



## SEMANTIC ADEQUACY IN POPULAR SCIENCE TEXTS FOR YOUNG AUDIENCES: A COMPARATIVE ANALYSIS OF GOOGLE TRANSLATE, DEEPL, CHATGPT AND MANUS AI TRANSLATIONS

**Author:** Zokirjonova Madina Iqboljon qizi<sup>1</sup>

**Affiliation:** Independent Researcher, Namangan State Institute of Foreign Languages named after Ishoqxon Ibrat<sup>1</sup>

**DOI:** <https://doi.org/10.5281/zenodo.19694871>

### ANNOTATION

This study explores semantic adequacy in AI-based translations of popular science texts intended for young audiences. Translations produced by four systems are comparatively analyzed in terms of meaning accuracy, terminology, and completeness. The findings reveal differences in semantic performance and emphasize the importance of human supervision in ensuring reliable and precise translation.

**Keywords:** AI-based translation, semantic adequacy, popular science texts, young audience, scientific terminology, semantic analysis, translation quality, neural machine translation.

### INTRODUCTION

Recent advances in artificial intelligence have significantly transformed the field of translation, particularly in rendering popular science texts intended for young audiences. Such texts require not only linguistic accuracy but also semantic clarity and accessibility, as they aim to explain complex scientific concepts in an understandable way. AI-based translation systems such as Google Translate, DeepL, ChatGPT, and Manus AI are increasingly used for this purpose; however, their ability to preserve semantic adequacy in this specific genre remains a critical issue.

Semantic adequacy refers to the degree to which the meaning of the source text is accurately and completely conveyed in the target language. In scientific discourse, this concept becomes particularly critical, as even minor deviations in meaning or terminology may lead to misunderstanding or misinterpretation of key concepts [1]. This is especially relevant in interdisciplinary fields such as biology, where terms like *fungi*, *arthropods*, or *cuticle* carry specific and well-defined meanings.

Previous studies have shown that machine translation systems often struggle with domain-specific terminology, polysemy, and context-dependent meanings [2]. In particular, errors such as semantic narrowing, omission, or incorrect lexical choice may significantly reduce the quality of translation output. For instance, translating the biological term *fungi* as a narrower equivalent such as "mushrooms" represents a case of semantic narrowing that affects scientific accuracy.

Moreover, AI translation systems may differ in how they handle complex sentence structures, contrastive constructions, and pragmatic elements. While some systems prioritize fluency and readability, others attempt to preserve structural and

terminological fidelity. This variability in translation performance necessitates systematic comparative analysis.

Therefore, the present study aims to conduct a semantic analysis of translations produced by four AI-based systems—Google Translate, DeepL, ChatGPT, and Manus AI—based on selected excerpts from a popular science article published in *Frontiers for Young Minds* [3]. The analysis focuses on key aspects such as terminological accuracy, semantic equivalence, contextual adequacy, and completeness of translation.

The findings of this study are expected to contribute to a better understanding of the strengths and limitations of AI translation tools in handling scientific texts, and to highlight the importance of human oversight in ensuring semantic precision.

## RESEARCH METHODOLOGY

This study employs a qualitative comparative approach to analyze the semantic adequacy of AI-generated translations. The research focuses on four widely used machine translation systems: Google Translate, DeepL, ChatGPT, and Manus AI. These systems were selected due to their popularity, accessibility, and differing underlying architectures, which allow for a meaningful comparison of translation performance.

The data for analysis were collected from a popular science article titled “*The fungi that fight bugs: Nature's tiny heroes*”, published in *Frontiers for Young Minds* [3]. This source was chosen because it combines scientific terminology with accessible language intended for a young audience, making it particularly suitable for evaluating both terminological precision and contextual clarity.

A total of five sentences were selected from the source text. These sentences were chosen based on their semantic complexity and the presence of key linguistic features, including domain-specific terminology (e.g., *fungi*, *arthropods*), contrastive structures (e.g., *despite*, *although*), and descriptive explanations of biological processes. Each sentence was translated into Uzbek using the four selected AI systems.

The analysis was conducted based on several semantic criteria adapted from translation studies [1], including:

- semantic equivalence (accuracy of meaning transfer),
- terminological accuracy (correct rendering of scientific terms),
- contextual adequacy (appropriateness within context),
- completeness (absence of omissions),
- and lexical choice (selection of appropriate vocabulary).

