



An-Najah National University
Faculty of Graduate Studies

**A STUDY OF THE TRANSLATION
QUALITY ASSURANCE SYSTEM IN THE
SMARTCAT TRANSLATION SOFTWARE**

By
Ayah Mahmoud Demaidi

Supervisor
Prof. Abdelkarim Daraghme

**This Thesis is Submitted in Partial Fulfilment of the Requirements for the Degree
of Master of Applied Linguistics and Translation, Faculty of Graduate Studies,
An-Najah National University, Nablus, Palestine.**

2025

STUDY OF THE TRANSLATION QUALITY ASSURANCE SYSTEM IN THE SMARTCAT TRANSLATION SOFTWARE

This Thesis was defended successfully on 19/10/2024 and approved by:

Prof. Abdelkarim Daraghme

Supervisor



Signature

Dr. Jamal Nafie

External Examiner



Signature

Dr. Rami Qawariq

Internal Examiner



Signature

Dedication

To my dear father, the one who always looked up to me with pride and admiration, and encouraged me to reach this point.

To my caring mother, thank you for all your sacrifices that helped me become who I am.

To my brothers and sisters, my companions through thick and thin.

To my loyal husband, who tolerated my absences and the long hours I spent studying and who had my back and supported me through everything until I reached my goal.

To my second family, my husband's loving family, thank you for taking care of me and appreciating my efforts.

To my daughter who is yet to be born, who accompanied me through the pain of sleeplessness and fatigue, I dedicate this achievement to you as my first gift before you see the light.

Acknowledgement

First and foremost, all praise is due to Allah, the Most Gracious and the Most Merciful, for his abundant grace. The one who gave me enough strength, patience, and endurance to successfully complete this stage. I wouldn't have been able to get to this place if I hadn't relied on Him and asked for His help every time I felt scared and unsteady.

I would like to thank my wonderful supervisor, Professor Abdelkarim Daraghme, for his dedicated work and tremendous support throughout my master's journey. His valuable feedback and insightful information enriched my thesis and made it stand out.

I also extend my sincere gratitude to my professors in the English Department at An-Najah National University: Dr. Rami Qawariq, Dr. Ruqayya Harzallah, Dr. Ayman Nazzal, Dr. Sufyan Abu Ara, and Dr. Abdelkarim Zawawi. Each of them left an unforgettable mark on my academic development and helped me fulfill my dream.

Finally, I extend my heartfelt thanks to my family, friends, and colleagues. Their presence and encouragement made this journey lighter and more meaningful.

Declaration

I, the undersigned, declare that I submitted the thesis entitled:

A STUDY OF THE TRANSLATION QUALITY ASSURANCE SYSTEM IN THE SMARTCAT TRANSLATION SOFTWARE

I declare that the work provided in this thesis, unless otherwise referenced, is the researcher's own work, and has not been submitted elsewhere for any other degree or qualification.

Student's Name

Ayah Mahmoud Demaidi

Signature:



Date:

19/10/2025

List of Contents

Dedication.....	iii
Acknowledgement.....	iv
Declaration.....	v
List of Contents.....	vi
List of Tables.....	ix
List of Figures.....	x
Abstract.....	xi
Chapter One: Introduction and Literature Review.....	1
1.1 Background of the Study.....	1
1.2 Statement of the Problem.....	4
1.3 Purpose of the Study.....	5
1.4 Research Questions.....	5
1.5 Scope of the Study.....	6
1.6 Significance of the Study.....	7
1.7 Limitations of the study.....	8
1.8 Literature Review.....	8
1.8.1 Introduction.....	8
1.8.2 Quality Assurance vs. Quality Control.....	9
1.8.3 ISO 17100:2015 and its Role in Ensuring Translation Quality.....	11
1.8.4 ISO 18587 and its Role in Ensuring MT Quality.....	11
1.8.5 Translation Theories and their Applications in Quality Assurance.....	13
1.8.5.1 Introduction.....	13
1.8.5.2 Skopos Theory.....	14
1.8.5.3 Julian House model.....	16
1.8.6 Summary Comparison Table.....	17
Chapter Two: Methodology.....	19
2.1 Introduction.....	19
2.2 Research Design.....	19
2.3 Research Instrument.....	20
2.4 Theoretical Framework.....	20
2.4.1 Introduction to the Theoretical Framework.....	20
2.4.2 Çetiner’s Error Typology (2021).....	21
2.4.3 Theoretical Positioning of the Study and Its Conceptual Contribution.....	22

2.4.4 Summary	23
2.5 Data Collection	23
2.6 Validity and Reliability	26
2.7 Ethical Consideration.....	26
Chapter Three: Data Analysis.....	27
3.1 Introduction.....	27
3.2 MT and Segmentation Challenges	27
3.3 Error Detection and Classification Analysis.....	27
3.4 Data Classification and Applicability	28
3.5 Linguistic Errors	28
3.5.1 Errors within Çetiner’s Surface-level Category.....	28
3.5.2 Errors Beyond Çetiner’s Surface-Level.....	29
3.5.2.1 Stylistic Deviation.....	30
3.6 Terminological Errors.....	33
3.6.1 Errors within Çetiner’s Surface-level Category.....	33
3.6.2 Errors beyond Çetiner’s Surface-level.....	33
3.6.2.1 Lexical generalization of technical terms	34
3.6.2.2 Literal translation of conceptual terms	35
3.6.2.3 Terminological Conflation (Overlapping Concepts)	36
Figure 4 Distribution of terminological errors in MT using Smartcat.....	37
3.6.2.4 Compatibility of Smartcat Translation QA System with ISO 17100 and ISO 18587 Standards in Terms of Terminological Errors	38
3.7 Formatting and Tags Error.....	38
3.7.1 Errors within Çetiner’s Surface-level	38
3.7.2 Errors beyond Çetiner’s Surface-level.....	39
3.7.2.1 Formatting Errors Due to Linguistic Differences	39
3.7.2.2 Lack of Number Localization	39
3.7.2.3 Inclusion of English Terms in Arabic Translation.....	40
3.7.2.4 Compatibility of Smartcat QA System with ISO 17100 and ISO 18587 in Terms of Formatting Errors	40
3.8 workflow Errors	41
3.8.1 Errors within Çetiner’s Surface-level	41
3.8.2 Compatibility of Smartcat Translation QA System with ISO 17100 and ISO 18587 Regarding Workflow Errors	41
3.9 Evaluating TM: Quality Support	43
Chapter Four: Conclusion and Recommendation	44

4.1 Conclusion	44
4.1.1 Comparison with Previous Studies	44
4.1.2 Scientific Contribution.....	45
4.2 Recommendations.....	46
List of Abbreviations	47
References.....	48
الملخص.....	52

List of Tables

Table 1: Comparison between Smartcat QA features and functionalist translation theories	17
Table 2: Types of linguistic errors detected/ undetected by SC QA according to Çetiner's classification.....	29
Table 3: Linguistic errors beyond surface level appeared in the MT	32
Table 4: Types of terminological errors detected/ undetected by SC QA according to Çetiner's classification	33
Table 5: Types of formatting errors detected/ undetected by SC QA according to Çetiner's classification.....	39
Table 6: Examples of formatting and tag errors	40
Table 7: Types of workflow errors detected/ undetected by SC QA according to Çetiner's classification.....	41
Table 8: Examples of missing term types	42

List of Figures

Figure 1: Image of the confirm button that must be pressed after the translation of each segment	3
Figure 2: Image of the triangle used as an error alert	3
Figure 3: Image of the new version of Smartcat with the orange triangle as the only alert tool	4
Figure 4: Distribution of terminological errors in MT using Smartcat.....	37
Figure 5: Distribution of MT errors found within the studied text according to field type	38

A STUDY OF THE TRANSLATION QUALITY ASSURANCE SYSTEM IN THE SMARTCAT TRANSLATION SOFTWARE

By

Ayah Mahmoud Demaidi

Supervisor

Prof. Abdelkarim Daraghmeh

Abstract

This study examines the translation quality assurance system used in Smartcat, a computer-assisted translation platform used by professional translators to help facilitate and speed up the translation process. The error analysis uses Çetiner's model (2021), in which he classified translation errors into four main categories: linguistic, terminological, and formatting and workflow. Then the researcher compares the results with professional translation standards like ISO 17100 and ISO 18587. The researcher expands the scope of the analysis by adding errors beyond the surface level of Çetiner's classification. The study also explores how the translation memory tool interacts with the quality assurance system. The study adopts a descriptive, evaluative design, and a qualitative approach to identify, classify, and interpret errors. The severity of errors was measured according to their impact on the understanding and meaning of the sentence.

The results show that the highest percentage of errors in the Smartcat platform was terminological errors, followed by linguistic errors and then formatting errors, most of which were not detected by the translation quality assurance system. The results did not reveal any errors in the system related to the translation memory. The results also show that Smartcat translation quality assurance system was not fully compatible with ISO translation standards. The study recommends incorporating human expertise into the translation evaluation process of computer-assisted translation tools to ensure reliable results.

Keywords: Smartcat, Quality Assurance System, Cat Tools, Machine Translation, Translation Memory, ISO Standards.

Chapter One

Introduction and Literature Review

1.1 Background of the Study

The field of translation is a very broad field, and it cannot be limited to transferring words or sentences from one language to another. Its broad nature is clear as it engages not only with linguistic aspects, but also with cultural and technological aspects. Several years ago, the translator had no choice but to look at the text in hands and try to translate it in the most appropriate way, manually, without any interference or external assistance. The matter was limited to setting strategies and theories in mind and proceeding according to the translation objectives to obtain a professional translation output. With the rapid development of technology and internet, translation has witnessed an exceptional development, with technology now providing unprecedented assistance in the field of translation.

It has become necessary to research all available tools that help in the translation process and in fulfilling its purposes. As a translation company or as a freelance translator, it is sometimes unavoidable to call for assistance to save time and effort, especially when there is a need to keep pace with the increased workload at particular times when demand is high. Here comes the interesting role played by translation technology and Computer-Assisted Translation (CAT) tools, which have proven their effectiveness and worthiness in many situations, fields, and occasions. According to (Mohammed, Samad, & Mahdi, 2018, pp. 1-2), “Computer-assisted translation (CAT) ... has created new opportunities for translators that cannot be found in traditional ways.” These opportunities include higher productivity, better terminology selection and consistency, fewer errors, improved quality better project management, and the ability to leverage previous translations. In the same context, (Biau Gil, Leal, Mor, Mesa-Lao, Orozco, & Sanchez-Guon, 2010) explained that computer-assisted translation tools are used to save time, facilitate project management, and make use of previous translations. Currently, many platforms that facilitate the translation process have become widespread, including computer-assisted translation (CAT) tools such as SDL Trados, MemoQ, and WordFast. Among these tools, Smartcat, a cloud-based CAT tool, ranks highly and occupies a prominent position. The researcher specifically selected Smartcat because it is widely used by translators and

project managers, offering a comprehensive suite of features that integrate project management, collaborative translation, and artificial intelligence tools. Moreover, through her practical experience with the platform, the researcher identified a gap in its functionality, which motivated her to conduct an in-depth analysis and provide critical insights and development suggestions (Smartcat, n.d.).

Basically, Smartcat receives different supported file formats like (DOC, JPEG, PDF, JSON, XLIFF) and has a systematic operation when it comes to dealing with these files. Uploading the source text file, the translator has to identify the source language and the target language, and determine whether they want to get assisted by machine translation (MT) that draws on Google Translate and some other free and paid translation choices that are found on the advanced settings button that appears at the top of the same page. The MT sources include Google Statistical MT, Google Neural MT, Microsoft Translator with both statistical and neural engines, depending on the language pair, Yandex MT, Baidu Translate API, Deepl, Amazon Translate, and Modern MT. As an additional step, the translators could choose their role in the translation process, as the platform gives them the option to choose between the following cases: (AI translation and human post editing, human translation only, human translation and human post editing, or human translation, post editing, and proofreading). Also, they must choose whether they want to use any existing translation memory (TM) in case the new file to be translated is similar to a previous translation file. Smartcat breaks the file up into segments, based on the punctuations which mark when each sentence is finished. It also places the phrases, the titles, the headlines, and words that come alone in the file and do not need to be placed in a sentence, each in a separate segment. Having done all these steps, the translators have to translate each segment in the space provided next to it. Some icons appear in orange with numbers in some sentences; the translators must transfer them and place them in the same place where they were found in the original text, as they preserve the file format. Emerging as a notable player, Smartcat committed itself to harness its effort for the benefit of the translator, by facilitating translation workflows, enhancing productivity, and ensuring translation quality through the Translation Quality Assurance System (QA) it offers.

