OPEN ACCESS

ISAR Journal of Arts, Humanities and Social Sciences

Volume 4, Issue 11, 2026 | pp: 13-16

Abbriviate Title- ISAR J Arts Humanit Soc Sci

ISSN *(Online)*- 2583-9691

https://isarpublisher.com/journal/isarjahss

The Modal Handling of AI Translation Technologies: A Contrastive Study of Human and Machine Translations of Chinese Instruction Manuals into English

Zhang Ziyan[1*], Han Yang[2]

University of Shanghai for Science and Technology.

***Corresponding Author**

Zhang Ziyan

Article History

Received: 07.09.2026

Revised: 01.10.2026

Accepted: 24.10.2026

Published: 12.11.2026

Abstract:

The accuracy of AI translation technologies can be evaluated along one linguistic dimension at a time, and this article focuses on one such dimension, namely modality. Drawing on a corpus of Chinese electronic instruction manuals and their machine and human English translations, this study conducts a contrastive analysis to examine how AI translation systems process modality and evaluates relevant prior findings. The review proceeds in three steps. It first asks whether mainstream systems provide a dedicated procedure for modality and finds that they do not, so output quality still rests on statistical modelling. It then asks what kind of text a manual is and finds that manuals are organised around the directive function, which makes modal expressions appear frequently, especially imperatives. It finally compares the modal systems of the two languages and finds that Chinese and English do not divide modality in the same way, so a given modal meaning has no fixed formal counterpart. The strength of a marker is therefore a matter for the translator to judge in context. Comparisons of human and machine output have registered modal shifts and over-forceful warnings, but most of them remain commentaries on single examples. The only large-scaled quantitative study dates from the statistical MT era and has not been repeated since. A gap therefore remains. Before translating, current systems cannot establish which functional category a sentence belongs to, whether instruction, caution or prohibition, and so cannot choose modal strength from the range that category allows. By linking modal strength to functional categories, this study offers a criterion for future comparisons and a basis for revising machine-translated output before it reaches readers.

Keywords: AI translation technologies, modality, instruction manual, human–machine translation.

This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License

Cite this article:

Ziyan, Z., Yang, H., (2026). The Modal Handling of AI Translation Technologies: A Contrastive Study of Human and Machine Translations of Chinese Instruction Manuals into English. *ISAR Journal of Arts, Humanities and Social Sciences*, *4*(11), 13-16.

1. Introduction

Any claim about the accuracy of AI translation technologies can only be assessed dimension by dimension, and modality is one dimension that has not yet been compared in any systematic way. Modality is the grammatical resource for marking a speaker's attitude towards a proposition. In English it is carried mainly by modal auxiliaries such as must, should, ought to, may, might, can and could. It sets the force of an assertion or a demand, and a wrong choice can change what the sentence means, so that the reader ends up acting differently. In electronic instruction manuals almost every sentence asks the reader to do something or to leave it undone. Modal words are everywhere in these texts and sentence patterns repeat, which makes them a good place to look for problems and a convenient basis for comparison at the same time. Work so far has either asked whether translation technology can cope with modality at all, or described how modality is spread across manuals. Few studies have set human and machine translations side by side and judged them by one standard. This article reviews, in turn, how AI translation technologies handle modality, how modality is distributed in manuals, and how human and machine translations differ in their modal choices, before setting out the gap that remains.

2. How AI Translation Technologies Handle Modality

From statistical machine translation to large language models, no mainstream system has set aside a stage of its own for modality. Wang et al. (2016) [1] point out that conventional statistical MT draws no distinction between parts of speech and gives modal words "no special treatment" at all. Quality rests entirely with the statistical model, and such a model struggles to capture modality. Large language models changed how language is modelled but did nothing to fill that gap. Dai and Huang (2025) [2] argue that a model of this kind, in processing a text, "merely mimics the co-occurrence probabilities of language according to statistical regularities" and often lacks any real understanding. Modality, though, turns on a judgement about what the speaker means to do with the utterance, and co-occurrence probabilities alone will not deliver that. AI translation technologies therefore carry no design feature aimed at modality, and the only way to judge them is by what they actually produce.

