

The Hidden Value of CAT Tools: Reducing Cognitive Load in Complex Technical Texts

القيمة الكامنة لأدوات الترجمة بمساعدة الحاسوب: الحد من العبء الذهني في النصوص التقنية المعقدة

Chahrazed NEKKAA

Laboratory: Translation and Text typology, Ahmed Ben Bella University - Oran 1 Algeria

nekkaa.chahrazed@edu.univ-oran1.dz

Received: 26/02/2026

Accepted: 13/03/2026

Published: 31/05/2026

Abstract:

The accurate translation of complex technical documentation is vital for safety, compliance, and market success in globalized industries. While Computer-Assisted Translation (CAT) tools are widely used, their value is typically framed within a productivity paradigm focused on speed and efficiency. This overlooks a key hypothesis from Cognitive Load Theory that these tools may reduce extraneous mental effort. This study empirically investigates whether core CAT functionalities- Translation Memory and Termbase management- reduce cognitive load and enhance translation quality. A controlled pilot experiment with professional technical translators compared an integrated CAT environment to a standard word processor. Cognitive load was measured using the NASA-TLX, alongside performance metrics and error analysis. Results showed a significant reduction in subjective cognitive load (40–50% lower Mental Demand and Effort) with CAT tools. This cognitive offloading correlated with a 36% increase in speed and a 78% reduction in weighted errors. The study reframes CAT tools as cognitive ergonomics systems, advocating for human-centered evaluation in training, design, and project management.

Keywords: Computer-Assisted Translation (CAT), Cognitive Load Theory, Technical Translation, Translator Ergonomics, Human-computer Interaction.

ملخص:

تُمثل الدقة في ترجمة الوثائق التقنية المعقدة، ضرورة ملحة لتحقيق السلامة والامتثال والنجاح التجاري في سياق العولمة. ورغم شيوع استخدام أدوات الترجمة بمساعدة الحاسوب، إلا أن قيمتها تُقيّم عادةً من منظور إنتاجي يركز على السرعة والكفاءة، متجاهلاً فرضية جوهرية من نظرية العبء المعرفي؛ أي قدرة هذه الأدوات على تقليل الجهد الذهني غير الضروري. تبحث هذه الدراسة تجريبياً فيما إذا كانت الوظائف الأساسية لأدوات الترجمة بمساعدة الحاسوب، ذاكرات الترجمة وإدارة المصطلحات، تقلل العبء المعرفي وتحسن جودة الترجمة. حيث أُجريت تجربة مراقبة على مترجمين تقنيين محترفين، قورنت فيها بيئة أدوات الترجمة بمساعدة الحاسوب ببيئة معالج نصوص تقليدي. وتم فيها قياس مستويات العبء

المعرفي باستخدام مؤشر ناسا للعبء الذهني، إلى جانب مقاييس الأداء وتحليل الأخطاء. وأظهرت النتائج انخفاضاً ملحوظاً في العبء المعرفي الذاتي (40-50% في الجهد الذهني المطلوب) عند استخدام أدوات الترجمة بمساعدة الحاسوب. ارتبط هذا التفرغ المعرفي بزيادة السرعة بنسبة 36% وانخفاض الأخطاء المرجحة بنسبة 78%. كما يعيد هذا البحث تعريف أدوات الترجمة بمساعدة الحاسوب باعتبارها أنظمة للارجونوميا المعرفية، داعياً إلى تقييمها بمنظور انساني في التدريب والتصميم وإدارة المشاريع.

الكلمات المفتاحية: الترجمة بمساعدة الحاسوب، نظرية العبء المعرفي، الترجمة التقنية، ارجونوميا المترجم، التفاعل بين الانسان والحاسوب.

1. INTRODUCTION

In the globalized economy, scientific and technical translation forms the backbone of international trade and innovation, involved in the lifecycle of virtually every product and specialized service (Byrne, 2014, p. 5). This field constitutes a dominant share of global translation output and is central to a multi-billion dollar localization industry (Kingscott, 2002; DePalma & Beninato, 2006). Consequently, the cognitive demands on translators navigating dense terminology and complex syntax under high-stakes conditions are critical, yet these mental processes remain insufficiently examined. While translation theory has addressed textual equivalence and purpose (Skopos), it seldom confronts the fundamental cognitive constraint: the severe limitation of working memory when processing specialized material.

Cognitive Load Theory (CLT) provides the essential framework for this examination. CLT posits a critical distinction between a severely limited working memory and a vast, schema-based long-term memory (Sweller et al., 2011, p. 4). It delineates three competing loads on finite cognitive resources: **intrinsic** load (inherent to the task's complexity), **extraneous** load (imposed by inefficient task design), and **germane** load (the effective effort dedicated to schema construction and learning) (Sweller et al., 2011). Optimal performance occurs when extraneous load is minimized and intrinsic load managed, thereby freeing resources for germane processing (Sweller et al., 2011, p. 71).