A translation QA system in any computer-assisted translation tool refers to a set of software parameters, procedures, and standards that are programmed to ensure that the

translation is accurate and error-free. In Smartcat, the translation QA system encompasses numerous aspects, all working together to assess various perspectives and gauge the translation worth and its validity. In previous versions of Smartcat, during the translation process, pressing the confirm button located beside each segment (see figure 1), the translation QA system used to assess the translation accuracy out of a hundred, while simultaneously highlighting the presumed QA error.

Figure 1

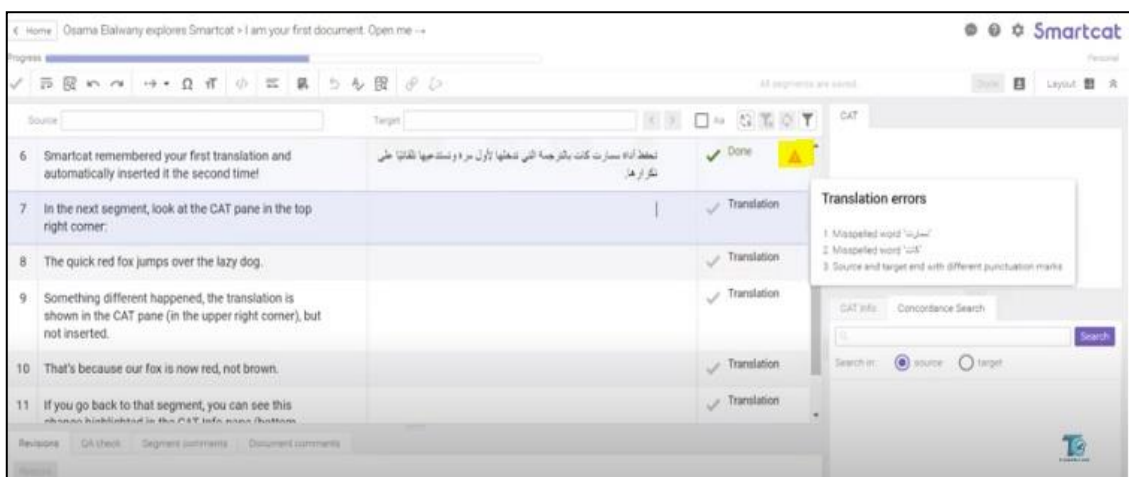
Image of the confirm button that must be pressed after the translation of each segment



It used to show the evaluation of the sentence by means of a triangle that appears next to the sentence (see figure 2).

Figure 2

Image of the triangle used as an error alert



Depending on the color of the triangle, we could know how correct the sentence is and the type of error it includes. The green triangle used to show that the translation is correct,

while the orange one used to show that there are some errors in the translation, which was going to be explained to the translator when they place the cursor on the triangle. However, if the triangle appears red, it means that the error was caused by the design and that the translator did not transfer the orange icons (tags) correctly. It is worth noting that the **new version** of Smartcat, which was used to collect the study data, relies on one small indicator: an orange triangle for non-critical warnings. The triangle shows the number of errors and what they are when the translator places the cursor on it. The researcher also noted that the new version has also seen the disappearance of the orange symbols that the translator should have transferred to the resulting text to maintain formatting, resulting in the red triangle no longer being present (see figure 3).

Figure 3

Image of the new version of Smartcat with the orange triangle as the only alert tool



However, given that this system relies on automated auditing, modern technology, and software tools that cannot replace humans, it is natural that the system's errors will be significant, which necessitates an examination to identify loopholes in this system.

1.2 Statement of the Problem

This study is based on personal experience. While translating some files between Arabic and English using Smartcat, the researcher noticed clear gaps in the performance of the tool's built-in translation QA system. It was found that the system did not always provide an accurate assessment of the translation quality. In some cases, it issued an error warning (via an orange or red triangle in the old version) and (via an orange triangle in the current one) in cases where the translation is linguistically and semantically correct. This case will be hereafter referred to as a false positive. In other cases, the translation contained

obvious errors, yet the system granted a 100% accuracy rate with the green triangle (in the old version) and issued no warning (in the current one), reflecting a flaw in the automated evaluation mechanism. This case will be hereafter referred to as a false negative. These observations raised concerns about the reliability of Smartcat translation QA system, especially when used to translate educational texts that require high precision and awareness of the educational context.

1.3 Purpose of the Study

A previous study that was conducted by (Alotaibi, 2020) has evaluated the usability of Smartcat among Arab translators based on its efficiency and usefulness, concentrating on some specific features the tool includes. Her study has also raised concerns about the need for assessing other features Smartcat has for a better translation experience, pointing out that developers should work more to enhance the translators' experience. Given that the translation QA system of Smartcat is underresearched, the researcher decided to conduct this study with the aim of identifying the many ways the tool's QA system may enhance the translation experience or distort it. This study aims at examining the role that the translation QA system of Smartcat plays in specific. The researcher's aim is to uncover this feature, to understand its potential, and to evaluate its impact on translation workflows and outcomes. By doing so, we seek to contribute to a broader understanding of the evolving role of Smartcat translation QA system and the pivotal role it plays in ensuring translation excellence, shedding light on some recommendations that may put Smartcat in the forefront of (CAT) tools.

1.4 Research Questions

In this study, the researcher seeks to find answers to the following questions:

1. Is Smartcat translation QA system able to detect linguistic, terminological, formatting, and workflow errors in children's educational texts? And if so, is it able to correctly classify and describe them?
2. How compatible is Smartcat translation QA system with ISO international translation quality standards (ISO17100 and ISO 18587)?

1.5 Scope of the Study

The overall objective of this study is to examine Smartcat translation QA system to determine its efficiency and effectiveness, and to know when errors occur. The study will examine Smartcat translation QA system in the context of translating an educational checklist for children performance in early stages in a Kuwaiti nursery, from English into Arabic. This text is an assessment file for children and their levels in several disciplines, including language, reading, writing, mathematics, scientific reasoning, history, geography, music, and art. Given that this research is an applied type of research, the researcher selected this text to analyze the errors in its translation outputs. This translation was part of an actual commission assigned to the researcher, which makes the research practice-driven and grounded in real professional work. This text is a true representation of an official educational document used in a real-life context, providing an opportunity to examine the translation QA system in Smartcat in a practical and direct way. This checklist had previously been manually translated by the researcher, without the use of any software or platforms, as part of a real translation project for the aforementioned nursery, along with dozens of other files. As she began preparing this study, the researcher found that the file itself would serve as a rich application model if translated using the Smartcat platform, with the goal of studying the tool's effectiveness and efficiency in improving translation and measuring its features within a semi-professional environment. The study doesn't examine the QA system of Smartcat in the context of other disciplines, higher educational levels, or languages other than those mentioned.

The study adopts the Çetiner framework, which divides translation errors into four types: linguistic, terminological, formatting, and workflow, and which was developed when comparing Memsource's Translation QA system with the international translation quality standards like ISO 17100, ISO 18587, SAE J2450, and LISA QA Model. The Çetiner classification will be adopted as a practical framework, and the researcher will compare the translation QA system of Smartcat with international standards such as ISO 17100 and ISO 18587. In other words, these international standards represent the theoretical reference for evaluating translation quality, while the Çetiner classification represents the practical framework followed. In addition, the researcher examines the relationship between Smartcat translation QA system and the use of TM tool, as it serves the research objectives. Including this technological factor, the analysis enables a deeper examination

of the translation QA system, making it possible to explain the errors it produces due to the use of technology and artificial intelligence. Secondly, this tool is the core of Smartcat operating principle, and overlooking it would create a research gap and would ignore critical errors in the system's performance.

This study focuses on examining the translation QA system in the latest version of Smartcat, version 2025. The system will be evaluated by fully machine-translating the text and observing the system's evaluation of the translation.

The translation was also limited to the free MT available within the Smartcat platform, specifically Google Translate, excluding paid MT engines such as DeepL or Amazon Translate, and other free MT. (Google Translate) has been chosen for testing because it is easily accessible and most widely used by freelance translators. Google Translate is also suitable for small business owners and demonstrates a good ability to handle educational texts. Moreover, most freelancers opt for free MT tools due to budget limitations and ease of use. Recently, (Sandwich, 2025) has posted that Google Translate is the largest and most widely used in the world. Every day, more than 500 people use it to easily translate an estimated 100 billion words. Additionally, Google Translate supports 109 languages and is used in more than 200 countries around the world. Statistics have shown that the Google Translate app has been downloaded on over a billion Android devices more than a billion times.

Studying free MT in particular is extremely important, as it is widely used by translators and students for reasons related to budget and speed. The researcher's selection of the free version reflects the practical realities targeted by the study and aligns with common usage patterns in real-world translation practices. Therefore, this choice was deemed appropriate given the applied nature of the research and its focus on realistic usage conditions. In this study, the researcher relies on her own assessment without calling in translation experts for examination. Therefore, the results should be interpreted accordingly.

1.6 Significance of the Study

The importance of this study stems from the continuous development in the field of translation in our current era. The proliferation of computer-assisted translation tools has paved the way for eliminating the human factor and reaching a stage where these tools

are sufficient to save time, effort, and money. To reach this stage, extensive studies should be conducted on all computer-assisted translation tools, along with their translation QA systems. This study analyzes the translation QA system used in Smartcat in the context of educational texts for children in early stages, such as nursery, from English into Arabic. The researcher found a severe shortage in such research related to modern translation. This study paves the way for future researchers to conduct studies that have to do with developing computer-assisted translation tools and improving the translation QA systems available therein.

1.7 Limitations of the study

It's worth noting some limitations of the study, which may later affect the generalizability of its findings. The researcher selected two texts specifically for kindergarten children and limited her analysis to these two texts. This may affect the generalizability of the results to different age groups or different fields. The researcher focused exclusively on examining Smartcat translation QA system from English to Arabic and did not explore translations in the opposite direction or in other language pairs. Additionally, the study was limited to examining free MT (Google Translate) and did not address other paid translations. It is worth noting that the evaluation was conducted by the researcher alone, based on her actual experience working in a translation office. Also, the evaluation and analysis focused on the linguistic aspects and did not extend beyond them to include any technical aspects of Smartcat translation QA tools. Finally, the study's results may not be easily generalized due to time constraints and human limitations.

1.8 Literature Review

1.8.1 Introduction

In this time of remarkable technological development, experts in every domain, including translation, work tirelessly with full dedication to push their fields forward. Since translation is currently witnessing a widespread and rapid global expansion, the developers of translation software have been working hard to bring the translation process to the peak of its prosperity and vitality. It is worth mentioning that one of the most important elements of successful service delivery, whether in the field of translation or in other domains, is the implementation of a robust QA system. The QA system appeared in the early twentieth century, initially applied in industrial contexts.

Several scholars have explored the concept of QA in depth. (Knight & de Wit, 1999, p. 31) define QA as “the policies, attitudes, actions, and procedures necessary to ensure that quality is being maintained and enhanced.” In the same context, (Sallis, 2002, p. 5) explains:

“Quality assurance became an issue with the advent of industrialization. Previously, craftsmen set and maintained their own standards. Mass-production changed the emphasis. Work broke down into narrow and repetitive tasks, taking away from the worker the possibility of self-checking quality.”

This quote illustrates that increased productivity has made it difficult for workers themselves to verify the quality of their work, creating the need for systems that help verify the quality of production, just like the translation quality tools in this study.

1.8.2 Quality Assurance vs. Quality Control

Considering this historical overview, it is important to distinguish between the concepts of QA and quality control, particularly in the field of translation. Quality control maintains control by evaluating performance during operations for operating forces. QA verifies control after operations and informs management, regulators, customers, and others. (Juran, Godfrey, Hoogstoel, & Schilling, 1999). This study adopts the concept of QA as defined by Woodhouse. According to (Woodhouse, 1999, p. 29), “The phrase quality assurance refers to the policies, attitudes, actions and procedures necessary to ensure that quality is being maintained and enhanced...”

