Tracing AI-Mediated Writing Tool Use Among Korean University Students: An Exploratory Screen-Recording Case Study

Janet Jaymin Kim[†]
Seoul National University

ABSTRACT

This exploratory case study uses process-tracing to examine how four Korean EFL university students utilize online dictionaries, machine translation, and ChatGPT during English argumentative writing. Analyzing screen recordings across three drafts, the study identifies specific tool-use purposes, sequences, and frequencies. Results reveal that tool-use trajectories are independent of English writing proficiency, reflecting idiosyncratic strategic choices and personal styles. Despite its potential, ChatGPT use remained minimal and surface-level due to study constraints, exposing a gap between AI's potential and actual application. As the instructional design prioritized restricting AI, participants' usage patterns demonstrated a tendency toward safe and basic routines, hindering the development of strategic digital writing skills. The study advocates for guided support to foster student agency and reflective practice. Ultimately, this process-oriented analysis provides a framework for curriculum designers seeking to integrate generative AI in ways that are both academically responsible and cognitively expansive.

Keywords: AI-mediated writing, process-tracing, Korean EFL university students, English argumentative writing, critical AI literacy

1. Introduction

Artificial intelligence (AI) is reshaping academic writing practices in English as a foreign language (EFL) education. Tools such as ChatGPT, DeepL, Google Translate, and Grammarly are increasingly used by L2 writers to support idea generation, translation, language refinement, and feedback. These applications, collectively referred to here as AI-mediated writing tools, encompass any digital resources that assist with writing tasks, including online dictionaries, translation software, grammar/spell checkers, and generative AI platforms. Such tools offer both

[†] Corresponding author: janetkimatsnu@gmail.com

pedagogical opportunities and ethical challenges in the writing classroom.

AI-mediated writing tools can offer L2 writers immediate support with vocabulary expansion, grammar correction, and the expression of ideas, contributing to improved writing fluency and confidence (Barrot, 2023; Kohnke, 2024; Su et al., 2023). They also have the potential to reduce teachers' workload by handling language-level issues, enabling more attention to content and rhetorical development (Hyland, 2025). Yet concerns persist regarding students' overreliance on AI-generated content and a lack of critical engagement. Research highlights issues such as the uncritical adoption of biased or inaccurate AI output (Warschauer et al., 2023), improper citation practices (Jarrah et al., 2023), and the challenge of detecting AI-generated text (Gao et al., 2023). These concerns underscore the need to move beyond outcome-based evaluation and investigate how students actually use AI-mediated writing tools during the writing process.

To date, most studies have relied on student surveys, interview data, or final drafts, offering limited insight into learners' real-time writing behaviors. As AI-mediated writing tools become embedded in students' writing routines, process-oriented methods are needed to examine how student writers navigate multiple resources, evaluate suggestions, and shift between tools. These real-time behaviors are often observed through screen capture or recording, an analysis tool that offers insights into learners' writing strategies as they emerge on screen (Hamel et al., 2015; Mroz, 2025).

This study addresses that gap by examining how Korean EFL university students interact with AI-mediated writing tools and online resources while writing argumentative essays, a central genre in academic writing. Drawing on screen-recorded data across three drafts, the study analyzes tool-use sequences, purposes, and switching patterns, showcasing learners of different English writing proficiency levels. The study aims to reveal individual differences in AI-mediated writing strategies and to identify pedagogical implications that promote responsible and ethical engagement with AI-mediated writing tools in academic writing.

2. Literature Review

To establish a clear empirical backdrop for the multi-tool behaviors analyzed in this study, the following literature review is organized by tool type. This structure traces the functional evolution of digital writing support, situating various resources

within the broader digital landscape to examine how the specific affordances and constraints of each technology have served as mediators of the L2 writing process.

2.1. Online dictionary use by EFL student writers

As foundational AI-mediated writing tools, online dictionaries have been extensively studied through comparative frameworks, evaluating monolingual versus bilingual formats (Gilquin & Laporte, 2021) and the transition from print to electronic media (Dziemianko, 2010). Beyond the format of the dictionary, researchers have examined how learners interact with specific parts of an entry, such as grammatical information (Chan, 2012) or definitions (Lew, 2010). These studies often use controlled experiments to measure how well students learn specific language features, including collocations or grammatical constructions (Chen, 2022).

In contemporary digital writing, online dictionaries function as immediate support tools within a larger digital environment (Lew & de Schryver, 2014; Shin et al., 2021). Rather than being isolated events, dictionary use is a recurring, individualized process that significantly impacts text quality (Gilquin & Laporte, 2021). Similarly, research on mobile learning suggests that students prioritize speed by relying on a limited repertoire of familiar applications to resolve vocabulary issues (Alzahrani, 2025).

However, simply having access to language-support tools such as online dictionaries does not guarantee better writing. Empirical evidence indicates that while low-proficiency learners can achieve significant improvements in writing accuracy and performance scores, these gains are typically restricted to surface-level and micro-level corrections (Chung & Ahn, 2022; Shin et al., 2021). Such writers often remain tethered to immediate lexical or orthographic needs, such as checking spellings or basic definitions, due to a lack of strategic competence and reference skills. This prevents them from effectively addressing higher-order rhetorical demands (Frankenberg-Garcia, 2020; Shin et al., 2021). Consequently, the effectiveness of dictionary use depends on the learner's ability to move beyond simple lookups through guided pedagogical routines (Frankenberg-Garcia, 2020).

2.2. Machine translation use by EFL student writers

Similarly, machine translation (MT) has transitioned from a prohibited aid to routine support for managing linguistic gaps (Chung & Ahn, 2022; Shin et al., 2021).

While MT significantly improves text accuracy, its impact on syntactic and lexical complexity remains inconsistent across proficiency levels (Chung & Ahn, 2022). In collaborative settings, students treat these tools as resources to resolve language problems, verify synonyms, and confirm usage while writing (Musk, 2022). However, this interaction is often complex, requiring students to negotiate and question the accuracy of the translation output.

Despite the popularity of MT, its impact on the final written product is mixed. While MT helps Korean college students produce text, the resulting essays often appear unfiltered rather than analytically developed (Baek & Rha, 2023). Grammatical and lexical errors frequently remain, suggesting that access to translation tools does not automatically lead to rhetorical sophistication or strong argumentation.

