

Received 14/03/2026

Accepted: 13/05/2026

AI-Mediated Translation and Vocabulary Learning: A Study of Critical Digital Literacy in the Algerian Plurilingual EFL Context

Amana HAMDI PACHA*¹ Dr. Faiza AIT ABDESLAM²

¹École Normale Supérieure de Bouzaréah (ENSB), Laboratoire de linguistique et sociodidactique du plurilinguisme (LISODIP), Algeria, amana.hamdi-pacha@ensb.dz

²École Normale Supérieure de Bouzaréah (ENSB), Laboratoire de linguistique et sociodidactique du plurilinguisme (LISODIP), Algeria, aitaabdeslam.faiza@ensb.dz

Abstract

This study investigates AI-mediated translation practices and critical digital literacy (CDL) among first-year EFL students at the École Normale Supérieure de Bouzaréah (ENSB), Algeria. Operating within a plurilingual context marked by the coexistence of Arabic, French, Tamazight, and English, students increasingly rely on AI-mediated translation for vocabulary learning. Drawing on a quantitative design, data were collected through a survey questionnaire (n=78) and an AI-translated vocabulary evaluation task (n=34), measuring self-reported AI translation practices and actual ability to identify and correct lexical inaccuracies in AI-generated output. The findings reveal that although students demonstrate general awareness of AI translation limitations, this awareness does not translate into effective critical evaluation. A disparity was found between self-reported verification behaviors and actual task performance, suggesting that current AI use promotes passive rather than reflective vocabulary acquisition. The study calls for explicit CDL instruction focused on developing concrete vocabulary evaluation skills in the Algerian plurilingual EFL context.

Keywords: AI-mediated translation, vocabulary learning, critical digital literacy, plurilingualism, Algerian EFL students.

1. Introduction

Vocabulary development constitutes a central component of learning English as a Foreign Language (EFL). In plurilingual settings such as Algeria, vocabulary learning does not occur in isolation from learners' existing linguistic inventory. Instead, learners draw extensively on Arabic, French, and Tamazight when attempting to understand or produce English lexical items. Within this plurilingual context, translation naturally functions as a bridge between languages, particularly for beginner university students whose lexical resources in English remain limited.

The increasing accessibility of generative AI translation tools has reconfigured translation-based vocabulary learning practices. Research by Boukheikh (2025) revealed that the skill mostly developed using AI among Algerian university learners was vocabulary, followed by

¹ Corresponding author : amana.hamdi-pacha@ensb.dz

writing and grammar. Unlike traditional dictionaries or bilingual word lists, AI tools generate immediate equivalents, explanations, and contextualized examples. While this may support simple and rapid lexical access, it also shifts part of the learning process from deliberate analysis to automated output. As a result, the learner's role risks becoming one of passive acceptance rather than active engagement. Studies conducted by Sebbah (2025), and Berrarbi and Amrane (2025) indicated that although AI tools can support EFL learning among Algerian university students, their frequent use occurs with limited critical digital literacy skills.

This shift raises important concerns for efficient vocabulary learning. English vocabulary is characterized by polysemy, collocational patterns, phrasal verbs, idiomatic expressions, and culturally embedded meanings. Literal translation across languages may fail to capture these nuances. While AI-assisted translation tools offer efficiency gains, research in the Algerian context highlights that their output still requires human supervision to ensure semantic accuracy and contextual appropriateness (Kobouch, 2025). When AI-generated translations are used without verification, learners may internalize forms that are semantically imprecise, contextually inappropriate, or unsuitable in register. The issue is not merely technological but pedagogical; students often receive no explicit formal guidance on how to evaluate AI output or how to integrate it responsibly into their vocabulary learning. Recent studies in the Algerian context highlight a lack of pedagogical training and structured guidance on the use of AI tools in higher education. Students often engage with these technologies without sufficient instructional support or clear frameworks for evaluating AI-generated output (Ghounane et al., 2024; Sebbah, 2025).

Within this context, critical digital literacy emerges as a central competence. This refers to the learners' ability to critically evaluate and verify AI-generated vocabulary for accuracy, appropriateness, and relevance before any receptive or productive use. Additionally, it involves using AI tools in an informed and responsible way. This means recognizing the limitations of the technology and the constant need for verification. Yahiaoui (2023) found that Algerian university students often lack the digital skills and competencies required to engage effectively with AI-generated content. In the absence of well-developed critical digital literacy, AI-mediated translation risks leading to passive lexical acquisition rather than reflective vocabulary development.

Despite the growing presence of AI tools in educational contexts, limited empirical research has examined how first-year EFL students in plurilingual settings actually use AI-mediated translation for receptive and productive vocabulary learning, and how they evaluate its output.

Research Aim

The present study aims to investigate the use of AI-mediated translation in receptive and productive vocabulary learning among first-year students at the École Normale Supérieure de Bouzaréah (ENSB) and how they demonstrate critical digital literacy when engaging with AI-generated vocabulary. By focusing on beginner learners' awareness and practices in a plurilingual context, the study seeks to provide insights into how AI translation currently

functions as a vocabulary learning tool and inform pedagogical practices that promote responsible, reflective, and context-sensitive vocabulary learning in the age of generative AI.

Significance of the Study

This study is significant for several reasons. First, it contributes to EFL vocabulary didactics by examining AI-mediated translation as a learning practice rather than as a purely technological tool. Second, it addresses a research gap concerning beginner university students, whose reliance on translation is particularly strong and whose vocabulary knowledge is still developing. Third, by focusing on students enrolled in a teacher-training institution, the study highlights an important pedagogical concern: these learners are not only language students but also future EFL teachers, whose current practices and awareness may influence their future teaching approaches, and thus future generations of learners.

The findings are expected to inform pedagogical decision-making by emphasizing the need for explicit instruction in AI use, vocabulary verification strategies, and evaluative practices. More broadly, the study aims to support the development of critical digital literacy in plurilingual EFL classrooms.

Research Questions

To achieve the research aim, the study is guided by the following research questions:

1. To what extent do first-year ENSB students use AI-mediated translation for receptive and productive vocabulary learning in English?
2. To what extent do these students demonstrate critical digital literacy when engaging with AI-generated vocabulary?
3. Does students' self-reported awareness of AI translation limitations align with their demonstrated ability to detect lexical errors in AI-generated output?

Research Methodology

To address the research questions, this study adopts a quantitative research design employing two complementary data collection instruments: a survey questionnaire and a vocabulary evaluation task, administered to first-year English department students at the École Normale Supérieure de Bouzaréah (ENSB). These instruments were designed to collect data on students' AI translation practices for vocabulary learning, their critical digital literacy when engaging with AI-generated lexical input, and the relationship between their self-reported awareness of AI limitations and their demonstrated ability to identify lexical inaccuracies in AI-translated output. This design is appropriate for identifying patterns of tool use, levels of reliance on AI translation, awareness of its limitations, and variations in students' evaluative and verification practices.

2. Literature Review

2.1. The Nature of Vocabulary Knowledge

In linguistics, vocabulary is defined as more than a simple inventory of isolated words. It encompasses a complex system of lexical units that vary in size, from single morphemes to multi-word units like phrasal verbs (*carry out*) and idiomatic expressions (e.g. *break the ice*). Because these units often function as single semantic entities, mastering English vocabulary involves understanding how these different patterns actually work together, rather than simply memorizing a list of word definitions.

2.1.1. Dimensions of Vocabulary Knowledge

A common misconception is that knowing a word involves only pairing a form with its meaning. However, Nation's (2001) framework identifies that comprehensive vocabulary knowledge consists of three interconnected dimensions: 'form' (spoken, written, and word parts), 'meaning' (concepts, referents, and associations), and 'use'. The "use" dimension is particularly critical for communicative competence, as it involves mastering grammatical functions, collocations (natural word pairings), and constraints such as register and frequency.

Daller et al. (2007) propose a similar model consisting of vocabulary breadth (the number of words a learner knows), vocabulary depth (the richness of knowledge about those words, including nuances), and vocabulary fluency (the speed and automaticity with which a learner can access their lexical store). This framework suggests that proficiency is not merely a matter of a large mental lexicon, but the ability to retrieve the appropriate word accurately and fluidly in various contexts.

Together, these frameworks suggest that vocabulary knowledge is multidimensional and context-sensitive, a complexity that poses significant challenges for AI-mediated translation, which often tends to prioritize form and surface meaning over the nuanced dimensions of use and contextual appropriateness.

2.1.2. Receptive vs. Productive Mastery

Vocabulary knowledge is typically categorized into receptive (passive) and productive (active) mastery. Receptive knowledge involves the ability to recognize and understand a word when it is encountered in reading or listening. In contrast, productive knowledge refers to the ability to retrieve and use the word correctly in speaking or writing. Generally, a learner's receptive vocabulary is significantly larger than their productive one, as the latter requires a much higher degree of retrieval and precision (Nation, 2013).

This distinction is particularly relevant to AI-mediated translation: students tend to use AI predominantly for receptive purposes, and as the findings of this study suggest, their ability to productively correct AI-generated vocabulary errors proved notably more challenging than simply identifying them, a pattern consistent with Nation's (2013) observation that receptive and productive competencies develop at different rates and require different types of engagement.

2.1.3. Challenges in EFL Vocabulary Acquisition

Acquiring English vocabulary presents several challenges for EFL learners (Nation, 2013). One primary difficulty is polysemy, where the same word form carries multiple meanings depending on the context. Furthermore, learners frequently struggle with multi-word units like phrasal verbs and idioms, where the overall meaning cannot be easily inferred from the individual components. Collocations also pose a significant obstacle, as learners often apply the linguistic habits of their first language to English, resulting in unnatural pairings. Finally, mastering register, which involves knowing what words are appropriate for specific settings, is essential to avoid pragmatic failure (Schmitt, 2008). Folse (2004) corroborates these challenges and further argues that they are frequently overlooked in EFL vocabulary instruction.

These challenges are especially apparent in plurilingual contexts like Algeria, where learners navigate multiple linguistic systems. Because students are accustomed to the vocabulary structures of Arabic and French, these pre-existing systems often interfere with the acquisition of English vocabulary, leading to a variety of cross-linguistic inconsistencies.

2.2. Plurilingualism

Plurilingualism refers to an individual's integrated and dynamic linguistic repertoire developed across multiple languages (Council of Europe, 2001). Unlike multilingualism, which describes the societal coexistence of separate languages, plurilingualism focuses on how individuals draw strategically on interconnected linguistic resources in communication. Languages in the plurilingual mind interact simultaneously rather than functioning as isolated systems. This distinction is significant for language pedagogy, as it shifts the focus from managing separate language systems to leveraging an integrated repertoire as a learning resource.

This perspective is reflected in the concept of translanguaging, where speakers flexibly use their full linguistic repertoire to construct meaning (García & Wei, 2014). Similarly, Coste et al. (2009) describe plurilingual competence as an integrated system rather than a simple accumulation of separate languages. Drawing on multilingual acquisition research, Cenoz and Gorter (2015) and Cook (2016) argue that this integration differentiates L2 users from monolingual speakers and reframes cross-linguistic influence as a resource rather than interference. It is also associated with enhanced metalinguistic awareness that can positively support language learning (Jessner, 2006).