In the context of translation QA, translation quality tasks fall into two categories. The first is linguistic quality: correct grammar, spelling, punctuation, and natural style. The second is formatting quality: preserving original formatting, detecting errors like double spaces or inconsistent translations, and following the project glossary, which doesn't require deep knowledge of the target language. (Makoushina, 2007). From this perspective, the importance of computer-assisted translation tools becomes evident in facilitating these tasks. With regard to computer-assisted translation tools, (Torres-Hostench, Gil, & Leal, 2010, p. 255) state that “CAT tools have been integrated into the translation process because they save time, support multilingual project management, enable the reuse of previous translations, simplify file handling and formatting, standardize translation workflows, and promote professional specialization.” This quote

illustrates that CAT tools have become a key player in translation processes, improving their efficiency and organizing project management, especially multilingual projects.

Despite the growing use of CAT tools in the translation industry, this remains a relatively new area of study, especially when it comes to Smartcat translation QA system. Therefore, there is a scarcity of research specifically focused on it. Delving deeper into previous studies related to the Smartcat localization tool, the researcher found that no studies have been carried out to examine its translation QA system. However, the researcher was able to find two studies addressing the translation QA systems of other CAT tools. (Petrova, 2019, p. 91) examined five of the most prominent CAT tools in terms of their translation QA systems which are QA Distiller, Bench. Verifika, ErrorSpy and Ltb. As she attempted to clarify the issue of the resulting errors checked by the QA system of each CAT tool, she stated that during the initial phases of their evolution, “Translation QA tasks were grouped into two categories: grammar, involving spelling, punctuation, and fluency; and formatting, including detection of unnecessary spaces, redundant punctuation, and glossary inconsistencies.”

Another study by (Çetiner, 2021) examined another CAT tool called Memsource and analyzed the QA system of that tool in relation to these professional quality standards. The error analysis in his research categorizes errors into four main groups, with an additional custom check integrated into these classifications. These four groups are Linguistic, Terminology, Formatting and Tags, and Workflow. Each category includes subheadings aligning with specific error checks and corresponding functions. Based on the fact that the translation QA system is useful in the translation process and since it is adopted by freelance translators and translation companies when it comes to using CAT tools, research into all the translation QA systems for all CAT tools should continue. Building on this, the researcher aims at discovering the errors checked by Smartcat translation QA system and classifying them into different categories in an attempt to identify the gaps that may appear in this system, especially that, after testing it, the researcher noticed that sometimes a mistake warning still pops up even when the translation is flawless. This can be achieved through adopting the ISO 17100 and ISO 18587 international translation standards as a standard reference and following the application of the Çetiner classification of translation errors related to computer-assisted translation tools.

1.8.3 ISO 17100:2015 and its Role in Ensuring Translation Quality

Attention must be paid to the mission of the International Organization for Standardization (ISO, a non-governmental organization that was created in 1947 in Switzerland. This organization has 165 members who are experts in all fields from all over the world. These members meet in an attempt to agree on and develop international standards that aim to improve the quality of services across various industries. These standards are updated every five years to ensure that they remain up to date. In 2011, ISO set up a committee on translation standards, including ISO 17100 for translation services and ISO 18587 for the MT post-editing. (Popiołek, 2020)

ISO 17100 standard classifies the professional competencies that all translators must possess into six categories. Firstly, translation competence, which refers to the translator's ability to understand the source text and transfer it accurately into the target language, and their ability to problem solving, implementation of client instructions, and adherence to project specifications. Secondly, linguistic and textual competence, which refers to the translator being fluent in both the source and target languages, having extensive knowledge of texts, distinguishing between their types, and dealing with them during translation to ensure the translation appears more natural. Third, research and information acquisition competence, which refers to the translator being able to locate any information they need and not being unable to translate specialized words. Fourth, cultural competence, which refers to the translator being aware of the differences in customs and cultures between the source and target languages. Fifth, technical competence, which refers to the translator being experienced in computer-assisted translation programs, translation memories, and machine and electronic translation dictionaries. Finally, Domain competence, which refers to the translator being familiar with translating a specific type of specialization, such as medicine or sports, and being able to use specialized terminology accurately and professionally. (Karabıyık, 2019).

1.8.4 ISO 18587 and its Role in Ensuring MT Quality

This standard was developed to ensure the quality of MT output after post-editing. It ensures MT quality through six points. First, it specifies what MT output should be: linguistic and grammatical accuracy, clarity, error-freeness, adherence to style, and similarity to human translation. Second, it defines the skills and professional

competencies of editors, who must be linguistically, culturally, technically, and specializedly competent. Third, it divides the editing process into three stages: pre-editing, during editing, and post-editing. The process first assesses the text's translatability, then applies the criteria, followed by the revision phase. Fourth, the standard distinguishes between light and full editing, assessing whether the translation requires minor or substantial modifications, depending on the project's purpose. Fifth, the standard requires editors to provide feedback on the MT performance to continuously improve it. Finally, it requires editors to adhere to client instructions and project requirements, using correct language, appropriate linguistic style, and correct grammar ((ISO), 2017).

While the main goal of using MT and TM is to facilitate the translation process and provide assistance to the translator, they may be a flaw themselves and the main reason for disrupting the match and recognition operations in the system, and thus for causing it to produce wrong results. Previous studies have shown that although some professional translators recognize the benefits of computer-assisted translation tools, they are still not satisfied with the quality of these services and even report that they cause annoyance and confusion during work. (O'Brien, Ehrensberger-Dow, Hasler, & Connolly, 2017). In the same context, (Han, 2020, p. 3) stated, "Even if the rationale behind machine translation is the simplification of translation with artificial intelligence, human intervention cannot be substituted. The translator needs to give the final touch to ensure desired results." This is true since MT may display a defect in the QA system, making the human translator better placed to differentiate between the meanings of different words, which may be similar in spelling while different in meaning, and between names and adjectives, as well as many other issues. According to (Poibeau, 2017, p. 12), "a translator must be at ease with the two languages involved, and he must have special skills to reformulate a source language in a target language that does not have the same wording or the same structure. These kinds of skills are not directly available to machines". (Taravella & Villeneuve, 2013, p. 63) support the view that "Translation automation leads to human-machine interaction ranging from entirely human translation—translation craftwork—to entirely automated translation. Fully automated translation is in use but not yet good enough for widespread public use." From their side, (Craciunescu, Gerding-Salas, & Stringer-O'Keeffe, 2004, p. 4) clarified that "It is important to stress that automatic translation systems are not yet capable of producing an immediately useable text, as languages are

highly dependant on context and on the different denotations and connotations of words and word combinations”.

The second feature of Smartcat, which is the TM, will be also examined through examples later on. Alan & Wright (2014, p. 662) defined TM as "a database of paired text segments, where Segment B is a translation of Segment A. Translators use TMs to ‘remember’ the content of past translations." On the other hand, (Zerfass, 2002, p. 49) defines TM systems as “tools that ‘memorize the translations made by a human translator,’ noting that most TM systems consist of ‘a database that stores the original text along with its translation – a database of segment pairs.’”.

Šanca (2018, p. 17) argued that "a TM does not only provide fuzzy (approximate) matches, but also full (100%) matches and in some tools even in-context (101%) matches that take the preceding and succeeding segments, i.e., the context, into consideration"; accordingly, the research into whether TM contributes to improving the quality of Smartcat translation QA system or causes errors is still required for further exploration.

1.8.5 Translation Theories and their Applications in Quality Assurance

1.8.5.1 Introduction

In the context of studying Smartcat translation QA system, it is important to consider translation theories that have addressed the concept of QA from different perspectives. The study of translation quality cannot be based solely on linguistic and formal criteria; rather, the communicative function of the translated text and the context in which it is used must also be taken into account. Among these important theories, functional theories rank at the forefront. According to Zheng (2018, p. 623) “Functionalism” means focusing on the function of texts and translation”. Functional theories focus on the purpose and communicative function of the translated text, making them partially compatible with Smartcat translation QA system. The researcher will utilize this partial compatibility, highlighting the differences to determine the system's effectiveness in achieving translation quality standards from a functional perspective.

A discussion of these important theories will help us understand translation quality from a broader perspective, as each of these theories provides a different understanding of translation quality. The researcher will compare the operating principles of two important

theories related to translation quality with the operating principles and mechanisms of Smartcat translation QA system. This discussion will help provide a deeper understanding of whether the system takes into account the functional and communicative considerations of the text or whether it merely aims to preserve its form. In other words, the researcher will be able to prove the system's shortcomings, if there are any, by comparing its translation QA mechanism with major translation theories.

This analysis will help determine whether the translation QA system at Smartcat takes into account the functional and communicative aspects of translated texts, especially in an educational context when the texts are used by children and teachers.

1.8.5.2 Skopos Theory

The Skopos theory is a translation approach developed in Germany in the late 1970s by the German linguist and researcher Hans J. Vermeer. This theory marked a turning point between theories that focused on the formal linguistic aspects of translation and theories that emphasize the functional, contextual, and social aspects of translation. The main idea of this theory is that a good translation is one that successfully performs its communicative function in the target culture and environment and takes into account the target audience. In other words, the measure of translation quality, according to Skopos theory, is the extent to which it fulfills its function and properly reaches the target audience, rather than preserving the form of the source text (Schäffner, 1998).

Vermeer argues that “Linguistics alone won’t help us; first, because translating is not merely and not even primarily a linguistic process. Secondly, because linguistics has not yet formulated the right questions to tackle our problems. So let’s look somewhere else.” as cited in (Nord, 1997, p. 10).

Applying Skopos theory to quality assurance in the educational context

It is important to consider Skopos theory in this study because the translated texts under investigation are used in a specific environment and benefit specific individuals. These texts are used in a learning environment that benefits both teachers and kindergarten children. Therefore, the communicative function of the translated text is to support the educational process and facilitate classroom communication.

According to Skopos theory, indicators of translation quality in an educational context go beyond linguistic correctness to include the following:

- The suitability of the text for the target audience and age group.
- The clarity and ease of understanding of the instructions by children.
- The achievement of the educational objective of the text.
- The use of simple language appropriate to the educational context.

These indicators reveal functional and educational dimensions of translation quality that Smartcat's translation QA system cannot systematically and automatically monitor. This is because it focuses primarily on language and form. Even if a translator attempts to establish and adhere to specific standards and indicators, Smartcat's translation QA system, as an automated system, cannot verify whether the translator has complied with those standards.

Compatibility with Smartcat

There are points of agreement between the Skopos theory and the translation QA system of Smartcat. The Skopos theory emphasizes that the goal of translation is to fulfill the purpose of that translation, whereas the project managers at Smartcat can agree with the translators to consider the purpose of the translation instead of focusing on transferring the content literally and preserving the form. Another key element to consider is the focus on the target audience, as the Skopos theory prioritizes the target audience in its translation approach. Conversely, Smartcat can achieve the same goal through collaborative translation and agreeing on a translation mechanism that targets a specific audience.

As for the differences, the QA system in Smartcat relies more on linguistic standards than on functional standards. Additionally, Smartcat focuses on identifying linguistic and normative errors that may not necessarily be considered mistakes by Skopos theory. It can be said that there is a partial overlap between the main premises of the Skopos theory and Smartcat QA system, especially when Smartcat is used in a flexible and customized manner. However, the fundamental difference lies in that Smartcat tends to evaluate translation based on normative standards more than on functional or goal-oriented considerations, as in Skopos theory.

1.8.5.3 Julian House model

The Julian House model is one of the most important translation models. Considered as a functional linguistic assessment model, it compares the source text and the target text with a focus on meaning and context. It evaluates the translation not only in terms of linguistic equivalence but also in terms of its functional appropriateness and cultural context. House explains the difference between two types of translation within this model: covert translation (implicit translation) and overt translation (explicit translation). An implicit translation appears to be an original text formulated in the target language in the target culture. An overt translation occurs when a text appears to be unoriginal in the target language and culture, but rather appears related to the source language and culture, and the target audience recognizes it as unoriginal (House, 1997).

Quality indicators in House's model and the educational context

- Conveying the meaning correctly
- Text coherence and clarity
- The text's suitability to the culture of the target audience
- Achieving the intended goal of the original text

These indicators underscore the importance of the functional and cultural aspects of translation quality, which are particularly important in educational texts. However, the translation QA system in the Smartcat platform does not fully capture these aspects through automated verification processes, which focus primarily on linguistic accuracy.

