imagery remains interpretable after necessary grammatical adaptation, whereas paraphrase is used where direct transfer would risk obscuring the immediate narrative meaning. Idiomatic substitution is comparatively rare under the adopted coding criteria because it requires a functionally comparable conventional English expression, while omission is a selective category reserved for alignments in which no independent target rendering is required for the relevant contribution to remain recoverable from context. Recurring idioms further show that the same source expression can receive different English realizations when its grammatical role, collocational environment, or discourse function changes. The study contributes a transparent occurrence-level classification of idiom translation in *To Live* and demonstrates how quantitative strategy patterns can be connected to context-sensitive multidimensional interpretation without treating any single strategy as intrinsically superior.

Index Terms *To Live*, four-character idioms, Chinese–English literary translation, translation strategies, eco-translatology, parallel corpus, occurrence-based analysis

I. Introduction

Literary translation is not limited to lexical replacement; it mediates between linguistic systems, cultural knowledge, textual conventions, and reader expectations [1]. This mediation becomes particularly visible in the translation of idioms, whose meanings may depend simultaneously on conventionalized form, figurative imagery, discourse function, and cultural association [2], [4]. Chinese four-character idioms are particularly relevant because their formal compactness can package semantic, rhetorical, and cultural information that is not always recoverable through word-for-word substitution [3]. Translators must consequently decide whether a source expression can retain its imagery in English, whether its contextual meaning should be restated, whether a conventional target-language expression can perform a comparable function, or whether part of the source wording can remain im-

plicit because the surrounding translation already conveys its contribution. These alternatives make four-character idioms a productive object for examining the interaction between formal preservation, semantic accessibility, cultural specificity, and literary context [5].

In the present study, the term *four-character idiom* is used operationally for dictionary-attested or otherwise lexicographically recognized conventionalized four-character units identified in the Chinese source text. The term does not refer to arbitrary four-character strings, nor does it assume that every retained item belongs to a single narrowly defined historical subtype of Chinese idiomatic expression. This operational definition is important because the purpose of the corpus is not to classify the full lexicographic taxonomy of Chinese phraseology; rather, it is to investigate how a consistently screened set of compact conventionalized expressions is rendered in an

English literary translation.

Yu Hua's *Huozhe* was first published in book form in 1993 [6], and Michael Berry's English translation, *To Live*, was published by Anchor Books in 2003 [7]. The novel provides a productive site for translation analysis because culturally embedded language is woven into everyday speech, family relations, rural experience, social change, humor, grief, and narrative characterization. Existing scholarship has approached *To Live* from several translation-related perspectives. Zhao examines the translation of Chinese passive constructions through quantitative and qualitative analysis [13]; Zhang and Sun study cultural vacancy in Berry's translation from the perspective of relevance theory [15]; and Wang, Amini, and Tan Ai Lin analyze loss, gain, and equivalence in the translation of culture-loaded words [16]. Zhang and Lou apply eco-translatology to culture-loaded words in Portuguese translations of the same novel [14], while Chen applies eco-translatology to Berry's English translation and emphasizes adaptive choices across linguistic, cultural, and communicative dimensions [17]. These studies establish that *To Live* has already attracted sustained attention as a site of intercultural mediation; therefore, the contribution of the present study lies not in claiming an unexplored text, but in narrowing the analytical object and making the strategy classification more explicit.

Research on Chinese idiom translation provides a second line of relevant scholarship. Kang and Yang review major difficulties and strategies in the English translation of Chinese four-character idioms [3]. Zhu conducts a corpus-oriented study of four-character idioms in Chinese opera and emphasizes context-sensitive use of literal, free, abridged, and borrowing strategies [12]. Corpus-based translation studies more generally provide a methodological basis for moving beyond isolated examples by examining aligned recurrent evidence and making the analytical unit explicit [8], [9], [18]. Research methodology in translation studies likewise stresses the importance of matching claims to the evidence generated by the selected design [19]. Taken together, this literature suggests two recurring methodological challenges: strategy labels may overlap across studies, and a distinct idiom may be treated as though it always receives the same translation strategy even when its individual occurrences function differently in context.