The effectiveness of MT largely depends on explicit instruction. For example, Korean middle school students significantly improved their spelling and grammatical accuracy after receiving guidance on how to effectively utilize Google Translate for revisions (Lee & Oh, 2025). Without such training, students often use the tools superficially. This need for post-editing literacy applies to advanced writers as well; even professional researchers must rely on a deep understanding of linguistic and rhetorical conventions to bridge the gap between a grammatically correct translation and an academically appropriate text (Comelles & Laso, 2025).

2.3. Generative AI use by EFL student writers

Generative AI tools like ChatGPT have fundamentally expanded the scope of digital writing support, enabling dialogic interaction for brainstorming and drafting. Studies indicate that this shift can lead to measurable improvements in writing performance. For instance, AI-assisted revisions often receive significantly higher scores for vocabulary, grammar, and organization compared to unassisted drafts (Song & Song, 2023). Additionally, using AI during the pre-writing stage can help learners generate a wider range of ideas, leading to higher-quality texts and increased engagement with the task (Hwang et al., 2025; Nguyen et al., 2025).

However, the value of generative AI tools depends on when and how learners use them. Research indicates that student interaction with generative AI shifts during writing, often moving from broad idea generation in early drafts to targeted linguistic checks during revision (Hwang et al., 2025). Because a polished final product can mask a wide spectrum of engagement, ranging from deep critical reflection to

unreflective copy-pasting, identifying behavioral and emotional patterns is key to understanding how students process AI feedback (Koltovskaia et al., 2024).

Despite functional benefits, integrating generative AI raises concerns about authorship, academic integrity, and blind reliance on the tool (Barrot, 2023; Warschauer et al., 2023). Addressing these risks requires shifting from banning the tools to developing critical AI literacy. Meaningful engagement depends on the learner's ability to evaluate AI output and maintain their own voice (Wang & Wang, 2025). Consequently, successful implementation relies on teacher guidance that aligns the technology with the students' specific needs (Hyland, 2025; Kohnke et al., 2023; Liu & Chao, 2018).

2.4. Gaps in AI-mediated L2 writing research and rationale for the present study

Despite the advances of AI-mediated writing tools, critical gaps remain regarding how these technologies are used together. Most notably, there is little empirical work capturing the complexity of tool use in argumentative writing, a genre essential for academic literacy. While recent studies have examined interactions with specific AI tools (Hwang et al., 2025; Koltovskaia et al., 2024; Nguyen et al., 2025), few have visualized how learners navigate across multiple resources, coordinating online dictionaries, machine translation, and generative AI, within a single task.

Furthermore, there is a lack of longitudinal data capturing how AI-mediated behaviors evolve across multiple drafts within individual writing trajectories. While language ability is often assumed to dictate strategic tool use, existing evidence remains inconclusive regarding how specific English writing proficiency levels characterize a learner's ability to evaluate AI suggestions, switch between resources, or regulate the writing process over time. Given the inherently idiosyncratic nature of the writing process, research must move beyond broad generalizations to examine the unique ways individual learners navigate digital environments from an ecological perspective.

To address these gaps, a process-oriented case study approach is required to trace how learners manage a suite of online resources—including online dictionaries, machine translation (MT), and ChatGPT—during drafting, revising, and editing. Unlike product-based analyses, process-oriented research employs screen recording as a non-intrusive lens to capture the temporal dynamics and cognitive moves often lost in final texts, identifying specific linguistic hurdles and tracing how L2 writers negotiate with digital artifacts to evaluate, incorporate, or reject AI suggestions

(Hamel et al., 2015; Mroz, 2025). Building on this need, the present study investigates the individual writing processes of four Korean EFL university students across three drafts of an English argumentative essay. The research questions guiding the study are as follows:

1) What are the temporal patterns and functional purposes of AI-mediated writing tool use (i.e., online dictionaries, machine translation, and ChatGPT) among four Korean EFL university writers across different English writing proficiency levels?

2) How do the sequential routines and switching behaviors between these tools differ across the drafting stages for four Korean EFL university writers across different English writing proficiency levels?

3. Methods

3.1. Study context and participants

The study was conducted in an English Composition course for English Education majors at a research-oriented university in Seoul. Students wrote their final paper on a self-selected topic within one of the genres covered in class, completing three drafts on their own laptops. To regulate AI use and facilitate consistent, supervised screen-recording, three 70-minute in-class writing sessions were conducted where in-situ assistance was provided.

Purposeful sampling was adopted to select cases rich in information and likely to yield insights into the research questions (Patton, 2015). Four students who wrote in the same genre and closely followed the screen-recording guidelines were purposefully selected to control for genre-related variation and to include writers at various English writing proficiency levels. Participants' English writing proficiency was determined by a diagnostic writing task completed in Week 1. Texts were evaluated using Jacobs et al.'s (1981) analytic scoring rubric, which assesses content, organization, vocabulary, language use, and mechanics. Total scores were used to group students into four English writing proficiency levels based on the overall score distribution, enabling comparison of tool-use behaviors across levels. Table 1 presents their background information, with pseudonyms used to protect participants' identities.

Table 1. Participants' background information and writing topic

Participant	Year	English writing proficiency	Genre	Topic
Jiwan	Sophomore	Advanced	Argumentative essay	Cancel culture
Suha	Sophomore	Upper-intermediate	Argumentative essay	English speaking assessment in Korean secondary school
Minjung	Senior	Intermediate	Argumentative essay	Native-level English proficiency requirements for Korean non-native English teachers
Hankyu	Sophomore	Lower-intermediate	Argumentative essay	AI as research tool

3.2. Data collection and analysis

Data were obtained from three primary sources: (1) background survey questionnaires, which included consent forms and statements agreeing not to use AI for content generation or idea organization, intended to encourage original idea development and prevent overuse of AI; (2) Zoom screen recordings of each student's in-class writing sessions (totaling 842 minutes); and (3) the three drafts of each participant's final paper. No prior training on AI-mediated writing tools was provided to ensure participants' tool-use reflected natural writing habits. The screen recordings captured real-time writing activity, including application switching, online searches, and interactions with AI-mediated tools.