Collectively, these studies converge on a view of plurilingualism as a dynamic and integrated competence that supports rather than hinders language learning. For Algerian EFL learners, this means that their Arabic, French, and Tamazight repertoires are not obstacles to English vocabulary acquisition but potential resources.

2.2.1. Plurilingualism in the Algerian Context

Algeria has a distinctive sociolinguistic landscape defined by the coexistence and interaction of multiple language systems. It is organized around three main spheres. The Arabic sphere includes Standard Arabic, the first official language since 1963, and Algerian Arabic/Darija, used daily by 70–75% of the population. The Tamazight sphere comprises various indigenous

varieties spoken by 25–30% of the population and granted official status in 2016. The foreign language sphere is dominated by French, which continues to shape political and economic domains, alongside a growing presence of English among younger generations (Ouahmiche et al., 2017; Rouabah, 2022).

Plurilingualism in Algeria extends beyond demographic diversity and is structurally embedded in everyday educational and professional practices (Rouabah, 2022). While Standard Arabic holds official status, Algerian Arabic (Darija) dominates daily interaction, and French continues to occupy a quasi-co-official role, particularly in scientific and technical higher education. Language use and policy in Algeria remain closely tied to ideological debates surrounding identity, history, and globalization (Messekher & Miliani, 2020). Within this complex linguistic landscape, learners constantly navigate between their mother tongue, the former colonial language, and the emerging global lingua franca, English, a dynamic that has been described as a “de-facto plurilingual tendency” (Benrabah, 2014; Boukreris, 2015). This landscape is currently undergoing a significant shift. Recent national reforms, including the introduction of English in primary schools and its growing adoption as a medium of instruction in higher education, signal a gradual move away from the historical dominance of French in favor of English (Tihal, 2025).

Overall, the studies reviewed in this section converge on a shared understanding: Algeria's linguistic landscape is complex, dynamic, and ideologically charged, characterized by the unequal coexistence of Arabic, French, Tamazight, and an increasingly prominent English, with each language occupying a distinct but shifting role in education, daily life, and national identity. Together, these dynamics suggest that Algerian EFL learners do not approach English vocabulary learning as a clean slate, but through the lens of a complex, layered linguistic repertoire. While the sociopolitical and ideological dimensions of this landscape are acknowledged, the scope of this research focuses strictly on the linguistic and educational aspects of EFL vocabulary acquisition. It centers on how the complex plurilingual background of Algerian university students influences their approach to EFL vocabulary learning.

2.2.2. Implications for EFL Vocabulary Learning:

Plurilingualism plays a significant role in English vocabulary acquisition among Algerian learners. First, learners with academic proficiency in French benefit from lexical similarities between French and English, which facilitate cognate recognition and positive transfer (Ouahmiche et al., 2017). Research on third language acquisition shows that learners draw on previously acquired languages when learning additional ones, allowing cross-linguistic influence to support vocabulary development (Cenoz, 2001). In the Algerian context, both Arabic and French function as linguistic resources that help structure the initial stages of English learning (Rothman, 2011).

Second, plurilingual practices in Algeria encourage learners to engage in natural translanguaging strategies. The interdependence hypothesis suggests that conceptual knowledge developed in one language can facilitate comprehension in another (Cummins, 2007). Similarly, learners can rely on their existing linguistic knowledge to infer meaning, recognize patterns, and make informed lexical guesses (Hall & Cook, 2012). Contemporary

perspectives therefore view cross-linguistic engagement as a strategic resource rather than interference.

Overall, English vocabulary learning in Algeria occurs through interaction with learners' existing linguistic repertoires. Their plurilingual background makes cross-linguistic strategies, particularly translation, a natural and productive component of vocabulary development.

2.3. Translation as a Tool for Vocabulary Learning

The role of translation in language teaching has long been a subject of scholarly debate. Hall and Cook (2012) observe that what has developed in language teaching is "a divide between those who have regarded monolingual classrooms as a given and those who teach using the learners' own language," yet note that "this division, which, in many contexts, might arguably be characterised as one between theory and practice, may be coming to an end, and that the existence and advantages of using the learners' own language in class are increasingly recognised" (p. 278).

In plurilingual environments, translation is increasingly recognized as a purposeful cognitive strategy rather than a sign of linguistic weakness. By translating unfamiliar English terms into a more dominant language in their repertoire, such as Arabic or French, Algerian students establish more efficient mental associations. Folse (2004) states that "learners retain words better when the words are initially learned with a first language translation than when they are accompanied by a picture, an L2 explanation, or an L2 example sentence" (p. 161). This process of cross-linguistic mapping is crucial for mastering the relationship between the form and meaning of English vocabulary, which helps students transition more effectively from understanding words to using them actively (Laufer & Girsai, 2008).

Translation is a natural and inevitable component of second language acquisition. Cook (2010) argues that preventing learners from using their own language is often counterproductive because cognitive research demonstrates that learners automatically activate translations in their mind when processing new words (Kroll & Stewart, 1994). Empirically, translation has been shown to be highly effective; studies indicate that groups using translation often outperform those relying solely on reading or context when it comes to recalling and using new words (Laufer & Girsai, 2008).

The success of translation-based learning is rooted in the depth of cognitive processing it requires. Prince (1996) found that words learned through translation are often retained better than those learned through definitions in the target language because the process forces learners to create more elaborate memory traces between their known languages and the target language.

For receptive knowledge, Nation (2001) notes that translation is a highly efficient way to learn EFL vocabulary meaning because it builds directly on what the learner already knows. When it comes to productive knowledge, the act of translating a word from the learner's native language into English helps learners notice the gaps in their knowledge. This processing effort creates a more durable memory of the word's form than just seeing it used in a text (Laufer & Girsai, 2008). Furthermore, tools like bilingual word cards have been

documented as one of the most efficient ways to learn vocabulary, allowing for much higher learning rates than incidental exposure alone (Webb & Nation, 2017).

Taken together, this body of research consistently supports translation as a legitimate vocabulary learning strategy. However, these findings were established in the context of deliberate, human-guided translation. The growing shift toward AI-mediated translation raises the important question of whether those benefits still apply when the cognitive effort of translation is largely automated.

2.3.1. The Shift toward AI-Mediated Translation

The emergence of Large Language Models (LLMs) has transformed translation from a manual task into an automated, AI-mediated process. Unlike traditional tools, Generative AI provides context-aware translations that mimic human fluency, shifting the learner's role from a "constructor" of meaning to an "editor" of AI-generated output. While this shift offers instant support, research suggests it can lead to "cognitive offloading," which is the process of using external tools or technologies to support or extend cognitive processing by reducing internal mental effort in tasks such as memory, reasoning, or problem-solving (Hellmich & Vinall, 2023). In language learning contexts, excessive reliance on such tools may reduce learners' engagement in deeper processing required for long-term retention. Lee et al.'s (2026) systematic review of generative AI in language classrooms confirms this concern, identifying passive consumption of AI-generated content as a primary pedagogical risk that can undermine vocabulary depth.

In the Algerian context, this shift is characterized by complex social representations, where teachers and students view AI with a mix of pedagogical hope and professional apprehension regarding its integration into the classroom (Bouzidi & Zarif, 2025). Furthermore, the transition to AI-mediated tools introduces a new layer of complexity known as the "AI tongue-twister," which refers to the difficulties AI systems face in disentangling multilingual inputs and capturing their contextual and structural nuances (Benfareh, 2025). In African contexts, AI translation performs more reliably with well-represented languages than for underrepresented ones (Alabi et al., 2022; Qin et al., 2025). For Algerian learners, this means stronger vocabulary support when translating from French or Standard Arabic, but significantly reduced accuracy when working with dialectal Arabic or Tamazight varieties, potentially creating unequal learning outcomes based on learners' linguistic backgrounds.

While AI simplifies the translation process, it introduces a significant pedagogical risk. Passive overreliance on automated output often bypasses the mental effort necessary to master vocabulary, leading to a variety of challenges related to linguistic accuracy, cultural nuances, and pragmatic appropriateness.

2.3.2. Challenges of AI-Mediated Translation in EFL Vocabulary Learning

While the speed of Generative AI provides learners with instant access to linguistic equivalents, this convenience often encourages superficial learning. As Jeon et al. (2025) show, AI systems trained mainly on monolingual data often overlook the natural translanguage practices plurilingual learners use when processing vocabulary across languages. Another fundamental challenge underlying AI translation errors lies in the nature

of language itself. As Bowker and Buitrago-Ciro (2019) explain, "each approach to machine translation has strengths and weaknesses, but an overarching challenge is that language is full of ambiguities at both the word level and the level of larger linguistic units (e.g., phrases or sentences)" (p. 53). This inherent ambiguity means that AI translation tools, despite their growing sophistication, remain prone to producing inaccurate or contextually inappropriate vocabulary, particularly when handling complex lexical features such as idioms, collocations, and culturally embedded expressions.

Because AI outputs appear fluent and seem reliable, students frequently overlook the nuanced linguistic and cultural gaps that exist between English and their existing repertoire. This over-reliance leads to various challenges in both receptive and productive language use.

2.3.2.1. Linguistic and Structural Constraints

A significant hurdle in vocabulary acquisition is that a single word form often carries vastly different meanings depending on the context in which it appears. This is clearly demonstrated in cases of polysemy, where a term like *branch* might be understood by a learner only as a "tree limb," while its use as a "department" or "specialized division" is ignored. Similarly, homonymy presents a risk; a student might decode *fine* as a state of being "well" (*bikhayr*), remaining oblivious to its meanings as "thin" or a "financial penalty" (*gharama*). Without contextualized exposure, learners treat these words as static units, leading to significant receptive misunderstandings (Nation, 2013).

Beyond individual meanings, translation often fails to capture the natural pairings known as collocations. Students frequently apply the structural habits of Arabic or French to English, resulting in phrasing that sounds unnatural to native speakers. For instance, a learner might produce *pass an exam* when they actually mean to *take* or *sit for* an exam. Similarly, the use of *take a decision*—influenced by the French *prendre une décision*—clashes with the conventional English *make a decision*. This also applies to descriptive pairings; while a student knows *heavy* in the context of a *heavy bag*, they may not realize that *heavy rain* requires a different conceptual association (*amtar ghazira*) rather than a literal transfer of the word for weight (Schmitt, 2008).

Furthermore, lexico-grammatical units like phrasal verbs cannot be decoded through literal word-for-word substitution. As Folse (2004) explains, the meaning of the verb shifts entirely based on the particle attached to it. A learner may understand the verb *look* in isolation but struggle to distinguish between *look for* (to search) and *look into* (to investigate). Because these combinations function as single semantic units, relying on the translation of individual components often leads to a total loss of the intended communicative intent.