The present study addresses these problems by focusing specifically on four-character idioms in *To Live* and distinguishing four mutually exclusive operational categories at the occurrence level: literal translation, paraphrase, idiomatic substitution, and omission. The design separates a *distinct idiom* from an *occurrence* of that idiom. This distinction is necessary because the same lexical item may recur in different narrative environments and may be translated differently when its grammatical role, collocational pattern, or discourse function changes. The study therefore combines a descriptive quantitative account of strategy frequencies with a qualitative explanation of representative cases.

The study addresses the following research questions:

- 1) What is the distribution of literal translation, paraphrase, idiomatic substitution, and omission across the 229 aligned translation occurrences in the specialized corpus?
- 2) How can representative translation choices be interpreted across the linguistic, cultural, and communicative dimensions of eco-translatology?
- 3) What kinds of context-sensitive variation are observable when recurring four-character idioms receive different English treatments?

II. Theoretical and Analytical Framework

Eco-translatology conceptualizes translation as a process of adaptation and selection within a translational environment that includes linguistic, cultural, communicative, and contextual constraints [10]. Hu's later systematic account develops this framework through the concepts of selective adaptation, adaptive selection, and multidimensional transformation [11]. In the present study, eco-translatology is used as an interpretive framework rather than as a substitute for operational coding. The four strategy categories describe the observable relationship between an aligned source expression and its English treatment, whereas the three eco-translatological dimensions are used to explain how that treatment functions in its local context. Maintaining this distinction reduces the risk of using broad theoretical terminology as a replacement for explicit textual criteria.

The linguistic dimension concerns lexical choice, grammatical restructuring, rhetorical form, imagery, collocation, and target-language fluency. The cultural dimension concerns whether a source-associated image or cultural implication is preserved, modified, replaced, made less explicit, or omitted. The communicative dimension concerns pragmatic force, local coherence, narrative intelligibility, and the degree to which the translated wording appears textually transparent in its immediate context. Because the present study does not include a reader-response experiment, statements about accessibility or interpretive effort are treated as text-based analytical judgments rather than direct measurements of reader cognition. This distinction keeps the theoretical interpretation consistent with the evidence actually generated by the corpus analysis.

III. Corpus and Analytical Procedure

A. Corpus Scope and Unit of Analysis

The study compares the Chinese source text of Yu Hua's *Huozhe* with Michael Berry's English translation *To Live* [6], [7]. A specialized parallel corpus was constructed by reading the Chinese source text, identifying candidate four-character expressions, checking their conventionalized idiomatic status against lexicographic resources, and aligning each confirmed source occurrence with its English treatment in the corresponding narrative context. Idiom status was cross-checked against identifiable reference works including the *Hanyu Chengyu Cidian* (*Dictionary of Chinese Idioms*) [20], the *Xin Shiji Hanyu Chengyu Cidian* (*New Century Dictionary of Chinese Idioms*) [21], and the second edition of the *Hanyu*

Da Cidian [22]. Ordinary four-character strings that did not function as conventionalized expressions were excluded. The use of several lexicographic resources reduces dependence on any single dictionary’s inclusion policy while keeping corpus membership tied to explicit lexical evidence.

The resulting dataset contains 153 distinct idiom types and 229 aligned translation occurrences. The occurrence, rather than the distinct idiom type, is the primary quantitative unit. Repeated appearances of the same source idiom were retained as separate observations because an expression can change grammatical role, discourse function, collocational environment, or contextual implication across the narrative. Each occurrence was evaluated as a source–target alignment rather than inferred from the idiom type alone. Distinguishing type-level and occurrence-level information therefore prevents a recurrent idiom from being assigned a single permanent strategy when its target treatment varies. This procedure is consistent with corpus-based translation research emphasizing clearly defined units, aligned evidence, and recurrent patterns [8], [9], [18].

For analytical purposes, an occurrence was evaluated in relation to the smallest surrounding narrative context needed to determine the function of the source expression and the contribution of the English rendering. Where no independent target-language expression corresponded lexically to the source idiom, the surrounding English passage was still examined so that omission could be distinguished from translation by paraphrase or semantic redistribution. The resulting analysis is descriptive: it reports how the coded corpus was classified and then interprets selected cases, rather than claiming that the observed percentages represent Chinese–English literary translation as a whole.