3. Modality in Instruction Manuals

3.1 A High Density of Modal Sentences

Instruction manuals are organised around the directive function, so modal elements are dense in them, and imperatives most of all. Jin (2008) [3] describes the English of electrical and electronic manuals as accurate, concise and objective, and notes that it "frequently uses imperatives, with the predicate generally in the base form and no subject". Of the seven sentences she counted in a PHILIPS monitor installation guide, six are imperatives. Chen and Xiao (2026) [4], working on manuals for exported household appliances, come to the same conclusion. Such texts lean heavily on imperatives and on inanimate subjects. Zhang (2012) [5] offers a functional account. A manual has both to pass on product information and to guide consumers in using the product. It serves an informative and a directive function at once, and the latter explains why most sentences appear as directives. Modal sentences are therefore thickly distributed in manuals. Besides imperatives they include declaratives built with must, should and may, and warnings such as 切勿 (never) and 严禁 (strictly prohibited), which makes manuals a suitable site for testing how AI translation technologies cope with modality.

3.2 Different Modal Strength, Different Meaning

Modal strength is not just a matter of how forcefully a sentence is worded. It is part of what the sentence means. Halliday's modal value names the systemic category behind it, and strength is what that value does in a particular sentence. Wang et al. (2016) [1] use three Chinese sentences that share aproposition. They read 大卫去学校了, 大卫一定去学校了 and 大卫可能去学校了, rendered as *David went to school, David must have gone to school and Maybe David went to school*. The first states a fact. The other two add the speaker's attitude, the third with the least certainty. Change the modality and the translation of the whole sentence changes with it. Xia and Zhang (2017) [6] pin the difference to modal value.Obligation tends to be marked by higher-value items such as must, should and ought to, which carry " a strong suggestion, command or prohibition ", while possibility is marked by lower-value items such as may, can, might and could. Value corresponds to modal strength. A choice that slips one notch can leave a sentence that still reads smoothly but no longer means the same thing. Render 必须 as 可以 and the reader will act quite differently. Shifts of this kind never surface as grammatical errors, and holistic scoring is unlikely to catch them. Only a sentence-by-sentence comparison of modality will.

4. How Human and Machine Translations Differ in Modality

4.1 Chinese and English Modal Markers Are Not Fully Equivalent

Chinese and English do not divide modality in the same way, and a given modal meaning has no fixed formal counterpart across the two languages. The translator judges from context. The baby monitor manual studied by Yang and Tian (2024) [7] makes the point. 请勿, 请确保 and 请 come out as *Do not, Ensure that and please*, three markers with no formal relation to each other, and it falls to the translator to decide which value each occupies. Lin and Cao

(2021) [8] counted the 53 imperatives in a Dyson reader ' s manual. Of the 36 affirmative sentences, 21 are translated literally, without adding请, and most of these come from assembly and setup. Another 15 have no please in the source but gain one in translation, and most of these come from cautions. Of the 17 negative sentences, 16 take 请勿 and one takes 不可. None uses 禁止, which was felt to sound too blunt. How a given pattern is rendered depends on the function of the section it appears in. Modal translation cannot be run off a word list. The translator has first to work out what the sentence is doing, then pick the strength that fits. That step is still performed by people, and it is where machines are most likely to fail.

4.2 Observation Stops Short of Measurement

Research has noticed that machine output drifts in modality, but it stops at the commentary on single examples and never turns the drift into a variable that can be measured. Lü and Wang (2026) [9] compare the official translation of an electronic instruction manual with the ChatGPT and DeepSeek versions. On 打孔前，请确保避开墙内预埋的水电线路, the official version reads *When drilling holes, avoid …*, which sends the reader straight to the action, while the machine versions read *ensure that ... to avoid ...*, with potential danger and hazards added to sharpen the consequence. The authors conclude that the machine output has "greater pragmatic warning force" but "may make the tone too heavy". The direction is plausible, but the evidence is one example, with no frequency and no distribution. Zhang et al. (2025) [10] do build Directive Intensity Deviation into their error annotation scheme. The category covers a command or warning that is too weak or too strong for the communicative intention, which comes close to illocutionary force. But the category occurs once in the Google output and once in the ChatGPT output, at 0.40% and 2.17% respectively, and the article comments only in passing on the Google case, where *your studio display becomes the studio display*, without tying it to the function of the sentence. The reason for this over-strengthening is not far to seek. The probabilistic selection described in Section 2 picks by co-occurrence and never asks what the sentence is asking of the reader [2]. Yet "too strong" and "too weak" remain impressions. With no shared scale of value behind them and no frequency count, they cannot say which notch is wrong, or in which sentences.