In the context of texts aimed at children, what is more important than grammatical accuracy in translation is that the meaning is conveyed smoothly and easily to the child, as if they were listening to simple, everyday conversations. The child should not feel that the text is translated from another language; rather, the translation must be appropriate to the child's educational and cultural environment. Achieving a translation of this caliber requires human judgment and functional evaluation, as it is a task that automated translation QA systems cannot perform correctly on their own.

Compatibility with Smartcat

The Julian House model and the translation QA system in Smartcat both focus on comparing the source text and the target text and emphasize linguistic errors such as meaning, grammar, coherence, and the use of correct terminology. The Smartcat translation QA system aligns with the second type of translation in House's model, which is overt translation, where the text is preserved in terms of language, meaning, and form. Meanwhile, it differs from the first type of translation, which is covert translation, in which Smartcat does not inherently aim to create a text that seems like a new text distinct from the source text unless this is the goal desired by project managers and clients.

1.8.6 Summary Comparison Table

The following table summarizes the key similarities and differences between the two selected translation quality theories/models and the Smartcat QA System:

Table 1

Comparison between Smartcat QA features and functionalist translation theories

Theory/Model	Focus	Compatibility with Smartcat QA	Differences
Skopos Theory	Purpose/function of translation	Can align when the purpose is prioritized by project managers	Smartcat relies more on linguistic/normative criteria
House's Model	Linguistic-functional equivalence	Smartcat aligns with overt translation and linguistic comparison	Does not support covert translations or cultural adaptation unless there are specific instructions by project managers.

In conclusion, these two theories are functional theories. Functional theories focus on the function and purpose of translation in the target language and culture.

According to Nord (2014, p. 1) "Functionalist means focusing on the function or functions of texts and translations. Functionalism is a broad term for various theories that approach translation in this way." This section explains that although Smartcat provides partial support for functional considerations, its automated systems for ensuring translation quality do not fully meet the complex requirements of functional translation theories. Such theories require a complex and sensitive workflow that differs from that of

computer-assisted translation tools that rely on automated assistance. (Sdobnikov, 2019) points out that the functional approach to translation requires the translator to understand the communicative context in which the text occurs and analyze the contextual data to determine the tasks the translator must accomplish. He also emphasizes that translation is a purposeful process aimed at writing a text that serves the same purpose for the target audience. This also supports the argument that the automated systems for ensuring translation quality in the Smartcat platform prioritize linguistic form over communicative purpose which has been emphasized in functional translation theories.

This approach helps in evaluating Smartcat translation QA system in terms of whether it actually helps in understanding the context and purpose of the translation or whether it merely checks, corrects and evaluates based on the words and external form.

Chapter Two

Methodology

2.1 Introduction

The overall objective of this chapter is to illustrate the methodological framework used by the researcher to examine the translation QA system used in the Smartcat tool. The subsections that follow this introduction provide a detailed explanation of the research design, research instrument, data collection, analysis methods, validity and reliability tools, ethical considerations, and theoretical framework. Given the nature of the study, which seeks to analyze MT in the Smartcat tool and examine the effectiveness of the evaluation of the translation QA system, a qualitative approach was adopted to measure the ability of the translation QA system in Smartcat to detect errors selected and analyzed. The aim of using a qualitative approach was to determine whether Smartcat translation QA system was able to accurately detect errors and to determine the validity of the feedback provided by the system, especially in educational texts for early age groups, such as nursery-related texts.

2.2 Research Design

This study adopted the qualitative approach by manually detecting, classifying, and interpreting errors. Descriptive evaluative design will be employed in this study. The study is deemed descriptive as it describes the assessment process done by Smartcat translation QA system after the translation of any segment, and in the sense that it aims to identify the gaps in this tool and explain why they occur. However, it is considered evaluative as well because it provides an evaluation of the whole process done by the QA system of Smartcat after checking the results. The researcher follows this assessment process in order to make a complete integrated evaluation of Smartcat translation QA system and to know how successful it is and when it fails.

The study was based on an in-depth analysis of a single case in which the researcher translated an educational checklist for children in early childhood (nursery) using the Smartcat tool from English to Arabic, using free MT (Google Translate) to discover potential errors that the translation QA system makes when dealing with this type of texts.

First, the researcher reviewed the notifications issued by Smartcat translation QA system regarding a specific error and examined the feedback it provided. The researcher then evaluated these notifications as valid errors and false notifications. Second, the researcher manually searched for errors in the MT, categorized them, and determined whether the translation QA system detected them. The researcher evaluated Smartcat translation QA system in terms of false positives and false negatives.

2.3 Research Instrument

The study relied on one main tool to examine the translation QA system used in Smartcat. The tool is a readily available technological tool, Smartcat (computer-assisted translation software). As previously mentioned, this tool includes three auxiliary sub-tools: MT, TM, and the translation QA system, which is the main focus of the thesis. All of these tools work together to form a comprehensive system for detecting translation errors, such as formatting issues, omissions of words, and the addition of other words. In addition to examining Smartcat translation QA system, a quick comparison was conducted using a table between the system and its counterpart on the Memsource platform. The two systems were compared in terms of their ability to detect errors suggested by Çetiner.

2.4 Theoretical Framework

2.4.1 Introduction to the Theoretical Framework

This study is based on Çetiner's error classification as a theoretical framework used to analyze translation errors in a technical context. Çetiner developed this classification based on a comparison that he made between the QA system of Memsource and some international standards for translation, two of which the researcher will also use as a basic standard reference. This classification is characterized by being objective and appropriate for the computer-assisted translation environment. Although the researcher did not modify the basics of the classification itself or add new error types, she expanded the scope of the analysis to include a technological factor that may affect the translation QA system in Smartcat, which is the TM tool. In addition, the researcher added an analytical dimension in which she discussed linguistic and terminological errors beyond Çetiner's surface-level.

2.4.2 Çetiner's Error Typology (2021)

The study draws on comparing Smartcat QA system with the ISO 17100 and ISO 18587 standards by applying Çetiner' errors typology (2021) to highlight the possible errors that can be detected immediately by the system. First, Çetiner divided the translation errors appearing on the Memsource platform into four categories. Then, he designed a table in which these categories were listed, with subdivisions showing the subcategories of errors under each main category. He added a column to explain the reason behind the error of each type to justify why it is considered an error in the first place.

This QA overview, introduced by Çetiner resulted in four categories of errors, which are Linguistic, Terminology, Formatting and Tags, and Workflow. The linguistic errors from Çetiner' errors typology include (Empty target, extra numbers in the target, Inconsistent translations, leading and trailing spaces, maximum target segment length in % of the source, maximum target segment length in characters, missing numbers, multiple spaces, repeated words, spelling, target text identical with source one and trailing punctuation, unedited NT/MT fuzzy match, and unedited TM fuzzy match).

As for the terminological errors, they include (missing terms, forbidden terms, and not allowed). Concerning the errors related to tags, the researcher will consider the errors resulting from (Empty tag content, inconsistent tag content, nested tags, no text between the tags, tags/formatting are missing or invalid, tags-joined segments, and XLIFF tags). It should also be pointed here that the researcher will be adopting only one subcategory from the fourth type of error introduced by Çetiner, which is (not confirmed segments) as the rest two (unresolved conversations, and custom checks) don't align with the research objectives, as discussing them would lead into a context far removed from the translation itself. Later, the researcher will present a table of Çetiner's main and sub-categories of translation errors and answer the question of whether the Smartcat translation QA system is able to detect these errors or not.

Since this is an applied study, the researcher has extended the analysis beyond Çetiner's classification. This suggestion is based on her hands-on experience using the Smartcat platform, which showed some practical aspects that were not addressed in the original model. So, to tailor the framework further to the Smartcat interface, the researcher expanded the scope of the analysis to examine the potential relationship between

Smartcat translation QA system and the platform's use of TM. Based on her experience as a translator working for a translation company, the researcher translated the entire file using MT and collected translation errors to determine whether there were any errors not covered by Çetiner's classification. She then reviewed these errors to determine whether the translation QA system would be able to detect them. All these aspects were taken into consideration to frame an in-depth study that may end with useful results related to the translation QA system of Smartcat.

To deepen and enhance the analysis of errors through the main table introduced by Çetiner, the researcher added more details about each error. The researcher did not just present the error and its type, and explain its cause, but she also discussed the method of detecting the error and the source of the error.

Regarding the method of detection, it will determine whether the error was detected manually or automatically. In other words, it will tell whether the error was detected by the researcher herself, or by Smartcat through the system. It will also tell whether the translation QA system gives an error notification while the sentence is error free. Following this, the source of the error will be added. The aim of the researcher's additions to Çetiner's model lies in making the table a precise logical analysis tool, not just a list.

The proposed structure for categorizing translation errors in this study includes a space for each error example (the source text example, the erroneous translation, the corrected version, the specific sub-category, the detection method, and the probable source of error).

After collecting and analyzing the examples according to Çetiner's classification, the extent to which the results and translation outputs comply with the quality requirements of ISO 17100 and ISO 18587 will be evaluated.

2.4.3 Theoretical Positioning of the Study and Its Conceptual Contribution

This study bridges the gap between translation theories, such as Skopos theory and House model, which focus on the function, meaning, and impact of translation on the reader, and translation applications, or what are called computer-assisted translation tools, which rely on modern technology for QA, such as the Smartcat platform and its translation QA system. It aims to determine whether this application adheres to the principles of these

theories. Functional theories maintain that translation quality is assured when function, meaning, and context are preserved. A good translation, according to them, meets the intended purpose, conveys meaning to the target audience that must understand and comprehend the target text, and belongs to the cultural and educational environment of the text. This is not the case with the Smartcat application, which gives first place to linguistic correctness and pays great attention to formal aspects such as spelling, punctuation, formatting, and repetition of words. By comparing the Smartcat translation QA system with functional theories and international quality standards like ISO 17100 and ISO 18587, this study reveals a clear difference in the definition of good translation between the theories and the system. Herein lies the conceptual problem, which goes beyond being merely a technical error in the platform. The first conceptual contribution of this study relates to expanding on Çetiner's (2021) framework of translation errors by distinguishing between two levels of quality in technical translation. The first level is technical quality, which has to do with simple errors that computers can detect. The second level is deep quality, which relates to deep errors that only a human can detect. This distinction is not explicitly present in previous studies, which constitutes a conceptual contribution.

2.4.4 Summary

The translation error analysis was based on Çetiner's model for classifying errors because it is clear, categorizes errors in a scientific way, and depends on comparing the QA system of CAT tools with the international ISO standards for translation services, making it suitable for use in computer-assisted translation environments. The model was adopted in its original form, but the analysis was developed by exploring two technological factors in Smartcat that may affect the assessment process in the system and assist in the research process. In addition, the researcher added linguistic and terminological translation errors beyond his classification. This was done to tailor the model more specifically to this study and to directly link it to the evaluation of the Smartcat translation QA system.

2.5 Data Collection

The researcher's data collection method is based on testing the assessment process that Smartcat translation QA system uses in the translation of a file from the Reggio Kuwaiti Nursery, which follows the "Reggio Emilia" approach that originated in Italy after World

War II. This file is an educational checklist for evaluating the performance of children in the KG1 stage during the 2024/2025 academic year, based on the nursery's academic standards. The title of this text is "Kuwait Reggio Center KG1 Standards" and it contains 3183 words. It has special design elements and formatting, and it has a variety of vocabulary that puts the researcher in the heart of the topic. The educational standards in this checklist cover all areas, including language, reading, writing, mathematics, scientific reasoning, social and emotional learning, physical education, history, geography, music, and art. The text is organized using a table consisting of rows and columns. Within this table, there were seven main headings that introduce the educational subjects, such as social and emotional learning, language and literacy, math, scientific reasoning and the physical world, history and geography, music, and art. Five of these sections are followed by rows that outline learning goals, which are numbered and grouped into subcategories.