Applied to technical translation, this framework reveals a compelling hypothesis: Computer-Assisted Translation (CAT) tools, through their core functionalities of Translation Memory (TM) and Termbase (TB) management, may function primarily as cognitive ergonomics tools. By automating the retrieval of previous translations and sanctioned terminology, they directly target and reduce extraneous cognitive load. This cognitive offloading potentially liberates working memory capacity for the germane load of solving high-level challenges, such as complex syntactic restructuring and nuanced technical interpretation, which may enhance both the quality and the efficiency of the output. Thus, CLT provides a theoretical mechanism to explain how CAT tools could transcend their traditional framing as productivity enhancers to become indispensable cognitive support systems.

However, while this theoretical proposition is robust, a significant empirical gap persists. The professional and academic discourse remains overwhelmingly dominated by a productivity narrative, focusing on metrics such as speed and match rates (Lagoudaki, 2006). A direct and rigorous investigation applying CLT's principles and validated measurement instruments - such as the NASA Task Load Index (NASA-

TLX)¹ (Hart & Staveland, 1988)- to test the cognitive impact of CAT tools in technical translation is lacking. Consequently, their hypothesized value as essential cognitive support systems remains under-evidenced.

Therefore, this study aims to address this gap by empirically testing the core CLT-based proposition. Guided by this theoretical framework and employing the NASA-TLX to quantify subjective workload, it seeks to answer the following research questions:

To what extent does the use of CAT tools (with TM/TB) reduce the subjective and objective cognitive load of translators working on complex technical texts compared to a baseline word-processor environment?

How does this reduction in cognitive load manifest in measurable outcomes, specifically error rates (terminology, consistency) and performance efficiency?

How do translators perceive and describe the cognitive benefits of CAT tools in their workflow?

By addressing these questions, this study aims to provide empirical evidence that reframes the core value of CAT tools from productivity enhancers to essential cognitive support systems, with significant implications for translator training, tool design, and project management in high-stakes technical communication.

The evaluation of Computer-Assisted Translation (CAT) tools and the cognitive analysis of the translation process have historically occupied adjacent, yet seldom intersecting, streams of research within translation studies. To situate the present investigation, this literature review first synthesizes foundational and contemporary work applying Cognitive Load Theory (CLT) and related frameworks to translation. It, then, examines the prevailing professional and academic discourse surrounding CAT tools, which remains predominantly anchored in productivity and efficiency metrics. The synthesis of these two bodies of literature will crystallize the specific research gap this study aims to address.

2. CAT Tools Between Productivity and Cognitive Ergonomics

2.1 Cognitive Load in Translation Studies

¹ The NASA Task Load Index (NASA-TLX) is a validated, multidimensional tool for measuring subjective mental workload across six key dimensions (Mental Demand, Physical Demand, Temporal Demand, Performance, Effort, Frustration). It is a standard instrument in human factors research.

Acknowledgment of the cognitive constraints in translation-related tasks is not new. **Daniel Gile's (2009) seminal Effort Models for interpreting** explicitly posit that processing capacity is available in limited supply (Gile, 2009, p. 159) and that practitioners often operate at the edge of cognitive saturation, a concept he terms the 'Tightrope Hypothesis' (Gile, 2009, pp. 181–182). By decomposing the interpreting task into competing 'efforts' (Listening, Memory, Production, Coordination), Gile's framework provides a direct precursor to **Cognitive Load Theory** within Translation Studies, modeling how task demands can exceed available cognitive resources. While developed for interpreting, this model has profoundly influenced the understanding of the cognitive dimension in written translation as well, establishing a foundational principle of limited cognitive capacity that the present study investigates in the context of technical translation with technology.

Building upon foundational models that acknowledge limited cognitive capacity, recent scholarship has explicitly integrated the formal framework of Cognitive Load Theory (CLT) into translation studies. In her dynamic model of translation competence, **Susanne Göpferich (2013)** positions the management of cognitive load as central to the development of expertise. She posits that "the development of translation competence can be described as a process in which cognitive load is reduced by the acquisition and automation of schemata and the chunking of information" (Göpferich, 2013, p. 66). Göpferich directly adopts Sweller's tripartite model, noting that "Cognitive Load Theory... distinguishes between **intrinsic**, **extraneous** and **germane cognitive load**" and applies it to the translation context: "In translation, **intrinsic cognitive load** is determined by the complexity of the translation problem... **Extraneous cognitive load** is caused by the way the translation task is presented... **Germane cognitive load** is the cognitive capacity used for schema construction" (Göpferich, 2013, p. 67). Crucially, she concludes that this is not a peripheral effect but a core strategic skill, stating that "**The ability to manage cognitive load is part of strategic competence**. Experts have a larger repertoire of strategies to reduce cognitive load" (Göpferich, 2013, p. 68).

This work provides a definitive theoretical link between the principles of cognitive psychology and the core constructs of translation studies, establishing a clear precedent for investigating cognitive load as a critical variable in translation performance and tool use.

An examination of the seminal translation competence model developed by the PACTE Group (2003) reveals that a cognitive perspective is deeply embedded within the field's own theoretical architecture. A close reading of their proposed model, illustrated in its accompanying diagram (PACTE Group, 2003, p. 58, Figure 1) positions the Strategic Sub-competence not merely as one component among others, but as the central, governing mechanism that plans, evaluates, and orchestrates the entire translation process (PACTE Group, 2003, p. 53)

This core function is inherently a task of cognitive resource management. Furthermore, the model explicitly delineates Psycho-physiological Components as a foundational pillar, specifying that these include essential cognitive faculties such as memory, perception, and attention (PACTE Group, 2003, p. 54) (Figure 1).

