From Detection to Revision: Identifying Coherence Errors in Chinese–English MT of Journalism via Thematic Progression

Chao Tang
New Oriental Education & Technology Group

Author Note
The author declares no conflicts of interest to disclose. Correspondence concerning this article should be addressed to Chao Tang, New Oriental Education & Technology Group, Guangzhou, Guangdong 510055, China.
Email: 617236364@qq.com

Abstract

Machine translation (MT) of Chinese-English journalistic texts frequently suffers from discourse-level coherence errors that are undetectable by sentence-level metrics. This study proposes a theory-driven method to identify and revise such errors using thematic progression (TP) as a diagnostic framework. Taking the machine-translated version of *Ideas of Journalism in Contemporary China* produced by ChatGPT as an example, we extracted 156 clauses and identified two dominant TP patterns in the source text (constant theme and derived theme), and documented pattern-specific MT errors as follows: ambiguous reference and improper thematic shift under the constant pattern, and cohesion loss and information focus shift under the derived pattern. Based on these error types, we develop three post-editing strategies specification, combination, and amplification--each designed to repair a specific coherence failure while preserving the original TP structure. The findings demonstrate that TP theory provides an objective, replicable heuristic for detecting coherence errors in MT output and guiding targeted revision. This study contributes a practical framework for post-editing training and offers implications for discourse-aware MT evaluation.
Keywords: thematic progression, coherence errors, Chinese-English machine translation, journalism texts, post-editing strategies

1. Introduction

Journalistic discourse, as a functional genre, exhibits distinctive patterns of information organization. Unlike literary or conversational texts, news discourse prioritizes the efficient transmission of factual information, often employing structured formats such as the inverted pyramid, where information is arranged in descending order of importance. This organizational

Received 4 January 2026; revised 16 February 2026; accepted 3 March 2026; available online 25 March 2026; Version of Record 30 March 2026.

Citation: Tang, C. (2026). From detection to revision: Identifying coherence errors in Chinese–English MT of journalism via thematic progression. *Journal of Language*, *2*(1), 132-147. https://doi.org/10.64699/26TEZN9415

logic, however, does not imply loose coherence; rather, news discourse relies on well-defined thematic structures to guide readers through the logical progression of events and arguments. The coherence of a news text is not merely a matter of grammatical cohesion but is fundamentally driven by how themes and rhemes connect across clauses and paragraphs. Consequently, when such texts are translated, preserving the original coherence becomes as critical as lexical and syntactic accuracy.

The challenge of maintaining coherence is particularly acute in Chinese–English translation. Chinese is a topic-prominent language, where the topic often appears at the clause-initial position and the subject may be omitted (zero anaphora), relying on implicit semantic relations to maintain discourse flow (Li & Thompson, 1981). English, by contrast, is subject-prominent, requiring explicit subjects and overt cohesive devices such as pronouns, conjunctions, and lexical repetition. This structural divergence means that when translating from Chinese to English, the translator—whether human or machine—must constantly infer implicit subjects and thematic relations from the source text and render them in the explicit grammatical structures required by English. Moreover, empirical studies have shown that while Chinese and English share similar types of cohesive devices, their frequency of use differs significantly across the two languages. Therefore, translation requires systematic adjustments to these devices to approximate the original text and avoid awkward "translationese" (Weng & Wang, 2020).

Machine translation (MT), despite recent advances in neural models and large language models such as ChatGPT, continues to struggle with discourse-level coherence. When processing Chinese journalistic texts, MT systems frequently produce sentences that are grammatically correct but discursively fragmented. Taking the machine-translated version of *Ideas of Journalism in Contemporary China* as an analytical example—a representative academic journalistic text that systematically traces the historical development of Chinese journalistic thought—we identified four recurring types of coherence errors in ChatGPT's output: ambiguous reference, improper thematic shift, cohesion loss, and information focus shift. These errors are not random; they systematically arise from the MT engine's inability to recognize and preserve the source text's thematic progression patterns. In other words, the MT system fails to track how information flows from one clause to the next, often resetting the theme at sentence boundaries or introducing vague referring expressions that break the logical chain. Such errors are particularly problematic for academic journalism, a genre that demands precision, logical rigor, and faithful transmission of ideological positioning.