In this study, AI-mediated writing tools are broadly defined as online resources using computational or AI-driven processes to support the composition, revision, and refinement of texts. This category encompasses online dictionaries (OD), machine translation (MT), and generative AI systems, such as ChatGPT. These classifications guided the inductive coding process, where every tool-use instance identified in the screen recordings was categorized accordingly. Online dictionaries included monolingual English–English and bilingual Korean–English dictionaries, while MT tools comprised Google Translate, Papago, and DeepL. Although MT tools also rely on AI, they were separated from generative AI systems to distinguish translation from generative text production. ChatGPT was the only generative AI tool observed and was therefore coded as GPT. Table 2 summarizes the three tool types.

Table 2. Subcategories of AI-mediated writing tools observed in the study

Type	Code	Description
Online dictionaries	OD	English–English or Korean–English dictionaries (e.g., Oxford, Naver English dictionary), thesauri
Machine translation	MT	Google Translate, Papago, DeepL
Generative AI	GPT	ChatGPT

Participants' purposes for using AI-mediated writing tools were coded based on the observable actions taken after each tool consultation. Following established process-tracing methods (Hwang et al., 2025; Su et al., 2023), coding was intention-based when writers provided explicit prompts (e.g., "check grammar") and inference-based when purpose had to be interpreted from subsequent revisions (e.g., lexical adjustments). These codes were applied consistently across OD, MT, and GPT, with the final code determined by the linguistic function of the resulting change to the draft.

To capture changes across three drafts, tool-use purposes were categorized and coded, drawing on frameworks established in prior studies (Hwang et al., 2025; Lee & Oh, 2025). Table 3 presents the dominant purposes of AI-mediated writing tool use, categorized as translation (TRANS), source searching (SS), grammar correction (GC), word search for lexical adjustments (WS), formatting (FM), and mechanics checking (MCH).

Table 3. Coding scheme of tool-use purposes across three drafts

Code	Description	Code	Description
TRANS	To translate between Korean and English	SS	To search for sources to cite
GC	To correct grammar	WS	To search better word choices
FM	To format according to APA	MCH	To check mechanics (punctuation, spacing)

Although drafting, revising, and editing can occur fluidly across drafts, this study treated the first, second, and third drafts as analytically distinct stages to enable systematic comparison of tool-use patterns over time. This approach does not assume that each draft corresponds to a single cognitive phase; rather, it reflects the task's temporal structure and the observable behaviors captured in the screen recordings.

Coding tool-use purposes related to each draft allowed the analysis to trace how students' engagement with AI-mediated writing tools evolved across revisions while acknowledging the recursive nature of composing.

All screen-recorded tool-use events were timestamped in milliseconds, converted to minutes for readability, and organized chronologically to examine switching patterns, frequency, and duration. Quantitative indicators, such as total time and number of tool-use events, were cross-analyzed with qualitative observations of writing behaviors. Patterns were compared across the three drafts and between higher- and lower-proficiency writers to identify individual differences in AI-mediated writing processes.

To ensure coding reliability, intra-coder agreement was assessed over two rounds of coding six months apart. Ten percent of the dataset was recoded in the second round, and agreement was computed using Cohen's Kappa (κ = 0.82), indicating strong consistency across rounds.

The coded actions were transformed into transition maps using Python scripts to visualize the recursive writing paths. In these maps, nodes represent a specific tool and purpose across Drafts 1–3, with their size reflecting total engagement time, while edges indicate chronological transitions, with thickness representing switching frequency.

4. Results

4.1. Overview of AI-mediated tool-use duration across participants

To provide an initial cross-case view of how students engaged with AI-mediated writing tools during writing, the total duration of online dictionary use, machine translation, and ChatGPT activity was calculated for each participant across the three drafts. This overview captures broad differences in the amount and distribution of tool use before turning to the individual case analyses in Section 4.2.

The total time participants spent using AI-mediated writing tools across the three drafts showed notable individual variation, as illustrated in Table 4 and Figure 1. Jiwan, the most advanced participant, and Hankyu, the lower-intermediate participant, used tools briefly (approx. 1.4% of writing time), whereas Suha and Minjung spent significantly more time (7.86% and 9.63% respectively). These profiles highlight individualized approaches: Suha relied heavily on ChatGPT, while

Minjung showed the strongest dependence on dictionaries. Crucially, these differences indicate that tool-use duration and distribution varied by individual writer rather than English writing proficiency level.

Table 4. Total duration and percentage of AI-mediated tool use by participant

Participant (Proficiency)	GPT (%)	OD (%)	MT (%)	Total min. (%)
Jiwan (Adv.)	0.67 (0.32)	2.25 (1.07)	0.00 (0.0)	2.92 (1.39)
Suha (Up-int.)	13.28 (6.33)	2.35 (1.12)	0.87 (0.41)	16.50 (7.86)
Minjung (Int.)	6.92 (3.29)	13.17 (6.27)	0.13 (0.06)	20.22 (9.63)
Hankyu (Low-int.)	0.00 (0.00)	2.77 (1.32)	0.20 (0.10)	2.97 (1.4)

Note. GPT = ChatGPT; OD = online dictionaries; MT = machine translation tools. Percentages indicate each tool's share of the total writing time (210 min) across three drafts.

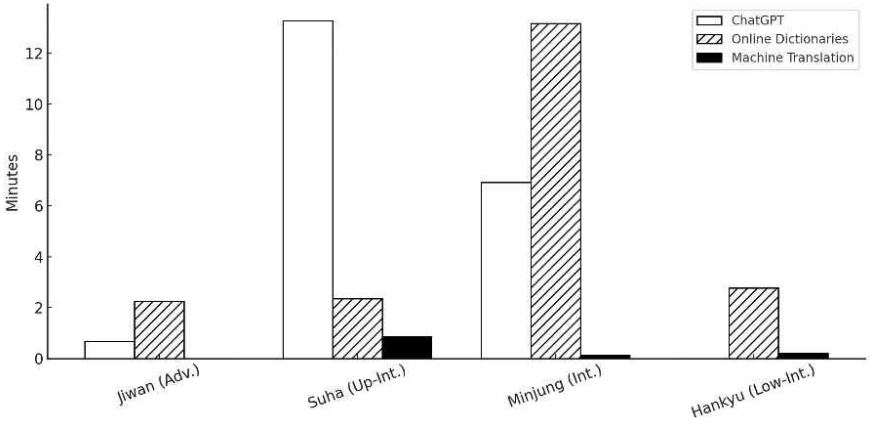

Figure 1. Distribution of AI-mediated tool-use time across participants.

