



ARCHIE: Exploring Language Learner Behaviors in LLM Chatbot-Supported Active Reading Log Data with Epistemic Network Analysis

Steve Woollaston*

Kyoto University
Kyoto, Japan
s.m.woollaston@gmail.com

Brendan Flanagan

Kyoto University
Kyoto, Japan
flanagan.brendanjohn.4n@kyoto-u.ac.jp

Patrick Ocheja

Kyoto University
Kyoto, Japan
ocheja.patrick@gmail.com

Yuko Toyokawa

Kyoto University
Kyoto, Japan
toyokawa.yuko.2j@kyoto-u.ac.jp

Hiroaki Ogata

Kyoto University
Kyoto, Japan
ogata.hiroaki.3e@kyoto-u.ac.jp

Abstract

With the increasing integration of technology in education, chatbots and e-readers have emerged as promising tools for enhancing language learning experiences. This study investigates how students engage with digital texts and a purpose-built chatbot designed to promote active reading for EFL students. We analysed student interactions and compared high-proficiency and low-proficiency English learners. Results indicate that while all students perceived the chatbot as easy to use, useful, and enjoyable, significant behavioural differences emerged between proficiency groups. High-proficiency students exhibited more frequent interactions with the chatbot, engaged in more active reading strategies like backtracking, and demonstrated less help seeking behaviours. Epistemic Network Analysis revealed distinct co-occurrence patterns, highlighting the stronger connection between navigation and review behaviours in the high-proficiency group. These findings underscore the potential of chatbot-assisted language learning and emphasise the importance of incorporating active reading strategies for improved comprehension.

CCS Concepts

• **Applied computing** → **Interactive learning environments; Computer-assisted instruction; Learning management systems.**

Keywords

chatbot, e-readers, EFL, active reading, epistemic network analysis (ENA), log data, backtracking

ACM Reference Format:

Steve Woollaston, Brendan Flanagan, Patrick Ocheja, Yuko Toyokawa, and Hiroaki Ogata. 2025. ARCHIE: Exploring Language Learner Behaviors in LLM Chatbot-Supported Active Reading Log Data with Epistemic Network Analysis. In *LAK25: The 15th International Learning Analytics and Knowledge Conference (LAK 2025)*, March 03–07, 2025, Dublin, Ireland. ACM, New York, NY, USA, 13 pages. <https://doi.org/10.1145/3706468.3706556>

1 Introduction

Reading plays a fundamental role in second language acquisition, serving as a primary means through which learners encounter new vocabulary, grammar structures, and language patterns in context. Exposure to authentic language through reading helps build vocabulary, improve reading comprehension skills, and enhance overall language proficiency [7]. Engagement with a variety of meaning-focused input at the appropriate level is an important part of second language learning [19]. Nation [19] highlights the crucial role of comprehensible input, where learners are exposed to language they can understand, even if it contains some unknown elements, allowing for incidental vocabulary acquisition and language development. Both intensive and extensive reading are crucial for language development. However, selecting materials that optimally match each student's reading level can be a major challenge for teachers. Additionally, the level of support each student needs to understand a text can vary significantly, requiring different degrees of scaffolding, which adds to the workload of already busy teachers [5].

The utilisation of LLM-powered chatbots in language learning has opened up exciting new possibilities for personalised and engaging learning experiences. Complementing traditional classroom settings, chatbots can offer personalised instruction tailored to individual needs, provide unlimited, tireless, and on demand access to interactive learning, and create a private and non-judgemental learning environment [14]. Despite their limitations, namely the potential to generate incorrect information (hallucination), LLM-powered chatbots have remarkably improved the state of the art in chatbot technology, offering more nuanced human-like understanding and response generation [29]. Furthermore, e-reading software has enabled the easy and efficient distribution of texts to students.



This work is licensed under a [Creative Commons Attribution International 4.0 License](https://creativecommons.org/licenses/by/4.0/).

LAK 2025, Dublin, Ireland

© 2025 Copyright held by the owner/author(s).

ACM ISBN 979-8-4007-0701-8/25/03

<https://doi.org/10.1145/3706468.3706556>

Not only does this software allow learners to engage with texts in a more dynamic way, offering features such as annotation, highlighting, and vocabulary support, but it can also provide rich data to teachers and researchers about student reading activity, such as reading time at the page level, navigation behaviour, written memos, and highlighted content. The granularity offered by these systems allows analysis of reading not previously possible [24].

This study advances educational chatbot research by investigating an LLM-based chatbot, ARCHIE (Active Reading CHatbot for Interactive Engagement). Uniquely, we combine chat and reading log data for a more comprehensive understanding of learner engagement with the text and the chatbot. Archie was designed to provide personalised, one-on-one reading support, reducing teacher workload and promoting active reading. By capturing learner perceptions via the exTAM we assess the tool's usability, usefulness, and engagement potential, crucial factors in language acquisition. This research explores learner interactions with Archie and the e-reading software Bookroll, comparing the behavioural patterns of different levels of English proficiency using epistemic network analysis (ENA). These insights can help refine tools like Archie and inform both classroom practice and future research in second-language acquisition by highlighting distinct engagement patterns across proficiency levels.

Research questions:

- (1) How do learners interact with a chatbot for supporting active reading and a digital reading system together?
- (2) What are the learners' experiences and attitudes towards a chatbot for supporting active reading, and is there a difference between high and low English proficiency learners?
- (3) How do high-proficiency English learners' chat and reading behaviours compare with those of low-proficiency learners?

2 Related Work

2.1 Active Reading

Active reading, characterised by deeper engagement, questioning, review, predicting, and making connections, is an important part of language learning. While there is no singular, universally agreed-upon definition of active reading, numerous studies have shown a positive relationship between the use of these strategies (e.g., SQ3R, a reading method that involves Surveying, Questioning, Reading, Reciting, and Reviewing) and reading comprehension, vocabulary acquisition, and critical thinking [1, 4]. Technology can facilitate and promote these behaviours. For example, enabling annotation and highlighting in e-readers encourages active note taking and making connections with parts of the text relevant to the reader.

Backtracking in particular has been shown to be particularly important for reading comprehension. Backtracking involves intentionally rereading and reviewing previously encountered sections of the text [20]. In digital reading environments, such as on computers, tablets, and smartphones, backtracking often involves scrolling or paging backwards to check information or clarify understanding. Haverkamp and Bråten [13] found that strategic backtracking during reading was a significant predictor of text understanding among undergraduate students, even when controlling for factors such as working memory and reading comprehension skills. Similarly,

this time in English as a Foreign Language (EFL) learners, Yin et al. [28] analysed ebook reading log data and found that students who frequently return to previous pages tended to have better learning outcomes.

2.2 Chatbots and Reading

Chatbots have gained significant attention in the field of education and second language learning; even more so since the introduction of LLMs [26]. Recent research has explored the application of chatbots in supporting reading engagement and comprehension in EFL learners. Behforouz and Al Ghaithi [2] found that an interactive WhatsApp bot which provided instructions and exercises to Omani students led to significant improvements in their reading skills. Another study compared children's reading comprehension in human-to-human and chatbot-assisted conversations, finding no significant difference in comprehension scores between the two groups, suggesting that chatbots can be as effective as human teachers in guiding reading activities and promoting reading comprehension [15]. Research indicates that chatbots designed for reading support can effectively reduce foreign language reading anxiety among high school students [30]. This reduction in anxiety can foster a positive feedback loop, encouraging further reading and language acquisition. These studies highlight the potential of chatbots to enhance traditional instruction by offering additional interactive exposure to the target language and creating opportunities for practice.