To diagnose and repair these coherence errors, this study adopts thematic progression (TP) theory as the analytical framework. Originating from the Prague School and later systematized by Halliday (1994, 2014) within Systemic Functional Linguistics, TP theory describes how the theme (the starting point of a clause) and the rheme (the new information presented about the theme) relate across clauses to create textual coherence. Daneš (1974) further proposed several thematic progression patterns, among which two are particularly relevant to expository and argumentative texts: the constant theme pattern, where the same theme is repeated across multiple clauses, each introducing a new rheme; and the derived theme pattern, where the themes of individual clauses are derived from a higher-level hypertheme (e.g., a paragraph topic). By comparing the TP structure of the source text with that of the MT output, post-editors can pinpoint exactly where and why coherence breaks occur. Thus, TP theory provides a systematic, objective heuristic for detecting coherence errors—an alternative to the vague intuition-based advice commonly found in post-editing guidelines.

The existing research on MT post-editing has largely focused on error typologies based on accuracy, fluency, and terminology, with coherence often mentioned but rarely operationalized (Koponen, 2016; O'Brien, 2011). Post-editing guidelines typically offer general recommendations such as "improve readability" or "ensure logical flow" without

 Vol.2 No.1 2026
DOI: 10.64699/26TEZN9415

providing a shared metalanguage to diagnose coherence breakdowns. Furthermore, most studies on post-editing have concentrated on English–Chinese translation, paying relatively less attention to Chinese–English translation, especially for academic journalistic texts. The present study addresses these gaps by proposing a theory-driven, replicable method for detecting coherence errors in Chinese–English MT of journalism and developing targeted post-editing strategies. Three research questions are launched:

1) What dominant thematic progression patterns are typical of the source text—*Ideas of Journalism in Contemporary China*?

2) What types of coherence errors does ChatGPT produce when translating these patterns into English, and how are these errors pattern-specific?

3) What post-editing strategies can systematically repair each error type while preserving the original TP structure?

Rather than describing the translation process, we adopt a problem-oriented approach. First we identify pattern-specific MT errors through TP annotation, then propose three post-editing strategies—specification, combination, and amplification—each designed to repair a specific error type. To sum up, some guidelines will be explored to fix the error problems in the specific case in hope that they can

2. Literature Review

The study of discourse coherence has long been a central concern in linguistics and translation studies. Within this broad field, thematic progression (TP) theory, as developed by Daneš (1974) and grounded in Halliday's (1994, 2014) theme-rheme model, offers a systematic framework for analyzing how information flows across clause and sentence boundaries. According to this framework, the theme serves as the point of departure of a clause, while the rheme presents new information about that theme. Coherence emerges from regular patterns of thematic progression—most notably the constant theme pattern, where the same theme recurs across multiple clauses, and the derived theme pattern, where themes are generated from a higher-level hypertheme. These patterns have been widely applied to textual analysis and, to a lesser extent, to translation evaluation, where they help assess whether the target text preserves the original information structure (Baker, 2018). However, the application of TP theory to machine translation (MT) evaluation remains rare, and its potential as a diagnostic tool for coherence errors is largely unexplored.