The data in Table 5 are illustrated in Figure 2 to better visualize individual variations in each participant's time allocation across AI-mediated tool purposes. Jiwan, the most advanced participant, used AI-mediated writing tools mainly for translation with little engagement in other functions. In contrast, Suha at the upper-intermediate level used the tools for a wider range of purposes, including extensive grammar and word-search support and substantial source searching. Minjung, the intermediate participant, also demonstrated a varied pattern with considerable time devoted to translation, source searching, and text-level polishing

through mechanics and formatting. Hankyu, the lower-intermediate participant, engaged with tools only briefly, with relatively small amounts of time spent on translation and grammar support. Overall, the figure shows that Suha and Minjung used AI-mediated writing tools in more diverse and sustained ways, reflecting broader functional engagement than the more restricted patterns observed for Jiwan and Hankyu. These findings underscore that tool-use purposes reflected personalized writing strategies rather than group-level commonalities.

Table 5. Duration and percentage of AI-mediated tool-use purposes by participant

Participant (Proficiency)	TRANS	SS	GC/WS	MCH/FM	Total min. (%)
Jiwan (Adv.)	1.88	0.00	0.37	0.67	2.92
	(0.90)	(0.00)	(0.17)	(0.32)	(1.39)
Suha (Up-int.)	3.82	3.68	8.23	0.77	16.50
	(1.82)	(1.75)	(3.92)	(0.37)	(7.86)
Minjung (Int.)	5.80	1.42	9.52	3.48	20.22
	(2.76)	(0.68)	(4.53)	(1.66)	(9.63)
Hankyu (Low-int.)	2.43	0.00	0.53	0.00	2.97
	(1.16)	(0.00)	(0.25)	(0.00)	(1.41)

Note. TRANS = translation; SS = search for sources; GC/WS = grammar check/word search; MCH/FM = mechanics/formatting. Percentages represent the share of total writing time (210 min) spent on each tool-use purpose.

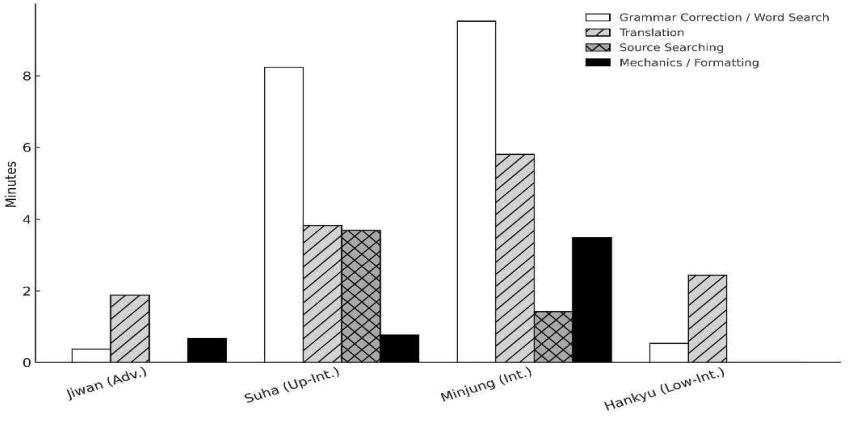

Figure 2. Purposes of AI-mediatead tool use by participants.

Results reveal clear individual variation in tool use as participants exhibited unique technological fingerprints regardless of proficiency level. Suha and Minjung demonstrated sustained and multifunctional engagement while Jiwan and Hankyu used tools briefly for limited purposes. The quantitative variations in engagement duration and the distribution of tool-use purposes define these four distinct profiles, which serve as the empirical basis for the in-depth case studies in Section 4.2.

4.2. Individual patterns of AI-mediated writing

While Section 4.1 summarized differences in the overall duration and distribution of AI-mediated writing tool use, this section presents a case-by-case analysis of how each participant engaged with specific tool-use purposes across the three drafts. These analyses draw on transition maps that depict moment-by-moment shifts between tool-purpose categories. In the maps, nodes represent individual tool-use instances categorized by draft (D1–D3), tool type (OD, MT, or GPT), and purpose. The numbers inside each node denote the chronological sequence, while line thickness corresponds to the frequency of transitions between these activities.

4.2.1. Jiwan (Advanced): Minimal and locally targeted consultations

Jiwan's writing process was defined by remarkably low tool-use duration, totaling only 2.92 minutes across all three drafts (Table 6). This minimal and locally targeted approach is further reflected in the sparse interactions shown in his transition map as shown in Figure 3. Most of Jiwan's cumulative tool-use duration occurred during Draft 2, where he utilized ChatGPT for formatting (0.67 min) alongside online dictionary-based translation (0.40 min). The map reveals a simple, bidirectional exchange between translation and word-search functions, with transitions occurring only once. While he briefly engaged with ChatGPT for formatting in the second draft, this did not trigger a wider sequence of tool engagement. By Draft 3, his activity was limited to a single, yet longest, instance of dictionary translation (1.15 min). This pattern suggests that Jiwan utilized tools only to resolve specific linguistic hurdles while maintaining extensive periods of independent drafting.

Table 6. Jiwan's chronological tool-use patterns and purposes across drafts

Draft	Tool	Purpose	Min. (%)
1	OD	GC/WS	0.17 (5.8)
	OD	TRANS	0.33 (11.3)
2	GPT	MCH/FM	0.67 (23.0)
	OD	TRANS	0.40 (13.7)
	OD	GC/WS	0.20 (6.8)
3	OD	TRANS	1.15 (39.4)
Total			2.92 (100)

Note. Tools: GPT = ChatGPT; OD = online dictionary. Purposes: GC/WS = grammar check/word search; MCH/FM = mechanics/formatting; SS = source search; TRANS = translation.