2.3 Log Data and ENA

Automatically collected log data from digital tools and platforms, such as e-readers, offer valuable insights into learners' reading processes and their interaction with texts. Navigation patterns, annotations, and highlights provide a detailed view of how learners engage with a given text. This behavioural data can be analysed to better understand the learners' level of engagement with the material [12]. Chatbot logs, including messages sent and received, can also provide insight into the specific areas that learners may struggle with or seek clarification on. Analysis of the interactions with the chatbot can reveal the learning behaviours and perhaps shed light on cognitive processes [22].

Epistemic Network Analysis (ENA) is an increasingly popular technique used to analyse log data with the aim of identifying patterns in learning behaviour. ENA is a technique that uses network visualisation, where nodes can represent behaviours and the edges indicate the relative connections between them. This approach enables researchers to see which behaviours co-occur more often, identify clusters of related behaviours, and compare learner and average group behaviour. Furthermore, each individual log which contributed to these connections can be interrogated, closing the 'interpretive loop' and ensuring the analysis and subsequent interpretation is grounded and meaningful [23]. In a study investigating learning behaviours on a video learning platform, Mohammadhasan and Mitrovic [18] created networks to visualise and compare the connections between different actions (e.g. watching videos, pausing, making comments, and receiving nudges to make comments etc.). Their findings revealed distinct patterns between learner types

(classified by the ICAP framework which categorises learning activities as Interactive, Constructive, Active, or Passive based on cognitive engagement levels) [6], and between early starters (interaction throughout the summer) and late starters (interaction closer to deadlines) on the platform. For example, Constructive learners were found to engage in more reflective commenting without the need for reminders, whereas Active and Passive learners relied more heavily on nudges. Early adopters demonstrated a more focused approach, engaging in one task at a time, while late starters exhibited more frequent task switching. These insights, derived from ENA, highlight its potential to uncover nuanced behavioural patterns and inform the design of more personalised and effective learning tools.

This study examines how learners engage with digital texts and a purpose-built chatbot designed to promote active reading for English language learning. It combines Bookroll, an e-reader software providing data on learner reading behaviour, and Archie, a chatbot fostering deeper engagement with texts through a variety of reading activities. By merging the log data from both datasets, we hope to contribute to the current understanding of how learners interact with digital texts and chatbots.

3 Methodology

3.1 Participants

The sample consisted of eight first year students (age 15-16) classes (250 total) from a public Japanese high school known for its academic performance. The reading task and interaction with Archie was assigned as summer vacation homework. Three students opted out of their data being used and were excluded from the analysis.

3.2 Materials

18 short informational texts (e.g. The History and Significance of Vending Machines), each approximately two pages long (ranging from 412 - 698 words; 575 words on average). With permission, texts were sourced directly from past practice papers of the Center Test for Japanese University Admissions, a standardised exam widely used by universities to evaluate students' English proficiency. These texts are rated CEFR A1-B1 Level. Each text also had several multiple choice comprehension questions about its content on the final page. These were fed to the LLM as exemplar questions. Students could also answer these questions in Bookroll, but this data is not utilised in the current study.

3.3 System Overview

3.3.1 Archie Chatbot. Archie is a simple chatbot (Figure 1) designed to facilitate active reading and engagement with preloaded texts. The chatbot was developed using Gradio and was powered by Gemini-1.5-Pro-001. For safe use, strict filtering of the LLM's inputs and outputs was enforced. All interactions with Archie were recorded for later analysis in a database.

Archie provides the following features:

- **Language selection:** Converse with Archie in English or Japanese.
- **Text selection:** Choose from the 18 preloaded texts
- **Activity selection:** 12 activities available (described below)

- **Retry:** Where the response is unsatisfactory, learners can ask the chatbot to generate a new answer
- **Undo:** Allows learners to delete the chatbot's response and their message to retry that part of the conversation
- **Examples:** Common questions and commands are available to facilitate ease of use
- **Like / dislike AI response:** Learners can indicate whether they like or dislike the chatbot's responses
- **Copy button:** Allows learners to copy their messages or the chatbot's responses.

The system prompt (see Appendix) for Archie defines how the chatbot interacts with learners. It greets them, introduces an activity, and follows predefined instructions for various learning tasks (e.g., cloze passages, comprehension questions, grammar exercises). The system prompt instructs the LLM to provide concise, personalised feedback, guiding learners without giving direct answers. It is asked to adjust language complexity to the learner's proficiency and keep them focused on the task. The chatbot uses excerpts from the text, encourages rereading for better understanding, and helps learners with hints if they struggle. Twelve reading activities were provided to facilitate active reading (see Appendix for activity prompts):

- (1) **Cloze Passage:** practices vocabulary and contextual understanding
- (2) **Comprehension Questions:** evaluates grasp of text content
- (3) **Creative Writing:** open-ended activity to foster writing skills and creativity with immediate feedback
- (4) **Discussion:** open-ended dialogue on any aspects or themes from the text
- (5) **Fact vs. Opinion:** develops analytical skills and critical thinking with excerpts from the text
- (6) **Grammar Focus:** explicit reference to grammar found in the texts and how it works
- (7) **Role Play:** learner and chatbot take on personas from the text to practise dialogue
- (8) **Sentence Reordering:** encourages deeper understanding of the text through reorganization
- (9) **Sentence Translation:** develops bilingual proficiency and vocabulary
- (10) **Summarisation:** practices ability to synthesise information concisely
- (11) **True or False:** assesses identification of key information from the text
- (12) **Vocabulary Matching:** practice and expand vocabulary knowledge.

3.3.2 Bookroll. Bookroll is an interactive reading application for learners to read texts on device, create memos, highlight text, create bookmarks, and answer quiz questions [10]. Teachers are able to set texts and observe learner activity. All activity in the system is recorded as log data which can be used to analyse learner engagement with the reading material. For this study, only data related to access, navigation, highlighting (marker), and memo creation were utilised.

Figure 1: Archie chatbot interface

3.4 Procedure

3.4.1 Induction to Bookroll and Archie. Students were already familiar with the e-reading software Bookroll, having used it previously for other tasks. Archie was introduced by their teachers when the summer vacation homework was assigned in late July. Students were instructed to read a minimum of six (of the 18) texts, trying each reading activity in Archie at least once.

3.4.2 Post-survey. A post-survey was made available to students immediately after the experiment ended in late August. It included the extended Technology Acceptance Model (exTAM), questions about which activities they found useful and enjoyable, and open-ended questions about the strengths of Archie and potential improvements for the chatbot. The exTAM [27], adapted for chatbot evaluation, is a well-established model for assessing affective factors that influence learners' attitudes and behavioural intentions toward using interactive learning technologies [11]. The questionnaire contains 17 questions where respondents rated their agreement with statements on a seven-point Likert scale, ranging from *Strongly Disagree* to *Strongly Agree*. Five factors are calculated: **Perceived Ease of Use** (PEOU), **Perceived Usefulness** (PU), **Perceived Enjoyment** (PE), **Attitude** (ATT), and **Intention to use** (INT).