Parallel to this theoretical development, empirical research on MT has increasingly recognized that while neural models and large language models such as ChatGPT achieve high scores on sentence-level metrics (e.g., BLEU, COMET), they consistently fail to maintain discourse-level coherence. Documented failures include dangling anaphora, broken topic chains, and illogical discourse connectives (Meyer et al., 2020; Voita et al., 2019). For the Chinese–English language pair, the challenge is particularly acute. Chinese is a topic-prominent language that frequently omits subjects and relies on implicit topic chains, whereas English is subject-prominent, requiring explicit subjects and overt cohesive devices (Li & Thompson, 1981). Recent evaluations of ChatGPT confirm this pattern: while the model performs well on generic texts, it struggles with discourse-sensitive tasks, often “flattening” Chinese topic chains into English subject-verb-object structures without maintaining thematic continuity (Jiao et al., 2023; Wang et al., 2023). For news texts specifically, empirical studies have shown that Chinese and English differ significantly in their use of cohesive devices, requiring systematic adjustments in translation (Weng & Wang, 2020). Yet, most of these studies diagnose coherence problems at a general level; they do not offer a systematic method to classify coherence errors by their underlying discourse structure.

This gap is directly relevant to the field of post-editing. Post-editing research has developed robust error typologies based on accuracy, fluency, and terminology (Krings, 2001; O’Brien, 2011), and has measured the cognitive effort required to correct MT output (Koponen, 2016). However, coherence is rarely operationalized in these frameworks. Existing post-editing guidelines—such as those provided by TAUS (2016)—offer general advice like "improve readability" or "ensure logical flow", but they lack a shared metalanguage to diagnose why a text feels incoherent or where the coherence failure originates. Moreover, the vast majority of post-editing studies have focused on English–Chinese translation, with Chinese–English translation receiving far less attention, especially for the genre of academic journalism (Tang & Chen, 2025). This imbalance is significant because academic journalism places a premium on logical precision, explicit argumentation, and faithful transmission of ideological stance—all of which depend on coherent thematic progression.

Thus, a clear research gap emerges. While TP theory provides a well-established descriptive framework for discourse coherence, and while MT research has documented numerous coherence failures, no study to date has systematically linked specific TP patterns to specific MT coherence errors in Chinese–English journalistic translation, nor has any study developed theory-driven post-editing strategies based on such error patterns. The present study aims to fill this gap by using TP theory as a diagnostic grid: first identifying which TP patterns characterize the source text, then classifying MT errors according to the pattern they disrupt, and finally proposing targeted post-editing strategies—specification, combination, and amplification—each designed to repair a specific error type. In doing so, the study contributes a replicable, theory-informed method for coherence-oriented post-editing.

3. Case Study: Coherence Errors in ChatGPT's Translation of *Ideas of Journalism in Contemporary China*

This section presents a qualitative case study of coherence errors in ChatGPT's translation of a representative Chinese academic journalistic text. The reason why we choose qualitative approach is that it allows for in-depth, context-sensitive analysis of how thematic progression patterns are disrupted in MT output, generating insights that can inform post-editing practice (Yin, 2018). The section first defines the case and describes the analytical framework, then reports the identified error types with illustrative examples.

Case Definition and Source Text

The case was defined as the machine-translated excerpt of "*Ideas of Journalism in Contemporary China*" (Science Press, 2023) produced by ChatGPT. This source text was selected for three reasons: (a) it is a representative academic journalism text with dense argumentation, historical narrative, and explicit logical progression; (b) it contains culturally and politically specific terminology (e.g., "政治家办报" – newspaper run by statesmen), making it a challenging test for MT; and (c) at the time of the study, no English translation of the book had been published, eliminating potential training data contamination. The analyzed excerpt comprised 30 consecutive paragraphs (156 clauses) drawn from Chapter 3 ("Historical Evolution of Journalistic Ideas") and Chapter 4 ("What Is True News?"). These sections were selected because they contained both major TP patterns observed in the full text. The 30 paragraphs were taken consecutively from the beginning of each selected chapter (15 paragraphs from each chapter) without random sampling, as the purpose was in-depth qualitative analysis rather than statistical representation.

Machine Translation Generation

The MT engine used was ChatGPT (GPT-4, May 2024 version), accessed via the OpenAI API. A minimal prompt was used to simulate a typical user scenario: "Translate the following Chinese text into English. Do not add or omit any information." No stylistic, domain, or discourse-level instructions were provided. Each paragraph was translated separately to preserve original paragraph boundaries.