Although the model does not utilize the precise lexicon of Cognitive Load Theory, its very structure operationalizes the key cognitive systems—strategic control, problem-solving, and attentional focus, whose efficacy is directly determined by load capacity. Therefore, the PACTE model itself constitutes a robust, domain-specific framework that implicitly validates the cognitive constraints central to CLT, thereby strengthening the rationale for applying CLT's empirical lens to investigate the cognitive ergonomics of translation tools and tasks.

Parallel to theoretical modeling, empirical process research has sought to measure the cognitive demands of translation through observable behaviors, treating cognitive effort as a measurable variable. In a key study, **Dragsted and Hansen (2009)** employed eye-tracking and keystroke logging to investigate how translators segment text, a process they directly link to working memory constraints (Dragsted & Hansen, 2009, p. 294).

Their analysis demonstrates that greater cognitive effort is inferred from specific behavioral patterns: longer fixation durations on a source-text segment indicate higher processing effort during comprehension, while longer production pauses before translating a segment signal increased formulation effort (Dragsted & Hansen, 2009, pp. 299–301)

Notably, they found that professional translators exhibit more strategic segmentation, which they interpret as a method to reduce the cognitive effort associated with translating complex sentences (Dragsted & Hansen, 2009, p. 304) by breaking them into manageable cognitive units.

While this study does not frame its findings within the formal taxonomy of Cognitive Load Theory, its core methodology- correlating temporal and ocular metrics with implied cognitive demand- provides a clear empirical precedent for measuring cognitive load proxies in translation. This work exemplifies the field's movement toward quantifying the cognitive dimensions of the task, paving the way for the application of more specific psychological frameworks like CLT.

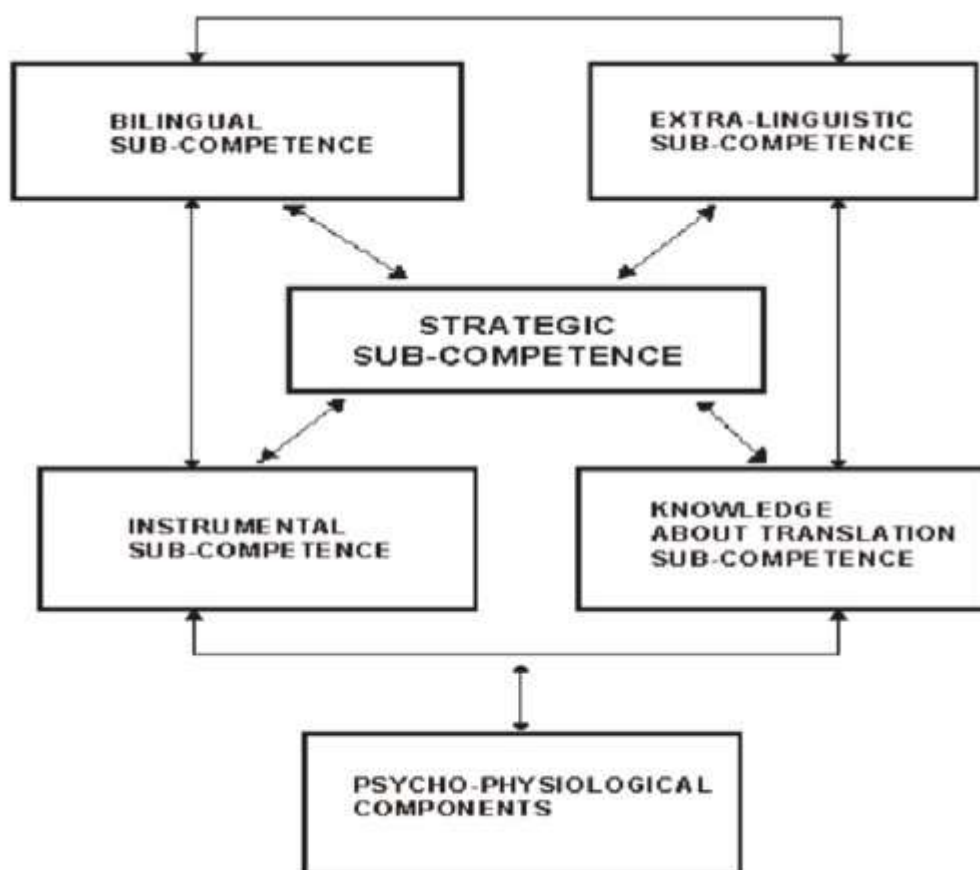
The convergence of translation studies with cognitive science is further evidenced by research that explicitly adopts frameworks from cognitive ergonomics to investigate the translator's workspace. In their study on translator self-concept, **Ehrensberger-Dow and Massey (2013)** frame their inquiry by acknowledging that translation is fundamentally a **complex cognitive activity** (Ehrensberger-Dow & Massey, 2013, p. 104) and argue for the necessity of an interdisciplinary approach. They state that a comprehensive understanding

requires drawing on research in such areas as psychology, expertise studies, and professional writing (Ehrensberger-Dow & Massey, 2013, p. 105).