3.5 Data Preparation

3.5.1 English Proficiency. Three English test scores were used to establish a baseline of English proficiency. Two of these were 50-minute end-of-term English exams created and administered by the school on June 6 and June 10, while the third was a 10-minute informal pretest created by the researchers and conducted on July 31. The pretest included 36 multiple-choice questions, with one grammar and one vocabulary question per text. The school-based exams were comprehensive assessments targeting intermediate English learners and covered a wide range of skills, including grammar, vocabulary, phonetics, reading, and listening comprehension, using a mix of multiple-choice questions, sentence restructuring, and phonetic transcription tasks. Although these tests effectively assess English accuracy and identify areas for improvement, they do not measure communicative fluency or real-world language use. The three scores were averaged for each student, and students were divided into two groups—low-proficiency and high-proficiency—split at the median.

3.5.2 Reading Behaviour Data Collection. The raw Bookroll operation data extracted from the database (between July 31st to August 23rd) were categorised into meaningful groups based on their function (shown in Table 1). Navigation actions NEXT and PREV for sequential page turns were given their own category as they were so numerous. otherNavigation was used for actions like opening a text (OPEN) or jumping between pages via thumbnails

(PAGE_JUMP). Actions related to the memo feature were grouped under MEMO, encompassing creating, deleting, editing, and navigating within memos (e.g., ADD_MEMO, MEMO_JUMP). The ADD_MARKER action was further differentiated based on an associated colour code within the data: red markers were classified as markImportant, and yellow markers as markDifficult (consistent with the Bookroll interface). All remaining actions not covered by the above categories were classified as otherBookroll.

3.5.3 Chatbot Logs. Chat interaction data extracted from the Archie chatbot database (July 31st to August 23rd) were categorised to facilitate the analysis of learner behaviour. Each interaction was classified into one of five categories: interactionControl, provideAnswer, askSeekHelp, offTask, and otherChat, as shown in Table 2. These categories were developed through incremental experimentation with the dataset. The classifications were performed using a local instantiation (for privacy) of the open source Llama3.1:70b large language model [9], which was provided with the entire conversation history, including the system prompt at the beginning of the conversation and the LLM’s response to the single learner message being classified. This is consistent with the ‘coding with context’ employed by Liu et al. [16]; however, in our case we did not provide annotated examples to the LLM. We also utilised Langchain’s EnumOutputParser to constrain the output to only one of these classifications with this prompt:

```
Classify the student's message intent based on the
following categories:
- Interaction Control: includes salutations such as
  greeting the chatbot, thank yous and
  appreciation, and commands or instructions like
  "next question", "let's finish" or requests to
  change the text or activity
- Off Task: messages unrelated to the task or text
- Provide Answer: answering questions about the task
  or text
- Ask Question: asking a question about the task or
  text
- Seek Help: asking for assistance, clarification,
  further explanation, guidance, or feedback; also
  includes expressing difficulty or confusion.
- Other: any other message not covered by the above
  categories

Conversation Context (includes student message to
classify and the AI response to the student): {
context}

Current student message: {message}
Activity Context: {activity}

Classify the intent of the current student message using
only one of the above categories.
```

To assess the accuracy of the automated classification, two researchers independently classified a random sample of 200 messages from the total dataset of 17,500 interactions. Prior to classification, the researchers discussed each category, with clear examples and edge cases discussed during calibration sessions. Categories were iteratively refined based on pilot classifications to ensure clarity and consistency. Initially, the categories askQuestion and seekHelp were distinct; however, upon inspection of the data and due to the difficulty in reliably differentiating between the two categories, as

well as the relatively low frequency of askQuestion (250 total instances), these were merged into the single category askSeekHelp, resulting in a total of 1,298 interactions in this category. Discrepancies between researcher classifications were resolved through discussion until a consensus was reached. The researchers achieved strong agreement in their classifications, with a Cohen’s Kappa of 0.95 (SE = 0.033, 95% CI [0.89, 1.00]). This high level of inter-rater reliability suggests that the rubric and human coding process effectively minimised subjective interpretation.

3.5.4 Data Pre-processing. To merge the Bookroll and Archie logs and ensure ENA corresponding Bookroll actions for each interaction with Archie, for each learner the chat_id for a given conversation was back and forward propagated onto the Bookroll actions for the same text name and within one hour. One hour was chosen to ensure the Bookroll actions are relevant to the conversation. This meant that some Bookroll actions do not have an associated chat_id and were ignored in subsequent analysis. Coding for ENA was done by setting a binary 1 or 0 for each of the codes utilised based on the action from the merged log data, as shown in Table 3.

4 Results

4.1 Activity and Language Preferences

Table 4 shows the number of interactions for each activity in Archie. An interaction is a message sent to and received from the chatbot. The minimum chat length for each activity was one (chat was abandoned after first message and response).

Comprehension Questions and *True or False* were the most popular activities with students, generating the highest number of total interactions (3979 and 3257 respectively) and conversations (276 and 195). *Cloze Passage* also had a high number of total interactions (3257). *True or False*, *Comprehension Questions*, and *Fact vs. Opinion* had the highest average conversation length (16.19, 14.54, and 12.94 respectively), suggesting they were the most engaging for students. *Creative Writing*, *Discussion*, *Role Play*, *Sentence Reordering*, and *Summarisation* were the least popular, with the lowest total interactions and shortest average conversation lengths. Students overwhelmingly preferred to interact with Archie in Japanese, with the chatbot’s language setting configured to Japanese for over 90% of all interactions. English was selected as the preferred language only for a small fraction of interactions, ranging from 5% to 10% for all activities.

4.2 Chatbot and Reading Behaviour

Table 5 presents the frequency and normalised distribution of coded student behaviours across all participants ($n = 225$), as well as within high-proficiency ($n = 125$) and low-proficiency ($n = 110$) groups. The analysis reveals that certain behaviour types, namely interactionControl, provideAnswer, nextPage, and previousPage, constitute the majority of student actions across both proficiency levels. However, notable differences emerge when comparing high and low-proficiency groups.

High-proficiency students demonstrate higher engagement with the text compared to low-proficiency students. For instance, text navigation (nextPage, previousPage, otherNavigation) makes up 39.4% of total their behaviour compared to 33.5% for low-proficiency

Table 1: Bookroll Codes

Code	Example (Bookroll operation)	Description
nextPage	NEXT	Go to next page
prevPage	PREVIOUS	Go to previous page
otherNavigation	PAGE_JUMP, OPEN, SEARCH, SEARCH_JUMP, BOOKMARK_JUMP, MEMO_JUMP	Navigate using functionality other than next and previous
memo	REDO_HW_MEMO, ADD_HW_MEMO, CHANGE_MEMO, UNDO_HW_MEMO, MEMO_TEXT_CHANGE_HISTORY, CLEAR_HW_MEMO	Write / edit a short annotation anywhere on the page
markDifficult	ADD_MARKER, rgb(255, 255, 0)	Draw a yellow rectangle marking text as difficult
markImportant	ADD_MARKER, rgb(255, 0, 0)	Draw a red rectangle marking text as important
otherBookroll	ADD_BOOKMARK, DELETE_BOOKMARK, DELETE_MARKER, CLOSE	These operations were classified as other-Bookroll as there were either too few or they were not relevant to the current study