Analytical Framework

Thematic progression: Thematic progression analysis followed Halliday's (2014) identification of topical theme—the first experiential element in a clause. Two patterns were central to the analysis, based on Daneš's (1974) typology: the constant theme pattern, where the same theme is repeated across multiple clauses, each introducing a new rheme; and the derived theme pattern, where themes of individual clauses are derived from a higher-level hypertheme (e.g., a paragraph topic). These patterns were selected because preliminary reading of the source text indicated they were the most frequent and most frequently disrupted by MT.

Error identification procedure. Two researchers independently annotated the source text for theme and rheme in each clause. They then examined the MT output clause by clause, comparing its thematic structure with that of the source. Disagreements were resolved through discussion. Through iterative close reading and pattern recognition (Braun & Clarke, 2006), four distinct coherence error types emerged inductively from the data. Frequencies are reported descriptively; no statistical inference is attempted.

Distribution of errors. Of the 156 annotated clauses, 43 coherence errors were identified. Table 1 shows the distribution by TP pattern and error type.

Table 1

Coherence Error Types by Thematic Progression Pattern

TP Pattern	Error Type	Frequency(n)	Percentage of total errors
Constant theme	Ambiguous reference	12	27.9%
Constant theme	Improper thematic shift	8	18.6%
Derived theme	Cohesion loss	14	32.6%
Derived theme	Information focus shift	9	20.9%
Total		43	100%

Errors in the constant theme pattern. Under the constant theme pattern, two error types were observed. The first was ambiguous reference, where a pronoun or noun phrase failed to clearly refer back to the intended antecedent theme. In Example 1, the source maintains "evaluative facts" as the constant theme, but ChatGPT's "they" could refer either to "evaluative

ISSN 2755-9335 (print) | ISSN 2755-9343 (online)
Vol.2 No.1 2026
DOI: 10.64699/26TEZN9415

facts" or to "assessments," creating referential ambiguity.

Example 1 (Source p. 129)

Chinese: 评价性事实是指对事件现状和趋势进行估价的事实，它与记者的观点交织在一起。记者的观点来自报道内容…

MT: Evaluative facts refer to assessments of the current state and trends of an event, and they are intertwined with the journalist's perspectives.

The second error type was improper thematic shift, where ChatGPT replaced the intended constant theme with a different grammatical subject, altering the propositional meaning. In Example 2, the source theme is "China's journalism industry", but ChatGPT shifts the theme to "the development of…", erroneously implying that the development is part of the Communist Party.

Example 2 (Source p. 100)

Chinese: 中国新闻事业作为共产党的重要组成部分，其发展过程自然与党的发展相辅相成。

MT: The development of China's journalism industry, as an integral part of the Communist Party of China, has naturally progressed in tandem with the Party's growth.

Errors in the derived theme pattern. Under the derived theme pattern, two error types were observed. The first was cohesion loss, where the hypertheme was lost or obscured, causing subsequent sub-themes to appear disconnected. In Example 3, the hypertheme is "Newspaper Run by Historians". ChatGPT's "This approach" is vague and does not clearly derive from the hypertheme, weakening the cohesive chain.

Example 3 (Source p. 74)

Chinese: "史家办报"的新闻人才观要求记者要具备忠实、准确地记录时代的"史家素养"，这成为中国新闻界对于新闻人才素质的一种共识和杰出新闻工作者的实践理念，

 Vol.2 No.1 2026
DOI: 10.64699/26TEZN9415

邓拓即是其中的典范。

MT: The idea of “Newspaper Run by Historians” asserts that journalists should possess the same dedication to faithfully and accurately recording their times as historians. This approach has become a widely accepted standard of journalistic excellence in China, exemplified by distinguished journalists like Deng Tuo.

The second error type was information focus shift, where the rheme of a source clause was repackaged as the theme of a subsequent MT clause, altering the argumentative emphasis. In Example 4, the source is a single rhetorical question whose rheme emphasizes “facing major issues and providing correct answers”. ChatGPT splits the sentence and shifts “theoretical vitality” into the theme of a new sentence, placing emphasis on vitality rather than on answering questions.