A further complication in the Algerian context is the presence of "false friends" (*faux amis*), where words appear identical across English and French but differ in meaning. For instance, a student may assume the English *lecture* refers to "reading" (*lecture* in French), whereas it actually denotes an oral presentation by an instructor. As Granger and Swallow (1988) note, these deceptive cognates create a false sense of familiarity that leads to persistent lexical errors.

2.3.2.2. Sociolinguistic and Cultural Barriers

Translation frequently fails to convey the sociolinguistic nuance required for appropriate communication, particularly regarding register. Translating *old* from the French *vieux/vieille* or Arabic *musinn* can create social friction; while the word is linguistically correct, calling someone "old" directly in English can sound blunt or rude. In formal or polite contexts, alternatives like *elderly* or *senior* are more socially appropriate. This demonstrates how a dictionary equivalent can be technically accurate in meaning but socially inappropriate for the intended audience (Nation, 2013).

Cultural gaps present a profound barrier because vocabulary is a reflection of a society's specific environment and history. Algerian concepts like *Baraka* (divine grace) or *Sadaqah* (voluntary charity) carry spiritual and collective weights that the English "blessing" or "charity" do not fully capture (Folse, 2004). Conversely, Western concepts such as "People of Color" are rooted in specific social histories that may not align with the Algerian context. These mismatches result in shallow output where the words are recognized, but the underlying cultural intent is mismatched.

Finally, idiomatic expressions represent a unique challenge because their meaning is figurative rather than literal. For example, the idiom *to be under the weather* is used to describe feeling ill or tired. However, a literal translation into Arabic or French would suggest a physical position beneath the climate or sky. Because these phrases have no direct semantic equivalents in the Algerian repertoire, literal translation results in nonsensical phrases that fail to achieve productive mastery (Laufer & Girsai, 2008; Nation, 2001).

In light of these various challenges, the current pedagogical landscape underscores the necessity of integrating Critical Digital Literacy to ensure that students remain active participants in the translation process rather than passive consumers of automated data when learning vocabulary.

2.4. Critical Digital Literacy

Critical digital literacy (CDL) stems from the broader concept of digital literacy, which Gilster (as cited in Pool, 1997, p. 9) defined as the ability to "assimilate the information, evaluate it, and then reintegrate it". This definition positions evaluation of digital content, rather than mere technical proficiency, as the core competency of the digitally literate user.

Critical digital literacy extends this foundation by adding a layer of judgment and interrogation: not simply whether the user can *find* or *use* digital information, but whether they can question it, assess its limitations, and detect its flaws. As Pangrazio (2016) argues:

While learning how to use and manipulate digital technology is important, without an understanding of the role humans play in questioning, challenging and therefore shaping this techno-social system, then the scope of digital literacy is limited. (p. 169)

Therefore, critical digital literacy can be defined as the ability to go beyond the functional use of digital tools and technologies, to actively question, evaluate, and critically verify the accuracy, reliability, and limitations of digitally produced content. It positions the user not as

a passive consumer of digital information, but as an informed and reflective evaluator capable of interrogating what digital systems and tools produce.

Central to CDL are two interconnected processes: evaluation and verification. Evaluation refers to the broader act of critically judging the quality, accuracy, and reliability of digitally produced content, while verification denotes the more specific and deliberate act of cross-checking that content against other sources or prior knowledge to confirm its correctness. In other words, CDL is the overarching critical disposition, evaluation is the process it entails, and verification is the concrete action it demands.

In the context of the present study, critical digital literacy is conceptualised as the ability to evaluate and verify the accuracy and reliability of AI-generated vocabulary translations, positioning EFL learners in the Algerian plurilingual context as active critical evaluators of machine-translated vocabulary rather than passive recipients of automated output.

2.4.1. Critical Digital Literacy and AI-Mediated Translation

In the context of AI-mediated translation, CDL serves as a vital safeguard against uncritical reliance on machine-generated output, a phenomenon known as cognitive offloading, where learners passively accept machine output, bypassing the mental effort required for long-term vocabulary retention. This is particularly evident in the Algerian EFL context, where Yahiaoui (2023) and Sebbah (2025) found that learners tend to over-rely on AI tools without critically engaging with or questioning the accuracy of the output they receive.

This is particularly significant given that AI translation tools, despite their growing sophistication, remain prone to errors of meaning, context, and lexical appropriateness (Kobouch, 2025). Learners who accept AI-translated vocabulary without critical evaluation risk internalising inaccurate or contextually inappropriate lexical items. CDL therefore equips EFL learners with the reflective capacity to evaluate, verify, and where necessary, reject or modify machine-produced vocabulary, rather than passively incorporating it into their linguistic repertoire.

Drawing on the machine translation literacy frameworks of Chapelle and Sauro (2017) and the AI competency models of Ng et al. (2021), this study identifies four core dimensions that constitute CDL in AI-mediated vocabulary acquisition:

2.4.1.1. Awareness of Limitations

The foundational dimension of CDL is the recognition that AI is an imperfect linguistic tool. Based on Chapelle and Sauro (2017), this involves a critical mindset where learners understand that AI often fails to navigate complex linguistic boundaries. Specifically, learners must be aware that AI may struggle with idiomatic expressions, phrasal verbs, and the nuances of register (choosing between formal and informal language).

2.4.1.2. Critical Evaluation

Once a translation is generated, the learner must act as a "linguistic auditor." This involves assessing the semantic precision and contextual fit of the suggested vocabulary (Hellmich & Vinall, 2023). Rather than accepting a word at face value, the learner evaluates whether the

AI-generated term accurately reflects the intended meaning within the specific context of their sentence.

2.4.1.3. Verification (Triangulation)

Verification represents the external action taken to ensure accuracy. Drawing on the concept of lateral reading (Wineburg & McGrew, 2019) and translation literacy (Bowker & Buitrago-Ciro, 2019), this pillar involves triangulating AI output with trustworthy, human-authored sources. This includes cross-referencing AI suggestions with academic dictionaries (e.g., Oxford, Cambridge) or linguistic corpora to confirm frequency and usage patterns.

2.4.1.4. Revision and Refinement

The final dimension is the exercise of linguistic agency (Niño, 2020). Here, the learner treats the AI output merely as a "first draft." This involves the active modification of the vocabulary to ensure it matches the learner's personal style and precise message. This transformative step ensures that the learners remain in active control of their communication, rather than passive recipients of machine-generated text.

3. Methodology

3.1. Research Design

This study adopts a quantitative research design employing two complementary data collection instruments. Quantitative data were collected through two instruments: a survey questionnaire and a vocabulary evaluation task. As Dörnyei (2007) observes, tests and questionnaire surveys constitute the most common data collection instruments in quantitative applied linguistics research.

In the first phase, a structured survey questionnaire was administered to 78 first-year ENSB students to capture self-reported AI translation practices for vocabulary learning and critical digital literacy awareness. In the second phase, a vocabulary evaluation task was administered to a subsample of 34 students to assess their actual ability to identify and correct lexical inaccuracies in AI-generated translations, allowing for a direct comparison between proclaimed awareness and demonstrated performance.

The difference in sample sizes across the two phases reflects the voluntary nature of participation and the greater time commitment and cognitive effort required for the performance-based task. Notably, both subsamples meet the minimum threshold recommended for normal distribution, as Hatch and Lazaraton (1991) argue that a sample of at least 30 participants is sufficient for this purpose. Creswell (2009) further acknowledges, when two data collection methods are employed within the same study, they are often unequal in their priority, which may result in unequal evidence and discrepancies that need to be carefully considered when interpreting findings. These limitations are acknowledged and addressed in the limitations section of this study.

3.2. Participants and Sampling

The target population of this study consists of 108 first-year students enrolled in the English department at the École Normale Supérieure de Bouzaréah (ENSB), all of whom are training to become future secondary school teachers. This population was purposefully chosen as it

involves beginner learners of a teacher-training institution who are not only language students but also future educators who may later integrate similar principles into their own classrooms.

Convenience sampling was employed for both data collection phases, as participants were selected based on their willingness to participate. From this population, 78 students (72.2%) completed the survey questionnaire and 34 students (31.5%) completed the vocabulary evaluation task, with both instruments distributed online across various student groups.

Analysis of survey responses further revealed that participants fell into three distinct instructional profiles regarding their exposure to guidance on AI translation tool use: those who had received teacher-led instruction ($n=19$, 24.4%), those who had engaged in self-initiated learning ($n=37$, 47.4%), and those who reported no explicit instruction of any kind ($n=22$, 28.2%). These groups subsequently formed the basis for a One-Way ANOVA.

3.3. Data Collection Instruments

The study employed two data collection instruments. The first was a structured, 14-item online questionnaire designed to measure two primary constructs: the prevalence of student engagement with AI-mediated translation for vocabulary learning and the application of critical digital literacy strategies when evaluating AI-generated output. The instrument comprised a mix of categorical questions and Likert-scale items, chosen to capture both the objective breadth of student usage patterns and the nuanced intensity of their critical evaluation practices.

The second instrument was a 12-item AI-translated vocabulary evaluation task administered online to a subsample of 34 students, designed to assess their actual ability to identify and correct lexical inaccuracies in AI-generated translations. The task comprised the six previously mentioned lexical categories (idiomatic expressions, phrasal verbs, collocations, etc.) with two sentences per category, one translated from Arabic and one from French into English. These two languages were selected as source languages given that they emerged as the most frequently used languages for AI-mediated vocabulary translation among participants in the survey phase. Each sentence contained a single AI-generated lexical error sourced from ChatGPT and Google Translate, reflecting common translation inaccuracies produced by these tools. Students were required to first identify the error and then provide a correction, allowing for a two-stage assessment of both recognition and productive lexical knowledge. A time limit of 12 minutes was allocated for task completion, corresponding to approximately one minute per item.

3.4. Reliability and Validity

The validity of both instruments was established through a review of relevant literature and expert consultation to ensure that all items accurately reflected the study's objectives and adequately captured the targeted constructs. Both the survey questionnaire and the vocabulary evaluation task were piloted prior to full administration, allowing for the refinement of items, the clarification of instructions, the determination of the appropriate number of items, and the establishment of the 12-minute time limit for the task. Reliability for the survey was assessed by calculating Cronbach's alpha (α) coefficients, which demonstrated an acceptable level of internal consistency for the Likert-scale items, with an overall coefficient of $\alpha = 0.702$.

3.5. Data Analysis

Data were processed using SPSS version 27 to analyze the participants' plurilingual linguistic backgrounds and vocabulary learning through AI translation patterns. The analysis employed a three-fold approach: descriptive statistics, including frequencies and percentages, were used to summarize categorical data. Additionally, Likert-scale items were averaged into composite scores to assess overall AI engagement and critical evaluation competencies. Finally, inferential statistics were employed, including independent and paired samples t-tests, Pearson correlation, One-Way ANOVA, and Cramér's V, to examine relationships between variables and determine whether statistically significant differences existed across groups. For the vocabulary evaluation task, descriptive statistics were used to summarize identification and correction scores across lexical categories and source languages, complemented by paired samples t-tests and a One-Way ANOVA to examine performance differences.