Table 2: Archie Codes

Code	Example (learner messages)	Description
interactionControl	“Hi Archie!”, “Let’s finish now”, “i want to move on cloze passage”, “What’s the next question?”	Salutations, responding to questions related to activity setup, giving instructions to the chatbot
provideAnswer	“1342”, “False”, “the adults were discussing SNS risks with young people whether study examined”, “Opinion”	Answering questions provided as part of the activity the learner has selected
askSeekHelp	“Give me a hint”, “I need more help”, “Which line?”, “what do you mean presence?”, “The meaning of furthermore?”	Includes seeking help, feedback, clarification, and expressing difficulty
offTask	“What’s your favourite food?”, “I’m sleepy”, “Speak Kansai dialect”, “I clean my room and my mom is happy”	Chat is not related to the task or text. Includes inappropriate messages.
otherChat	“Oh no! I was careless.”, “I see”, “Thank you”	Anything else

ncy students, indicating a greater tendency to engage with the text. Counter-intuitively, highly proficient students exhibit lower rates of provideAnswer (25.8% vs. 29.7%) as a proportion of their total behaviours; this may be explained more correct answers, and therefore less attempt on the same question. Conversely, behaviours such as interactionControl, offTask, and otherChat show minimal variation between proficiency groups, with high-proficiency students engaging in slightly more offTask interactions. Notably, highly proficient students exhibit a considerably higher total mean interaction count per learner (158.2 vs. 101.1), further supporting the observation that they adopt a more active and engaged approach to learning with the system.

4.3 Learner Appraisal

The exTAM survey results (Table 6) show there is generally positive attitudes and behavioural intentions toward the chatbot for both proficiency groups. The subscales demonstrated high reliability overall, despite several students consistently gave the same rating to all questions, a response pattern known as straightlining; this

suggests they may not have fully engaged with the survey. This was particularly evident in the ATT subscale, which includes a reverse-scored question, contributing to the lower reliability for that subscale. The sample size is large enough to prevent any significant impact on the survey results.

An independent sample t-test was conducted to compare the response patterns between high and low-proficiency students across the exTAM subscales. The results showed no statistically significant differences between the two groups for any of the subscales.

4.4 Epistemic Network Analysis

We applied epistemic network analysis [21] to our data using the ENA Web Tool (version 1.7.0) [17]. The unit of analysis in this study was all coded behaviour log data from Archie and Bookroll (within a chat_id) for a given learner. A moving window of four (current and three preceding logs within the same conversation) was utilised to construct the network model. This enabled each coded behaviour to be temporally connected to others. The window length was chosen based on a preliminary analysis of combined chat logs and reading

Table 3: Merged Bookroll and Archie Log Data with ENA Coding

Time	English level	chat_id	Behaviour	Learner	Archie	interactionControl	offTask	provideAnswer	askSeekHelp	otherChat	nextPage	previousPage	otherNavigation	memo	markDifficult	markImportant	otherBookroll
20:06:52	Low-proficiency	34aead31-6cb	OPEN			0	0	0	0	0	0	0	1	0	0	0	0
20:06:54			PAGE_JUMP			0	0	0	0	0	0	0	1	0	0	0	0
20:07:56			interaction Control	Hi Archie!	Hello Honoka! Today, we'll be working on Vocabulary Matching. Have you read Concerned Parent's and Social Adjustment?	1	0	0	0	0	0	0	0	0	0	0	0
20:08:08			interaction Control	Yes	Great! Match the English words from the text with their English definitions: A. penchant B. isolated C. delighted D. establish E. leadership 1. very happy 2. a strong liking for something 3. to create something that will last 4. being alone or separated from others 5. the ability to guide others. Tell me which definition matches which word. For example: A - 1, B - 2...	1	0	0	0	0	0	0	0	0	0	0	0
20:10:59			provide Answer	A_2,B_4,C_1,D_3,E_5	Excellent work, Honoka! You matched all the words correctly! Would you like to try Vocabulary Matching again? Or choose a different activity from above?	0	0	1	0	0	0	0	0	0	0	0	0

Table 4: Summary of Interactions

Activity	Total interactions	Mean length	SD	Median length	Max length	Total chats	English chat (%)
Cloze Passage	3144	12.29	11.59	9	75	263	7.22
Comprehension Questions	3979	14.54	14.19	11	96	276	5.23
Creative Writing	648	8.59	8.55	6	40	81	8.95
Discussion	674	7.46	7.72	5	56	89	5.34
Fact vs. Opinion	2008	12.94	9.77	11	67	156	6.03
Grammar Focus	1640	10.77	8.35	10	72	147	5.24
Role Play	749	8.11	6.15	6	35	95	7.61
Sentence Reordering	1301	8.28	6.93	6	40	140	6.99
Sentence Translation	1675	10.17	8.43	8	49	155	4.66
Summarisation	954	7.14	5.82	6	43	152	6.60
True or False	3257	16.19	10.14	13	59	195	6.23
Vocabulary Matching	1577	10.01	9.14	7	60	148	5.90

Table 5: Frequency and Distribution of Coded Student Behaviours by Proficiency Group

Behaviour	Total	%	Mean (<i>n</i> =235)		High (<i>n</i> =125)		Low (<i>n</i> =110)	
			Per learner	Per chat	Total	%	Total	%
interactionControl	7022	22.72	29.88	10.99	4354	22.01	2668	23.97
offTask	683	2.21	2.91	1.07	402	2.03	281	2.52
provideAnswer	8416	27.22	35.81	13.17	5107	25.82	3309	29.73
askSeekHelp	1298	4.20	5.52	2.03	868	4.39	430	3.86
otherChat	28	0.09	0.12	0.04	18	0.09	10	0.09
nextPage	5719	18.50	24.34	8.95	3852	19.47	1867	16.77
previousPage	4695	15.19	19.98	7.35	3209	16.22	1486	13.35
otherNavigation	1106	3.58	4.71	1.73	732	3.70	374	3.36
memo	600	1.94	2.55	0.94	489	2.47	111	1.00
markDifficult	263	0.85	1.12	0.41	155	0.78	108	0.97
markImportant	210	0.68	0.89	0.33	109	0.55	101	0.91
otherBookroll	873	2.82	3.71	1.37	487	2.46	386	3.47
Total	30913	100.00	131.54	48.38	19782	100.00	11131	100.00

Table 6: exTAM Subscale Ratings for English Proficiency Groups

Subscale	Proficiency	<i>n</i>	Mean	SD	Min	Max	Cronbach's α
PEOU	Low	57	4.55	1.41	1.00	7.00	0.91
	High	34	4.82	1.37	1.00	7.00	
PU	Low	57	4.40	1.27	1.00	7.00	0.95
	High	34	4.68	1.57	1.00	7.00	
PE	Low	57	4.51	1.55	1.00	7.00	0.95
	High	34	4.57	1.53	1.67	7.00	
ATT	Low	57	4.59	1.17	1.00	7.00	0.83
	High	34	5.03	1.33	2.00	7.00	
INT	Low	57	3.93	1.26	1.00	6.67	0.93
	High	34	4.11	1.48	1.00	7.00	

logs, which revealed that meaningful interaction sequences typically unfolded within this window length. Furthermore, a shorter or longer window length did not appear to be justified or significantly impact the outcome of the model. The learner networks were aggregated using binary summation, whereby the networks for a given line reflect the presence or absence of the co-occurrence between each pair of codes.

As there was no assumption of linearity, Spearman's ρ was calculated to assess model fit. The generated ENA model exhibited good fit (X-axis $\rho = 0.98$, Y-axis $\rho = 0.95$). The variance for the X-axis (MR1) = 0.27 and Y-axis (SD2) = 0.31.