Example 4 (Source p. 101)

Chinese: 那么，耳目喉舌论作为马克思主义新闻观指导的、与中华优秀传统文化和党的新闻实践相结合的、与时俱进的中国特色话语理论，应如何发挥其理论活力，以真正面对时代和实践提出的重大问题，做出符合中国实际和时代要求的正确回答？

MT: To effectively respond to the major challenges posed by the times and practice, how should this theory harness its theoretical vitality? It must offer correct answers that align with China’s realities and the requirements of the era.

Preliminary discussion of the case findings. The four error types are not random; they systematically reflect ChatGPT’s inability to track thematic information across clause boundaries. Under the constant theme pattern, the model fails to maintain a stable topic, either by introducing ambiguous referents or by shifting to an unintended grammatical subject. Under the derived theme pattern, the model loses sight of the paragraph-level hypertheme, either by obscuring the derivational link (cohesion loss) or by reorienting the informational emphasis (focus shift). These failures are consistent with the architectural limitations of large language

 Vol.2 No.1 2026
DOI: 10.64699/26TEZN9415

models. ChatGPT processes text autoregressively, predicting each token based on a fixed-length context window (up to 128K tokens in GPT-4). However, the model has no explicit mechanism to track discourse-level topic structures or thematic progression across sentence boundaries. It treats each sentence as a separate prediction task, often resetting the thematic anchor at the beginning of a new sentence—a phenomenon known as topic drift (Koto et al., 2022). Unlike human translators who can maintain a mental representation of the evolving topic chain, ChatGPT lacks long-term discourse memory. For Chinese–English translation, this limitation is magnified by Chinese's frequent zero anaphora (omitted subjects) and topic-prominent structure (Li & Thompson, 1981), which require the model to infer implicit subjects from previous clauses and maintain them across sentences. When the model fails to do so, it introduces ambiguous pronouns (Example 1) or shifts to unintended themes (Example 2). For derived patterns, the model's inability to recognize paragraph-level hyperthemes leads to cohesion loss (Example 3) and information focus shift (Example 4).

4. Post-Editing Strategies Derived from the Case Study

Based on the error types identified in the case study, this section proposes three post-editing strategies. Each strategy is theory-driven, targeting a specific coherence failure while preserving the original TP structure of the source text. The strategies are not claimed to be exhaustive but are directly derived from the observed error patterns and are intended to be transferable to similar Chinese–English academic journalism translations.

Strategy 1: Specification

Specification refers to the replacement of vague or underspecified referring expressions (pronouns, demonstratives, or generic nouns) with explicit noun phrases that match the intended theme from the source TP pattern. This strategy targets ambiguous reference (under the constant pattern) and cohesion loss (under the derived pattern). In Example 1, ChatGPT's "they" is ambiguous. Specification repairs this by writing "these facts" instead, making the constant theme explicit. In Example 3, ChatGPT's "This approach" is vague; specification replaces it with "This historian's literacy", restoring the derivational link to the hypertheme.

Strategy 2: Combination

Combination refers to merging two or more clauses or sentences into one to prevent unwarranted thematic resetting and to preserve the original rheme-theme transition. This strategy targets improper thematic shift (under the constant pattern) and information focus shift (under the derived pattern). In Example 2, ChatGPT shifts the theme from "China's journalism industry" to "the development of…", altering the proposition. Combination repairs this by restoring the original theme: "China's journalism industry, as an integral part of the Communist Party, has naturally developed in tandem with the Party". In Example 4, ChatGPT splits a single rhetorical question into two sentences, shifting the focus to "theoretical vitality". Combination recombines them into one sentence, keeping the emphasis on "addressing major issues and providing correct answers".