Regarding text formatting, the main headings are distinguished from the rest of the text by using bold and a large font. Numbered lists are used to classify learning standards. In addition, main phrases are italicized or written in a different color, for example, blue, to emphasize their importance. In terms of punctuation, the majority of sentences end with periods for being short and direct. Parentheses are used in cases of giving examples, and commas are used to separate elements of a sentence. Long paragraphs do not appear in the text, as it primarily relies on numbered as well as bulleted points. All of the previous characteristics make this specific text an appropriate choice for studying the performance of the translation QA system across various fields. The researcher can compare the system's evaluation of educational terminology versus scientific terminology, as well as compare the assessment of the translation of procedural language versus descriptive language. Additionally, research can be conducted on the system's evaluation of complex scientific sentences versus simple, easy sentences. Also, the presence of different formatting styles in the text contributes to examining the effectiveness of the translation QA system at Smartcat in identifying errors resulting from formatting and layout.

The researcher accessed the platform and translated the text between May and June 2025. The translation process required uploading the file to the platform, selecting the human translation and proofreading option, and reviewing the evaluation from the QA system for each segment where the translation is considered incorrect. During the translation process, the sentence was first translated by using (Google Translate) suggestions,

followed by a manual translation, in order to differentiate between the QA system evaluations of both translations. Then, the researcher gathered the errors identified by the Smartcat translation QA system and began classifying them according to Çetiners' error typology.

The researcher studied the workflow of the translation QA system of Smartcat to see if there are any gaps through testing the segments translation. This included examining the errors checked by the translation QA system after the translation of each segment in isolation from the rest of the file, on the one hand, and examining the checked errors considering the whole text and taking the context and other features in Smartcat into account, on the other hand. The researcher described and classified these errors into linguistics, terminology, formatting, and workflow. As previously mentioned, the errors were classified and analyzed according to Çetiner's error typology. Since the researcher has added analytical elements to the table, such as the method of detection (manually or via QA), and the source of the error, each error was analyzed from these perspectives. An evaluation of the results took place after a deep analysis of the data collected, so that the researcher was finally able to give some recommendations to enhance the use of the QA system of Smartcat. (Jiménez-Crespo, 2010, p. 195) points out, "Translation memory systems aim to reuse previously translated texts. Since the primary unit of first-generation translation memory systems is the sentence, such systems are only useful for text types where full-sentence repetition frequently occurs." So, to examine the TM, the researcher selected a text similar to the text under study and used the text under study as a TM within Smartcat.

It is important to note that the previous description in the first chapter of the Smartcat platform and the translation QA system it has was explained to provide a general background on the platform's development. In previous versions, the system issued a notification using a triangle (green, orange, red) representing the severity of the error or sentence correctness, along with a percentage out of 100. For this study, the data were analyzed using the latest version of Smartcat, 2025. In the current version, the green and red triangles and the percentage were removed. It is important to clarify this point, as the results of the data analysis will be considered in this context.

2.6 Validity and Reliability

Selecting an educational checklist as the study text contributed to achieving both validity and reliability in the study. In terms of validity, the researcher selected this authentic text as it enhances ecological validity. In addition, the text contains various and complex formats, such as numbered tables and lists, colored words, and underlined words. Such formatting is integral to the Smartcat tool's functionality, allowing for measuring the tool's actual performance. The text also contains specialized vocabulary, as the text is divided into several disciplines, such as language, reading, writing, physical education, history, geography, scientific reasoning, mathematics, music, and art. This allows for testing how the system deals with the outcome of the MT and evaluating Smartcat translation QA system, making the study more comprehensive and increasing the generalizability of the results.

In terms of reliability, the researcher used a previous classification framework, the Çetiner classification framework, and compared the results with international ISO standards, which reduces personal bias in the analysis. The researcher also expanded the analysis method to include one of Smartcat tools that assists in examining Smartcat translation QA system, like TM. Also, the presence of a unified mechanism for detecting errors within the Smartcat platform and its translation QA system (such as the orange triangle and feedback) helps researchers retest the results. Finally, the researcher documented all steps of the study and analysis to support the possibility of conducting the study again on other text types in the future.

2.7 Ethical Consideration

The researcher didn't write the study text herself. The study text used in this study was obtained from a children's nursery in Kuwait. It is an empty template with no personal information about children, parents, or staff. The researcher got Permission from the nursery director to use the text on condition that the nursery's name is provided and no personal information is revealed. The text itself did not contain any sensitive information that may affect any individual's privacy or cause them harm. The intellectual property rights of the original text were adhered to. There was no information forged or concealed, and the research was conducted with honesty and transparency.

Chapter Three

Data Analysis

3.1 Introduction

This section provides a detailed analysis of the data collected in this study to evaluate the performance of Smartcat translation QA system. In this section, examples of linguistic errors will be analyzed according to Çetiner's classification to determine whether Smartcat translation QA system is capable of detecting them. Beyond surface-level linguistic errors will then be examined to determine the accuracy of the translation QA system's assessment. The same analysis will be conducted for terminological errors, formatting errors, and workflow errors. After analyzing each type of error individually, Smartcat translation QA system standards will be compared with ISO translation standards such as ISO 17100 and ISO 18587 to determine their compatibility and compliance.

3.2 MT and Segmentation Challenges

As she uploaded the file to Smartcat, the researcher noticed that the platform does not always capture the functional meaning of sentences due to the nature of Smartcat sentence segmentation. Translating sentence by sentence in isolation from the full text, Smartcat makes translation errors. The researcher first uploaded the file and performed a full free MT (Google Translate). This step aimed to examine how Smartcat handles translation, what the relationship between sentence segmentation and translation accuracy is, and how Smartcat translation QA system evaluates the MT and whether it is related to the system's accuracy or inaccuracy.

3.3 Error Detection and Classification Analysis

The purpose of the MT experiment was to test whether Smartcat translation QA system was able to detect linguistic, terminological, formatting, and workflow errors and correctly classify them. To increase the credibility of the results, the experiment was practice-driven and focused on translating a real translation file under real working conditions.

Data analysis began by examining whether the system was able to detect the error. The MT errors that the system was unable to detect were then collected and compared to the correct translation proposed by the researcher, to determine the cause and type of error. The errors that the system was able to detect were then analyzed, identified, and classified to determine whether they were actually errors, or merely a glitch and to determine whether they were correctly classified.

3.4 Data Classification and Applicability

Eighteen types of errors were analyzed under Çetiner's classification, which included linguistic, terminological, formatting, and workflow errors. Of the 18 different Çetiner's errors, Smartcat translation QA system was able to detect only nine types, representing a score of 50%, while it was unable to detect nine errors, also representing a score of 50%. For an objective examination, the researcher identified some errors as "not applicable." The system's programming differs from previous versions and from MemSource, and the text under study did not include some of the cases in the Çetiner classification for initial examination. Conversely, the answer "yes" was given to errors that Smartcat translation QA system was able to detect, while the answer "no" was given to errors that it could not identify.

As for the cases the researcher considered "not applicable," most of them were related to the orange tags responsible for saving formatting in the previous version. These tags no longer exist in the current version under study, so they are simply described as "not applicable." The following subsections explain these types of errors, what was detected, and what was not. The collected data reveals that Smartcat translation QA system was able to detect simple formal errors while failed to detect more complex and difficult errors.

3.5 Linguistic Errors

3.5.1 Errors within Çetiner's Surface-level Category

According to Çetiner's classification, linguistic errors include fourteen types. In the following table, the researcher indicates which errors were detected by Smartcat translation QA system and which were not.

Table 2

Types of linguistic errors detected/ undetected by SC QA according to Çetiner's classification

Main Category	Çetiner's Subcategory	Detected by Smartcat QA?
Linguistic	Empty target	No (prevent confirmation)
Linguistic	Extra numbers in the target	No
Linguistic	Inconsistent translations	Yes
Linguistic	leading and trailing spaces	Yes
Linguistic	maximum target segment length in % of the source,	No
Linguistic	maximum target segment length in characters	No
Linguistic	Missing numbers	Yes
Linguistic	Multiple spaces	Yes
Linguistic	Repeated words	Yes (only exact matches)
Linguistic	Spelling	Yes
Linguistic	Target text identical with source one	No
Linguistic	Trailing punctuation	Yes
Linguistic	unedited NT/MT fuzzy match	No
Linguistic	unedited TM fuzzy match	Yes

3.5.2 Errors Beyond Çetiner's Surface-Level

Çetiner's classification relied on the automated QA system to classify errors. Therefore, it focused on simple, superficial errors that can be detected by software. However, during the researcher's translation experience and her use of the Smartcat platform, a group of deeper linguistic errors was identified that may not be detected by automated proofreading. These errors go beyond the technical aspect to include stylistic and structural aspects. This observation aligns with the ISO 17100, as all of these aspects are within the scope of both translation competence as well as linguistic and textual competence. The translator must not only be grammatically sound, but also stylistically consistent, textually fluent, and contextually relevant. This section details four types of deviations that appeared in the MT output, namely inconsistent translations, part of speech errors, lexical and collocational errors, and hyper correction.

3.5.2.1 Stylistic Deviation

This subsection addresses errors related to structural inconsistencies within sentences or lists appearing in the MT available on the Smartcat platform, as well as errors that result from literal translation and cause a loss of function or stylistic effect. The researcher classified those errors into four main types: Inconsistent translation, part of speech errors, lexical and collocational errors, and hypercorrection. These proposed types of errors focus on errors that go beyond superficial issues, which emphasizes the need for strong translation and linguistic competencies, as shown in the ISO 17100.

1. Inconsistent Translation

The first observation has to do with the stylistic consistency between items in a list. At the beginning of the text, certain items are listed consecutively and are translated in the definite infinitive form, such as (الاستيعاب، الاعتناء، التواصل). Then, suddenly, the MT uses the verb "يعبر" instead of the definite infinitive "تعبير", which differs from the pattern and causes a stylistic flaw. This was not considered an error by the QA system; rather, it was an error discovered by the researcher. We note here that the translation of "expresses feelings" as "يعبر عن مشاعره" is considered correct in isolation from the text as a whole, but the researcher considered it incorrect, showing one stylistic weakness in Smartcat translation QA system.

2. Part of Speech Errors

In some of the MT outputs, the researcher noted that verbs indicating children's actions were translated in a wrong way. For instance, action verbs indicating a child's ability to perform an action, such as "draw a picture" or "dress independently" were translated into "ارسم صورة" or "ارتداء الملابس بشكل مستقل". The MT used the imperative form or the verbal noun in this case, thus distorting the original function of the original sentence. Instead of telling what the child can do, it becomes an instruction directed at the child or a nonfunctional linguistic expression. In this case, Smartcat translation QA system was unable to detect the error and deemed the sentences correct.

3. Lexical and Collocational Errors

Examining how the system translated the expression "Provide a rhyming word," the researcher found that the MT provided the translation as "قدم كلمة قافية". This translation is literal and not usual in the Arabic instructional context, not having functional clarity, particularly when directed at early childhood learners or language learners. The MT used an inappropriate verb with the expression "rhyming word". "كلمة قافية" is not an appropriate term in the Arabic language, and the word "قدم" is considered unnatural in school education, which would apply more straightforward words such as "اكتب" or "اذكر".

A more precise communicative translation would be "اذكر" or "اكتب كلمة تُشابهها في القافية" as this formulation is clearer to the learner and better serves the educational purpose. Despite this translation issue, Smartcat QA system did not classify this error as a flaw or inaccuracy, indicating a deficiency in the system's pedagogical and contextual sensitivity, as the assessment relied on linguistic form without considering the phrase's function within the educational context and without considering the combination of words in the sentence.

4. Hyper Correction

Also, the system deemed the translation of "use zippers" to "استخدم سحابات" incorrect, as it considered the word "سحابات" to be misspelled, even though this word is grammatically and orthographically correct. The researcher suggested a better functional translation, which is "يفتح ويغلق السحاب بنفسه".

Table of Examples:

Below is a table summarizing the above-mentioned linguistic errors that appeared in the MT.