The model for each group was calculated and projected onto the same space for comparison, as shown in Figure 2. For both groups, there is a strong relationship in co-occurrences between interactionControl and provideAnswer (denoted by the vivid line between the nodes), as well as between previousPage and nextPage, which is expected given the relative frequencies of these behaviours in Table 5.

A prominent cluster of nodes is observed in the top centre of the ENA space, comprising askSeekHelp, otherChat, markDifficult,

and offTask. This cluster appears to represent behaviours associated with student struggle or disengagement. The close proximity of these nodes suggests strong and consistent connections between these codes across the analysed units. Specifically, the cluster indicates that activities such as seeking help, engaging in unrelated chat, marking content as difficult, and being off-task frequently co-occur within the observed conversations or learning contexts. This pattern may reflect a tendency for students experiencing difficulty or disengagement to seek assistance, engage in off-topic discussions, or explicitly mark challenging content.

In contrast to the "struggle-disengage" cluster, a separate cluster located in the bottom left of the ENA space comprises markImportant, memo, and otherBookroll actions. This grouping suggests a pattern of active engagement with the learning materials characterised by a focus on organisation and knowledge retention. The close proximity of these nodes indicates that students who mark content as important also tend to utilise memo functions and engage in otherBookroll behaviours. This cluster also has closer proximity to the provideAnswer node, indicating higher engagement with the task.

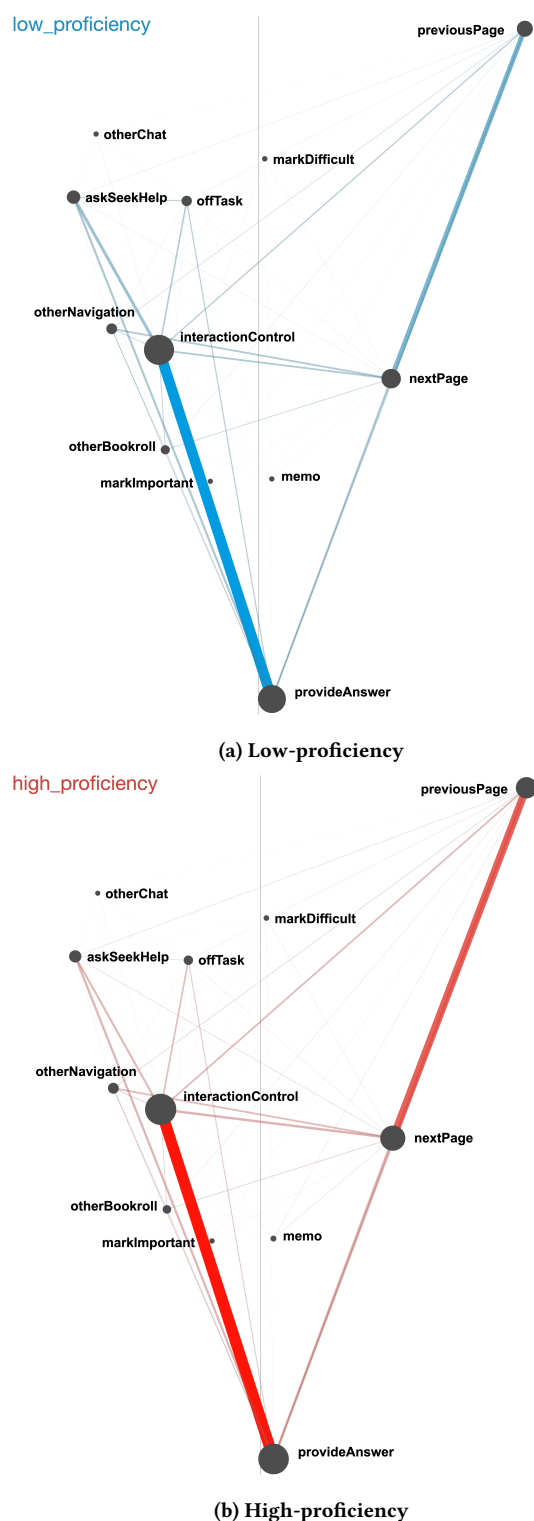


Figure 2: Comparison of low and high-proficiency models

To further investigate the behavioural distinctions between high and low-proficiency students, a comparison network was generated to highlight the differential co-occurrence patterns between the groups. Figure 3 illustrates this comparison, revealing a notable difference in the connection strength between `nextPage` and `previousPage`. Specifically, the co-occurrence of these two actions is significantly more pronounced in the high-proficiency group, suggesting a greater tendency for these students to engage in revisiting or backtracking through the learning materials.

A subtle difference in connection strength is also observed between `interactionControl` and `askSeekHelp` (-0.05), with a slightly stronger co-occurrence pattern evident in the low-proficiency group.

This suggests that low-proficiency students may exhibit a greater tendency to seek assistance from the chatbot. This pattern could indicate a stronger reliance on external guidance and support among low-proficiency students, potentially reflecting a need for more frequent clarification, feedback, or direction during the learning process. This is also evidenced in the chat log data with interactions such as “Help me start”, “I don’t understand”, “What are some examples?” and “What do the words mean?”.

To compare the high and low-proficiency group means, a two-sample Mann-Whitney test for non-parametric data was applied to the location of the group points in the projected ENA space. Along the X-axis (MR1), the test showed that the high-proficiency group (median = 0.07, $n = 125$) was statistically significantly different from the low-proficiency group (median = -0.05 , $n = 110$, $U = 8896.50$, $p < 0.01$, $r = -0.29$). Similarly, a Mann-Whitney test was conducted to compare the high and low-proficiency groups along the Y-axis (SVD2) of the projected ENA space. However, this analysis revealed no statistically significant difference between the groups ($U = 7043.00$, $p = 0.75$, $r = -0.02$). The high-proficiency group (median = -0.02) and low-proficiency students ($n = 110$) exhibited similar distributions along this dimension, indicating that this aspect of the behavioural patterns captured by the ENA model does not effectively discriminate between proficiency levels.

High-proficiency students exhibited higher engagement with the core interaction types in both Archie (`interactionControl`, `provideAnswer`) and Bookroll (`nextPage`, `previousPage`) compared to their low-proficiency counterparts. This suggests a more active and engaged approach to the task, characterised by frequent use of Archie, more text navigation, and a greater tendency for backtracking. To ensure that the observed differences reflect relative patterns of engagement rather than just the amount of data collected from each group, ENA applies normalisation in the model to account for differences in the total amount of data across students.

5 Discussion and Conclusion

This study investigated how high school students engaged with digital texts and a purpose-built chatbot, Archie, designed to facilitate and support active reading for English language learning. We sought to understand how learners interacted with both Archie and Bookroll (RQ1), their experiences and satisfactions with using Archie (RQ2), and how high-proficiency English learners’ chat and reading behaviour compared with low-proficiency learners (RQ3).