Strategy 3: Amplification

Amplification refers to adding short cohesive or logical markers (conjunctions, adverbs, or parenthetical phrases) to make implicit thematic or causal relations explicit, without altering the TP pattern. This strategy is particularly useful for derived pattern cohesion loss when the hypertheme is implicitly understood but not stated. For instance, in a passage where the hypertheme is "Party newspapers must be open", a fragmented MT output may read: "Party newspapers must assume the task of guiding work. They must be open". Amplification adds a conditional marker: "For Party newspapers to assume the task of guiding work, they must first be open", making the logical relation explicit and strengthening the thematic link.

Scope and Transferability of the Strategies

The three strategies were derived from a single case study and are therefore not claimed to be universal. However, because they are grounded in TP theory—a general framework for discourse coherence—they are likely transferable to other Chinese–English academic journalism texts that exhibit similar TP patterns. Post-editors using these strategies should first identify the dominant TP patterns in the source text, then check the MT output for the specific error types documented here, and finally apply the corresponding strategy. The strategies are not mutually exclusive; a single coherence failure may require a combination of specification and amplification, for example. They are intended as a heuristic toolkit rather than a rigid algorithm.

5. Discussion and Conclusion

This study set out to investigate whether thematic progression theory could serve as a systematic diagnostic tool for coherence errors in Chinese–English machine translation of academic journalism and whether theory-driven post-editing strategies could be derived from such error analysis. The case study of ChatGPT’s translation of Ideas of Journalism in Contemporary China yielded three main findings. First, the source text is dominated by two TP patterns: constant theme and derived theme. Second, MT errors are pattern-specific: ambiguous reference and improper thematic shift occur under the constant pattern, while cohesion loss and information focus shift occur under the derived pattern. Third, these error types can be repaired by three corresponding strategies—specification, combination, and amplification—each directly targeting a specific coherence failure.

The theoretical significance of these findings lies in demonstrating that TP theory provides an operationalizable heuristic for coherence evaluation. Unlike existing post-editing guidelines that offer vague advice such as “improve readability” or “ensure logical flow”, TP allows post-editors to ask precise questions: Does the target text maintain the same theme across this sequence? If not, is the shift justified? For a paragraph with a clear hypertheme, are the subsequent themes derived from it? These questions are teachable and can be applied consistently, making TP a valuable addition to post-editing training curricula. The study also contributes to MT evaluation by suggesting that TP consistency could be developed as an additional evaluation dimension, complementing sentence-level metrics such as BLEU and COMET.

Practically, the three strategies offer a replicable method for coherence-oriented post-editing. Specification repairs ambiguous reference and cohesion loss by making the intended theme explicit. Combination corrects improper thematic shifts and information focus shifts by merging fragmented sentences. Amplification adds missing cohesive links without altering the TP structure. These strategies are not ad hoc; they are directly derived from the error typology and are grounded in a well-established linguistic theory. For translator training programs that include a post-editing component, TP-based instruction provides a structured way to teach coherence revision, moving from intuition to principled analysis.

Several limitations must be acknowledged. The study is a qualitative case study based on a single source text (one academic journalism monograph) and a single MT engine (ChatGPT). The findings are therefore not statistically generalizable. However, in qualitative research, generalizability is replaced by transferability—the extent to which the findings can be applied to other contexts. Because the strategies are derived from TP theory, a general framework for discourse coherence, they are likely transferable to other Chinese–English academic journalism texts that exhibit similar TP patterns, but this remains to be tested empirically. Additionally, the sample size (156 clauses, 43 errors) is modest, and the annotation process, while reliable between two coders, is labor-intensive and not scalable for real-time post-editing without automation.

Future research should extend this approach in several directions. First, the TP-based diagnostic framework should be applied to other genres (e.g., literary, legal, conversational) and other language pairs (e.g., Chinese–German, Chinese–Japanese) to test its transferability.

Second, multiple MT engines (DeepL, Gemini, Claude) should be compared on TP preservation to determine whether the observed error patterns are specific to ChatGPT or general across systems. Third, larger-scale studies with statistical testing could confirm the frequency distribution of error types. Fourth, automated TP recognition tools, perhaps using fine-tuned language models, could be developed to assist post-editors in real time. Fifth, experimental studies could measure whether TP-informed post-editing reduces cognitive effort or improves translation quality compared to conventional editing.