Table 3

Linguistic errors beyond surface level appeared in the MT

ST Example	Erroneous Translation	Corrected Version	Subcategory	Detection Method	Source of Error
Conceptualize, care, communicate, express	استيعاب، اعتناء، تواصل، أن يعبر	استيعاب، اعتناء، تواصل، تعبير	Inconsistent translation	Manual Review	Segment isolation
draw a picture, dress independently	ارسم صورة، ارتد ملابسك	يرسم صورة، يرتدي ملابسه بنفسه	Part of speech error	Manual Review	Target text identical with source structure
Provide a rhyming word	قدّ مكلمة قافية	اكتب كلمة على نفس القافية	Lexical and collocational error	Manual Review	Syntactic simplification by MT
use zippers → use zippers	Considered error because of the word 'zippers'	يغلق ويفتح السحاب بنفسه	Hypercorrection	QA tool (false positive)	Lexical database limitation

5. The Compatibility of Smartcat QA System with ISO 17100 and ISO 18587 Requirements Regarding Linguistic Errors

Considering the examples of linguistic errors described above, the researcher found that Smartcat translation QA system does not comply with the international translation standards ISO 17100 and ISO18587. The system ignored two of the most important criteria of ISO standards for high-quality translation: translation competence and linguistic competence. For example, Smartcat translation QA system did not consider the inconsistent translation of "يعبر عن مشاعره" as incorrect, despite the presence of clear stylistic errors between the phrases. The system also failed to notice that the MT incorrectly conveyed the idea of describing the child's ability to perform a task, such as "ارسم صورة", and even converted it into the imperative form of the verb "ارسم صورة". Furthermore, the system did not succeed in noticing the structural deficiency in the translation of "قدم كلمة قافية" despite its lack of clear pedagogical connotation in an

educational context. Finally, the system considered the word "مسابات" incorrectly spelled, despite the fact that it's correct. This reflects a lack of understanding of the correct usage of words. One of the most important criteria that ISO standards call for is that the translator must be able to produce a linguistically and stylistically coherent text, and this is not available in the Smartcat translation QA system.

3.6 Terminological Errors

3.6.1 Errors within Çetiner's Surface-level Category

According to Çetiner's classification, terminological errors include three types. In the following table, the researcher indicates which errors were detected by Smartcat translation QA system and which were not.

Table 4

Types of terminological errors detected/ undetected by SC QA according to Çetiner's classification

Main Category	Çetiner's Subcategory	Detected by Smartcat QA?
Terminological	Missing terms	No
Terminological	forbidden terms	No
Terminological	not allowed	No

3.6.2 Errors beyond Çetiner's Surface-level

Çetiner's definition of terminological errors was limited to the three aforementioned types: missing words, forbidden words, and not allowed words. However, during her experiment with the Smartcat platform, the researcher observed terminological errors related to the failure to choose a term or expression that was contextually and functionally appropriate. It is worth noting that a narrower definition of terminological errors has been used in this section to broaden understanding, knowledge, and analysis. Terminological errors are those that result from the use of an equivalent in the target language that is inappropriate for the context and specialty. This expansion is not a departure from Çetiner's classification, but rather a critical and applied extension of it. Terminological errors in the narrow sense are more difficult to detect than linguistic errors, such as grammatical, spelling, and punctuation errors. The researcher compiled a comprehensive set of specialized terms in the file, which covers a wide range of subjects, such as

language, mathematics, scientific reasoning, history, geography, music, art, and physical education. She was able to collect translation errors in these terms that were generated by MT, and found that Smartcat translation QA system was unable to detect any of them. After collecting these terminological errors, the researcher categorized them into three types, which are: lexical generalization of technical terms, literal translation of conceptual terms, and terminological conflation (overlapping concepts).

3.6.2.1 Lexical generalization of technical terms

When translating, in many cases, translators neglect to convey the precise meaning appropriate to the context, resorting to the easier option of using a general word they believe will suffice. This is what the researcher observed regarding Smartcat MT. "phonemic awareness" was translated into "الوعي الصوتي," while the correct translation is "الوعي الفونيمي." This translation is terminologically incorrect because "الوعي الصوتي" refers to sounds in general, not phonemes. Also, "one-to-one correspondence" has been translated as "المطابقة الفردية," while in mathematics it is a term that means "الاقتران واحد لواحد." Considering the translation of "fingerplays", the word has been transferred into the target text by choosing a very general word, "ألعاب الأصابع", which does not express "أنشطة الأصابع الايقاعية", which are activities in which the fingers are used to produce certain rhythms. The same applies to the phrases "temporal words," "visual elements," and "written numerals." The first of these terms was rendered into the Arabic language as "كلمات زمنية" instead of using "مفردات زمنية", the second one was translated as "عناصر مرئية" rather than "عناصر بصرية", and the third one as "أرقام مكتوبة" instead of "أعداد مكتوبة". The term "observe" was used in the Arabic language as "راقب" instead of using the Arabic word "لاحظ", which is the more accurate term in scientific and educational environments. The word "راقب", though, has a security or military connotation. Another such example that helps in illustrating this issue is the translation of the word "motif" as "زخرف" instead of using "عنصر زخرفي متكرر". In art-related contexts, "motif" refers to a set of decorative elements repeated in a particular way for an artistic purpose. The use of the word "زخرف"

is a very general translation that refers to decoration in general. Another example is the translation of "pattern card" as "بطاقة نمط" instead of "بطاقة نمط بصري". In general terms, a pattern may be linguistic, behavioral, or motor, but in this context, the intended pattern is the pattern of shapes or colors, i.e., the visual pattern. An important example of this issue that has emerged in MT is the translation of "quantity words" into "كلمات الكمية" instead of "ألفاظ الكمية". The first translation choice is considered incorrect because it is a literal translation that suggests the use of ordinary words. The second choice of translation is more appropriate because it is usually used in the context of teaching mathematics. Another example worth noting is the translation of "Basic Geographic Concepts" into "مفاهيم جغرافية بسيطة" instead of "مفاهيم جغرافية أساسية". A single simple word here distorts the entire meaning of the sentence. In all of the examples mentioned and in many others, Smartcat translation QA system was unable to identify, detect, and classify any of the errors.

3.6.2.2 Literal translation of conceptual terms

Some translators make the mistake of translating literally without considering the purpose of the translation, choosing the closest linguistic dictionary meaning without taking the functional context into account. When analyzing the MT results of the text under study, the researcher noted that a number of terminological errors resulted from literal MT. The first example is translating "delayed desires" as "تأجيل الرغبات" instead of "تأجيل الإشباع". The literal translation error here is that desires cannot be delayed; a desire remains a desire, while the gratification of desires is what can be delayed. The second example is translating "hands-on investigation" as "استكشاف عملي" instead of "تحقيق عملي". Simply put, in the context of nurseries, children explore the texture of things with their hands, so the second option is more appropriate. The third example is the translation of "representational space" as "المساحة التمثيلية" instead of "الفضاء الرمزي". The second option is more appropriate in this context because it expresses the child's imaginary mental space while expressing things that exist in reality, while the first option is a weak literal translation. The last example is the translation of "geographic features" as "الميزات الجغرافية".

instead of “الملاح الجغرافية”. The word “ميزة” suggests something that is useful and has positive aspects, while the word “ملمح” means that there is a visual or physical characteristic. Another example is translating "straight edge" as a “حافة مستقيمة” instead of “مسطرة”. It should be observed here that Smartcat translation QA system did not consider any of the aforementioned translation errors as errors.

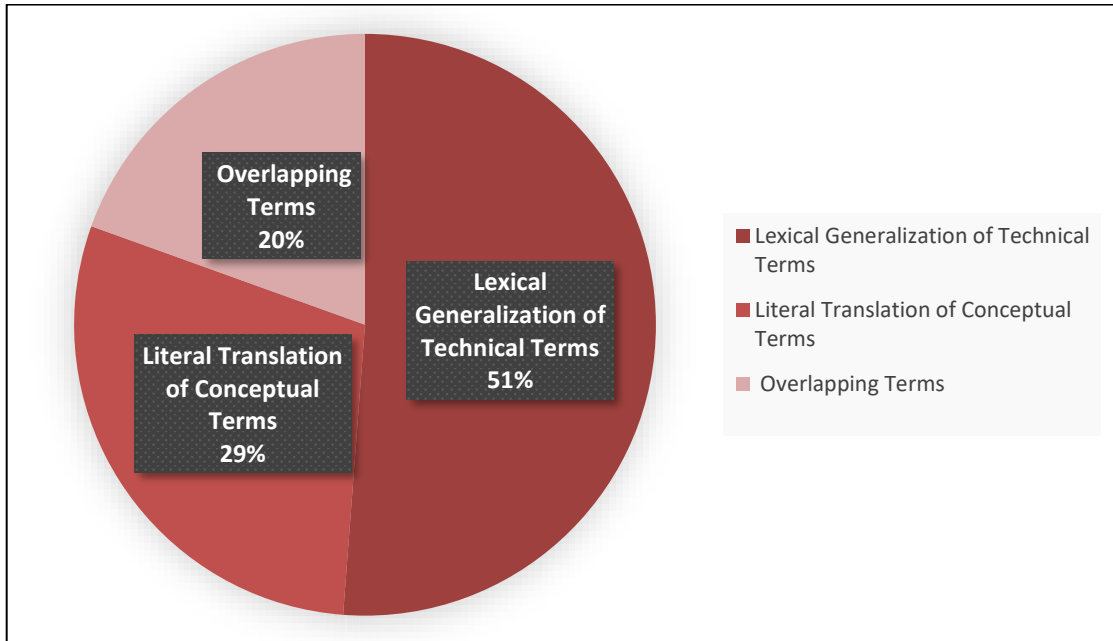
3.6.2.3 Terminological Conflation (Overlapping Concepts)

Sometimes, when translating closely related words within the same discipline, confusion and ambiguity occur, resulting in an inaccurate term being chosen. This phenomenon was observed in Smartcat MT in many examples. One of the such examples is the translation of both "rhythm" and "tempo" into "إيقاع". It must be pointed out here that there's an extremely huge difference in meaning between the two terms in a musical context. "rhythm" means "إيقاع" and it refers to the distribution of the sounds and clicks, while "tempo" means "السرعة الإيقاعية", that is, the speed at which a musical rhythm is executed. Another error that appeared in the MT occurred in the translation of the term "physical characteristics" into "خصائص جسدية" instead of "خصائص مادية". Choosing the first translation option is correct only if the context refers to humans or living beings. The context to be translated, however, referred to materials, making the second option more suitable. In any case, Smartcat translation QA system was unable to detect any of the errors mentioned above.

The researcher found 85 terms in the text under study. The accuracy and precision of the MT of these terms using Smartcat were examined. It was found that 44 of these terms were translated correctly, while 41 were incorrectly translated. The errors were divided into the three types mentioned above: lexical generalization, literal translation, and overlapping concepts. Most errors fell into the category of lexical generalization, accounting for 51.2 percent of the total. Literal translation errors accounted for 29.3 percent of the total. Errors related to overlapping terms accounted for 19.5 percent. (see pie chart 1). Smartcat translation QA system could not detect any of the 85 errors. This shows a clear deficiency in the system's operating principle.