RQ1: How do learners interact with a chatbot for supporting

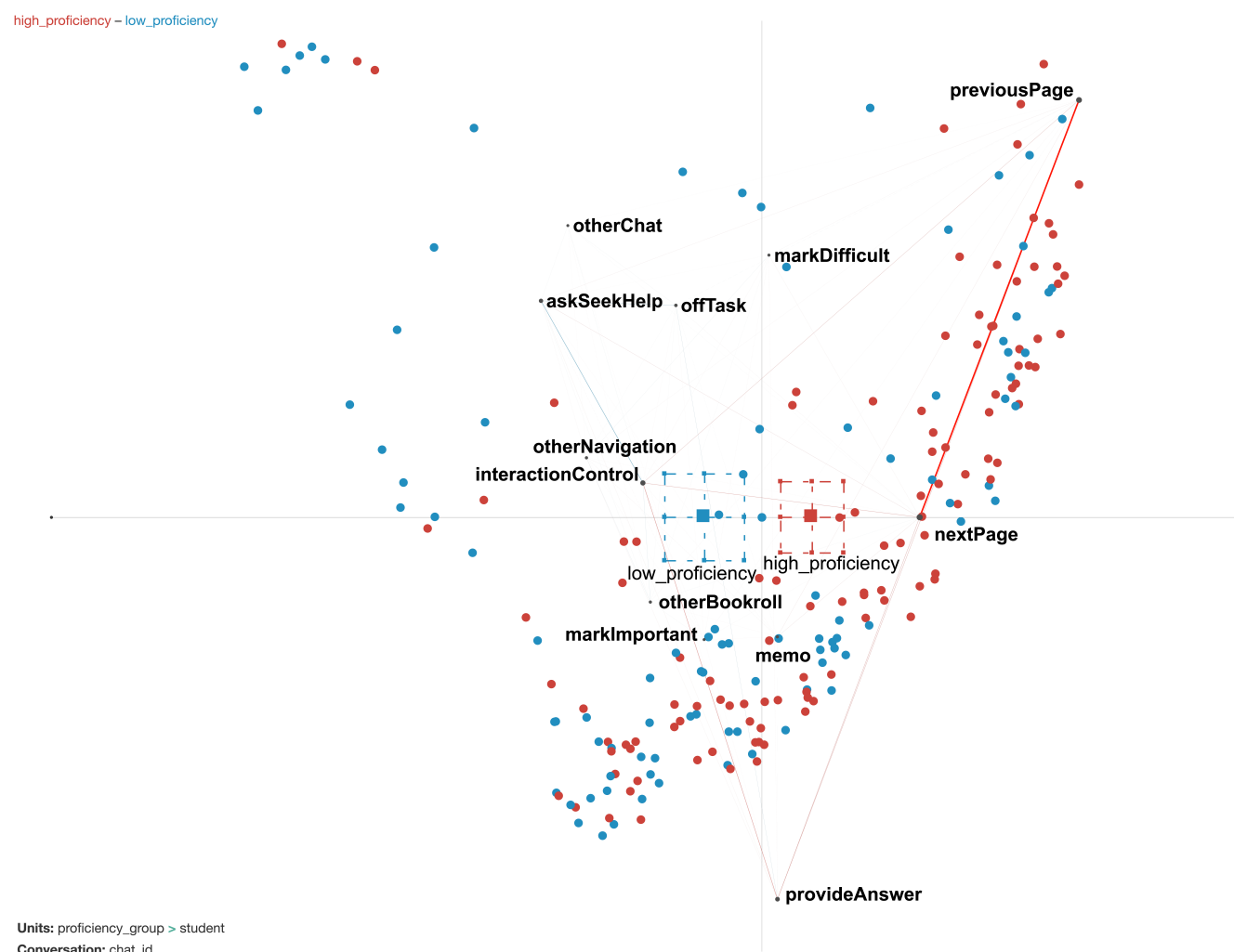


Figure 3: Comparison of co-occurrence patterns between high and low-proficiency groups

active reading and a digital reading system together? Regarding the student preferences, *Comprehension Questions* and *True or False* emerged as the most popular activities, potentially due to their structured nature, clear feedback, and familiarity to students. More open-ended activities such as *Role Play* or *Summarisation* were less popular. Interestingly, despite the purpose of the activities being to practise English, students overwhelmingly favoured interacting with Archie in Japanese, suggesting a prioritisation of comfort and fluency. This preference may indicate a need for explicit incentivisation of English usage with chatbots to improve interactive English skills.

RQ2: What are the learners' experiences and attitudes towards a chatbot for supporting active reading, and is there a difference between high and low English proficiency learners? The exTAM survey results indicated generally positive attitudes and behavioural intentions towards Archie, echoing findings from similar studies with EFL practice chatbots [25]. Students perceived Archie as easy to use, useful for their learning, and enjoyable

to interact with. These positive perceptions suggest that Archie successfully provided a supportive and engaging learning environment, potentially contributing to increased motivation and willingness to use the chatbot for future learning.

RQ3: How do high-proficiency English learners' chat and reading behaviours compare with those of low-proficiency learners? Analysis of coded behaviours revealed significant differences between high and low-proficiency students. High-proficiency students exhibited higher frequencies of *interactionControl*, *provideAnswer*, *nextPage*, and *previousPage* actions, indicating a more active and engaged approach. This aligns with previous research highlighting the importance of active reading strategies, such as backtracking and reviewing, for better comprehension and learning outcomes [13]. Notably, the higher relative frequency of *previousPage* actions among high-proficiency students provides direct evidence that higher levels of English proficiency is correlated with increased backtracking behaviour [28].

The ENA comparison network further emphasises this distinction, showing a stronger co-occurrence of `nextPage` and `previousPage` actions in the high-proficiency group, showing a greater tendency for these students to revisit and review the text. Additionally, a slightly stronger connection between `interactionControl` and `askSeekHelp` in the low-proficiency group suggests a higher reliance on external guidance and support. These students may benefit from explicit scaffolding and targeted interventions to help them develop their reading comprehension skills. The Mann-Whitney test confirmed the statistically significant difference between high and low-proficiency students along the X-axis (MR1) of ENA space, further supporting the observed distinctions between the behaviour configurations of high and low-proficiency students.

5.1 Limitations and Future Directions

This study has several limitations. The sample consisted of students from a single high school in Japan. The task and texts were narrow, and the students used the applications during their summer vacation, limiting the generalisability of the findings. The reliance on self-report data through the exTAM survey is subject to response biases. The coding scheme for behaviours in Archie could have been more nuanced, especially with regard to `interactionControl` and `askSeekHelp`. Moreover, we did not evaluate the correctness of student answers, which could have provide valuable insights into their learning processes. The choice to assign a single classification to each student's message, while simplifying analysis, may have overlooked important nuances in student-chatbot interactions. A multi-classification approach could capture more detailed insights into how learners engage with the chatbot. Similarly, Bookroll's log data only offers page-level navigation details, preventing an analysis of the specific text sections students were reading. Future research could incorporate eye tracking or more granular logging to better understand learners' reading patterns within a page. This study also did not examine the responses generated by Archie, which could provide insight into the effectiveness of the chatbot's feedback and guidance. Chatbots are also prone to the novelty effect, where initial enthusiasm fades over time [8]; future research should explore longer-term system use. For simplicity only examined message interactions with Archie were examined. Other logs (e.g. changing text or activity) were excluded, but could enhance our understanding of learner behaviour. Finally, future research could investigate the effectiveness and impact on learning outcomes of chatbots designed to support active reading. As a cross-sectional study, this research cannot make any claims about causation. While certain patterns, such as backtracking and proficiency, were observed, these represent correlations rather than causal relationships. Further longitudinal research and controlled experimental designs would be needed to investigate causal mechanisms.