In conclusion, this study has demonstrated that thematic progression theory offers a systematic, theory-driven method for detecting coherence errors in Chinese–English machine translation of academic journalism and for developing targeted post-editing strategies. By moving beyond sentence-level corrections to discourse-level coherence management, TP theory provides an objective, teachable heuristic for post-editing practice and evaluation. While the findings are based on a single case study, they establish a proof of concept that TP-based error detection and strategy formulation is feasible and productive. As MT systems continue to improve, discourse-level coherence will become an increasingly critical frontier, and linguistic theories such as TP will play an essential role in guiding both evaluation and revision.

References

Baker, M. (2018). *In other words: A coursebook on translation (3rd ed.)*. Routledge.

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, *3*(2), 77–101. https://doi.org/10.1191/1478088706qp063oa

Daneš, F. (1974). Functional sentence perspective and the organization of the text. *In F. Daneš (Ed.), Papers on functional sentence perspective* (pp. 106–128). Academia.

Halliday, M. A. K. (1994). *An introduction to functional grammar* (2nd ed.). Edward Arnold.

Halliday, M. A. K., & Matthiessen, C. M. I. M. (2014). *Halliday's introduction to functional grammar* (4th ed.). Routledge.

Jiao, W., Wang, W., Huang, J., Wang, X., Shi, S. & Tu, Z. (2023). Is ChatGPT a good translator? Yes with GPT-4 as the engine. https://doi.org/10.48550/arXiv.2301.08745

Koponen, M. (2016). Is machine translation post-editing worth the effort? A survey of research into post-editing and effort. *The Journal of Specialised Translation*, 25, 131–148. https://doi.org/10.26034/cm.jostrans.2016.303

Koto, F., Lau, J. H., & Baldwin, T. (2022). Discourse-aware prompting for narrative text generation. *Proceedings of the 2022 Conference on Empirical Methods in Natural Language Processing*, 2345–2360.

Krings, H. P. (2001). *Repairing texts: Empirical investigations of machine translation post-editing processes*. Kent State University Press.

Li, C. N., & Thompson, S. A. (1981). *Mandarin Chinese: A functional reference grammar.* University of California Press.

Meyer, M., Tiedemann, J., & Teich, E. (2020). *Evaluating discourse phenomena in neural machine translation*. Proceedings of the 12th Language Resources and Evaluation Conference, 2918–2927.

O'Brien, S. (2011). Towards predicting post-editing productivity. *Machine Translation*, *25* (3), 197–215. https://doi.org/10.1007/s10590-011-9096-7

TAUS. (2016). TAUS post-editing guidelines. TAUS BV.

Tang, C., & Chen, D. (2025). The machine translation coherent errors & post-editing strategies of *Ideas of Journalism in Contemporary China* (excerpt): A perspective of thematic progression theory [Master's thesis, Guangdong University of Finance and Economics].

Voita, E., Sennrich, R., & Titov, I. (2019). *When a good translation is wrong in context: Context-aware machine translation improves on coherence.* Proceedings of the 57th Annual Meeting of the Association for Computational Linguistics, 1198–1212.

Wang, L., Lyu, C., Ji, T., Zhang, Z., & Yu, D. (2023). *Is ChatGPT a good translator for Chinese–English news texts? A discourse coherence perspective.* Proceedings of the 2023 Conference on Machine Translation, 78–92.

Weng, Y., & Wang, J. (2020). 文学语篇机器翻译的特征与局限——汉语流水句人机英译对比研究 [Characteristics and limitations of machine translation of literary discourse: A comparative study of human and machine translation of Chinese run-on sentences]. 当代外语研究[*Contemporary Foreign Language Studies*] , (6), 128–137.

Yin, R. K. (2018). *Case study research and applications: Design and methods* (6th ed.). Sage.