4.3 Measuring Modal Strength

Modal strength need not be left to intuition. It can be counted. Wang et al. (2016) [1] set two groups of 115 sentences against each other, one with modal words and one without. In the Chinese-to-English direction, short sentences containing modal words averaged 3.66 against 4.05 for those without, and long sentences 3.37 against 3.88, the gap being wider for long sentences (0.51) than for short ones (0.39). Errors of negative polarity, where the translation turns what should be done into what should not be done or the reverse, came to 27 in the sentences with modal words and just 1 in those without. What matters here is less the finding than the design. Treat the presence of modal words as a grouping variable, count errors by type, and a difference in strength turns into a number that can be compared. All four systems tested (Moses, Baidu, Bing and Youdao) were statistical MT systems rather than large language models, and the grouping analysis has not been repeated since. Work on large language models can start from the same design.

5. Conclusion

One question is still unanswered, namely whether AI translation technologies can first decide what kind of modal meaning a sentence carries and then choose modal strength to match. On the object side, the modal sentences of manuals have never been counted on their own, and work on imperatives looks only at human translations. On the variable side, functional category and modal strength have never been crossed in a single design. Function is usually brought in afterwards, as explanation. On the coding side, "too strong" and "too weak" have no measurement standard behind them. Yet the modal sentences of manuals fall roughly into three functional categories, namely operating instructions, cautions and prohibitions, and the three make different demands on strength, as the counts of Lin and Cao (2021) [8] already hint at this. Assembly and setup are mostly translated literally, cautions mostly gain a 请, negative sentences mostly take 请勿, and 禁止 is barely used. A correspondence between functional category and modal strength can therefore be observed and coded. Only when a system assigns the category first, and then draws its modal markers from the range that category allows, will modal translation be accurate. Current systems align sentence by sentence and pick by probability, but they stop short of this step of classification. The present study accordingly treats functional category and modal value as two crossed coding dimensions and compares human and machine translations along both.

References

1. Wang, L., Zhang, Q., Zhang, D., et al. (2016). An analysis of error types in the machine translation of Chinese modal words. In Proceedings of the 12th China Workshop on Machine Translation (CWMT 2016), 116–125. (in Chinese)
2. Dai, G., & Huang, B. (2025). Cognitive features of large language model translation and the path of its development. Chinese Translators Journal, 46(4), 122–130. (in Chinese)
3. Jin, F. (2008). Features and techniques in the English translation of electrical and electronic product manuals. Modern Business Trade Industry, (11), 271–272. (in

Chinese)

4. Chen, M., & Xiao, K. (2026). Strategies for translating export home appliance manuals into English: A functionalist Skopos perspective. Chinese Science & Technology Translators Journal, 39(2), 38–41. (in Chinese)
5. Zhang, Y. (2012). Linguistic features and translation techniques of different types of product manuals. The Science Education Article Collects, (1), 133–134. (in Chinese)
6. Xia, J., & Zhang, H. (2017). Interpersonal function in the English translation of drug package inserts. Journal of Suzhou University, 32(1), 81–84. (in Chinese)
7. Yang, J., & Tian, D. (2024). Translating electronic communication product manuals: A cognitive linguistics perspective. Taste and Classics, (9), 78–80. (in Chinese)
8. Lin, J., & Cao, Q. (2021). Translating imperatives in electronic product manuals from the perspective of Skopos theory. Popular Literature and Art, (18), 115–116. (in Chinese)
9. Lü, M., & Wang, L. (2026). A comparative study of electronic product manual translation in the age of artificial intelligence: Official and machine translations. English Square, (25), 17–21. (in Chinese)
10. Zhang, Z., Abdullah, S. N. S., Abdullah, M. A. R., et al. (2025). Translation quality in an evolving paradigm: Neural machine translation and large language models in technical domains. International Journal of English Language Education, 13(2), 52–80.