B. Strategy Coding

Each occurrence was assigned to one of four mutually exclusive strategy categories. Table 1 summarizes the operational decision rules. The categories were designed to distinguish formal-image retention, non-idiomatic semantic restatement, target-language idiom replacement, and absence of an independent lexical rendering.

Table 1: Operational criteria used to classify translation occurrences

Strategy	Operational decision rule
Literal translation	The English rendering retains the principal contextual meaning together with a recognizable component of the source structure, image, or rhetorical configuration. Necessary grammatical or lexical adjustment does not by itself exclude the occurrence from this category.
Paraphrase	The contextual meaning is restated in ordinary English, while the distinctive source image or idiomatic form is not independently reproduced as such.
Idiomatic substitution	An established English idiom, proverb, or conventional phrase is used to perform a semantic and pragmatic function comparable to that of the source idiom.
Omission	The source idiom receives no independent lexical rendering because its contribution is recoverable from nearby target-text material or because an additional rendering would largely repeat information already expressed.

The coding hierarchy was applied as follows. If the recognizable source image or rhetorical configuration remained present, the occurrence was classified as literal translation even when English grammar required adjustment. If the source image was not retained and the meaning was expressed through ordinary non-idiomatic English, the occurrence was classified as paraphrase. If the source item was replaced by an established conventional English expression, the occurrence was classified as idiomatic substitution. Omission was reserved for cases in which no independent target expression represented the source idiom and the relevant meaning was recoverable from the surrounding target passage. This hierarchy prevents the same occurrence from being counted simultaneously under more than one strategy.

C. Quantitative and Qualitative Analysis

The quantitative stage counted all coded occurrences and calculated each strategy as a proportion of the 229-occurrence total. Category-specific counts of distinct idioms were also retained to show how many different lexical items appeared under each strategy; unlike occurrence counts, these type-level counts can overlap because the same recurring idiom may receive different strategies in different contexts. The qualitative stage then examined representative occurrences in their narrative environments and interpreted the relevant linguistic, cultural, and communicative adjustments.

The study uses explicit coding rules to improve analytical transparency, but the classification remains interpretive, especially near the boundary between literal translation and paraphrase. The analysis was conducted as a descriptive single-coder study; therefore, no inter-coder reliability coefficient is reported or implied. The resulting frequencies should consequently be interpreted as the outcome of the operational scheme applied in this case study rather than as an independently validated annotation benchmark. Likewise, no reader-response experiment was conducted, so observations concerning target-language accessibility refer to textual properties and contextual interpretability rather than directly measured reader behavior.

IV. Translation Strategies for Four-Character Idioms in *To Live*

The corpus contains 153 distinct four-character idiom types represented by 229 translation occurrences. Because some idioms recur in different contexts, the number of occurrences exceeds the number of distinct expressions. In addition, the category-specific counts of distinct idioms overlap because a recurring idiom can appear under more than one strategy when its English treatment changes across contexts. Occurrence counts, by contrast, are mutually exclusive. As shown in Table 2, the coded corpus contains 120 instances of literal translation (52.40%), 94 instances of paraphrase (41.05%), 8 instances of idiomatic substitution (3.49%), and 7 instances of omission (3.06%). The four occurrence counts sum to 229 and the percentages sum to 100.00% after rounding to two decimal places. Literal translation and paraphrase jointly

account for 93.45% of the aligned occurrences, indicating that the dominant pattern lies between retention of recognizable source imagery and direct semantic restatement in ordinary English. The examples below are used to explain unambiguous manifestations of the coding criteria; they do not constitute a second recoding of the full dataset.

Table 2: Distribution of translation strategies in the 229-occurrence corpus

Strategy	Distinct types ^a	<i>N</i>	%
Literal translation	85	120	52.40%
Paraphrase	67	94	41.05%
Idiomatic substitution	8	8	3.49%
Omission	7	7	3.06%
Total	153 ^b	229	100.00%

^aThe category-specific distinct-idiom counts are not additive because a recurring idiom can appear under different strategies in different contexts. ^bThe corpus contains 153 unique idiom types overall.