Figure 4

Distribution of terminological errors in MT using Smartcat

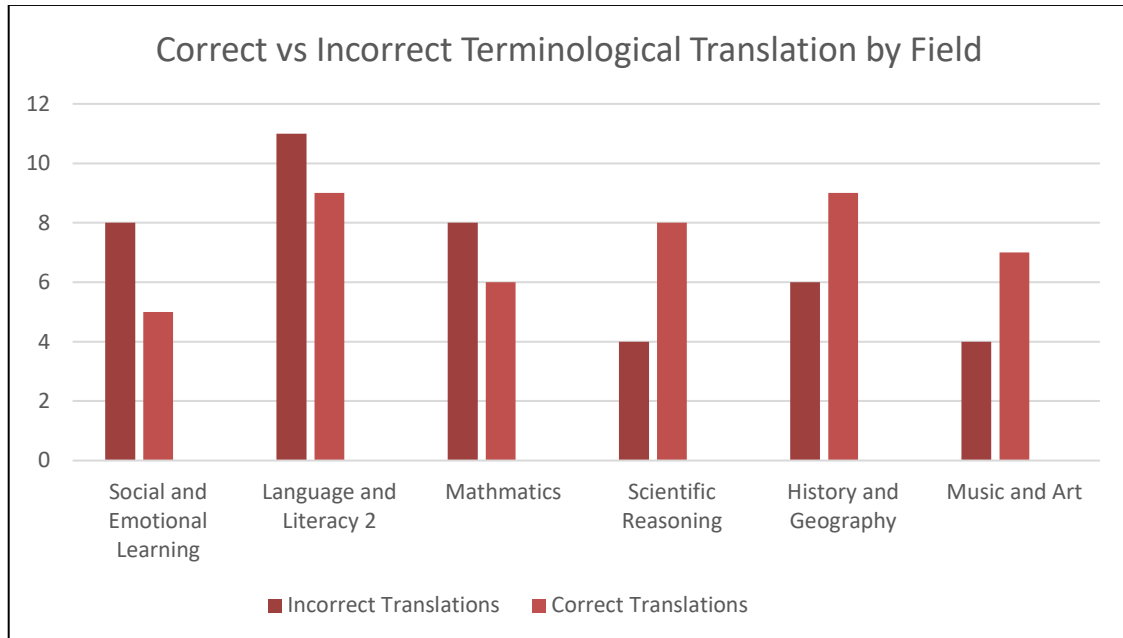


These error types were distributed across the disciplines included in the text under study, and an analysis was conducted to measure the accuracy rate of MT in each discipline. The number of terms in the field of social and emotional learning was thirteen, eight of which were translated incorrectly, meaning that the error rate in translating terms in this discipline was 61.5%. In the field of language and literacy, the number of terms was twenty, eleven of which were translated inaccurately, representing a 55% error rate. In the field of mathematics, the number of terms was fourteen, eight of which were mistranslated, representing a 57.1% error rate. Scientific reasoning included twelve terms, four of which were mistranslated, representing a 66.7% error rate. For history and geography, the examination involved fifteen terms of which six were mistranslated, representing a 40% error rate. Finally, in the music and art discipline, the number of terms that were found in the text was eleven, four of which were mistranslated, representing a 36.4% error rate. These results show a disparity in percentages by discipline, where arts and sciences are less likely to suffer an error rate than more difficult disciplines with more difficult vocabulary, such as language and mathematics. This suggests another problem facing MT across disciplines. It is worth mentioning that despite the diversity of texts that are being studied and the variety and frequency of errors, the translation QA system

suffered from a clear deficiency in detecting and identifying these errors. Not only was the system unable to classify the errors, but it was also unable to detect any of them at all.

Figure 5

Distribution of MT errors found within the studied text according to field type



3.6.2.4 Compatibility of Smartcat Translation QA System with ISO 17100 and ISO 18587 Standards in Terms of Terminological Errors

The results of the analysis of terminological errors in the text revealed clear shortcomings in Smartcat translation QA system when compared to ISO 17100 and ISO 18587. While a high-quality translation under ISO 17100 requires translators to be specialized in terminology selection and distinguish between different contexts and text types, the system was unable to detect any of the 41 terminological errors identified during the data analysis. ISO 18587 also stipulates that a MT review process requires verifying the absence of terminological problems and addressing and detecting them if they exist.

3.7 Formatting and Tags Error

3.7.1 Errors within Çetiner's Surface-level

According to Çetiner's classification, formatting errors include five types. In the following table, the researcher indicates which errors were detected by Smartcat translation QA system and which were not.

Table 5

Types of formatting errors detected/ undetected by SC QA according to Çetiner's classification

Main Category	Çetiner's Subcategory	Detected by Smartcat QA?
Formatting and Tags	Empty tag content	Not applicable
Formatting and Tags	Inconsistent tag content	Not applicable
Formatting and Tags	Nested tags	Not applicable
Formatting and Tags	No text between the tags	Not applicable
Formatting and Tags	tags/formatting are missing or invalid	Not applicable

3.7.2 Errors beyond Çetiner's Surface-level

The researcher found a small number of examples of formatting and spelling errors, but they were widespread throughout the text. She included only one example for each case.

3.7.2.1 Formatting Errors Due to Linguistic Differences

Punctuation Placement

The first formatting error noted by the researcher relates to coordination and the natural differences between the two languages. For example, the researcher noticed that Smartcat translation QA system considers any sentence ending in a full stop as an incorrect sentence simply because the full stop in English has moved from the right of the sentence to the left in Arabic. An orange triangle appeared, indicating that the reason was a punctuation difference (despite the presence of the same punctuation mark). It is worth mentioning that the QA system was able to successfully identify the extra spaces between punctuation marks and words everywhere they occurred.

3.7.2.2 Lack of Number Localization

English Numbers not Arabized

The researcher noted that the MT often renders English numbers as they are, without Arabizing them or adjusting their formatting to suit the Arabic style, which weakens the formal consistency of the translated text. However, Smartcat translation QA system was unable to identify such an error. The researcher found that the Smartcat QA system was remarkably successful in detecting mistranslations of numbers and dates from the source text to the translated text.

3.7.2.3 Inclusion of English Terms in Arabic Translation

Flagging Necessary English Terms as Errors

The researcher noticed that the system flagged the addition of English words or letters to the translation as a translation error, even though in certain cases, such additions are necessary, especially in translating grammar rules involving elements from the source language. For example, in translating a rule with a (verb + ing) expression, one has to use this English expression as it is, because the (ing) form represents a specific grammatical rule that should not be translated word by word into Arabic without losing its educational significance. forbidding including these terms by the system reduces the accuracy of the translation and unfairly evaluates it.

Table 6

Examples of formatting and tag errors

ST Example	Erroneous Translation	Corrected Version	Subcategory	Detection Method	Source of Error
(full stop) moved in Arabic	Considered a mistake because there is a period at the beginning of the sentence.	The translation is correct	Trailing punctuation	QA tool alert (false positive)	Right-to-left punctuation handling
1.1	1.1	1.1	Script mixing	QA tool alert (false negative)	Local settings
Verb+ing	ing + الفعل	The translation is correct	English term inclusion	QA tool alert (false negative)	Local settings

3.7.2.4 Compatibility of Smartcat QA System with ISO 17100 and ISO 18587 in Terms of Formatting Errors

After analyzing the results of formatting errors in Smartcat, and how Smartcat translation QA system handled them, it became clear that there is partial compatibility between the system and ISO 17100 and ISO 18587 standards. ISO 17100 emphasizes that the target language text should be similar to the source language text in terms of formatting, form, numbers, and the order of paragraphs, tables, and lists. ISO 18587 emphasizes that any errors related to spaces, punctuation, and order should be removed, and that the resulting text should resemble the untranslated original human text. Conversely, Smartcat

translation QA system successfully detected errors such as extra spaces and errors in numbers, dates, and punctuation, but it failed to consistently and adequately address natural differences between languages. It also failed to accept the presence of a foreign term in an Arabic text, which is sometimes necessary.

3.8 workflow Errors

3.8.1 Errors within Çetiner’s Surface-level

From Çetiner’s classification, the researcher included only one workflow error. In the following table, the researcher indicates whether this error was detected by Smartcat translation QA system.

Table 7

Types of workflow errors detected/ undetected by SC QA according to Çetiner's classification

Main Category	Çetiner’s Subcategory	Detected by Smartcat QA?
workflow	Unconfirmed segments	Yes

In the fourth type of errors in Çetiner's classification, the researcher considered one case: (unconfirmed segment). It is worth noting that this case was considered not applicable because the Smartcat system does not allow translation evaluation of a single segment without it being confirmed.

3.8.2 Compatibility of Smartcat Translation QA System with ISO 17100 and ISO 18587 Regarding Workflow Errors

In the analyzed example, Smartcat translation QA system fully complies with ISO 17100 and ISO 18587 standards. These standards require that any translated segment be complete before evaluation can be performed. By the same token, Smartcat translation QA system refuses to evaluate any segment until it has been confirmed.

The results of the analysis of the collected data revealed clear shortcomings in Smartcat translation QA system. The collected errors included linguistic errors, terminological errors, formatting, and workflow errors according to Çetiner's classification. The system failed to detect most of these errors, while displaying an error notification for correct sentences.

Çetiner’s classification shows three types of terminological errors, one of which, missing terms, is considered a single error. The researcher developed this error through in-depth analysis to include five subtypes, which are missing terms (contextual age reference), missing terms (context appropriate), missing terms (contextual intent), missing terms (educational instruction), and finally missing terms (audience appropriate). The following table shows the difference between these five:

Table 8

Examples of missing term types

Type of Error	Explanation	Resulting in
missing terms (contextual age reference)	The MT removed a term related to the target age group.	Misleading the target group
missing terms (context appropriate)	The MT ignored a term that fit the specialized context of the text.	loss of terminological precision
missing terms (contextual intent)	The MT did not convey the author's intention within the context, although the translation was linguistically correct.	loss of tone and implied meaning
missing terms (educational instruction)	The MT ignored a term that indicates an instruction or educational request.	Weak educational tone
missing terms (audience appropriate)	The MT used a term that is not appropriate for the target audience.	Confusing the target audience

As previously mentioned, the researcher relied on the Çetiner 2022 classification as the primary framework for error analysis and classification. To make the framework more suitable for the nature of the study, which relies on artificial intelligence and modern technologies, the researcher expanded the scope of analysis to see how the QA system interacts with the TM.

This analytical expansion is not a real modification of the framework used, but rather represents an attempt by the researcher to adapt existing theoretical models to contemporary translation environments, where modern technology can be widely integrated into translation studies.

3.9 Evaluating TM: Quality Support

The results of the qualitative data analysis showed that the TM effectively serves Smartcat translation QA system and has nothing to do with its shortcomings. The researcher used the same children's performance checklist file that had been translated before, but this time as a TM. A similar file was chosen from the same Kuwaiti Reggio Nursery, but for KG1. The researcher translated the second file with Smartcat, using the first file as the TM. The researcher focused on the segments that demonstrated the possibility of using the TM and noticed that the TM retrieved the previously translated sentences in the first file as is, even if there was a percentage of mismatch between the old and new sentences. It is worth noting here that the translation QA system displays the percentage of match out of 100 between the old and new sentences and warns if the match is not complete. When the translator presses the "Confirm" button without making any changes, the system issues a warning (Confirm Fuzzy Match without Edit), and an orange triangle appears indicating the cause of the error (Unedited Fuzzy Match). However, if the translator continues with the edit and then clicks "Confirm," the warning does not reappear, and the orange triangle with the error cause disappears.

The results of translating entire sentences using the TM show that it did not cause any errors in Smartcat translation QA system, and rather, it has proven effective in supporting consistency and detecting subtle inconsistencies. If there are any errors, it is due to the translator's reliance on the TM entirely, ignoring mismatch notifications and warnings, and not modifying the sentence. In all cases, as long as there is a mismatch, the system continues to issue an error warning. This careful analysis highlights the importance of integrating human effort and technical tools within the system.

Chapter Four

Conclusion and Recommendation

4.1 Conclusion

This study aimed to examine the translation QA system in Smartcat to determine whether it was able to correctly assess translations, detect errors, and correctly classify them if they were detected. It also aimed to determine the reasons for misclassifying some errors. In addition, the study focused on whether TM played a role in the system's success in assessing errors and whether it was possible to rely on Smartcat translation QA system without resorting to human effort. The study also aimed to determine the extent to which Smartcat translation QA system complies with ISO 17100 and ISO 18587 international translation standards.

The results showed that Smartcat translation QA system suffers from significant shortcomings in evaluation, as it fails to detect most linguistic, terminological, and formatting errors. The results also showed that the system may flag translation errors even when the translation is correct and valid in all respects. Furthermore, the results did not show any contribution of the TM to errors in Smartcat translation QA system. Rather, the findings demonstrated that this tool functioned as a supportive feature by issuing warnings related to the match rate between the previously translated sentence and the new one. In addition, the results showed that Samratcat's translation QA system did not fully comply with the international translation standards ISO 17100 and ISO 18587. It showed partial compliance in terms of maintaining form and detecting minor errors such as punctuation, extra spaces, order, and numbers, while still requiring human intervention to review the translation output, especially since the system failed to maintain stylistic consistency, functional meaning, and cultural appropriateness. It is also important to note that the nature of Smartcat work, which is based on isolating sentences and phrases into separate segments, creates context isolation. That is, we cannot blame Smartcat translation QA system for its inability to recognize context.

4.1.1 Comparison with Previous Studies

These findings align with those of Çetiner in his study of the translationQA system on the Mimsorce platform, regarding the limitations of QA systems in computer-assisted translation tools. These systems are often unable to detect and classify errors accurately

and do not adhere to international translation standards. Nevertheless, the research into Smartcat differs from that of Çetiner because it gives a profound analysis of the causes of these classification errors and investigates the role of translation memory, acting as a catalyst, which enables the QA system to function correctly. Therefore, this study offers a more comprehensive analysis and deeper insight into the quality assessment mechanisms in Smartcat.