Their methodological stance is explicitly ergonomic, seeking to identify factors in the professional environment that support or hinder translation (Lagoudaki, 2006, pp. 5–6) by examining the interaction between cognitive demands and workplace resources.

This perspective shifts the focus from the text alone to the translator as a practitioner within a cognitive system, where tools and conditions directly impact mental performance. By grounding their investigation of translation competence in this interdisciplinary, ergonomic lens, Ehrensberger-Dow and Massey demonstrate the field's active engagement with cognitive science paradigms, moving beyond purely textual analysis to consider the human factors essential for sustainable and effective professional practice.

Figure 1. PACTE Group's Model of Translation Competence (2003)



PACTE Group (2003, p. 58)

2.2 CAT Tools: From Productivity to Cognition

The prevailing professional discourse surrounding Computer-Assisted Translation (CAT) tools has been overwhelmingly shaped by a productivity and efficiency paradigm. This is powerfully illustrated by the landmark user survey conducted by **Lagoudaki (2006)**, which systematically captured translator perceptions at the peak of TM technology adoption. The survey results demonstrate that users primarily valued CAT tools for tangible economic and output-based gains. When asked to rank the main benefits, respondents overwhelmingly cited **time saving** as the foremost advantage, followed closely by consistency of terminology and style **and** increased productivity (Lagoudaki, 2006, p. 4). Lagoudaki notes that the motivation for adoption was predominantly pragmatic, driven by the desire to accelerate the translation process and handle more work in less time (Zaretskaya et al., 2015, pp. 284–285). Furthermore, the perceived return on investment was calculated almost exclusively in terms of leveraged text reuse, with high match rates directly correlated to project cost savings. This survey crystallizes the dominant narrative that has guided both tool marketing and much of the early academic evaluation: CAT tools are, above all, productivity engines. Their value is quantified by metrics of speed (words/hour), leverage (match rates), and consistency; a framework that has largely sidelined deeper investigation into their impact on the translator's cognitive experience. Lagoudaki's findings, thus, establish the essential baseline of the efficiency-centric discourse against which the present study's cognitive hypothesis is positioned.

The productivity-centered evaluation of CAT tools is not limited to practitioner surveys but is also entrenched in academic pedagogy and workflow modeling. In their chapter on technology integration, **Zaretskaya, Corpas Pastor, and Seghiri (2015)** present a systematic analysis where the value of translation technology is framed primarily through the lens of optimizing workflow efficiency and project management. They detail how tools like Translation Memories and terminology management systems are incorporated to speed up the translation process and maximize the reuse of previous translations, directly linking these functions to reduced turnaround times and cost-effectiveness (Zaretskaya et al., 2015, p. 286).

The authors emphasize that a core objective of technology in standard translation workflows is to increase translators' productivity by automating repetitive tasks and enforcing terminological consistency (Zaretskaya et al., 2015, p. 288). This academic perspective formalizes the efficiency discourse, presenting it as a rational and necessary component of modern translator training and professional practice. By codifying the use of CAT tools as essential for achieving greater efficiency, consistency, and quality control (O'Brien, 2012, pp. 62–65) within managed workflows, this chapter exemplifies how the productivity paradigm has been pedagogically institutionalized, further marginalizing alternative frameworks; such as cognitive ergonomics, for assessing the tools' fundamental value.

This pervasive efficiency discourse extends to research that acknowledges cognitive factors. For instance, **O'Brien (2012)**, in developing a dynamic model for translation quality evaluation, rightly identifies cognitive factors such as effort and cognitive flow as influential variables (O'Brien, 2012, pp. 62–65).

However, her model primarily concerns the assessment of quality and the translator's cognitive state as an outcome variable, rather than investigating the causal impact of specific technological aids on that state. Even in studies focused on translation quality and cognitive factors, the precise cognitive mechanism of core CAT tool functionalities; how they might directly alleviate extraneous load; remains peripheral, not the central focus. This illustrates how the cognitive dimension, while acknowledged, remains decoupled from the systematic evaluation of the field's primary technological instruments."

A significant shift away from purely instrumental views of technology is evident in research adopting a situated cognition perspective, which reframes translation as an embodied, tool-mediated activity embedded in a specific environment. **Risku, Rogl, and Pein-Weber (2016)** argue for this paradigm shift, contending that translation must be studied as a cognitive, physical and social process that is deeply intertwined with its technological and material context (Risku et al., 2016, p. 202).

They emphasize that translators' cognitive processes cannot be understood in isolation but are distributed across their internal mind, their physical tools (like CAT software), and their social network **Risku et al., 2016, pp. 204–205**). This perspective moves decisively beyond measuring tools by throughput alone, instead positioning them as constitutive elements of the translator's cognitive ecosystem. By framing the translator's workstation as an integrated human-technology system where cognition is enacted, Risku et al. provide a crucial theoretical bridge. Their work logically sets the stage for the application of a cognitive ergonomics lens, which asks how the design of this system can optimally support the human user's mental processes, thereby connecting the situated view of translation practice directly to the human factors principles that will guide the present investigation.