A. Literal Translation

Literal translation is used here in a functional rather than mechanically word-for-word sense. A rendering belongs to this category when it preserves the principal contextual meaning together with a recognizable component of the source image, structure, or rhetorical organization. Necessary changes in English syntax or lexical selection do not by themselves make the translation non-literal. This interpretation is consistent with broader accounts of idiom translation that distinguish preservation of semantic and figurative content from strict formal copying [3], [4].

Example 1. Source idiom: *ben niao xian fei*, literally “the clumsy bird flies first” [6]. Berry’s rendering is: “I still need to fly more, but as they say, the clumsy bird flies first” [7].

The source expression presents early effort through an avian metaphor: a less capable bird compensates by beginning before others. In the narrative, Fugui uses the image self-deprecatingly while also expressing persistence and determination. The English rendering preserves both the bird image and the contrast between clumsiness and early effort. Linguistically, the metaphor is accommodated within ordinary English syntax; culturally, the source image remains visible rather than being replaced by an unrelated target expression; and communicatively, the surrounding context makes the practical implication inferable without explanatory expansion. The occurrence therefore satisfies the operational criterion for literal translation.

Example 2. Source idiom: *duan zi jue sun*, literally “to leave no sons or descendants” [6]. Berry’s rendering includes: “In the Xu family, I am the lone remaining flame. The Xu family won’t have any sons or grandchildren if I leave” [7].

The source expression invokes the extinction of a family line. The target passage retains explicit kinship reference to sons and grandchildren and supplements it with the metaphor of a “lone remaining flame.” Linguistically, the source meaning is restructured into an accessible English formulation; culturally, the importance attached to lineage

remains salient; and communicatively, the emotional pressure of possible family extinction is conveyed without an extended explanatory gloss. The example shows that literal translation, as operationalized here, can include grammatical and lexical adaptation while preserving the source expression’s central conceptual image and cultural implication.

Example 3. Source idiom: *shan zhen hai wei*, literally “mountain delicacies and sea flavors” [6]. Berry’s rendering is: “Back then, delicacies from the mountains and seas could be bought with a bag of rice” [7].

The source expression refers to rare or luxurious foods and organizes its image through the paired domains of mountains and seas. Because this configuration remains interpretable in English, the translation preserves the source image while selecting the natural English head noun “delicacies.” Under the operational criteria adopted in this study, the occurrence is classified as literal translation rather than idiomatic substitution: the source image remains recognizable and is not replaced by a separate established English idiom. Linguistically, the wording is adapted for fluency; culturally, the mountain-and-sea image remains perceptible; and communicatively, the meaning is recoverable from the immediate context.

B. Paraphrase

Paraphrase gives priority to contextual meaning when direct transfer of the source form or image would be awkward, opaque, or unnecessarily marked in English. Unlike idiomatic substitution, paraphrase uses ordinary target-language wording rather than replacing the source expression with an established English idiom. Baker identifies paraphrastic treatment as an important response when idiomatic form cannot be carried across without distortion or unnaturalness [4].

Example 4. Source idiom: *zuo chi shan kong*, literally “to sit and consume until the mountain is empty” [6]. Berry’s rendering is: “After another two or three months, all the money will have been spent” [7].

The source idiom figuratively represents the exhaustion of accumulated resources through continued consumption. In the local narrative context, the relevant consequence is the depletion of available money. The English sentence states that consequence directly rather than reproducing the mountain image. Linguistically, figurative packaging is replaced by an ordinary declarative formulation; culturally, the source-specific metaphor becomes less visible; and communicatively, the immediate financial consequence is foregrounded. The occurrence therefore exemplifies paraphrase rather than literal translation or idiomatic substitution.

Example 5. Source idiom: *quan pan tuo chu*, literally “to bring out the whole tray,” conventionally meaning to disclose everything [6]. Berry’s rendering is: “There was never a man who told me everything I wanted to know as he did” [7].

The target sentence communicates openness and completeness without reproducing the source expression’s figurative packaging. The wording is ordinary English and does not

present itself as a target-language idiom. Linguistically, the expression is reformulated; culturally, the source metaphor is backgrounded; and communicatively, the interpersonal meaning remains explicit. The occurrence therefore fits the operational definition of paraphrase.