The results of this study also supports the view held by (Zita, Ildikó, & János, 2024), highlighting the necessity of human intervention in translation and proofreading processes, particularly in large or specialized projects. In their view, despite advancements in artificial intelligence technologies, these technologies cannot grasp the linguistic and cultural complexities that determine the quality of a translation. At the end of the research, it can be said that the researcher reached the standpoint that machines cannot be fully relied upon in the translation, proofreading, and evaluation process, no matter how advanced they are.

4.1.2 Scientific Contribution

The scientific contribution of this study stems from the following:

1. The study represents the first comprehensive assessment of the translation QA system on the Smartcat platform, hence contributes a valuable addition to the field of computer assisted translation, placing the study as one of the leading contributors in this area.
2. Broadening the scope of the discussion through linguistic and terminological analysis that goes beyond Çetiner's surface-level observation, supported with examples that are important to understand the full range of translation issues in educational texts.
3. Shedding light on how the translation memory functions in platforms that rely on computer-assisted translation tools, demonstrating that translation memory helps in enriching the QA process.
4. Highlighting how local settings and language patterns affect the accuracy of the translation QA system.

5. Expanding the academic discussions on the limitations of machine-based translation QA systems by demonstrating that human intervention is essential and that complete reliance on such systems is insufficient.

4.2 Recommendations

The recommendations stem from the limitations of the study found by the researcher and are recommended to be taken into consideration for the future development of the translation QA system in Smartcat:

1. It is recommended that future research be conducted using broader educational texts for different age groups to find generalizable results regarding the quality of Smartcat translation QA system in this field. It is also recommended that research be conducted to evaluate the Smartcat QA system using diverse scientific texts across all disciplines to determine whether the system is more efficient and effective in one field than another.
2. It is recommended that future studies be conducted to evaluate the QA system for Arabic-to-English translation and compare the system's efficiency and effectiveness with the results demonstrated in this study, which examined the quality of the system when translating from English to Arabic. It is also recommended that studies be conducted to examine the Smartcat translation QA system in all languages.
3. It is recommended to use the various paid and free translation options available through the Smartcat platform, not just Google Translate, to test the effectiveness of a QA tool. This is to ensure a fair assessment of the translation QA system and to help assess its capabilities, given the varying capabilities of each tool.
4. It is recommended to conduct similar studies based on human evaluations by translation experts and compare them with the evaluation of Smartcat translation QA system. This way, the results will be more objective, and free of bias.
5. It is recommended that studies be conducted to examine Smartcat translation QA system from a software and technical perspective to assist in system development. The current study limited itself to examining the system from a translational and linguistic perspective, that is, from the perspective of the user and translator, rather than the programmer and developer.

List of Abbreviations

Abbreviation	Meaning
CAT	Computer-Assisted Translation
MT	Machine Translation
QA	Quality Assurance
TM	Translation Memory

References

- Alan, M. K., & Wright, S. E. (2014). Translation Memory. In C. Sin-wai, *Routledge Encyclopedia of Translation Technology* (p. 662). London, UK: Routledge.
- Alotaibi, H. (2020). Computer-Assisted Translation Tools: An Evaluation of Their Usability among Arab Translators. *Applied Sciences* , 11.
- Biau Gil, J. R., Leal, P. C., Mor, A. M., Mesa-Lao, B., Orozco, M., & Sanchez-Guon, P. (2010). TRACE: Measuring the Impact of CAT Tools on Translated Texts. In M.-L. Gea-Valor, I. García-Izquierdo, & M.-J. Esteve, *Linguistic and Translation Studies in Scientific Communication* (p. 255). Frankfurt am Main, Germany: Peter Lang.
- Çetiner, C. (2021). Evaluating Quality Assurance (QA) Checks in CAT tools in Relation to Professional Quality Standards. *Tam Metin Bildiriler* (p. 403). Ekaterinburg, Russia: ECLSS (Euroasian Conference on Language and Social Sciences).
- Craciunescu, O., Gerding-Salas, C., & Stringer-O'Keeffe, S. (2004). Machine Translation and Computer-Assisted Translation: a New Way of Translating? *Translation Journal* , 4.
- Han, B. (2020). Translation, from pen-and-paper to computer-assisted tools (CAT Tools) and machine translation (MT). *The 14th International Conference on Interdisciplinarity in Engineering (INTER-ENG 2020)The 14th International Conference on Interdisciplinarity in Engineering (INTER-ENG 2020)* (p. 3). Târgu Mureş, Romania: MDPI.
- House, J. (1997). *Translation Quality Assessment: A Model Revisited*. Tübingen, Germany: Gunter Narr Verlag.
- (ISO), I. O. (2017). *ISO 17100:2015 Translation services — Requirements for translation services*. Geneva, Switzerland: ISO.
- Jiménez-Crespo, M. A. (2010). The effect of Translation Memory tools in translated Web texts: Evidence from a comparative product-based study. *Linguistica Antverpiensia* , 195.
- Juran, J. M., Godfrey, A. B., Hoogstoel, R. E., & Schilling, E. G. (1999). *Juran's quality handbook (fifth edition)*. New York: McGraw-Hill.
- Karabıyık, S. (2019). Applicability of the ISO 17100:2015 quality standard in designing a training program for future medical translators: The case of Turkey. *transLogos Translation Studies Journal* , 7-16.
- Knight, J., & de Wit, H. (1999). *Quality and Internationalisation in Higher Education*. Paris: OECD.
- Makoushina, J. (2007). Translation Quality Assurance Tools: Current State and. *Translating and the Computer* 29 (p. 2). London, UK: Aslib.

- Mohammed, O., Samad, S. S., & Mahdi, H. (2018). A Review of Literature of Computer-Assisted Translation. *Language in India* , 1-2.
- Nord, C. (1997). *Translating as a Purposeful Activity Functionalist Approaches Explained*. London: Routledge.
- Nord, C. (2014). *Translating as a Purposeful Activity: Functionalist Approaches Explained*. London, UK: Routledge.
- O'Brien, S., Ehrensberger-Dow, M., Hasler, M., & Connolly, M. (2017). Irritating CAT Tool Features that Matter to Translators. *Hermes: Journal of Language and Communication in Business* , 145.
- Petrova, V. (2019). Translation Quality Assessment Tools and Processes in Relation to CAT Tools. *Proceedings of the Human-Informed Translation and Interpreting Technology Workshop (HiT-IT 2019)* (p. 91). Shoumen, Bulgaria: Incoma Ltd.
- Poibeau, T. (2017). *Machine translation*. Cambridge, MA: MIT Press.
- Popiołek, M. (2020). ISO 20771:2020 overview and legal translator competence requirements in the context of the European Qualifications Framework, ISO 17100:2015 and relevant research. *Lingua Legis* , 8-9.
- Sallis, E. (2002). *Total Quality Management in Education*. London and New York: Routledge.
- Šanca, F. (2018). *The Use of CAT Tools in University Translation Courses: A Case Study Based on Teaching with Memsources*. Prague, Czech Republic: Faculty of Arts, Charles University (Univerzita Karlova).
- Sandwich, S. (2025, n.d. n.d.). *Google Translate Statistics & Trends*. Retrieved January 13, 2025, from SEO Sandwich: <https://seosandwich.com/google-translate-statistics/#:~:text=Google%20Translate%20serves%20over%20500%20million%20users%20every,access%20Google%20Translate%20via%20mobile%20devices.%20%28Source%3A%20Statista%29>
- Schäffner, C. (1998). Skopos theory. In M. Baker, *Routledge encyclopedia of translation studies* (p. 117). London, UK : Routledge.
- Sdobnikov, V. V. (2019). Translation studies today: Old problems and new challenges. *Russian Journal of Linguistics* , 298-299.
- Smartcat. (n.d., n.d. n.d.). *AI-Powered CAT Tool*. Retrieved February 2, 2025, from Smartcat: <https://www.smartcat.com/cat-tool/>
- Taravella, A., & Villeneuve, A. O. (2013). Acknowledging the needs of computer-assisted translation tools users: the human perspective in human-machine translation. *The Journal of Specialised Translation* , 63.
- Torres-Hostench, O., Gil, J. R., & Leal, P. C. (2010). TRACE: Measuring the Impact of CAT Tools on Translated Texts. In M. L. Gea-Valor, I. García-Izquierdo, & M. J.

Esteve, *Linguistic and Translation Studies in Scientific Communication* (p. 255). Bern, Switzerland: Peter Lang AG, International Academic Publishers.

Woodhouse, D. (1999). Quality and quality assurance. In J. Knight, & H. de Wit, *Quality and internationalisation in higher education* (p. 29). Paris: OECD.

Zerfass, A. (2002). Evaluating Translation Memory Systems. *LREC 2002: Language Resources in Translation Work and Research* (p. 49). Canary Islands, Spain: ELRA.

Zheng, W. (2018). Introduction of Functionalism and Functional Translation Theory . *Advances in Social Science, Education and Humanities Research* (p. 623). Paris, France: Atlantis Press.

Zita, H., Ildikó, T., & János, F. (2024). MemoQ, a CAT tool used in technical translation courses at Debrecen University. *Humán Innovációs Szemle* , 39.



جامعة النجاح الوطنية
كلية الدراسات العليا

دراسة نظام ضمان جودة الترجمة
في برنامج الترجمة سمات كات

إعداد
آية محمود ضميدي

إشراف
أ. د. عبد الكريم دراغمة

قدمت هذه الرسالة استكمالاً لمتطلبات الحصول على درجة الماجستير في اللغويات التطبيقية و الترجمة
بكلية الدراسات العليا في جامعة النجاح الوطنية في نابلس، فلسطين.

2025

دراسة نظام ضمان جودة الترجمة في برنامج سمارت كات

إعداد

آية محمود ضميدي

إشراف

أ. د. عبد الكريم دراغمة

الملخص

تهدف الدراسة إلى فحص نظام ضمان جودة الترجمة المتوفر في منصة سمارتكات، وهي منصة ترجمة بمساعدة الحاسوب يستخدمها المترجمون المحترفون حيث تساعدهم في تسهيل عملية الترجمة وتسريعها. لتحليل الأخطاء، استخدمت الباحثة نموذج تشيتر (2021)، الذي يصنف أخطاء الترجمة إلى أربع فئات رئيسية: الأخطاء اللغوية، والأخطاء المصطلحية، وأخطاء التنسيق وأخطاء سير العمل. تم إجراء مقارنة للنتائج بمعايير الترجمة المهنية مثل ISO 17100 و ISO 18587. كما وسّعت الباحثة نطاق التحليل بإضافة أخطاء تتجاوز المستوى السطحي لتصنيف تشيتر. استكشفت الدراسة أيضًا كيفية تفاعل أداة ذاكرة الترجمة مع نظام ضمان الجودة. كما اعتمدت الدراسة تصميمًا وصفيًا وتقييميًا ومنهجًا نوعيًا لتحديد الأخطاء وتصنيفها وتفسيرها. تم قياس شدة الأخطاء وفقًا لتأثيرها على فهم الجملة ومعناها.

أظهرت النتائج أن أعلى نسبة من الأخطاء في منصة سمارتكات هي الأخطاء المصطلحية، تليها الأخطاء اللغوية، ثم الأخطاء ذات العلاقة بالتنسيق والتي لم يستطع نظام ضمان جودة الترجمة الكشف عن معظمها. لم تكشف النتائج عن أي أخطاء في النظام مرتبطة بأداة ذاكرة الترجمة. بالإضافة إلى ذلك، أظهرت النتائج أن نظام ضمان جودة الترجمة في سمارتكات لم يكن متوافقًا تمامًا مع معايير الترجمة ISO. توصي الدراسة بدمج الخبرة البشرية في عملية تقييم الترجمة باستخدام أدوات الترجمة بمساعدة الحاسوب لضمان نتائج موثوقة.

الكلمات المفتاحية: سمارتكات، نظام ضمان الجودة، أدوات الترجمة بمساعدة الحاسوب، الترجمة الآلية، ذاكرة الترجمة، معايير ISO.