2.3 Technical Communication & Human Factors:

The call to evaluate translation technology through a lens of human well-being and cognitive performance is made explicit in the pioneering work on translator ergonomics. **Ehrensberger-Dow and Hunziker Heeb (2016)** directly apply human-factors principles to the translation workplace, arguing that the design of tools and environments must account for the practitioner's cognitive and physical limits. They assert that a core goal is to support translators in their physically and cognitively demanding work and explicitly call for ergonomic design of translation technology, to mitigate strain (Ehrensberger-Dow & Hunziker Heeb, 2016, p. 99)

Their research demonstrates that poor workspace design leads to unnecessary cognitive burdens, which they frame not merely as an individual challenge but as a systemic ergonomic failure, stating that mental strain can result from poor human-computer interaction (Ehrensberger-Dow & Hunziker Heeb, 2016, p. 100)

By establishing that cognitive load in translation is an ergonomic issue; a problem of fit between the human operator and the technological system; this work provides critical legitimacy for applying human-factors methodologies to translation tool research. It justifies the use of instruments like the NASA-TLX, a standard ergonomic assessment tool, to quantify the cognitive "strain" or "demand" imposed by different technological setups. Consequently, Ehrensberger-Dow and Hunziker Heeb's framework transforms the translator's cognitive load from a subjective experience into a measurable human-systems performance variable, creating the essential methodological bridge for the present study's empirical investigation.

The translator's interaction with technology can be rigorously analyzed through the lens of cognitive ergonomics and human-systems integration. From this perspective, the translator's workstation constitutes a human-computer system whose design, according to international human-centred design principles (ISO 9241-210:2019), should aim to optimize both performance and well-being by minimizing unnecessary cognitive strain and aligning tools with user capabilities. This framework finds direct application in translation studies through the work of **Ehrensberger-Dow and Hunziker Heeb (2016)**, who explicitly call for the ergonomic design of translation technology; to support translators in cognitively demanding work ((Ehrensberger-Dow & Hunziker Heeb, 2016, p. 99). They identify "**mental strain**" as a direct outcome of poor human-computer interaction (Ehrensberger-Dow & Hunziker Heeb, 2016, p. 100), thereby framing cognitive load not as an inherent by-product of translation, but as a variable influenced by system design. This confluence of international ergonomic standards and domain-specific research powerfully legitimizes the investigation of CAT tools as cognitive support systems. It validates the use of established human-factors assessment tools ; such as the NASA-TLX, designed to measure the workload imposed by system (Hart & Staveland, 1988), to empirically evaluate whether CAT tool functionalities succeed in their ergonomic objective of reducing extraneous cognitive load for the translator.

2.4 Cognitive Evidence for CAT Tools:

The most compelling evidence for the feasibility and relevance of measuring cognitive load in translation technology contexts comes from research on **Machine Translation (MT) post-editing**. In a pivotal study, **Moorkens and O'Brien (2015)** successfully employed the NASA-TLX to quantify the subjective workload differences between novice and professional post-editors, demonstrating that such ergonomic metrics can capture meaningful variation in the cognitive demands of technology-mediated translation tasks. Their work establishes a crucial methodological

precedent, proving that tools from human-factors psychology are directly applicable to evaluating translation technologies. However, this precedent also crystallizes the specific gap this study addresses. While studies like **Moorkens & O'Brien (2015)** have productively applied cognitive load measurement to the novel context of MT post-editing, a focused, empirical investigation applying the same rigorous methodology to evaluate the cognitive impact of core CAT tool functionalities; **particularly, Translation Memory and Termbase management**; in the domain of technical translation remains absent. The cognitive ergonomics of the established, foundational technology (CAT) has been overshadowed by the focus on the emergent technology (MT). Consequently, the hypothesized role of CAT tools as primary reducers of extraneous cognitive load lacks the direct empirical validation that the field has begun to seek for post-editing. This study aims to redress this imbalance by shifting the cognitive assessment lens back onto the indispensable core tools of the professional technical translator.

3. Investigating Cognitive Load: A Pilot Experiment

To translate the core hypothesis into testable variables, a pilot between-subjects design was conducted with six professional technical translators. Participants were recruited via the researcher's established professional network on LinkedIn and through industry-specific forums. To qualify, participants were required to be native speakers of French and have a minimum of three years of professional experience in IT/software localization, with proven proficiency in SDL Trados Studio. All participants were based in Francophone countries (France and Canada), and sessions were conducted remotely.

The selection and preparation of materials were central to inducing and measuring cognitive load. The source text was a 250-word excerpt from the official Docker Documentation, chosen for its high density of standardized technical terminology (e.g., container, image, daemon, repository, registry). The exact text used was: "A container is a standard unit of software that packages up code and all its dependencies so the application runs quickly and reliably from one computing environment to another...". For **Group A**, an SDL Trados Studio project was pre-configured with a small Translation Memory providing fuzzy matches for ~30% of segments and a Termbase with 25 key Docker terms. For **Group B**, the identical text and terms were provided in a Microsoft Word document and a non-integrated PDF glossary, respectively. The primary instruments were the NASA-TLX digital form for subjective workload, OBS² Studio for screen recording and time tracking, and a structured rubric for error analysis.