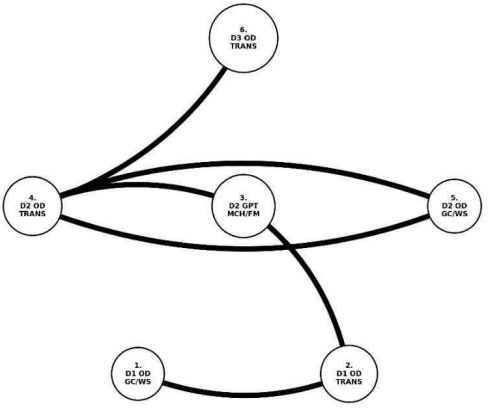

Figure 3. Jiwan's transition map of tool-use purposes and patterns across drafts.

Note. Node labels indicate sequence_draft_tool_purpose (e.g., 1_D1_OD_GC/WS); node size reflects total engagement time; edge thickness indicates switching frequency. **Drafts:** D1 = Draft 1; D2 = Draft 2; D3 = Draft 3. **Tools:** GPT = ChatGPT; OD = online dictionary. **Purposes:** GC/WS = grammar check/word search; MCH/FM = mechanics/formatting; SS = source search; TRANS = translation.

4.2.2. Suha (upper-intermediate): Purpose-driven multi-functional engagement

The integration of generative AI serves as the defining feature of Suha's writing

trajectory, evolving significantly as she progressed through the drafting stages (see Table 7). This expansion of tool functionality is most evident in her transition map (see Figure 4), where the linear patterns of her initial draft give way to more complex interactions. While her early process in Draft 1 remained anchored in traditional dictionary and translation tools (18.9% combined), her strategy shifted toward a recursive and highly deliberate approach in Draft 2. Within this middle stage, which accounted for 60.3% of her total 16.50-minute engagement, ChatGPT functioned as a dynamic workspace for both linguistic refinement and source evaluation. The back-and-forth transitions between grammar checks (GPT_GC, GPT_GC/WS) and source searching (GPT_SS) illustrate a highly recursive process; specifically, nearly half of her total engagement (8.23 minutes) was dedicated exclusively to these grammar and word search functions. When combined with source searching (GPT_SS), these recursive functions highlight her ability to bridge the gap between evidence-seeking and refining expression. As she reached the final polishing stage in Draft 3, her workflow streamlined once more, focusing primarily on source verification and mechanical checks (16.6% combined) to finalize her work.

Table 7. Suha's chronological tool-use patterns and purposes across drafts

Draft	Tool	Purpose	Min. (%)
1	OD	TRANS	2.35 (14.2)
	MT	TRANS	0.78 (4.7)
2	GPT	GC	3.75 (22.7)
	GPT	GC/WS	4.48 (27.2)
	GPT	SS	1.71 (10.4)
3	GPT	TRANS	0.60 (3.6)
	GPT	MCH	0.77 (4.7)
	GPT	SS	1.97 (11.9)
	MT	TRANS	0.09 (0.5)
Total			16.50 (100)

Note. **Tools:** GPT = ChatGPT; MT = machine translation; OD = online dictionary. **Purposes:** GC/WS = grammar chec/word search; MCH = mechanics; SS = source search; TRANS = translation.

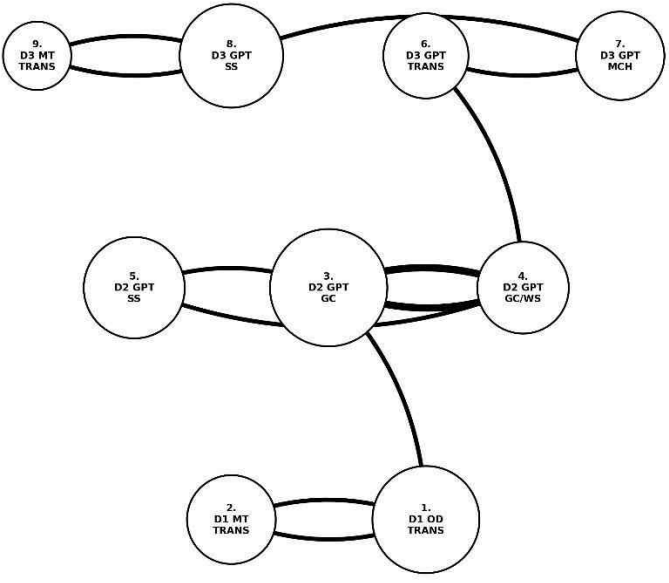

Figure 4. Suha's transition map of tool-use purposes and patterns across drafts.

Note. Node labels indicate sequence_draft_tool_purpose (e.g., 1_D1_OD_TRANS); node size reflects total engagement time; edge thickness indicates switching frequency. **Tools:** GPT = ChatGPT; MT = machine translation; OD = online dictionary. **Purposes:** GC/WS = grammar chec/word search; MCH = mechanics; SS = source search; TRANS = translation.

4.2.3. Minjung (intermediate): Dense, recurring cycles of refinement

Of all the participants, Minjung's writing process is distinguished by high-density interaction and recurring loops of tool engagement, totaling 20.22 minutes (see Table 8). This multi-layered complexity is most evident in her transition map (see Figure 5), which displays the most networked pattern among all participants. While the first two drafts relied on concentrated online dictionary use (OD_TRANS and OD_GC/WS; 62.1% combined), Draft 3 evolved into a dense 7.67-minute web of constant cycling. The map highlights this intensity through high-frequency bidirectional transitions between translation (OD_TRANS), word-level refinement (GPT_GC/WS), mechanics and formatting (GPT_MCH/FM), and source searching (GPT_SS). Rather than using tools in isolated bursts, Minjung utilized the AI-mediated environment as a continuous linguistic scaffold, deeply integrating these resources into her cognitive drafting for iterative layers of refinement.