5.2 Implications

The positive student perceptions of Archie suggest that chatbots can be successfully integrated into language learning environments to provide engaging and personalised support. This aligns with previous research highlighting the potential of chatbots to enhance motivation, reduce anxiety, and offer individualised feedback [3]. Furthermore, the observed differences in engagement and reading

strategies between high and low-proficiency students underscore the importance of explicitly encouraging and teaching active reading behaviours such as backtracking and reviewing [28]. Educators can leverage chatbots and e-readers to encourage these behaviours through targeted activities and feedback mechanisms. The insights gained from ENA can inform the design of more effective chatbot interventions and personalised learning pathways. By identifying specific behavioural patterns associated with different performance levels, educators can tailor chatbot interactions and e-reader functionalities to better support individual learner needs.

The integration of chatbots and e-readers into the classroom presents transformative opportunities for both teachers and students. For educators, these technologies can significantly alleviate the workload by providing personalised, on-demand support for learners, enabling teachers to focus on higher-order tasks. Chatbots, like Archie, offer tailored guidance that caters to diverse proficiency levels, ensuring that students receive support that aligns with their specific needs. Additionally, the data-rich interactions captured by these tools, such as log data from e-readers and chat transcripts, provide invaluable insights into student behaviours, engagement patterns, and learning needs. Teachers can leverage this information to inform instructional strategies, identify students who may require additional support, and refine their pedagogical approaches.

This study offers a novel contribution to the growing body of research on both chatbot-assisted EFL learning and e-reader usage. By integrating these two areas, our research provides unique insights into how learners engage with and learn from these technologies when used together. Our findings demonstrate that the combined use of chatbots and e-readers can create a rich and engaging learning environment that caters to diverse learners. Specifically, our study reinforces the established link between backtracking behaviour and reading comprehension, and—through the application of ENA—reveals distinct behavioural patterns that differentiate high and low-proficiency English language learners.

Acknowledgments

This work was partly supported by JSPS Grant-in-Aid for Scientific Research (B) JP20H01722 and JP23H01001, (Exploratory) JP21K19824, 22KJ1914, (A) JP23H00505, SIP JPJ012347, and NEDO JPNP20006.

References

- [1] Aziza M Ali and Abu Bakar Razali. 2019. A review of studies on cognitive and metacognitive reading strategies in teaching reading comprehension for ESL/EFL learners. *English Language Teaching* 12, 6 (May 2019), 94–111.
- [2] Behnam Behforouz and Ali Al Ghaithi. 2024. Investigating the effect of an interactive educational chatbot on reading comprehension skills. *International Journal of Engineering Pedagogy* 14, 4 (May 2024), 139–154.
- [3] Serge Bibauw, Wim Van den Noortgate, Thomas François, and Piet Desmet. 2022. Dialogue systems for language learning: A meta-analysis. *Language Learning & Technology* 26, 1 (2022), 1–24.
- [4] Amelia Biringkanee. 2018. The use of SQ3r technique in improving students' reading comprehension. *ELS Journal on Interdisciplinary Studies in Humanities* 1, 2 (June 2018), 218–225.
- [5] Anna C.-S. Chang and Willy A. Renandya. 2017. Current practice of extensive reading in Asia: Teachers' perceptions. *The Reading Matrix: An International Online Journal* 17, 1 (2017), 40–58.
- [6] Michelene T H Chi and Ruth Wylie. 2014. The ICAP framework: Linking cognitive engagement to active learning outcomes. *Educational Psychologist* 49, 4 (Oct. 2014), 219–243.
- [7] Kit U Chio. 2009. Reading and second language acquisition. *HKBU Papers in Applied Language Studies* 13 (2009), 153–174.

- [8] Emmelyn A J Croes and Marjolijn L Antheunis. 2021. Can we be friends with Mitsu? A longitudinal study on the process of relationship formation between humans and a social chatbot. *Journal of Social and Personal Relationships* 38, 1 (Jan. 2021), 279–300.
- [9] Abhimanyu Dubey, Abhinav Jauhri, Abhinav Pandey, et al. 2024. The Llama 3 herd of models. *arXiv [cs.AI]* (July 2024).
- [10] Brendan Flanagan and Hiroaki Ogata. 2018. Learning analytics platform in higher education in Japan. *Knowledge Management & E-Learning: An International Journal* 10, 4 (2018), 469–484.
- [11] Andrina Granić and Nikola Marangunić. 2019. Technology acceptance model in educational context: A systematic literature review. *British Journal of Educational Technology* 50, 5 (Sept. 2019), 2572–2593.
- [12] Carolin Hahnel, Dara Ramalingam, Ulf Kroehne, and Frank Goldhammer. 2023. Patterns of reading behaviour in digital hypertext environments. *Computer Assisted Learning* 39, 3 (June 2023), 737–750.
- [13] Ymkje E Haverkamp and Ivar Bråten. 2024. The role of strategic backtracking when reading digital informational text for understanding. *Literacy Research and Instruction* 63, 2 (April 2024), 176–191.
- [14] Weijiao Huang, Khe Foon Hew, and Luke K Fryer. 2022. Chatbots for language learning—Are they really useful? A systematic review of chatbot-supported language learning. *Journal of Computer Assisted Learning* 38, 1 (Feb. 2022), 237–257.
- [15] Chen-Chung Liu, Mo-Gang Liao, Chia-Hui Chang, and Hung-Ming Lin. 2022. An analysis of children's interaction with an AI chatbot and its impact on their interest in reading. *Computers & Education* 189, 104576 (Nov. 2022), 104576.
- [16] Xiner Liu, Jiayi Zhang, Amanda Barany, Maciej Pankiewicz, and Ryan S. Baker. 2024. Assessing the Potential and Limits of Large Language Models in Qualitative Coding. In *Proceedings of the Conference on Educational Data Mining (EDM)*. Educational Data Mining Society, 89–103.
- [17] C L Marquart, C Hinojosa, Z Swiecki, B Eagan, and D W Shaffer. 2018. Epistemic network analysis (version 1.7. 0)[Software]. Available from app. *epistemicnetwork.org* (2018).
- [18] Negar Mohammadhassan and Antonija Mitrovic. 2022. Discovering Differences in Learning Behaviours During Active Video Watching Using Epistemic Network Analysis: Third International Conference, ICQE 2021, Virtual Event, November 6–11, 2021, Proceedings. In *Advances in Quantitative Ethnography*, Barbara Wasson and Szilvia Zörgő (Eds.). Communications in Computer and Information Science, Vol. 1522. Springer International Publishing, Cham, 362–377.
- [19] Paul Nation. 1996. The four strands of a language course. *TESOL in context* 6 (June 1996), 7–12.
- [20] Ivana Roncencic. 2021. Eye-tracking in second language reading. *International Journal of Innovation, Creativity and Change* (2021), 416–432.
- [21] David Williamson Shaffer, University of Wisconsin-Madison, USA, A R Ruis, and University of Wisconsin-Madison, USA. 2017. Epistemic network analysis: A worked example of theory-based learning analytics. In *Handbook of Learning Analytics*. Society for Learning Analytics Research (SoLAR), 175–187.
- [22] Deepti Sharma, Archana B Saxena, and Deepshikha Aggarwal. 2024. Harnessing AI chatbots for personalized learning: Implications and strategies for education. *Journal of Informatics Education and Research* 4, 2 (2024), 2642–2645.
- [23] Yuanru Tan, Zachari Swiecki, A R Ruis, and David Shaffer. 2024. Epistemic network analysis and ordered network analysis in learning analytics. In *Learning Analytics Methods and Tutorials*. Springer Nature Switzerland, Cham, 569–636.
- [24] Yuko Toyokawa, Rwitajit Majumdar, Taisho Kondo, Izumi Horikoshi, and Hiroaki Ogata. 2024. Active reading dashboard in a learning analytics enhanced language-learning environment: effects on learning behavior and performance. *Journal of Computers in Education* 11, 2 (June 2024), 495–522.
- [25] Steve Woollaston, Brendan Flanagan, Patrick Ocheja, Yiling Dai, and Hiroaki Ogata. 2024. TAMMY: Supporting EFL Translation Practice with an LLM-powered Chatbot. In *Proceedings of the 32nd International Conference on Computers in Education (ICCE 2024)*. Asia-Pacific Society for Computers in Education (APSCE), 585–594.
- [26] Steve Woollaston, Brendan Flanagan, and Hiroaki Ogata. 2024. Chatbots and EFL learning: A systematic review. In *Joint Proceedings of LAK 2024 Workshops*. 89–98.
- [27] Xiaoyu Wu and Yuan Gao. 2011. Applying the extended technology acceptance model to the use of clickers in student learning: Some evidence from macroeconomics classes. *American Journal of Business Education* 4 (June 2011), 43–50.
- [28] Chengjiu Yin, Masanori Yamada, Misato Oi, Atsushi Shimada, Fumiya Okubo, Kentaro Kojima, and Hiroaki Ogata. 2019. Exploring the relationships between reading behavior patterns and learning outcomes based on log data from E-books: A human factor approach. *International Journal of Human-Computer Interaction* 35, 4–5 (March 2019), 313–322.
- [29] Zhihui Zhang and Xiaomeng Huang. 2024. The impact of chatbots based on large language models on second language vocabulary acquisition. *Heliyon* 10, 3 (Feb. 2024), e25370.
- [30] Shuyan Zheng. 2024. The effects of chatbot use on foreign language reading anxiety and reading performance among Chinese secondary school students. *Computers and Education: Artificial Intelligence* 7, 100271 (Dec. 2024), 100271.