3.6. Ethical Considerations

Prior to data collection, participants were informed of the study's purpose. Participation in both the survey questionnaire and the vocabulary evaluation task was entirely voluntary, and no personal identifying information was collected, ensuring full anonymity. Both instruments were distributed online, and submission was taken as implicit indication of informed consent.

4. Analysis of the Findings

This section presents the findings of the study across two data collection instruments. The first subsection reports the results of the survey questionnaire (n=78), addressing participants' plurilingual linguistic backgrounds, AI translation usage patterns, and self-reported critical digital literacy awareness and verification behaviors.

The second subsection presents the findings of the vocabulary evaluation task (n=34), examining participants' demonstrated ability to identify and correct lexical inaccuracies in AI-generated translations. Together, the two subsections address the research questions concerning the extent of AI-mediated translation use in vocabulary learning, students' demonstration of critical digital literacy, and the relationship between self-reported awareness and actual performance.

4.1. Survey Questionnaire Findings

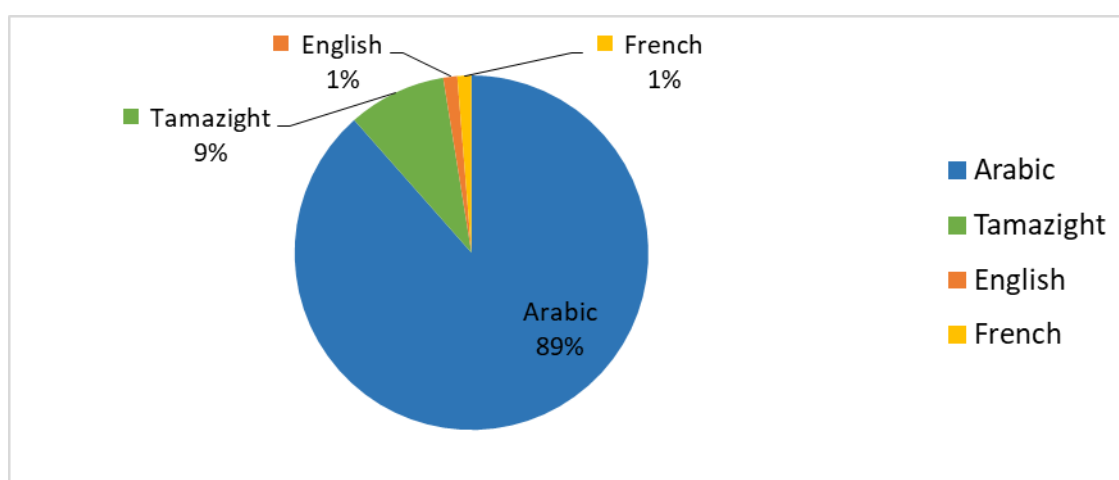


Figure N°01: Distribution of Respondents by Native Language

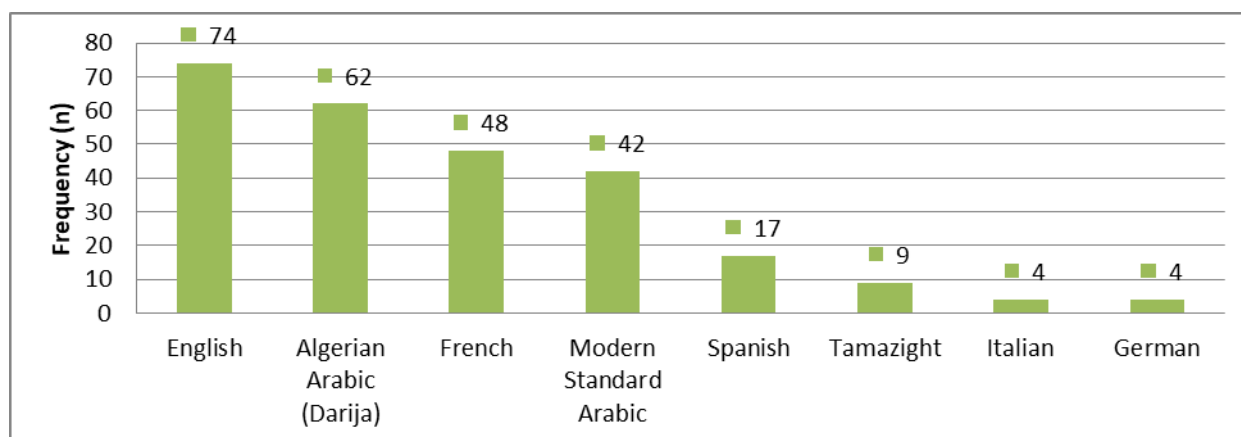


Figure N°02: Other Languages Spoken by Respondents

The results of the survey show that the majority of the 78 participants are native Arabic speakers (88.5%), with a Tamazight-speaking minority (9.0%), reflecting the typical sociolinguistic makeup of Algeria.

In terms of additional languages spoken, nearly all participants reported using English (94.9%) and Algerian Arabic/Darija (79.5%), with French also widely represented (61.5%). Modern Standard Arabic was reported by 53.8% of participants. Smaller numbers indicated proficiency in Spanish (21.8%), Italian (5.1%), and German (5.1%), languages that are offered as possible subjects in Algerian secondary schools.

These results confirm that participants operate within a rich plurilingual context, which is significant for this study as it suggests that AI-mediated translation practices may draw on multiple source languages.

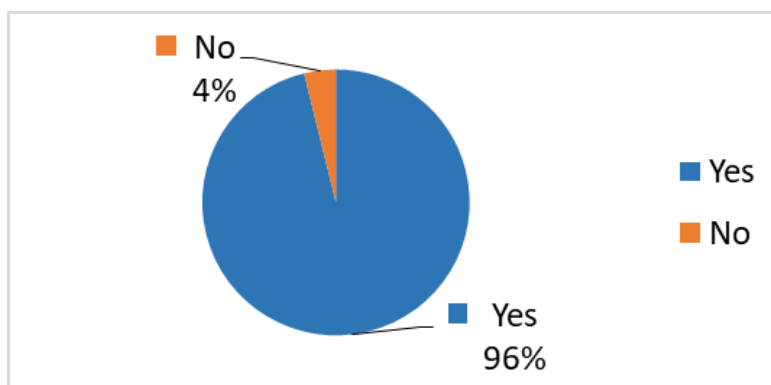


Figure N°03: Resorting to Translation for New/Difficult Words

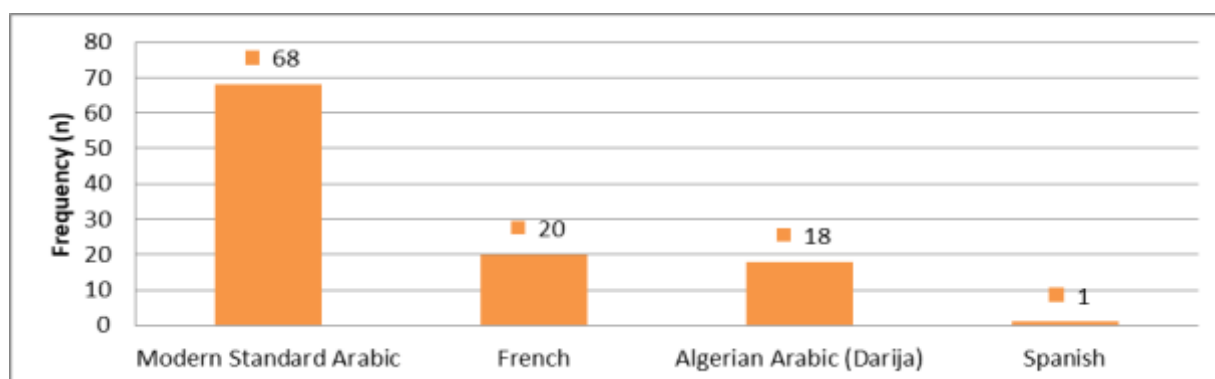


Figure N°04: Languages Used for Translation

The results show that the vast majority of participants resort to translation when encountering new or difficult English words (96.2%), indicating that translation is a near-universal vocabulary learning strategy within this sample.

Regarding the languages used for translation, Modern Standard Arabic was by far the most commonly reported medium (87.2%), followed by French (25.6%) and Algerian Arabic/Darija (23.1%), with English and Spanish rarely used (2.6% and 1.3% respectively). A crosstabulation analysis in SPSS further revealed that Arabic native speakers demonstrated the most diverse translation habits, drawing on Modern Standard Arabic (85.5%), French (24.6%), and Darija (20.3%), while Tamazight speakers showed a narrower range, relying almost exclusively on Modern Standard Arabic. This preference for standardized over dialectal varieties is noteworthy, as it likely reflects the stronger representation of Modern Standard Arabic and French in AI training data. Research suggests that AI translation tools perform more reliably with well-represented languages than with underrepresented ones such as Tamazight and dialectal Arabic (Alabi et al., 2022; Qin et al., 2025), which may result in unequal vocabulary learning outcomes depending on participants' native linguistic backgrounds. However, further research with larger, more balanced groups is needed to confirm these patterns.