Special attention was given to common types of semantic errors, such as semantic narrowing, omission, and incorrect interpretation [2]. For example, the mistranslation of the term *fungi* as a narrower lexical equivalent was categorized as semantic narrowing, while the absence of entire clauses was treated as omission.

The results of the analysis were presented in tabular form and supported by qualitative commentary. This combined approach allows for both systematic comparison and in-depth interpretation of the strengths and weaknesses of each AI translation system.

For analytical purposes, the translation systems were divided into two categories:

1. Classical neural machine translation tools (Google Translate and DeepL)
2. Generative AI-based translation tools (ChatGPT and Manus AI)
- 3.

System	Model	Version	Date	Platform
ChatGPT	GPT-4	free	12.02.2026	<a href="https://chatgpt.com/">https://chatgpt.com/</a>
Google Translate	NMT	free	12.02.2026	<a href="https://translate.google.com/">https://translate.google.com/</a>
DeepL	NMT	free	12.02.2026	<a href="https://www.deepl.com/en/translator">https://www.deepl.com/en/translator</a>
Manus AI	LLM-based	free	12.02.2026	<a href="https://manus.im/">https://manus.im/</a>

All systems were provided with the same English source text under identical conditions. During the translation process, the following standard input prompt was used: “Translate the following scientific text from English into Uzbek. Preserve scientific lexics and provide explanatory glosses in parentheses where necessary for young readers.” The experiments were conducted on February 12, 2026, through the official web interfaces of the systems. No APIs or paid versions were used. This approach ensured equal conditions for the comparative evaluation of the translation systems. The translations produced by these systems were organized into two comparative tables. Each translation was evaluated in relation to the **contextually adequate Uzbek equivalent** in order to determine its lexical accuracy and semantic appropriateness.

The analysis focused on identifying common translation strategies and errors, including **literal translation, transliteration, semantic shifts, and contextual inaccuracies**. In addition, a statistical evaluation was conducted to determine the percentage of adequate translations produced by each system.

This methodological approach allows for a clearer comparison between classical neural machine translation technologies and modern generative AI models, highlighting their respective strengths and limitations in translating popular science vocabulary intended for young audiences.

## RESULTS AND DISCUSSION

This section presents a comparative semantic analysis of translations produced by four AI systems—Google Translate, DeepL, ChatGPT, and Manus AI. Each system is evaluated individually based on semantic equivalence, terminological accuracy, contextual adequacy, and completeness.

### 1. Google Translate

Google Translate demonstrates generally stable performance in terms of preserving the overall meaning of the source text. In most cases, the system successfully conveys the core message and maintains the structure of the original sentences. It also performs relatively well in maintaining completeness, as no major omissions were observed.

However, several issues related to **terminological inconsistency** were identified. For example, the biological term *fungi* was translated inconsistently, alternating between broader and narrower equivalents. This resulted in cases of **semantic narrowing**, which reduced scientific accuracy. Similarly, terms like *arthropods* were sometimes rendered in less precise forms, affecting terminological clarity.

Despite these limitations, Google Translate maintains a good balance between accuracy and readability, making it a reliable, though not fully precise, translation tool.

## 2. DeepL

DeepL demonstrates the highest level of **semantic adequacy** among the analyzed systems. It consistently preserves both the meaning and structure of the source text, including complex grammatical constructions such as contrastive clauses (*although, despite*).

In terms of **terminological accuracy**, DeepL shows strong performance, particularly in translating domain-specific terms like *fungi* and *entomopathogenic fungi* correctly and consistently. Additionally, it effectively conveys cause-and-effect relationships and maintains logical coherence across sentences.

Minor limitations include occasional generalization of certain scientific terms and slight stylistic rigidity. However, these do not significantly affect the overall semantic quality.

Overall, DeepL provides the most balanced and reliable translations, particularly for scientific texts requiring precision.

## 3. ChatGPT

ChatGPT stands out for producing **fluent and natural-sounding translations**, often adapting the text to improve readability and accessibility. This makes it particularly suitable for general audiences.

However, this strength is accompanied by several semantic limitations. The system tends to **simplify or reinterpret** certain expressions, which can lead to partial loss of meaning. For instance, some translations showed **semantic shifts**, where the original meaning was slightly altered to produce a more natural output.