C. Idiomatic Substitution

Idiomatic substitution is reserved for cases in which the target text employs an established English idiom, proverb, or conventional phrase that performs a comparable semantic and pragmatic function. Because sufficiently close conventional correspondences are limited across languages and cultures, this strategy is much less frequent than literal translation or paraphrase in the present corpus.

Example 6. *Source idiom: huo bu dan xing, literally “misfortune does not come alone” [6]. Berry’s rendering is: “As they say, misfortunes never come singly. Jiazhen is already so ill; my only hope is that Youqing will recover soon” [7].*

The source idiom expresses the accumulation of adversity. The English proverb “misfortunes never come singly” is itself a conventional target-language expression with a closely comparable semantic and pragmatic function. The source wording is not reproduced structurally; instead, an established English formulation supplies the proverbial force. Linguistically, one conventional expression replaces another; culturally, source-specific form gives way to a target-language conventional resource; and communicatively, the cumulative sense of adversity remains prominent. This occurrence is therefore a clear example of idiomatic substitution.

D. Omission

Omission is used only when a source idiom receives no independent lexical rendering and its relevant semantic or pragmatic contribution is recoverable from the surrounding target text, or when an additional rendering would substantially repeat information already expressed. Omission is therefore an alignment-level category and should not be inferred merely because the target sentence is shorter or structurally different. In the coded corpus, 7 of 229 occurrences (3.06%) meet this operational criterion, making omission the least frequent strategy alongside the similarly infrequent idiomatic-substitution category.

A stand-alone target quotation is insufficient to demonstrate an absence claim reliably: an omission judgment requires inspection of the complete aligned source occurrence and its surrounding target context. For that reason, this qualitative section does not present a borderline target passage as if it independently proved omission. The seven omission-coded alignments remain included in the quantitative distribution in Table 2, and the complete coded dataset described in the Data Availability statement provides the appropriate level of evidence for row-level verification. This treatment avoids conflating omission with compressed translation, semantic redistribution, or paraphrase.

E. Contextual Variation in Recurring Idioms

Occurrence-level analysis is particularly useful when a source idiom recurs because it prevents the analyst from assigning one fixed realization to a lexical type regardless of changing grammatical or discourse function. The following paired cases address Research Question 3 as qualitative evidence of context-sensitive variation; they are not used to estimate the frequency of every possible strategy transition in the corpus.

Example 7. *The recurrent source expression mei wan mei liao, literally “without finish or end,” occurs in contexts involving both continuing weather and repeated speech [6]. Berry renders the relevant meanings as, respectively, “The unceasing rain began while the rice ears were still green and only beginning to fill” and “As I hurried her to bed, she kept insisting that she had fallen asleep” [7].*

In the first context, the expression modifies continuous rain, and “unceasing” provides a relatively direct lexical realization of duration. In the second, the same source expression relates to repeated speech, for which the verbal construction “kept insisting” is more natural in English than mechanically repeating an adjectival equivalent. The contrast shows that grammatical role and collocational environment can alter the form of the English realization even when the recurrent source expression is unchanged. The first occurrence retains the durative quality lexically, whereas the second foregrounds iteration through grammatical restructuring.

Example 8. *The recurrent source expression xi xi ha ha evokes cheerful laughter or lighthearted behavior [6]. In one context Berry foregrounds Youqing’s behavioral state in the wording “Youqing remained carefree,” while in another the translation uses the laughter-related wording “The people in the town laughed heartily” [7].*

The two renderings highlight different aspects licensed by the same source expression. The first makes the contextual implication of carefree behavior explicit, while the second foregrounds the audible action of laughter. Both are meaning-oriented English realizations, but their lexical focus differs in response to the immediate narrative environment. The example therefore supports an occurrence-level account of contextual variation without requiring the stronger claim that every recurrent idiom necessarily changes strategy category.

Table 3 summarizes how the representative evidence connects the strategy categories and contextual variation with the three eco-translatological dimensions. The table is interpretive rather than a second coding system: strategy labels identify source–target relations, whereas the three dimensions describe the principal textual consequences of those relations.