Table 8. Minjung's chronological tool-use patterns and purposes across drafts

Draft	Tool	Purpose	Min. (%)
1	OD	TRANS	2.13 (10.6)
	OD	GC/WS	5.60 (27.7)
2	OD	GC/WS	1.91 (9.4)
	OD	TRANS	2.91(14.4)
3	MT	TRANS	0.13 (0.7)
	OD	TRANS	0.62 (3.1)
	GPT	GC	0.35 (1.7)
	GPT	MCH/FM	3.48 (17.2)
	GPT	GC/WS	1.66 (8.2)
	GPT	SS	1.43 (7.1)
Total			20.22 (100)

Notes. **Tools:** GPT = ChatGPT; MT = machine translation; OD = online dictionary. **Purposes:** GC/WS = grammar check/word search; MCH/FM = mechanics/formatting; SS = source search; TRANS = translation.

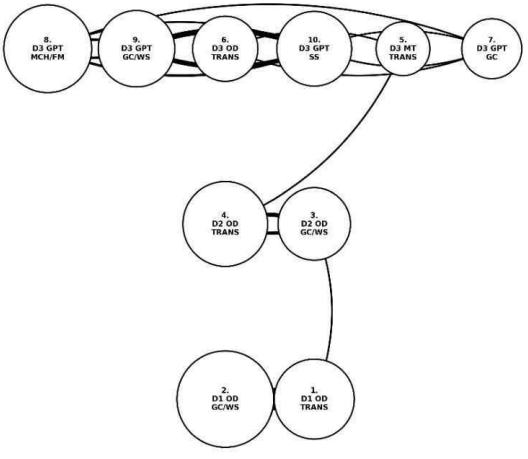

Figure 5. Minjung's AI-mediated tool-use transition map and functional sequences.

Note. Node labels indicate sequence_draft_tool_purpose (e.g., 1_D1_OD_TRANS); node size reflects total engagement time; edge thickness indicates switching frequency. GPT = ChatGPT; MT = machine translation; OD = online dictionary. GC/WS = grammar check/word search; MCH/FM = mechanics/formatting; SS = source search; TRANS = translation.

4.2.4. Hankyu (lower-intermediate): Sparse and linear tool-use sequences

Hankyu's engagement involved interactions with digital tools totaling 2.97 minutes across the writing process (Table 9). As illustrated in Figure 6, the writing process followed a linear structure, with engagement restricted almost exclusively to basic translation and word-search functions. The tool-use activity was concentrated primarily in the initial drafting phase (Draft 1), where dictionary-based translation (OD_TRANS) and word-searches (OD_GC/WS) accounted for 85.5% of the total tool time. Within this initial stage, the map reveals a small degree of repetitive movement between these two functions though this did not develop into more complex, multi-purpose sequences. Across subsequent drafts, transitions remained infrequent and unidirectional, with no evidence of connecting translation work to higher-order functions like source searching. This pattern indicates that the tools were utilized as discrete aids for linguistic retrieval rather than as a comprehensive workspace for iterative revision or content generation.

Table 9. Hankyu's chronological tool-use patterns and purposes across drafts

Draft	Tool	Purpose	Min. (%)
1	OD	TRANS	2.00 (67.4)
2	OD	GC/WS	0.53 (18.0)
	MT	TRANS	0.13 (4.5)
3	OD	TRANS	0.23 (7.8)
	MT	TRANS	0.07 (2.3)
Total			2.97 (100)

Notes. **Tools:** MT = machine translation; OD = online dictionary. **Purposes:** GC/WS = grammar check/word search; TRANS = translation.

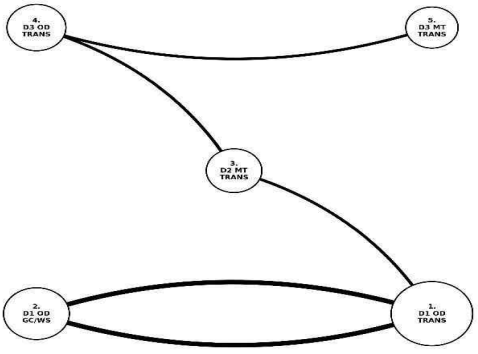

Figure 6. Hankyu's AI-mediated tool-use transition map and functional sequences.

Note. Node labels indicate sequence_draft_tool_purpose (e.g., 1_D1_OD_TRANS); node size reflects total engagement time; edge thickness indicates switching frequency. **Tools:** MT = machine translation; OD = online dictionary. **Purposes:** GC/WS = grammar check/word search; TRANS = translation.

Across the four cases, the transition maps illustrate individualized and non-linear trajectories, highlighting substantial cross-case variation in both the duration and manner of AI tool use. While participants like Jiwan and Hankyu exhibited short, isolated episodes of engagement, others like Suha and Minjung demonstrated sustained and more complex movement across functions. These patterns ranged from a reliance on a single tool to more diversified interactions across ChatGPT, online dictionaries, and machine translation for purposes such as grammar support, translation, and text-level refinement. Ultimately, these idiosyncratic representations underscore the highly individualized nature of AI-mediated writing processes, providing a descriptive foundation for the interpretive discussion that follows.

5. Discussion

5.1. Idiosyncratic tool-use trajectories and individualized affordances

One of the most salient observations across the four cases was the high degree of individuality in tool-use routines. Contrary to the assumption that technological sophistication might scale with linguistic ability, neither the total duration of tool use nor the complexity of transition patterns increased consistently with English

writing proficiency. This highlights that digital writing behaviors are highly individualized, shaped by personal needs and preferences. Such variation aligns with prior research suggesting that technology use is influenced more by individual and contextual factors than by proficiency alone (Musk, 2022; Wang & Wang, 2025).

In this study, while ChatGPT provided various functional possibilities, the actual application of these tools varied significantly among individuals. This variability aligns with research suggesting that strategy use stems from individualized trajectories rather than group-level commonalities (Yoon, 2016). Such findings reflect the complex, dynamic, and idiosyncratic nature of tool-use behaviors observed in L2 writing (Sung & Chang, 2024).

5.2. Task constraints and the narrowed scope of affordance utility

Across all cases, participants engaged with ChatGPT in a very controlled and narrow way, focusing almost exclusively on grammar and vocabulary. While other studies have seen EFL students use AI successfully for brainstorming and generating ideas (Hwang et al., 2025; Lee et al., 2025; Su et al., 2023), the cautious approach in this study was largely a result of the guidelines on generative AI use. Because the students were specifically instructed to avoid using AI for generating content or organizing their ideas, the instructional environment restricted the perceived utility of the tool.