More critically, instances of **omission** were observed, including the complete omission of a sentence in one example. This represents a significant reduction in semantic completeness and is considered a major translation error [2].

Thus, while ChatGPT excels in fluency, it is less reliable in preserving full semantic and terminological accuracy in scientific contexts.

## 4. Manus AI

Manus AI demonstrates relatively strong performance in maintaining the structural integrity of the source text. It often preserves sentence structure, logical relations, and completeness. In some cases, it also shows good handling of specific scientific terms, such as accurate rendering of *arthropods*.

However, a major limitation of this system is **terminological inaccuracy**, particularly in translating key terms like *fungi*. The frequent use of narrower lexical equivalents results in semantic narrowing, which reduces the scientific validity of the translation.

Additionally, minor inconsistencies in lexical choice and occasional omissions of specific details were observed. Despite these issues, Manus AI shows potential, especially in handling sentence structure, but requires improvement in domain-specific terminology.

The system-based analysis reveals that each AI translation tool has distinct strengths and weaknesses. DeepL provides the most semantically accurate and consistent translations, while Google Translate offers stable but less precise outputs. ChatGPT excels in fluency but suffers from omissions and semantic shifts. Manus AI demonstrates structural strength but struggles with terminological accuracy.

The overall findings of the analysis are summarized in the following Table.

Criteria	Google Translate	DeepL	ChatGPT	Manus AI
Semantic equivalence	Medium	High	Low–Medium	Medium
Terminological accuracy	Medium	High	Medium	Low–Medium
Contextual adequacy	High	High	Medium	Medium
Completeness	High	High	Low	Medium
Overall performance	Good	Very good	Variable	Moderate

Overall, the analysis demonstrates that DeepL provides the most balanced performance in terms of semantic adequacy, particularly in preserving both meaning and structure. Google Translate also performs reliably, though with occasional terminological inconsistencies. ChatGPT shows strengths in fluency and readability but tends to simplify or omit important information in some cases. Manus AI demonstrates potential in handling certain scientific terms but suffers from inconsistencies in terminological accuracy.

## CONCLUSION

The present study has examined the semantic adequacy of translations produced by four AI-based systems—Google Translate, DeepL, ChatGPT, and Manus AI—based on selected excerpts from a popular science text. The analysis focused on key aspects such as terminological accuracy, semantic equivalence, contextual adequacy, and completeness.

The findings reveal that while all systems are capable of conveying the general meaning of the source text, significant differences exist in their ability to preserve scientific precision. Among the analyzed tools, DeepL demonstrated the highest level of semantic adequacy, consistently maintaining both the meaning and structural relationships of the original text. Google Translate also showed relatively stable performance, although it occasionally exhibited terminological inconsistencies.

ChatGPT, on the other hand, proved effective in producing fluent and natural translations, but at the cost of semantic precision. Instances of simplification, reinterpretation, and omission were observed, which reduced the overall accuracy of the translated content. Manus AI showed strengths in preserving sentence structure and certain domain-specific elements, but frequent terminological inaccuracies—particularly in translating key scientific terms—limited its effectiveness.

A recurring issue across several systems was the improper handling of domain-specific terminology, such as the translation of *fungi*, which in some cases resulted in semantic narrowing. Additionally, complex sentence structures involving contrast and cause-effect relationships posed challenges for some AI systems, leading to partial meaning loss or distortion.

These findings suggest that, despite their growing capabilities, AI translation tools are not yet fully reliable for translating scientific texts without human supervision. Accurate translation in such contexts requires not only linguistic

competence but also domain-specific knowledge and sensitivity to semantic nuances.

Therefore, it is recommended that AI-generated translations be carefully reviewed and post-edited by human experts, particularly when used in academic or scientific settings. Future research may expand the dataset and include a broader range of text types to further investigate the performance of AI translation systems.

In conclusion, while AI tools offer valuable support in translation tasks, their limitations highlight the continued importance of human involvement in ensuring semantic accuracy and terminological precision.

## REFERENCES

1. Newmark, P. (1988). *A Textbook of Translation*. Prentice Hall.
2. Koehn, P. (2020). *Neural Machine Translation*. Cambridge University Press.
3. Frontiers for Young Minds. (n.d.). *The fungi that fight bugs: Nature's tiny heroes*.