V. Discussion

The distribution in Table 2 shows a pronounced concentration in literal translation and paraphrase, which together account for 93.45% of the 229 coded occurrences. This concentration is more informative than a simple opposition between “literal” and “free” translation. Literal translation, as operationalized

Table 3: Eco-translatological interpretation of representative strategy types

Examples	Strategy	Linguistic dimension	Cultural dimension	Communicative dimension
1–3	Literal translation	Source imagery or rhetorical organization remains recognizable after grammatical and lexical adaptation.	Source-associated imagery or culturally marked concepts remain comparatively visible.	Context supports interpretation without requiring extensive explanation.
4–5	Paraphrase	Source idiomatic form is replaced by ordinary English wording that states contextual meaning directly.	Some source-specific imagery becomes less explicit.	Immediate narrative meaning is foregrounded and textual transparency is increased.
6	Idiomatic substitution	A conventional English proverb or idiom replaces the source formulation.	Source form is replaced by a target-language conventional resource.	Comparable pragmatic force is conveyed through a familiar English expression.
Aggregate category	Omission	No independent lexical rendering is assigned where the aligned target context already carries the relevant contribution.	The source idiom's specific form is not independently retained.	Contextual recoverability permits reduction without requiring a redundant separate rendering.
7–8	Contextual variation	Grammatical role, collocation, and discourse function motivate different English realizations of recurrent items.	The degree and manner of source-form visibility can change across occurrences.	Local narrative requirements shape which aspect of meaning is foregrounded.

here, permits grammatical adaptation while retaining recognizable source imagery, whereas paraphrase removes the distinctive idiomatic packaging and states contextual meaning in ordinary English. Separating these categories makes it possible to distinguish image retention from semantic restatement without treating target-language idiomatic substitution as merely another form of “free translation.”

The pattern is compatible with research showing that Chinese four-character idioms require flexible combinations of strategies rather than a single default procedure [3], [12]. At the same time, the present findings refine that observation by showing that established English idiom replacement is comparatively uncommon in this corpus: only 8 of 229 occurrences (3.49%) were classified as idiomatic substitution. Within the study's coding logic, this category requires more than approximate semantic similarity; the English expression must also function as a recognizable conventional phrase with comparable pragmatic force. Literal translation and paraphrase are applicable across a broader range of contexts because they do not require a preexisting conventional match between the two languages.

Omission is similarly infrequent, accounting for 7 occurrences (3.06%). Its interpretation requires particular caution because omission is an absence claim at the alignment level. A shorter target sentence does not by itself establish omission, and semantic compression or redistribution must first be excluded. Reserving the category for cases in which no independent target realization is required and the relevant contribution remains contextually recoverable makes the four-category scheme more conservative and reduces the risk of counting paraphrase as omission. The quantitative result therefore indicates that explicit source-to-target realization is overwhelmingly more common than complete non-realization of an idiom as an independent unit in this corpus.

From an eco-translatological perspective, the strategy distribution can be interpreted as a series of context-sensitive trade-offs rather than a hierarchy in which one strategy is inherently superior. As summarized in Table 3, literal translation is most viable when linguistic restructuring does not make the source-associated image opaque; paraphrase becomes appropriate when immediate narrative meaning can be communicated more transparently without the original idiomatic form; idiomatic substitution depends on the availability of a functionally comparable target convention; and omission

depends on strong contextual recoverability [10], [11]. The linguistic, cultural, and communicative dimensions therefore interact: preserving cultural imagery may require linguistic adaptation, while maximizing local fluency may reduce the visibility of a source image. The framework is useful because it explains these interactions without converting them into a single universal rule.

The findings also complement previous studies of *To Live*. Research on passive constructions, culture-loaded words, cultural vacancy, and adaptive selection has already shown that Berry's translation repeatedly negotiates source-specific linguistic and cultural information with English-language textual requirements [13]–[17]. The present analysis adds a narrower occurrence-based perspective by focusing specifically on four-character idioms, distinguishing 153 distinct lexical types from 229 translation occurrences, and applying mutually exclusive operational definitions. This narrower focus avoids overstating novelty while providing a more precise account of one recurrent linguistic feature of the novel.