By restricting higher-order tasks to prevent over-reliance (Barrot, 2023; Yan, 2023), the framework effectively capped the students' ability to use the tool's full potential. This created a significant gap between what generative AI is capable of and how it was actually used in the classroom (Warschauer et al., 2023). Regardless of their English writing proficiency, students viewed the AI through a lens of functional fixedness, treating it like a digital dictionary or editor rather than a collaborative partner. This hesitation also reflects a desire to protect their own authorial voice and avoid textual homogenization (Sun, 2026). Students often see deep AI involvement as a risk to their own agency, choosing safe and linear routines to maintain ownership of their work. When the primary focus in an academic setting is on preventing overuse, AI application stays tethered to these traditional routines, which can accidentally prevent students from developing the strategic skills they need for digital writing (Yan, 2023).

5.3. Pedagogical implications for AI literacy

The results of this study underscore that strategic competence in AI-mediated writing does not develop automatically as a byproduct of writing proficiency. This shows that we need a new kind of instruction that moves beyond just setting rules and instead helps students discover more effective ways to learn. Educators should model how AI can handle surface-level tasks, which helps reduce the cognitive load on the student and allows them to focus more on their actual meaning and arguments (Lee et al., 2025; Su et al., 2023). Such an ecological approach requires teachers to proactively foster learner agency by helping students perceive and realize the diverse affordances available in the digital writing environment (Liu & Chao, 2018).

Since students who receive explicit training on digital tools tend to produce deeper and more varied revisions (Lee & Oh, 2025), teachers should show students how to use these tools critically throughout the entire writing process. Such instruction should be grounded in the development of AI literacy, which requires learners to not only use the tools but also to critically evaluate AI-generated outputs and navigate the ethical complexities of AI integration. This includes using scaffolded activities that encourage students to revisit and improve their content at different stages, while also reflecting on why they chose to use the tool in a certain way. As these choices are often invisible in a final draft, tasks such as reflective notes or peer discussions can help make these strategies more clear. Teachers can also demonstrate how to double-check AI output and how to rewrite AI suggestions so the student's own voice remains clear. Ultimately, institutions should move toward providing clear guidelines that explain how to use AI responsibly, such as for grammar support or brainstorming with proper credit. This approach helps students grow from passive users into reflective writers who can use technology to support, rather than replace, their academic development (Liu & Chao, 2018; Wang & Wang, 2025; Warschauer et al., 2023).

5.4. Limitations and directions for future research

While this study provides detailed insights into the chronological and functional patterns of AI-mediated tool use, several methodological limitations must be acknowledged. First, the findings are based on a small sample of four participants at a single university in Seoul. The findings prioritize in-depth exploration of specific digital behaviors over universal claims for all L2 writers.

The most significant methodological constraint is the absence of stimulated recall or interview data to triangulate with the behavioral logs. Although screen-recording data allowed for a detailed tracking of participant activity, these digital traces alone do not fully reveal the underlying cognitive rationales. Without qualitative retrospective data, the reasons behind specific strategic choices remain a matter of inference. It is difficult to determine with certainty whether the patterns observed across different levels of English writing proficiency were driven by strict adherence to task constraints, a lack of digital literacy, or varying levels of trust in AI-generated output.

To build on these findings, future research should incorporate triangulation methods such as think-aloud protocols, eye-tracking, or reflective notes. Such approaches would help bridge the gap between observed digital actions and the internal thoughts that drive students to switch between different tools. Furthermore, investigating writing tasks with fewer pedagogical restrictions might reveal a wider range of how students actually utilize AI when granted more agency in an authentic writing context.

6. Conclusion

This study explored how four Korean EFL university students engaged with online resources and AI-mediated writing tools across three drafts of an English argumentative essay. By analyzing screen-recorded writing processes, this research highlights how tool-use behaviors are profoundly individualized and non-linear, shaped more by idiosyncratic strategic choices and personal writing processes than by English writing proficiency alone. These findings demonstrate that strategic competence in AI-mediated writing is an emergent process rather than a static skill. By visualizing recursive cycles of refinement through moment-by-moment process tracing, this study provides a finer-grained view of the writer's cognitive moves that traditional surveys or final-product analyses often miss.

The results reinforce the need for a shift in pedagogy, moving beyond simple language instruction toward supporting AI literacy alongside traditional academic writing skills. While the small sample size and specific classroom setting mean these results are not universal, the use of screen recording successfully captured digital behaviors that are typically invisible in final texts. Future research should continue to explore AI use in more flexible, naturalistic writing contexts and across different genres. Methods such as stimulated recall, keystroke logging, or eye-tracking would further illuminate the internal decision-making that occurs when learners compose

with digital support.

Ultimately, effective AI-mediated writing depends not just on linguistic ability, but on a writer's capacity to make informed, reflective decisions about tool use. As generative AI becomes increasingly integral to academic settings, supporting learners in developing a thoughtful and ethical engagement with these tools is essential. Fostering this critical literacy will ensure that technology enhances student agency and independence, leading to sustained growth in L2 writing development.

References

Alzahrani, M. (2025). The impacts of mobile apps on English vocabulary development of Afghan EFL students. *Pamir International Journal of Social Science*, *1*(1), 1-13.

Baek, J. Y., & Rha, K. H. (2023). Investigating EFL college students' English writings produced by AI-based machine translation. *Foreign Languages Education*, *30*(2), 1-24.

Barrot, J. S. (2023). Using ChatGPT for second language writing: Pitfalls and potentials. *Assessing Writing*, *57*, Article 100745.

Chan, A. Y. W. (2012). Cantonese ESL learners' use of grammatical information in a monolingual dictionary for determining the correct use of a target word. *International Journal of Lexicography*, *25*(1), 68-94.

Chen, Y. (2022). The effect of learning conditions on collocation gains: A case study of task-based dictionary use instruction. *Lexikos*, *32*, 1-31.

Chung, E. S., & Ahn, S. (2022). The effect of using machine translation on linguistic features in L2 writing across proficiency levels and text genres. *Computer Assisted Language Learning*, *35*(9), 2239-2264.

Cohen, J. (1960). A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, *20*(1), 37-46.