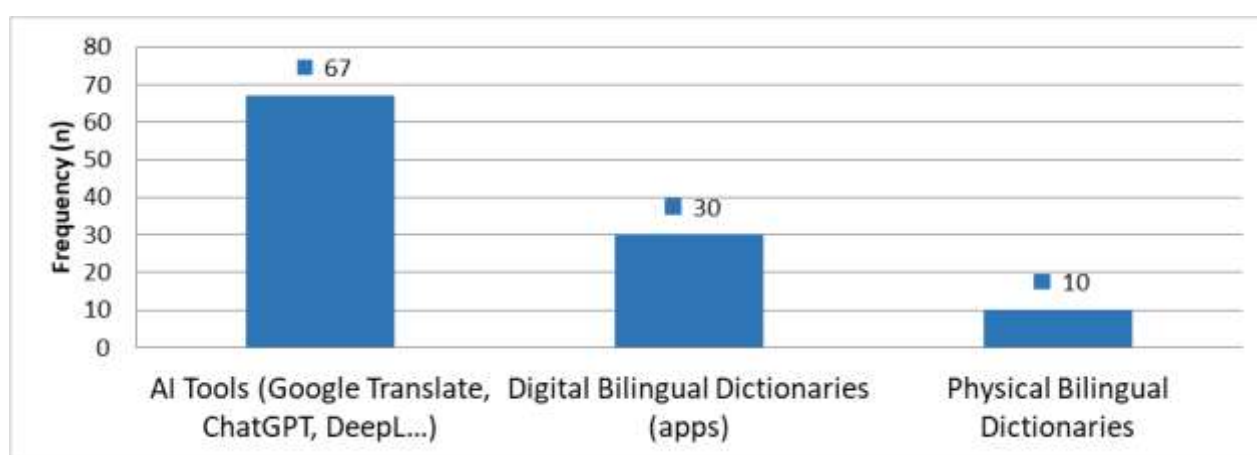


Figure N°05: Translation Tools Used by Respondents

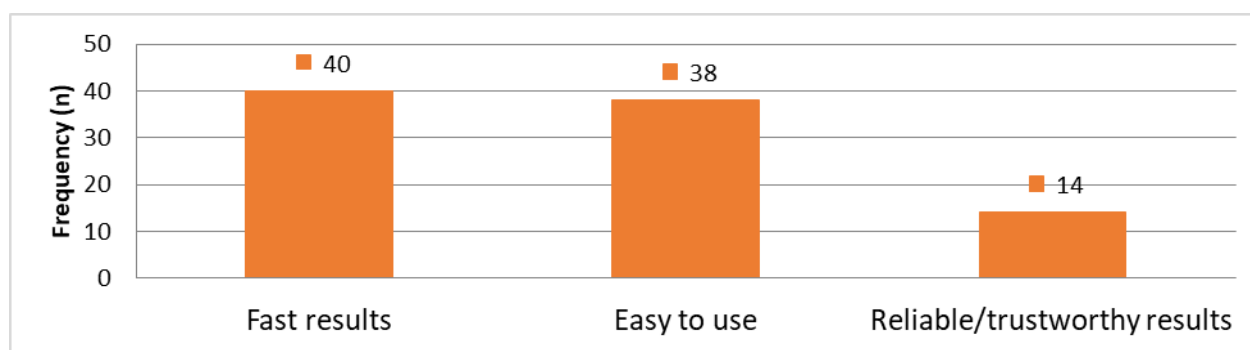


Figure N°06: Reasons for Choosing a Specific Translation Tool

The results show that AI tools such as Google Translate, ChatGPT, and DeepL are by far the most widely used translation tools among participants (85.9%), followed by digital bilingual dictionaries (38.5%) and physical dictionaries (12.8%). This clear preference for AI tools points to a significant shift away from traditional dictionary use in vocabulary learning practices.

When asked why they choose their preferred tool, participants most frequently cited speed (51.3%) and ease of use (48.7%), while reliability was a less common motivation (17.9%). This suggests that participants' tool selection is driven primarily by convenience rather than concern for accuracy, which raises important questions about the quality of vocabulary output they receive and their level of critical engagement with AI-generated translations.

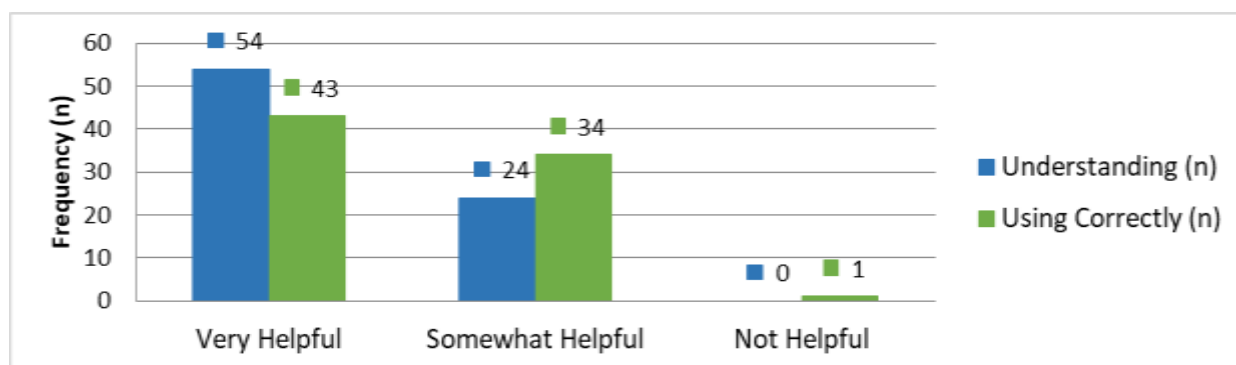


Figure N°07: Perceived Helpfulness of AI Translation Tools

The results show that participants generally view AI translation tools positively. For understanding new words (receptive vocabulary), the majority found AI tools very helpful (69.2%), with the remaining 30.8% finding them somewhat helpful. Ratings were slightly lower for using words correctly (productive vocabulary), with 55.1% reporting AI as very helpful and 43.6% as somewhat helpful, while one participant found it not helpful at all. This gap suggests that while AI translation tools effectively support word recognition and comprehension, their contribution to productive vocabulary mastery remains more limited.

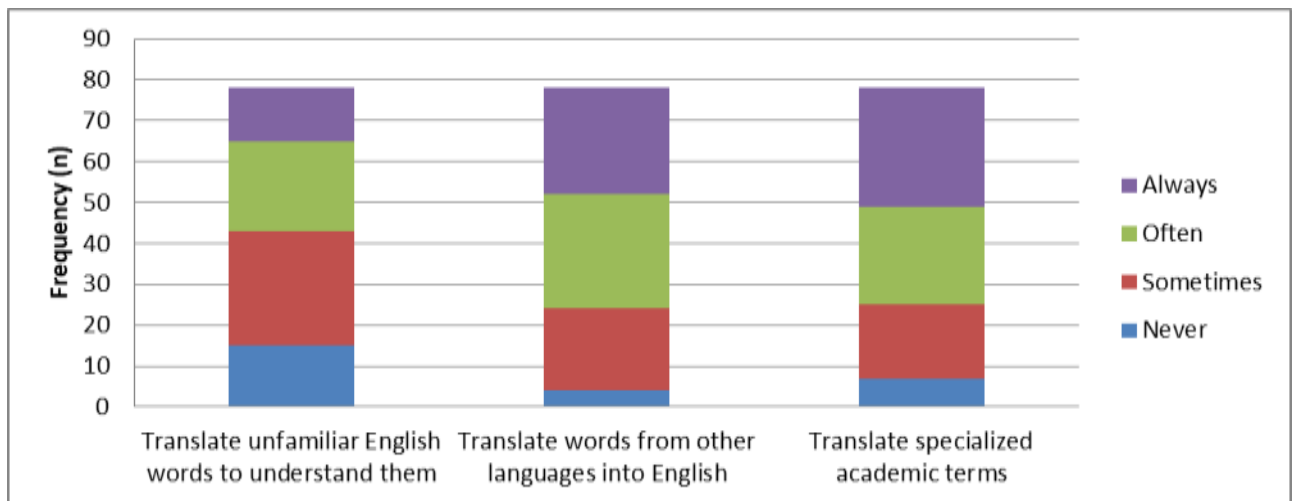


Figure N°08: Frequency of AI Use for Translation Tasks

The results show that participants most frequently use AI to translate words from their native languages into English (69.2% often or always) and to understand specialized academic terms (68.0% often or always), while using AI to understand unfamiliar words in general was somewhat less frequent (44.9% often or always). This suggests that participants tend to turn to AI when facing specific vocabulary challenges rather than using it on a regular basis.

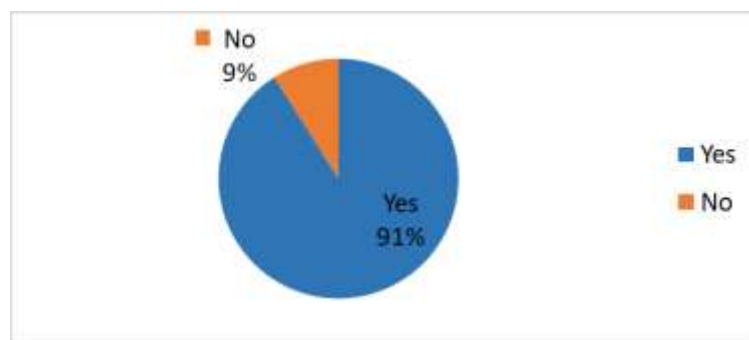


Figure N°09: Awareness of AI Translation Limitations

The results indicate that the large majority of participants are aware that AI translation tools can sometimes produce incorrect or inaccurate vocabulary (91.0%), with only 9.0% reporting no such awareness. While this high level of awareness of AI translation limitations is significant, it does not necessarily translate into critical checking behavior, as will be discussed in the following sections.

Lexical Feature	Never	Sometimes	Often	Always
Idiomatic Expressions	9	52	12	5
Phrasal Verbs	18	39	14	7
Cultural Appropriateness	14	43	19	2
Social Register	23	29	21	5
Collocations	21	31	19	7
Polysemy	20	30	17	11

Table N°01: Frequency Ratings for Perceived AI Failure Across Lexical Features (N=78)

The findings indicate a consistent recognition of AI translation limitations across all six examined linguistic features. Idiomatic expressions emerged as the most problematic category, with 67.9% of participants reporting occasional failures and 15.4% noting frequent or constant inaccuracies. A similar result was observed regarding phrasal verbs and cultural appropriateness, where over 60% of respondents acknowledged frequent AI failures. While participants also reported failures concerning social register, collocations, and polysemy, a notable proportion reported never noticing AI failure in these areas, suggesting limited awareness of these more subtle lexical features.

These findings are particularly concerning in an EFL vocabulary learning context, as the features most affected by AI inaccuracies are among the most difficult for learners to master. Without critical verification, repeated exposure to incorrect AI translations of idioms, phrasal verbs, or culturally inappropriate expressions could lead to defective and inaccurate vocabulary knowledge over time.

An independent samples t-test was conducted to determine whether self-reported awareness of AI translation limitations significantly influenced students' ability to perceive AI failures across specific lexical features. The results revealed no statistically significant difference between students who reported awareness of limitations ($M = 2.44$, $SD = 0.51$) and those who did not ($M = 2.39$, $SD = 0.35$), $t(76) = 0.071$, $p = .943$, $d = .114$. The Cohen's d value indicates a negligible effect, suggesting that reported awareness accounted for virtually no practical difference between the two groups. These findings indicate that mere conceptual awareness of AI limitations does not inherently translate into higher critical digital literacy. This indicates a potential gap between acknowledging potential inaccuracies and the actual recognition of subtle lexical failures, suggesting that active critical training may be required to improve digital literacy beyond basic awareness.

Behavior	Never	Sometimes	Often	Always
Checking semantic precision	5	18	26	29
Evaluating contextual fit	3	17	27	31
Cross-checking with trusted sources	11	29	22	16
Treating AI output as first draft	7	25	22	24

Table N°02: Frequency of Critical Checking Behaviors When Using AI Translations (N=78)

The results show that participants engage in critical verification to varying degrees. Checking semantic precision was the most frequently reported behavior (70.5% often or always), followed by evaluating contextual fit (58.9%) and treating AI output as a first draft (59.0%). Cross-checking with trusted external sources such as dictionaries or teachers was the least practiced, with only 48.7% reporting doing so often or always and 14.1% reporting never engaging in this behavior at all.