The recurrent examples are methodologically important for a related reason. A type-level inventory alone can conceal variation because it encourages an analyst to associate one idiom with one preferred realization. Examples 7 and 8 show that the same source expression can be realized differently when its grammatical role, local collocation, discourse function, or narratively foregrounded meaning changes. The evidence presented here is qualitative and does not estimate the prevalence of every possible transition among strategy categories. It nevertheless demonstrates why occurrence-level coding is appropriate for the research questions of this case study and why recurrent types should not automatically inherit a single strategy label from one example.

Several limitations define the appropriate scope of the conclusions. First, the corpus is drawn from one novel and one English translation, so the percentages are descriptive of this dataset and should not be generalized to Chinese literary translation as a whole. Second, strategy boundaries, especially between literal translation and paraphrase, involve interpretive judgment even when explicit rules are used. Because the study uses a single-coder design, no inter-coder agreement statistic is available; independent recoding would be required to establish annotation reliability. Third, the study does not directly measure target-reader comprehension, processing effort, or aesthetic response, so statements about accessibility

remain text-based interpretations. Fourth, although the complete coded alignment is available from the corresponding author, the article itself presents selected examples rather than all 229 occurrence-level records. These limitations constrain the scope of inference but do not alter the descriptive purpose of the present case study.

VI. Conclusion

This study provides an occurrence-based parallel-corpus account of how Chinese four-character idioms are translated in Michael Berry's English version of Yu Hua's *To Live*. By distinguishing 153 distinct idiom types from 229 aligned occurrences and applying four mutually exclusive operational categories, the analysis addresses a methodological problem that can arise when broad labels such as "free translation" are used inconsistently. In the coded corpus, literal translation is the most frequent strategy (52.40%), followed by paraphrase (41.05%), while idiomatic substitution (3.49%) and omission (3.06%) are comparatively rare. The numerical distribution therefore places most occurrences between retention of recognizable source imagery and direct restatement of contextual meaning.

The qualitative analysis shows that these strategies are best understood as context-sensitive choices rather than as a hierarchy of correct and incorrect procedures. Literal translation is effective when source imagery remains sufficiently interpretable after necessary English restructuring. Paraphrase is useful when preserving the source idiomatic form would make the immediate narrative meaning less transparent. Idiomatic substitution depends on the availability of an established English expression with comparable semantic and pragmatic force, whereas omission should be assigned only when alignment-level inspection shows that no independent rendering is required because the relevant contribution remains recoverable from context. Eco-translatology provides a useful explanatory framework for these choices because linguistic form, cultural representation, and communicative intelligibility interact rather than operate independently.

A key methodological implication is that recurring idioms can profitably be analyzed at the occurrence level rather than assigned a single fixed realization on the basis of type identity. The paired cases examined here show that grammatical role, collocation, discourse function, and locally salient meaning can motivate different English treatments of the same recurrent source expression. This conclusion is intentionally limited to the evidence generated by the present case study. Future research should test the stability of the four-category scheme across additional Chinese novels and translators, quantify strategy stability and switching among recurrent idioms, incorporate independent coding and agreement measures, and use reader-response evidence to examine whether textually inferred differences in interpretive burden correspond to actual target-reader comprehension. Comparative work across multiple translations of the same source text would also help distinguish translator-specific preferences from broader tendencies in Chinese–English literary translation.

Data Availability

Representative source–target examples and aggregate quantitative results are presented in this article. The complete coded dataset containing the 229 aligned translation occurrences is available from the corresponding author upon reasonable request for purposes of verification and scholarly reuse, subject to applicable copyright restrictions on the underlying literary texts.

Conflicts of Interest

The author declares no conflicts of interest.

Funding

This research was supported by the 2026 Annual School-Level Research Project of Xi'an Fanyi University, "Research on the English Translation of Shaanxi Shadow Puppetry Culture from the Perspective of Eco-Translatology" (Project No. 2026B64).