Comelles, E., & Laso, N. J. (2025). The impact of MT as a writing tool on EFL academic writing: A qualitative linguistic analysis. *Journal of Second Language Writing*, *69*, 101231.

Dziemianko, A. (2010). Paper or electronic? The role of dictionary form in language reception, production and retention of meaning and collocations. *International Journal of Lexicography*, *23*(3), 257-273.

Frankenberg-Garcia, A. (2020). Combining user needs, lexicographic data and digital writing environments. *Language Teaching*, *53*(1), 29-43.

Gao, C. A., Howard, F. M., Markov, N. S., Dyer, E. C., Ramesh, S., Luo, Y., & Pearson, A. T. (2023). Comparing scientific abstracts generated by ChatGPT to real abstracts with detectors and blinded human reviewers. *NPJ Digital Medicine*, *6*(1), 75.

Gilquin, G., & Laporte, S. (2021). The use of online writing tools by learners of English:

Evience from a process corpus. *International Journal of Lexicography*, *34*(4), 472-492.

Hamel, M. J., Séror, J., & Dion, C. (2015). *Writers in action: Modelling and scaffolding second language learners' writing process*. Toronto, Canada: Higher Education Quality Council of Ontario.

Hwang, H., Chang, X., & Sun, J. (2025). Generative AI is useful for second language writing, but when, why, and for how long do learners use it? *Journal of Second Language Writing*, *69*, 101230.

Hyland, K. (2025). Tomorrowland? Technical innovation and feedback on writing. *RELC Journal*. Advance online publication.

Jacobs, H. L., Zinkgraf, S. A., Wormuth, D. R., Hartfiel, V. F., & Hughey, J. B. (1981). *Testing ESL composition: A practical approach*. Newbury House Publishers.

Jarrah, A. M., Wardat, Y., & Fidalgo, P. (2023). Using ChatGPT in academic writing is (not) a form of plagiarism: What does the literature say? *Online Journal of Communication and Media Technologies*, *13*(4), e202346.

Kohnke, L. (2024). Exploring EAP students' perceptions of GenAI and traditional grammar-checking tools for language learning. *Computers and Education: Artificial Intelligence*, *7*, 100279.

Kohnke, L., Moorhouse, B. L., & Zou, D. (2023). ChatGPT for language teaching and learning. *RELC Journal*, *54*(2), 537-550.

Koltovskaia, S., Rahmati, P., & Saeli, H. (2024). Graduate students' use of ChatGPT for academic text revision: Behavioral, cognitive, and affective engagement. *Journal of Second Language Writing*, *65*, 101130.

Lee, H., & Oh, S.-Y. (2025). Effects of Google Translate instruction on young learners' L2 writing revisions. *The Journal of Asia TEFL*, *22*(1), 41-64.

Lee, Y.-J., Davis, R. O., & Choi, J.-I. (2025). Integrating generative AI into EFL writing: University students' strategies and perceptions. *Online Journal of Communication and Media Technologies*, *15*(4), Article e202541.

Lew, R. (2010). Multimodal lexicography: The representation of meaning in electronic dictionaries. *Lexikos*, *20*(1), 23-45.

Lew, R., & de Schryver, G.-M. (2014). Dictionary users in the digital revolution. *International Journal of Lexicography*, *27*(4), 341-359.

Liu, Q., & Chao, C. C. (2018). CALL from an ecological perspective: How a teacher perceives affordance and fosters learner agency in a technology-mediated language classroom. *ReCALL*, *30*(1), 68-87.

Mroz, A. P. (2025). Process research screen capture. In C. A. Chapelle (Ed.), *The encyclopedia of applied linguistics*. Wiley-Blackwell.

Musk, N. (2022). Using online translation tools in computer-assisted collaborative EFL writing. *Classroom Discourse*, *13*(2), 119-144.

Nguyen, L. Q., Le, H. V., & Nguyen, P. T. (2025). A mixed-methods study on the use of ChatGPT in the pre-writing stage: EFL learners' utilization patterns, affective engagement,

and writing performance. *Education and Information Technologies*, *30*, 10511-10534.
Patton, M. Q. (2015). *Qualitative research & evaluation methods: Integrating theory and practice* (4th ed.). Sage.
Shin, D., Kwon, S. K., & Lee, Y. (2021). The effect of using online language-support resources on L2 writing performance. *Language Testing in Asia*, *11*(4), 1-23.
Song, C., & Song, Y. (2023). Enhancing academic writing skills and motivation: Assessing the efficacy of ChatGPT in AI-assisted language learning for EFL students. *Frontiers in Psychology*, *14*, 1260843.
Su, Y., Lin, Y., & Lai, C. (2023). Collaborating with ChatGPT in argumentative writing classrooms. *Assessing Writing*, *57*, 100752.
Sun, Y. (2026). Resisting homogenization in EAL writing education: Translingual practices in the rise of generative AI. *TESOL Quarterly*.
Sung, S., & Jang, I. C. (2024). South Korean STEM graduate students' use of ChatGPT in self-initiated L2 writing: A process-tracing study. *Korean Journal of English Language and Linguistics*, *24*, 1415-1435.
Wang, C., & Wang, Z. (2025). Investigating L2 writers' critical AI literacy in AI-assisted writing: An APSE model. *Journal of Second Language Writing, 67*, 101187.
Warschauer, M., Tseng, W., Yim, S., Webster, T., Jacob, S., Du, Q., & Tate, T. (2023). The affordances and contradictions of AI-generated text for writers of English as a second or foreign language. *Journal of Second Language Writing*, *62*, 101071.
Yan, D. (2023). Impact of ChatGPT on learners in a L2 writing practicum: An exploratory investigation. *Education and Information Technologies*, *28*(11), 13943-13967.
Yoon, C. (2016). Individual differences in online reference resource consultation: Case studies of Korean ESL graduate writers. *Journal of Second Language Writing, 32*, 67-80.

Janet Jaymin Kim
Visiting Professor
Department of English Language Education
Seoul National University
1 Gwanak-ro, Gwanak-gu, Seoul 08826, Korea
E-mail: janetkimatsnu@gmail.com

Received: November 14, 2025
Revised version received: January 31, 2026
Accepted: March 24, 2026