The findings further reveal a divergence between strategic practice and conceptual awareness. Mean scores were derived from grand means computed in SPSS by averaging participants' responses across related survey items on a four-point Likert scale. A paired samples t-test confirmed that this divergence is both statistically significant and practically meaningful, with

participants scoring higher on strategy application ($M = 2.87$, $SD = 0.70$) than on awareness of AI translation limitations ($M = 2.15$, $SD = 0.58$), $t(77) = -7.918$, $p < .001$, $d = 0.897$. The large effect size indicates that while students possess a repertoire of checking strategies, these are not grounded in a deep critical understanding of AI's linguistic constraints. A Pearson correlation further revealed a weak positive relationship between awareness of AI translation limitations and strategy application ($r = .227$, $p = .046$), suggesting that while higher awareness is associated with slightly more frequent verification behavior, the relationship remains too weak to indicate that awareness alone drives critical practice. Consequently, pedagogical intervention at the ENSB level should shift from merely teaching how to use these tools toward fostering a more critical awareness approach.

Reason	n	%
It takes too much time	35	44,9%
No access to other resources	24	30,8%
I trust AI is always correct	21	26,9%
Don't know how to verify	14	17,9%
Don't think it's necessary	10	12,8%

Table N°03: Main Reasons for Not Verifying AI Translation Output

The results show that the perception that verification takes too much time was the most commonly reported reason for not checking AI translation output (44.9%), followed by lack of access to alternative resources (30.8%) and trust in AI accuracy (26.9%). A smaller proportion reported not knowing how to verify translations (17.9%) or not seeing the need to do so (12.8%).

These findings suggest that non-verification is driven more by practical barriers than by indifference, though the notable proportion of participants who simply trust AI output is concerning given the limitations identified in the previous section. Together, these reasons point to the need for practical verification strategies that are both accessible and time-efficient enough to be realistically integrated into students' learning habits.

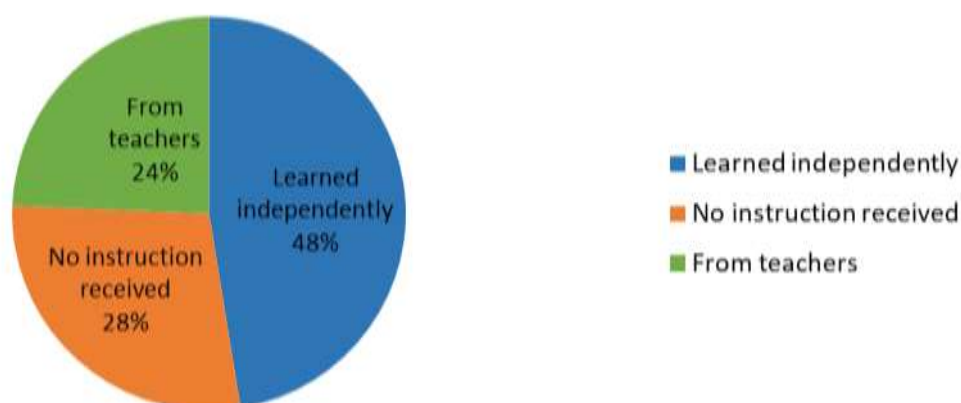


Figure N°10: Instruction Received on Effective Use of AI Translation Tools

The results reveal that the majority of participants learned to use AI translation tools on their own (47.4%), while 28.2% reported receiving no instruction at all. Only 19 participants

(24.4%) reported having received guidance from teachers, suggesting that formal instruction on the use of AI translation tools remains largely absent from the ENSB curriculum.

To evaluate the impact of explicit instruction on AI-mediated literacy, a One-Way ANOVA was conducted across three groups: those receiving teacher-led instruction, those engaging in self-initiated learning, and those reporting no explicit instruction. The results indicated no statistically significant differences in either Awareness_Grand_Mean ($F(2, 75) = .230, p = .795$) or Evaluation_Grand_Mean ($F(2, 75) = .372, p = .690$). Effect sizes were negligible ($\eta^2 = .006$ and $\eta^2 = .010$ respectively), suggesting that instruction type explained less than 1% of the variance in either awareness or evaluation scores. While these findings may suggest that current modes of instruction do not confer a measurable advantage in developing critical awareness of AI limitations or evaluative practices, the results should be interpreted with caution due to the small and uneven group sizes, which may limit the statistical power of the analysis. Since many participants are self-taught, independent use might not be enough to develop the critical skills needed to evaluate AI output. This suggests a potential need for more structured classroom guidance.

4.2.Vocabulary Evaluation Task Findings

Category	Arabic ID%	Arabic Corr%	French ID%	French Corr%	Overall Task%	Self-Reported%
Social Register	74.3%	65.1%	78.2%	68.4%	71.5%	70.5%
Cultural Differences	75.6%	66.8%	72.4%	62.7%	69.4%	82.1%
Phrasal Verbs	71.2%	62.4%	68.3%	58.6%	65.1%	76.9%
Idiomatic Expressions	65.8%	56.3%	61.4%	52.8%	59.1%	88.5%
Collocations	63.4%	54.7%	60.1%	51.3%	57.4%	73.1%
Polysemy	61.2%	52.6%	57.8%	48.9%	55.1%	74.4%
Overall	68.6%	59.7%	66.4%	57.1%	67.3%	77.6%

Table N°04: Test Results vs. Self-Reported Awareness of AI Failures (N=34 test, N=78 survey)

Table N°04 presents the results of the vocabulary evaluation task (n=34) across the six lexical categories examined in the survey (idiomatic expressions, phrasal verbs, collocations, polysemy, social register, and cultural differences) broken down by source language (Arabic and French), task stage (identification and correction), and overall performance. The results are presented alongside participants' self-reported awareness of AI failures in the same categories from the survey. Overall task performance stood at 67.3%, with identification scores (68.6% Arabic, 66.4% French) consistently exceeding correction scores (59.7% Arabic, 57.1% French) across all categories. Social register and cultural differences yielded the highest performance, while polysemy and collocations proved most challenging. Notably, self-reported awareness scores (77.6% overall) were consistently higher than actual task

performance across most categories, with the gap being particularly pronounced for idiomatic expressions, where self-reported awareness reached 88.5% against a task performance of 59.1%.

A paired samples t-test was conducted to compare participants' overall identification and correction scores. The results revealed a statistically significant difference between identification ($M = 8.12$, $SD = 1.59$) and correction scores ($M = 7.00$, $SD = 1.30$), $t(33) = 3.192$, $p = .003$, $d = 0.547$. The moderate effect size indicates that students were meaningfully better at identifying lexical inaccuracies in AI-generated translations than at correcting them, suggesting that recognition of errors does not automatically translate into productive lexical knowledge.

A second paired samples t-test was conducted to compare performance across the two source languages. The results indicated no statistically significant difference between Arabic source scores ($M = 7.71$, $SD = 1.27$) and French source scores ($M = 7.41$, $SD = 1.56$), $t(33) = 0.881$, $p = .385$, $d = .151$. The effect size was negligible, suggesting that source language did not significantly influence students' ability to detect lexical inaccuracies. Nevertheless, descriptive results indicated a slight tendency toward higher performance on Arabic-sourced items across most categories, which may warrant further investigation with a larger sample.

A One-Way ANOVA was conducted to examine whether performance differed significantly across the six lexical categories. The results indicated no statistically significant differences, $F(5, 198) = 1.494$, $p = .193$, $\eta^2 = .036$. The effect size was small, suggesting that lexical category explained only a limited proportion of variance in task performance. These findings should be interpreted with caution given the small sample size ($n = 34$), which limits statistical power. Descriptively, however, social register and cultural differences yielded the highest performance scores, while polysemy and collocations proved most challenging, a pattern consistent with participants' self-reported perceptions in the survey.

It is also worth noting that approximately 9-12% of participants skipped certain items entirely, while 6-9% indicated that the AI-translated sentence appeared correct to them despite task instructions explicitly stating that each sentence contained one lexical error. This suggests that a non-negligible proportion of students lacked the critical evaluative capacity to detect inaccuracies even when explicitly prompted, pointing to a fundamental gap in critical digital literacy beyond mere difficulty with correction.

While the uneven sample sizes preclude direct statistical comparison between the survey ($n=78$) and task ($n=34$) findings, a descriptive comparison offers useful insights. When considering at least occasional engagement, the vast majority of survey participants reported engaging in critical verification behaviors: evaluating contextual fit (96.2%), checking semantic precision (93.6%), and treating AI output as a first draft (91.0%). Task performance scores, however, reveal a more complex picture: while the overall task score stood at 67.3%, participants' correction score specifically ($M = 7.00$ out of 12, 58.3%) suggests that even when errors were identified, productive correction remained challenging. This divergence between self-reported verification behavior and demonstrated correction ability further

supports the gap between proclaimed critical digital literacy and its actual exercise, reinforcing the need for structured pedagogical intervention beyond self-directed awareness.

5. Discussion of the Findings

The findings of this study revealed that first-year ENSB students were active but largely uncritical users of AI-mediated translation for vocabulary learning. Operating within a rich plurilingual context, students drew on multiple source languages, predominantly Modern Standard Arabic (87.2%) and French (25.6%), when engaging with AI translation tools, reflecting the complex linguistic repertoire of Algerian EFL learners. While the vast majority relied on AI tools primarily for receptive vocabulary tasks, their use for productive purposes remained more limited. Although most students demonstrated a general awareness that AI can produce inaccurate output, this awareness did not translate into effective critical evaluation practice. Despite 96.2% reporting at least occasional verification behaviors, overall task performance stood at 67.3%, with correction scores dropping to 58.3%, indicating that self-reported critical engagement significantly overestimated actual evaluative capacity. Together, these findings suggest that current AI use promotes passive lexical acquisition rather than the deep, reflective engagement necessary for durable vocabulary development.

These findings were consistent with existing research on AI use among Algerian EFL students. Sebbah (2025) along with Berrarbi and Amrane (2025) found that Algerian university students frequently used AI tools with limited critical digital literacy skills, a pattern corroborated here: 85.9% of participants relied on AI tools, driven by speed (51.3%) and ease of use (48.7%) rather than accuracy. Boukhelkhal (2025) identified vocabulary as the skill most developed through AI among Algerian learners, aligning with the present study's finding that AI use was concentrated on receptive and productive vocabulary tasks. Yahiaoui (2023) identified a general deficit in digital literacy competencies among Algerian university students, and this study substantiated that finding specifically within vocabulary evaluation. Kobouch (2025) argued that AI translation output still requires human supervision to ensure semantic accuracy and contextual appropriateness, a concern directly evidenced here by students' inability to effectively correct AI-generated lexical errors. Bouzidi and Zarif (2025) noted students' uncritical trust in AI, reflected here in the 26.9% who did not verify output simply because they trusted AI to be correct. Collectively, studies conducted in the Algerian university context converge on a shared finding: EFL students use AI tools frequently but uncritically, and the present study both confirmed and extended this conclusion.