References

- [1] J. Munday, S. Ramos Pinto, and J. Blakesley, *Introducing Translation Studies: Theories and Applications*, 5th ed. London, U.K.: Routledge, 2022.
- [2] C. Fernando, *Idioms and Idiomaticity*. Oxford, U.K.: Oxford University Press, 1996.
- [3] H. Kang and Y. Yang, "A study on English translation of Chinese four-character idioms: Strategies and problems," *Linguistics and Culture Review*, vol. 6, no. 1, pp. 200–213, 2022, doi: 10.21744/lingcure.v6n1.2185.
- [4] M. Baker, *In Other Words: A Coursebook on Translation*, 3rd ed. London, U.K.: Routledge, 2018.
- [5] J. Franco Aixelá, "Culture-specific items in translation," in *Translation, Power, Subversion*, R. Álvarez and M. C.-Á. Vidal, Eds. Clevedon, U.K.: Multilingual Matters, 1996, pp. 52–78.
- [6] Y. Hua, *Huozhe [To Live]*. Wuhan, China: Changjiang Literature and Art Publishing House, 1993.
- [7] Y. Hua, *To Live*, M. Berry, Trans. New York, NY, USA: Anchor Books, 2003.
- [8] 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, The Netherlands: John Benjamins, 1993, pp. 233–250.
- [9] M. Baker, "Corpora in translation studies: An overview and some suggestions for future research," *Target*, vol. 7, no. 2, pp. 223–243, 1995, doi: 10.1075/target.7.2.03bak.
- [10] G. Hu, "Translation as adaptation and selection," *Perspectives: Studies in Translation Theory and Practice*, vol. 11, no. 4, pp. 283–291, 2003, doi: 10.1080/0907676X.2003.9961481.
- [11] G. Hu, *Eco-Translatology: Towards an Eco-Paradigm of Translation Studies*. Singapore: Springer, 2020.
- [12] X. Zhu, "Translation strategies of four-character idioms in Xiqu texts," *Argumentum*, vol. 20, pp. 541–557, 2024, doi: 10.34103/ARGUMENTUM/2024/32.
- [13] B. Zhao, "The use and translation of Chinese passive voice into English: A case study based on *To Live* by Yu Hua (1993) and its English translation," *British Journal of Applied Linguistics*, vol. 2, no. 2, pp. 10–19, 2022, doi: 10.32996/bjal.2022.2.2.2.
- [14] J. Zhang and Z. Lou, "Translation of culture-loaded words in the novel *Huozhe (To Live)* from the perspective of eco-translatology," *Cadernos de Tradução*, vol. 43, no. 1, pp. 1–22, 2023, doi: 10.5007/2175-7968.2023.e86532.
- [15] H. Zhang and B. Sun, "A translation study of cultural vacancy in the book *To Live* under the guidance of relevance translation theory," *Studies in English Language Teaching*, vol. 12, no. 1, pp. 76–86, 2024, doi: 10.22158/selt.v12n1p76.
- [16] Q. Wang, M. Amini, and D. Tan Ai Lin, "Loss, gain, and equivalence in the translation of culture-loaded words in *To Live: A Novel*," *International Journal of Language, Literacy and Translation*, vol. 8, no. 1, pp. 17–34, 2025, doi: 10.36777/ijolt2025.8.1.117.

- [17] S. Chen, "Adaptation and selection: A study on the eco-translational writing of Michael Berry's English translation of the novel *To Live*," in *Proceedings of the 2026 12th International Conference on Digital Humanities and Frontiers in Social Sciences (DHFSS 2026)*, Atlantis Press, 2026, pp. 1006–1017, doi: 10.2991/978-2-38476-611-6_109.
- [18] S. Laviosa, *Corpus-based Translation Studies: Theory, Findings, Applications*. Amsterdam, The Netherlands: Rodopi, 2002.
- [19] G. Saldanha and S. O'Brien, *Research Methodologies in Translation Studies*. London, U.K.: Routledge, 2014.
- [20] *Hanyu Chengyu Cidian [Dictionary of Chinese Idioms]*, Y. Jinchun and S. Mengmei, Chief Eds. Beijing, China: Shangwu Yinshuguan Guoji Youxian Gongsì, 2004.
- [21] W. Zongwen, Y. Zongyi, X. Luo, and H. Zheng, Eds., *Xin Shiji Hanyu Chengyu Cidian [New Century Dictionary of Chinese Idioms]*. Chengdu, China: Sichuan Cishu Chubanshe, 2006.
- [22] Z. Luo, Chief Ed., *Hanyu Da Cidian [Comprehensive Chinese Word Dictionary]*, 2nd ed. Shanghai, China: Hanyu Da Cidian Chubanshe, 2001.

• • •