A Chatbot System Prompt and Activity Instructions

```

system_prompt = f'''
    **Objective:** You are a helpful English learning
        chatbot named Archie:
    **Active **Reading **Chatbot for **Interactive **English the Alpaca. Assist
        Japanese high school students in understanding
        an English text through interactive English
        learning activities.

    **Text:** {text.content}
    **Teacher Notes:** {text.teacher_notes}

    **Chatbot Instructions:**
    1. **Greet the student by name and introduce the
        activity:**
        - "Hello {given_name}! Today, we'll be working on {
            activity_selection}. Have you read {text_name}?"
    2. **Activity Instructions:**
        Student instructions: {ACTIVITY_INSTRUCTIONS[
            activity_selection]['student_instruction']}
        Chatbot instructions: {ACTIVITY_INSTRUCTIONS[
            activity_selection]['chatbot_instruction']}
    3. **Provide feedback and further assistance:**
        - "Great job! You completed the activity! Would you
            like to try {activity_selection} again?"
    **Guidelines:**
        - Converse with the student in their preferred
            language ({preferred_language})
        - Keep the student on task; avoid unrelated
            discussions.
        - Be concise and concrete; limit responses to 50
            words maximum.
        - Personalise interactions using the student's name.
        - Vary responses and stay engaging to maintain
            student interest.
        - Provide feedback that is explanatory and concise,
            with corrections in bold.
        - Guide the student to the answer; avoid providing
            complete answers.
        - Be strict, but fair. Acknowledge correct answers
            and provide hints for incorrect answers.
        - It is important you are accurate in your feedback
            and guidance.
        - Use excerpts of the text rather than the full text.
        - When stuck, encourage the student read and reread
            the text for better understanding.
        - Students can read the text in Bookroll (accessible
            from Moodle).
        - If the student is struggling, provide hints, or ask
            questions to guide them to the answer.
    ...

    ACTIVITY_INSTRUCTIONS = {
        'Cloze Passage': {
            'student_instruction': 'Fill in the blanks in
                this passage: [passage]. What words fit?',
            'chatbot_instruction': 'Ask the student to fill
                in the blanks in the given passage. Provide
                hints if needed.'
        },
        'Comprehension Questions': {
            'student_instruction': "Here's a question to
                check your understanding: [question]. What's
                your answer?",

```

```

    'chatbot_instruction': "Ask comprehension
    questions based on the text. Provide
    feedback on the student's answers."
},
'Creative Writing': {
    'student_instruction': "Write an English story
    inspired by the text. Here's a prompt to
    start: [prompt].",
    'chatbot_instruction': 'Encourage the student to
    write a creative story in English based on
    the text. Provide a prompt and feedback.'
},
'Discussion': {
    'student_instruction': "Let's discuss the main
    ideas. Please write in English! Here's a
    prompt: [prompt]. Share your thoughts!",
    'chatbot_instruction': 'Facilitate a discussion
    about the main ideas of the text. Ask
    questions and provide feedback.'
},
'Fact vs. Opinion': {
    'student_instruction': 'Is this statement fact or
    opinion? (insert statement). Why?',
    'chatbot_instruction': 'Present five or more
    statements from or about the text and ask
    the student to identify them as facts or
    opinions. Provide feedback.'
},
'Grammar Focus': {
    'student_instruction': 'Identify the grammatical
    structure in this sentence: [sentence]. What
    grammar rule is applied? Here are some
    options: (insert three options - explain
    each options).',
    'chatbot_instruction': 'Ask the student to
    identify the grammatical structures in four
    different sentences. Provide feedback.'
},
'Role Play': {
    'student_instruction': 'Role-play this scenario
    from the text: [scenario]. What would your
    character say or do?',
    'chatbot_instruction': 'Guide the student in a
    role-play based on a scenario from the text.
    Provide feedback.'
},
'Sentence Reordering': {
    'student_instruction': 'Reorder these sentences
    to reflect the correct sequence: [sentences
    ].',
    'chatbot_instruction': 'Ask the student to
    reorder at least four sentences into the
    correct sequence. Provide hints and feedback
    .'
},
'Sentence Translation': {
    'student_instruction': 'Translate this sentence:
    (insert random important English sentence of
    appropriate level). Need help with a word
    or phrase? Just ask!',
    'chatbot_instruction': 'Ask the student to
    translate a sentence (English to Japanese).
    Provide assistance with vocabulary and
    grammar.'
},
'Summarisation': {
    'student_instruction': 'Summarise the text in a
    few sentences in English. Need help starting
    ? I can give tips!',
    'chatbot_instruction': 'Ask the student to
    summarise the text. Provide tips and
    feedback.'
},
'True or False': {
    'student_instruction': 'Based on the text, is
    this statement true or false? [first
    statement].',
    'chatbot_instruction': 'Present five or more
    statements and ask the student to determine
    if they are true or false. Provide feedback
    .'
},
'Vocabulary Matching': {
    'student_instruction': 'Match the English words
    from the text with their English definitions
    : [list].',
    'chatbot_instruction': 'Ask the student to match
    five words (A, B, C, D, E) with their
    definitions (1, 2, 3, 4, 5). Provide hints
    and feedback.'
}
}

```