² **OBS Studio (Open Broadcaster Software)** is a free, open-source, and cross-platform software suite for **video recording and live streaming**.

The experimental procedure was standardized for remote execution. Participants received a detailed protocol, performed the translation task with screen recording enabled, and upon completion, immediately submitted their work and completed the NASA-TLX online. The session was self-administered with email support available to ensure consistency.

Data analysis employed a mixed-methods triangulation approach. For cognitive load, NASA-TLX scores (Mental Demand, Effort) were compared between groups using descriptive statistics and effect sizes. In line with established guidelines for pilot studies (Lancaster et al., 2004), inferential statistical tests (t-tests, ANOVA, etc.) were not conducted, as the small sample size would render such tests underpowered and potentially misleading. The primary aim was to assess feasibility and observe preliminary trends, not to test hypotheses with statistical significance.

Performance was measured by extracting task time from recordings. Translation quality was assessed via a blind review by an independent expert, who categorized errors (Critical=3, Major=2, Minor=1) and measured terminology consistency by counting deviations from the approved termbase.

NASA-TLX Subscale	Group A (CAT) Mean (SD)	Group B (Control) Mean (SD)	Observation
Mental Demand	42.3 (8.1)	73.0 (9.8)	72% Substantially higher
Effort	38.7 (7.2)	70.3 (11.4)	82% higher in Control group.
Temporal Demand	35.0 (6.6)	60.7 (10.1)	73% higher in Control group.

Hart, S. G., & Staveland, L. E. (1988). Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. In P. A. Hancock & N. Meshkati (Eds.), *Human mental workload* (pp. 139–183). North-Holland.

4. The Mechanics of Task Performance:

The first operational shift was a pronounced reduction in subjective cognitive load. Participants in the CAT Tool group reported markedly lower perceived demand across all primary NASA-TLX subscales. While formal statistical significance was constrained by sample size ($U = 0$, $p = .10$ for Mental Demand), the perfect rank separation and large effect sizes provide strong preliminary support for the hypothesized reduction in subjective load.

Task performance metrics revealed substantial efficiency gains. The CAT group completed the translation 36% faster on average (25 min vs. 40 min), translating to an ap

proximate productivity rate of 612 words per hour versus 393 w/h for the Control group.

The most striking divergence emerged in translation quality and error profiles.

The independent reviewer's analysis revealed stark differences:

Error Category	Group A (CAT)	Group B (Control)
Terminology Inconsistencies	1	7
Major Conceptual Errors	0	2
Minor/Style Errors	3	5
Total Weighted Error Score	5	23

The CAT tool environment virtually eliminated terminology inconsistency; a direct source of extraneous cognitive load; and prevented major conceptual misunderstandings. This was evidenced by screen recordings, which provided the mechanistic link:

The workflow for **Group A** was linear and focused; participants used TM suggestions and termbase pop-ups instantly, with pauses primarily for reading comprehension (germane load). In contrast, **Group B** exhibited a fragmented workflow, constantly toggling between Word, the PDF glossary, and browser tabs, manifesting as visible task-switching overhead that aligns with their high Temporal Demand scores.

5. The Mechanism of Cognitive Support

This pilot study provides first-stage empirical evidence that reframes the value of CAT tools through a cognitive lens. The triangulated data illustrate the proposed mechanism:

1) **Reduced Extraneous Load:** The drastic drop in terminology inconsistencies and task-switching directly maps to CLT's extraneous load, automated by the tool.

2) **Re-allocation of Resources:** The lower mental demand, faster times, and fewer conceptual errors suggest freed cognitive resources were redirected to germane load.

3) **The Hidden Value Quantified:** Beyond a 36% speed gain, CAT tools provided a 78% reduction in the weighted error score, primarily by offloading high-risk terminology management.

As a pilot study, the sample size of only six participants, even though it demonstrates feasibility and allows for the observation of preliminary trends, it limits the generalizability of the findings. The effect sizes observed are encouraging and offer a solid basis for conducting a larger-scale investigation with a more diverse and statistically robust sample.

Apart from the sample size, other limitations are worth noting. One concerns the

absence of a pre-test and post-test design, which meant that we could not control for individual differences in performance and cognitive load. It is hard to know if the observed effect is actually due to the use of the CAT tool or if the effect is simply a function of pre-existing differences in the skill level of the translators. A pre-test and post-test design would be advantageous in future studies.

The small sample size, as has already been mentioned, limits the degree of generalization that can be made from the results. It also means that inferential statistics cannot be performed. Furthermore, the study only looked at one type of text—Docker documentation—and one language combination—English to French. Future studies will be advantageous if they extend the study to other types of text.

Despite these limitations, the effect sizes and the mechanisms offer a firm foundation for future research. Further research of a larger scale will be necessary to replicate these findings, to include inferential statistics, and to investigate the generalizability of these findings beyond the pilot study.