While previous Algerian studies established that AI tools were widely used for vocabulary learning as a general finding, none specifically examined the quality of students' critical engagement with AI-translated vocabulary at the lexical level. This study addressed that gap directly by focusing on the accuracy and appropriateness of AI-translated vocabulary across complex lexical features (such as idioms and phrasal verbs) that are particularly susceptible to AI inaccuracies. By combining a survey with a performance-based vocabulary evaluation task, it demonstrated empirically that students' self-reported verification behaviors significantly overestimated their actual evaluative capacity, a discrepancy that had not been previously documented in the Algerian EFL context. Moreover, the study showed that there was a notable gap between students' general awareness of AI limitations and their actual

ability to detect and correct AI-translated vocabulary errors, suggesting that awareness alone is insufficient for developing effective critical digital literacy. CDL instruction in the Algerian EFL context would therefore benefit from moving beyond raising general awareness and focusing on building concrete, vocabulary-specific critical evaluation skills.

6. Pedagogical Implications

The findings demonstrated that students' reliance on AI for vocabulary tasks often prioritized speed over accuracy. To address this, the ENSB curriculum could benefit from implementing explicit instruction on critical digital literacy when dealing with AI-mediated translation for vocabulary learning:

- **Explicit Lexical Instruction:** Incorporating direct training on complex EFL lexical features, such as idioms, phrasal verbs and collocations, could play a vital role in the curriculum. Developing an understanding of these nuances may enable students to better recognize when AI-generated vocabulary is unnatural or contextually incorrect.
- **From Passive to Active Inquiry:** It is suggested that training shift from teaching *how* to use AI to teaching *how* to challenge it. Under this approach, students are encouraged **to** treat AI output as a draft rather than an authoritative source, fostering a mindset of skepticism.
- **Verification Strategies:** Students may benefit from explicit training on verification strategies such as cross-referencing. This could be achieved through targeted classroom tasks that require them to actively evaluate, challenge, and refine AI-generated vocabulary rather than accepting it at face value.

7. Limitations of the Study

Several limitations should be considered when interpreting the findings of this study. The most significant methodological limitation concerns the uneven sample sizes between the two instruments: while 78 students completed the survey questionnaire, only 34 participated in the vocabulary evaluation task, which limits the comparability of findings across the two phases and reduces the statistical power of task-based analyses. Additionally, as both instruments were administered anonymously online, it was not possible to match individual survey responses to task scores, which precluded a direct individual-level statistical comparison between self-reported awareness and demonstrated performance. The survey data were inherently subject to social desirability bias, as participants may have overestimated their awareness of AI limitations or the frequency of their verification practices in order to appear more critically engaged than they actually were. At the analytical level, the One-Way ANOVA tests conducted on survey data involved small and uneven group sizes, which further limits the reliability of those comparisons and calls for cautious interpretation of the results. Moreover, the sample is drawn from a single institution, which limits the generalisability of the findings to other Algerian EFL contexts. Finally, the majority of participants were native Arabic speakers, which restricted the ability to draw meaningful conclusions about the specific translation behaviors and AI evaluation practices of Tamazight-speaking learners.

These limitations, while acknowledged, may nonetheless serve as valuable insights to guide future research design in this area.

8. Recommendations for Future Research

Future research should adopt a mixed-methods approach combining surveys with think-aloud protocols or classroom observations to capture how students process AI-generated vocabulary in real time. Longitudinal studies would further allow researchers to track how translation habits and critical skills evolve over time. Broadening the sample to include multiple institutions and a more diverse linguistic representation, particularly native Tamazight speakers, would also provide a more comprehensive understanding of AI-mediated translation practices in the Algerian EFL context.

9. Conclusion

This study investigated the use of AI-mediated translation for vocabulary learning among first-year EFL students at the École Normale Supérieure de Bouzaréah (ENSB), focusing on how these plurilingual learners navigate their linguistic inventory including mainly Arabic, French, Tamazight and English when engaging with AI-translated vocabulary. By analyzing the extent of AI usage, the level of critical digital literacy applied, and the relationship between students' awareness of AI limitations and their ability to detect lexical inaccuracies, the research aimed to uncover how this technology influences the development of English lexical proficiency in a formal academic environment.

Regarding the research questions, the findings first revealed that students rely heavily on AI for receptive vocabulary tasks, while its contribution to productive vocabulary remains limited. Second, although participants showed broad awareness of AI limitations, their critical evaluation of AI-generated output remains superficial, with difficulties identifying inaccuracies in complex lexical features such as idioms, register, and collocations. Third, the findings revealed that students' awareness of AI limitations does not significantly translate into the ability to detect and correct lexical inaccuracies, confirming that awareness alone is insufficient without structured guidance. Together, these findings suggest that current AI use promotes passive rather than reflective vocabulary acquisition.

Pedagogical practices should therefore shift toward explicit critical digital literacy instruction, equipping students with practical strategies to critically evaluate AI-translated vocabulary and identify its lexical limitations. Ultimately, supporting students in transitioning from passive consumers to active, critical users of AI-translated vocabulary is essential for fostering the deep, reflective engagement necessary for accurate and effective EFL vocabulary mastery.

Bibliography

Alabi, J. O., Adelani, D. I., Mosbach, M., & Klakow, D. (2022). Adapting pre-trained language models to African languages via multilingual adaptive fine-tuning. In *Proceedings of the 29th International Conference on Computational Linguistics* (pp. 4336–4349). International Committee on Computational Linguistics.

- Benfareh, A. (2025). The AI tongue-twister: Disentangling the algorithmic underpinnings of multilingual AI. *Revue Algérienne Des Sciences Du Langage*, 10(1), 285–297.
- Benrabah, M. (2014). Competition between four "world" languages in Algeria. *Journal of World Languages*, 1(1), 38–59. <https://doi.org/10.1080/21698252.2014.893676>
- Berrarbi, A., & Amrane, N. (2025). Enhancing EFL learners' critical thinking through human communication and human discussion activities in the era of AI. *Journal of Studies in Language, Culture and Society (JSLCS)*, 8(1), 266-281.
- Boukhelkhal, O. (2025). Algerian EFL students' perspectives and practices on AI-enabled learning at the University of Medea. *The Journal of Studies in Language, Culture and Society (JSLCS)*, 8(1), 93-110.
- Boukris, L. (2015). Plurilingualism and language policy: The case of Algeria. *Revue de Traduction et Langues/Journal of Translation and Languages*, 14(1), 94–99.
- Bouzidi, S., & Zarif, B. (2025). Les représentations sociales de l'intelligence artificielle chez les enseignants algériens: Perceptions, craintes et perspectives d'intégration pédagogique. *Passerelle*, 14(1), 142-165.
- Bowker, L., & Buitrago, J. (2019). *Machine translation and global research: Towards improved machine translation literacy in the scholarly community*. Emerald Publishing. <https://doi.org/10.1108/9781787567214>
- Cenoz, J. (2001). The effect of linguistic distance, L2 status and age on cross-linguistic influence in third language acquisition. In J. Cenoz, B. Hufeisen, & U. Jessner (Eds.), *Cross-linguistic influence in third language acquisition: Psycholinguistic perspectives* (pp. 8–20). Multilingual Matters. <https://doi.org/10.21832/9781853595509-002>
- Cenoz, J., & Gorter, D. (2015). Towards a holistic approach in the study of multilingual education. In J. Cenoz & D. Gorter (Eds.), *Multilingual education: Between language learning and translanguaging* (pp. 1–15). Cambridge University Press.
- Chapelle, C. A., & Sauro, S. (Eds.). (2017). *The handbook of technology and second language teaching and learning*. Wiley-Blackwell. <https://doi.org/10.1002/9781118914069>
- Cook, G. (2010). *Translation in language teaching: An argument for reassessment*. Oxford University Press.
- Cook, V. (2016). *Second language learning and language teaching* (5th ed.). Routledge. <https://doi.org/10.4324/9781315883113>
- Coste, D., Moore, D., & Zarate, G. (2009). *Plurilingual and pluricultural competence* (English version). Council of Europe.
- Council of Europe. (2001). *Common European framework of reference for languages: Learning, teaching, assessment*. Cambridge University Press.

- Creswell, J. W. (2009). *Research design: Qualitative, quantitative, and mixed methods approaches* (3rd ed.). SAGE Publications.
- Cummins, J. (2007). Rethinking monolingual instructional strategies in multilingual classrooms. *Canadian Journal of Applied Linguistics*, 10(2), 221–240.
- Daller, H., Milton, J., & Treffers-Daller, J. (2007). *Modelling and assessing vocabulary knowledge*. Cambridge University Press. <https://doi.org/10.1017/CBO9780511667268>
- Dörnyei, Z. (2007). *Research methods in applied linguistics: Quantitative, qualitative, and mixed methodologies*. Oxford University Press.
- Folse, K. S. (2004). *Vocabulary myths: Applying second language research to classroom teaching*. University of Michigan Press. <https://doi.org/10.3998/mpub.23925>
- García, O., & Wei, L. (2014). *Translanguaging: Language, bilingualism and education*. Palgrave Macmillan. <https://doi.org/10.1057/9781137385765>
- Ghounane, N., Rahmani, A., & Al-Zubaidi, K. (2024). Exploring Algerian EFL master's students' attitudes toward AI-giarism. *Indonesian Journal of Social Science Research*, 5(2), 444–459. <https://doi.org/10.11594/ijssr.05.02.07>
- Granger, S., & Swallow, H. (1988). *False friends: a kaleidoscope of translation difficulties*. *Le Langage et l'Homme*, 23(2), 108–120.
- Hall, G., & Cook, G. (2012). Own-language use in language teaching and learning. *Language Teaching*, 45(3), 271–308. <https://doi.org/10.1017/S0261444812000067>
- Hatch, E., & Lazaraton, A. (1991). *The research manual: Design and statistics for applied linguistics*. Newbury House.
- Hellmich, E. A., & Vinall, K. (2023). Student use and instructor beliefs: Machine translation in language education. *Language Learning & Technology*, 27(1), 1–27. <https://doi.org/10.64152/10125/73525>
- Jeon, J., Li, W., Tai, K. W. H., & Lee, S. (2025). Generative AI and its dilemmas: Exploring AI from a translanguaging perspective. *Applied Linguistics*, 46(4), 709–717. <https://doi.org/10.1093/applin/amaf049>
- Jessner, U. (2006). *Linguistic awareness in multilinguals: English as a third language*. Edinburgh University Press. <https://doi.org/10.1515/9780748626540>
- Kobouch, A. (2025). Human-AI Hybrid Translation Systems in Algeria: Efficiency, Policy, and Practice from a Government Translator's View. *Journal of Languages and Translation*, 5(3), 32–44.
- Kroll, J. F., & Stewart, E. (1994). Category interference in translation and picture naming: Evidence for asymmetric connections between bilingual memory representations.