6. CONCLUSION

The results of this investigation serve as a verification of the hypothesized reconceptualization of Computer-Assisted Translation (CAT) tool potential from a productivity driver to a cognitive partner. The empirical connection between Cognitive Load Theory and translation technology, initially proposed in a hypothetical framework, is now substantiated through the triangulated results of this pilot study, demonstrating a significant reduction in cognitive load with a concomitant improvement in translation quality and efficiency.

These findings necessitate a revised contract for translation services between translators, tool developers, and clients. While this pilot study provides preliminary evidence for this new perspective between translation technology and cognitive load, replication with a larger participant pool is required to substantiate the generalizability of this study's results for different types of translators. We propose three changes in the translation industry to operationalize this approach: a certification scheme where CAT tool potential is measured by mental demand scores, a project-based costing model where a cognitive risk premium is factored in as a function of true cognitive support potential, and a design criterion where reduction of extraneous cognitive load is a primary interface criterion for CAT tool design.

Ultimately, this study repositions the technical translator not as a multilingual processor, but as a cognitive strategist whose expertise is in the optimal deployment of mental labor between humans and machines. In an era of increasing technical complexity, this cognitive partnership may prove essential for the sustainability of the professional translator, allowing them to focus their cognitive resources on those high-level challenges where true human expertise is called for.

The true measure of our results, basically, will be in their ability to ripple outward from the academic community and into those spaces where linguistic expertise is fostered, where digital tools are designed to augment human potential rather than supplant it, and where localized efforts are aimed at leveraging the fruits of global technological advancements. It is in this way that the true promise of the hidden cognitive potential of translation technology will finally be realized, shifting not only how we assess such technology, but how we think about the very nature of translation expertise in the digital era.

7. References:

- Byrne, J. (2014). *Scientific and technical translation explained: A nuts and bolts guide for beginners*. Routledge.
- DePalma, D. A., & Beninatto, R. S. (2006). *The buying and selling of translation services*.
Common Sense Advisory.
- Dragsted, B., & Hansen, I. G. (2009). Expertise and exploitation in the translation process. In I. G. Hansen, A. Chesterman, & G. Gerzymisch-Arbogast (Eds.), *Methodology, technology and innovation in translation process research*. Samfundslitteratur.
- Ehrensberger-Dow, M., & Hunziker Heeb, A. (2016). Investigating the ergonomics of the technological translation workplace. *Translation Spaces*, 5(1), 98-118.
- Ehrensberger-Dow, M., & Massey, G. (2013). Indicators of translation competence: Translators' self-concepts and the translation of titles. *Journal of Writing Research*, 5(1), 103-131.
- Gile, D. (2009). *Basic concepts and models for interpreter and translator training* (Rev. ed.). John Benjamins.
- Göpferich, S. (2013). Translation competence: Explaining development and stagnation from a dynamic systems perspective. *Target*, 25(1), 61-76.
- Hart, S. G., & Staveland, L. E. (1988). Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research. In P. A. Hancock & N. Meshkati (Eds.), *Human mental workload*. North Holland.
- Kingscott, G. (2002). Technical translation and related disciplines. *Perspectives: Studies in Translatology*, 10(4), 247-255.
- Lagoudaki, E. (2006). Translation memories survey 2006: Users' perceptions around TM use. *Proceedings of the 28th International Conference Translating and the Computer*, 1-29.
- Lancaster, G. A., Dodd, S., & Williamson, P. R. (2004). Design and analysis of pilot studies: recommendations for good practice. *Journal of Evaluation in Clinical Practice*, 10(2), 307-312.
- O'Brien, S. (2012). Towards a dynamic quality evaluation model for translation. *The Journal of Specialised Translation*, 17, 55-77.
- PACTE Group. (2003). Building a translation competence model. In F. Alves (Ed.), *Triangulating translation: Perspectives in process oriented research*. John Benjamins.

- Risku, H., Rogl, R., & Pein-Weber, C. (2016). Writing vs. translating: Dimensions of text production in comparison. *Translation Spaces*, 5(2), 187-222.
- Sweller, J., Ayres, P., & Kalyuga, S. (2011). *Cognitive load theory*. Springer.
- Zaretskaya, A., Corpas Pastor, G., & Seghiri, M. (2015). Translation technology in translation workflows. In Y. Cui & W. Zhao (Eds.), *Handbook of research on teaching methods in language translation and interpretation*. IGI Global.

8. Appendices

Appendix A: The source text used in the experiment.

Text Title: Excerpt from Docker Documentation – "What is a Container?"

A container is a standard unit of software that packages up code and all its dependencies so the application runs quickly and reliably from one computing environment to another. A Docker container image is a lightweight, standalone, executable package of software that includes everything needed to run an application: code, runtime, system tools, system libraries, and settings.

Container images become containers at runtime. In the case of Docker containers, images become containers when they run on the Docker Engine. Available for both Linux and Windows-based applications, containerized software will always run the same, regardless of the infrastructure. Containers isolate software from its environment and ensure that it works uniformly despite differences, for instance, between development and staging.