- Journal of Memory and Language*, 33(2), 149–174.
<https://doi.org/10.1006/jmla.1994.1008>
- Laufer, B., & Girsai, N. (2008). Form-focused instruction in second language vocabulary learning: A case for contrastive analysis and translation. *Applied Linguistics*, 29(4), 694–716. <https://doi.org/10.1093/applin/amn018>
- Lee, S., Choe, H., Zou, D., & Jeon, J. (2026). Generative AI (GenAI) in the language classroom: A systematic review. *Interactive Learning Environments*, 34(1), 335–359. <https://doi.org/10.1080/10494820.2025.2498537>
- Messekher, H., & Miliani, M. (2020). The language situation of twenty-first-century Algeria: Navigating the ideology. In Y. H. Zoubir (Ed.), *The politics of Algeria: Domestic issues and international relations* (pp. 150–163). Routledge. <https://doi.org/10.4324/9780429447495-11>
- Nation, I. S. P. (2001). *Learning vocabulary in another language*. Cambridge University Press.
- Nation, I. S. P. (2013). *Learning vocabulary in another language* (2nd ed.). Cambridge University Press.
- Ng, D. T. K., Leung, J. K. L., Chu, S. K. W., & Qiao, M. S. (2021). Conceptualizing AI literacy: An exploratory review. *Computers and Education: Artificial Intelligence*, 2, 100041. <https://doi.org/10.1016/j.caeai.2021.100041>
- Niño, A. (2020). Exploring the use of online machine translation for independent language learning. *Research in Learning Technology*, 28, 2402. <https://doi.org/10.25304/rlt.v28.2402>
- Ouahmiche, G., Beddiaf, A., & Beddiaf, A. (2017). Reflections on the linguistic landscape and the prospects of English language teaching in Algeria. *International Journal of Language and Linguistics*, 5(3-1), 15–23.
- Pangrazio, L. (2016). Reconceptualising critical digital literacy. *Discourse: Studies in the Cultural Politics of Education*, 37(2), 163–174. <https://doi.org/10.1080/01596306.2014.942836>
- Pool, C. R. (1997). A new digital literacy: A conversation with Paul Gilster. *Educational Leadership*, 55(3), 6–11.
- Prince, P. (1996). Second language vocabulary learning: The role of context versus translations as a function of proficiency. *The Modern Language Journal*, 80(4), 478–493. <https://doi.org/10.1111/j.1540-4781.1996.tb05468.x>
- Qin, L., Chen, Q., Zhou, Y., Chen, Z., Li, Y., Liao, L., Li, M., Che, W., & Yu, P. S. (2025). A survey of multilingual large language models: Patterns, challenges and opportunities. *Patterns*, 6(1). <https://doi.org/10.1016/j.patter.2024.101118>

- Rothman, J. (2011). L3 syntactic transfer selectivity and typological determinacy: The typological primacy model. *Second Language Research*, 27(1), 107–127. <https://doi.org/10.1177/0267658310386439>
- Rouabah, S. (2022). Multilingualism in Algeria: Educational policies, language practices and challenges. *Journal of the British Academy*, 10(s4), 21–40. <https://doi.org/10.5871/jba/010s4.021>
- Schmitt, N. (2008). Instructed second language vocabulary learning. *Language Teaching Research*, 12(3), 329–363. <https://doi.org/10.1177/1362168808089921>
- Sebbah, L. (2025). Exploring Algerian EFL students' familiarity, use and attitudes towards generative artificial intelligence tools in education. *Journal of Languages & Translation*, 5(1), 1–21.
- Tihal, W. (2025). The shift from French to English in Algeria: A welcome change or controversial move? *Education and Linguistics Research Journal*, 11(1). <https://doi.org/10.5296/elr.v11i1.22800>
- Webb, S., & Nation, P. (2017). *How vocabulary is learned*. Oxford University Press.
- Wineburg, S., & McGrew, S. (2019). Lateral reading and the nature of expertise: reading less and learning more when evaluating digital information. *Teachers College Record: The Voice of Scholarship in Education*, 121(11), 1-40. <https://doi.org/10.1177/016146811912101102>
- Yahiaoui, K. (2023). Littératie numérique sur le terrain universitaire en Algérie: Réalités et défis. *Revue Plurilingue: Études des Langues, Littératures et Cultures*, 7(2), 57-70.

Appendices

Appendix 1: The Survey Questionnaire

Questionnaire: AI-Mediated Translation and Vocabulary Learning: A Study of Critical Digital Literacy

This survey investigates how university students use AI translation tools (such as ChatGPT, Google Translate, or DeepL) for learning English vocabulary and how they evaluate the results. This survey is completely anonymous — your name and answers are not recorded — so please answer as honestly as possible.

Q1. What is your native language(s)? *

- ☐ Arabic
- ☐ Tamazight
- ☐ French
- ☐ Other: _____

Q2. Which other languages do you speak? (multiple answers) *

- ☐ Modern Standard Arabic
- ☐ Algerian Arabic (Darija)
- ☐ Tamazight
- ☐ French
- ☐ English
- ☐ Other: _____

Q3. Do you resort to translation when you encounter new or difficult English words? *

- ☐ Yes
- ☐ No

Q4. Which languages do you use for translation? (multiple answers) *

- ☐ Modern Standard Arabic
- ☐ Algerian Arabic (Darija)
- ☐ Tamazight
- ☐ French
- ☐ Other: _____

Q5. Which tools do you usually use to translate? *

- ☐ Physical Bilingual Dictionaries
- ☐ Digital Bilingual Dictionaries (e.g., applications)
- ☐ AI tools (e.g., Google Translate, ChatGPT, DeepL...)
- ☐ Other: _____

Q6. Why do you choose this specific tool? *

- ☐ Fast results
- ☐ Easy to use
- ☐ Reliable / trustworthy results
- ☐ Other: _____

Q7. Do AI translation tools help you understand the meaning of new English words? *

- ☐ Very helpful
- ☐ Somewhat helpful
- ☐ Not helpful

Q8. Do AI translation tools help you use new English words correctly? *

- ☐ Very helpful
- ☐ Somewhat helpful

- Not helpful

Q9. How often do you use AI translation tools for the following? *

Statement	Never	Sometimes	Often	Al-ways
I use AI to translate unfamiliar English words to understand them.	☐	☐	☐	☐
I use AI to translate words from Arabic, French, or Tamazight into English.	☐	☐	☐	☐
I use AI to translate and understand specialized academic terms.	☐	☐	☐	☐

Q10. Are you aware that AI tools can sometimes provide incorrect or inaccurate vocabulary?

- Yes
- No

Q11. How often do you notice AI fails to capture the correct: *

Statement	Never	Sometimes	Often	Al-ways
Idiomatic expressions (e.g., "bite the bullet")	☐	☐	☐	☐
Phrasal verbs (e.g., "look for" vs. "look into")	☐	☐	☐	☐
Cultural appropriateness	☐	☐	☐	☐
Social register (formal vs. informal tone)	☐	☐	☐	☐
Collocations (e.g., "make" a decision, not "take" a decision)	☐	☐	☐	☐
Polysemy (e.g., "heart" the organ vs. "heart" the center)	☐	☐	☐	☐

Q12. How often do you perform the following? *

Statement	Never	Sometimes	Often	Al-ways
I check if the AI suggestion is semantically precise (exactly what I want to express).	☐	☐	☐	☐
I evaluate whether the translation fits the context of my specific sentence.	☐	☐	☐	☐
I cross-check with trusted sources (e.g., dictionaries or teachers).	☐	☐	☐	☐
I treat the AI output as a first draft and modify it to match my style.	☐	☐	☐	☐

Q13. If you don't always check AI translations, why not? (multiple answers) *

- I trust that AI is always correct.
- I don't know how to verify translations.
- It takes too much time.
- I don't have access to other resources (dictionaries, teachers).
- I don't think it's necessary.
- Other: _____

Q14. Have you received any instruction on how to use AI tools effectively for language learning? *

- Yes, from my teachers.
- Yes, I learned on my own.
- No.

Appendix 2: AI-Translated Vocabulary Evaluation Task

Instructions: The following sentences were generated using AI translation from Arabic/French to English. Each sentence contains one lexical (vocabulary) error. Identify then correct the error in each sentence to sound more natural in English.

Example:

✗ He closed the light.

✓ Correction: He **turned off** the light.

1) She is such a kind person her heart is white.

.....

2) She told her mother: "I am not going to the party! That's a final point!"

.....

3) He came in from the window because the door was locked.

.....

4) John got in the bus heading to town.

.....

5) (*Talking to an older American stranger*): "Excuse me uncle, could you please tell me where the bus station is?"

.....

6) Children buy new outfits for the school return in September.

.....

7) He told his best friend: "It is my honor to invite you to my birthday party this Saturday".

.....

8) (*Business e-mail*): "I want to inform you that I will not be present tomorrow".

.....

9) The students did a lot of effort to finish the project on time.

.....

10) Sarah had to take a difficult decision regarding her career.

.....

11) He drank some heavy coffee to get enough energy for work.

.....

12) He assisted the conference and took notes.

.....

Appendix 3: AI-Translated Vocabulary Evaluation Task (Answer Key and Original Sentences)

No.	Category	Source Language	Original Sentence	AI-Generated Sentence	Error	Correction
1	Idiom	Arabic	هي إنسانة طيبة، قلبها أبيض	She is such a kind person, her heart is white.	heart is white	she has a pure/good heart/a heart of gold
2	Idiom	French	Elle a dit à sa mère : « Je n'irai pas à la fête ! point final ! »	She told her mother: "I am not going to the party! That's a final point!"	final point	And that's final!/ End of discussion!/ Period!
3	Phrasal Verb	Arabic	دخل من النافذة لأن الباب كان مغلقاً	He came in from the window because the door was locked.	came in from	came in through
4	Phrasal Verb	French	John est monté dans le bus en direction de la ville.	John got in the bus heading to town.	got in	got on
5	Cultural Differences	Arabic	عفواً عمي، هل يمكنك إخباري أين محطة الحافلات؟	(Talking to an older American stranger): "Excuse me uncle, could you please tell me where the bus station is?"	uncle	sir
6	Cultural Differences	French	Les enfants achètent de nouveaux vêtements pour la rentrée scolaire en septembre.	Children buy new outfits for the school return in September.	school return	back-to-school season
7	Social Register	Arabic	قال لصديقه المقرب: "يشرفني أن أدعوك إلى حفلة عيد ميلادي هذا السبت"	He told his best friend: "It is my honor to invite you to my birthday party this Saturday."	It is my honor to invite you	I'd like to invite you/ You should come to my..
8	Social Register	French	(E-mail professionnel): « J'aimerais vous informer que je ne serai pas présent demain. »	(Business e-mail): "I want to inform you that I will not be present tomorrow."	I want to inform you	I would like to inform you
9	Collocation	Arabic	بذل الطلاب جهداً كبيراً لإنهاء المشروع في الوقت المحدد	The students did a lot of effort to finish the project on time.	did a lot of effort	made a lot of effort
10	Collocation	French	Sarah a dû prendre une décision difficile concernant sa carrière.	Sarah had to take a difficult decision regarding her career.	take a decision	make a decision
11	Polysemy	Arabic	شرب قهوة ثقيلة ليحصل على طاقة كافية للعمل	He drank some heavy coffee to get enough energy for work.	Heavy coffee	Strong coffee
12	Polysemy	French	Il a assisté à la conférence et a pris des notes.	He assisted the conference and took notes.	Assisted	Attended