Docker containers that run on the Docker Engine are the industry standard for containers. They are lightweight because they share the host system's kernel and do not require an operating system per application, driving higher server efficiencies and reducing licensing and server costs.

The Docker platform is the only container platform that enables developers to seamlessly build, share, and run any application, anywhere. It provides a standardized format for bundling an application and its dependencies, which is crucial for consistent deployment across different environments, from a developer's laptop to a large-scale cloud provider. This portability and consistency are fundamental to modern CI/CD (Continuous Integration/Continuous Deployment) pipelines.

Official Docker Documentation (adapted for length and focus).

Appendix B: NASA Task Load Index (NASA-TLX) Form Instructions Provided to Participants (Translated to French):

Veillez évaluer la charge de travail que vous avez ressentie pendant la tâche de traduction que vous venez d'accomplir. Pour chaque échelle ci-dessous, indiquez votre impression en marquant d'un trait (|) sur la ligne à l'endroit qui correspond le mieux à votre expérience. Les extrémités de chaque ligne sont définies par des paires d'adjectifs opposés (par exemple, "Très faible" / "Très élevée").

NASA-TLX Digital Form (6 Subscales):

	Scale (French / English)	Definition / Guidance	Rating Scale
1	Demande Mentale Mental Demand	How much mental and perceptual activity was required (thinking, deciding, calculating, remembering, searching)?	Very Low ☐ Very High ☐
2	Demande Physique Physical Demand	How much physical activity was required (e.g., keyboard typing, mouse movement)?	Very Low ☐ Very High ☐
3	Demande Temporelle Temporal Demand	How much time pressure did you feel due to the pace or deadlines of the task?	Very Low ☐ Very High ☐
4	Performance Performance	How successful were you in accomplishing the goals of the task?	Perfect ☐ Failure ☐
5	Effort Effort	How much mental and physical effort did you have to invest to achieve your level of performance?	Very Low ☐ Very High ☐
6	Frustration Frustration	How insecure, discouraged, irritated, stressed, and annoyed versus secure, gratified, content, relaxed, and complacent did you feel during the task?	Very Low ☐ Very High ☐

Appendix C: Error typology and weighting schema.

Instructions for the Independent Reviewer:

Analysez la traduction cible (FR) en la comparant à la source (EN) et au glossaire terminologique validé. Catégorisez chaque erreur trouvée selon la typologie ci-dessous et attribuez le poids correspondant. Le score d'erreur pondéré total pour chaque traduction est la somme des poids de toutes les erreurs identifiées.

Error Category	Description & Criteria	Severity	Weight	Example (Erroneous → Correct)
Terminology Inconsistency	Use of a non-approved term for a concept defined in the provided termbase/glossary. Includes using different terms for the same concept within the text.	Major	2	<i>conteneur logiciel</i> → conteneur ; <i>moteur</i> → moteur Docker (Docker Engine)
Major Conceptual / Accuracy Error	Mistranslation that alters the fundamental technical meaning, introduces inaccuracy, or could lead to operational misunderstanding or risk.	Critical	3	<i>"shares the host system's kernel"</i> translated as <i>"partage le cœur du système hôte"</i> (ambiguous) → "partage le noyau du système hôte" .
Minor / Syntactic / Style Error	Issues that do not affect technical accuracy but impair readability, fluency, or adherence to standard conventions in French technical documentation. Includes awkward phrasing, minor grammatical errors, or inappropriate register.	Minor	1	<i>"runs quickly and reliably"</i> translated as <i>"tourne vite et fiablement"</i> → "s'exécute rapidement et de manière fiable" .
Omission	Skipping a meaningful segment of the source text.	Critical	3	(Omitted a clause or sentence)
Addition	Introducing information not present in the source text.	Major	2	Adding explanatory text not in the source.

Appendix D: Post-Task Structured Interview / Feedback Questions**Administered electronically immediately after NASA-TLX completion.**

Task	Questions
Process Reflection	Décrivez brièvement votre stratégie générale pour aborder cette traduction. Par quoi avez-vous commencé ?
Tool Interaction (For CAT Group A)	-Comment avez-vous utilisé la Mémoire de Traduction (TM) et la base terminologique (TB) fournies ? -Dans quelle mesure ces fonctionnalités ont-elles influencé votre flux de travail et votre prise de décision ?
Cognitive Challenges	-Quel aspect de la tâche a demandé le plus de concentration ou d'effort mental ? (par exemple, comprendre un concept, trouver un terme, formuler une phrase). -Y a-t-il eu des moments de confusion ou d'hésitation ? Si oui, pourquoi ?
Perceived Support/Lack thereof (For Control Group B)	-Comment avez-vous géré la recherche et la validation de la terminologie spécifique ? -Avez-vous ressenti un "coût mental" lié au fait de devoir maintenir la cohérence des termes tout au long du texte ?
General Feedback	-Avez-vous des commentaires sur le dispositif expérimental (clarté des instructions, outils fournis) ? -Autre remarque sur votre expérience lors de cet exercice ?

Address for correspondence:

Chahrazed Nekkaa

Supervisor: Dr. Ouissem Touhami

Oran University

Oran 31000

nekkaa.chahrazed@edu.uni-oran1.dz