Developing and Evaluating an Interactive Reading Tool with Teachers in the Loop: Action Research Approach

Mihwa Lee

Hector Research Institute of Education Sciences and Psychology, Universirty of Tübingen, Walter-Simon-Str. 12, 72072 Tübingen, Germany
LEAD Graduate School & Research Network, University of Tübingen, Europastr. 6, 72072 Tübingen, Germany

Abstract

Reading is an essential life skill and crucial for students' academic success. Especially, there has been an increasing necessity for students to read in English as a second language (L2) due to its global importance. However, teachers in schools often face challenges in providing interactive L2 reading experiences for a large number of students due to limited time and highly heterogeneous students, leading to L2 readers having few opportunities for meaningful, interactive reading practice with instant support. The rapid advancements of artificial intelligence (AI) in education have given rise to a number of opportunities for interactive and adaptive learning. Despite significant advancements in AI-powered educational tools, many language educators continue to view them with skepticism. This may stem from a perceived misalignment between teaching methods that educators find effective and the features or approaches offered by these technologies. As a result, the gap between educators' expectations and the capabilities of AI-driven solutions remains a point of concern. It is crucial to ensure that educational systems align with established theories and pedagogical insights, and to investigate them from multiple perspectives, including perceptions of the system, learning outcomes, motivation, and learning behavior to better design educational products. This article introduces a pedagogically grounded web-based intelligent computer-assisted language learning (ICALL) system, designed to enhance L2 reading experiences, developed using the Action Research design with teachers in the loop. The article details the system's development and provide an overview of ongoing and planned studies, which focus on different aspects of the ICALL system, examining learners' behaviors through interaction logs to further L2 learning research and improve educational tools.

Keywords

Reading comprehension, Language learning, Intelligent computer-assisted language learning (ICALL), Process data

1. Introduction

In today's increasingly globalized world, the increasing necessity for students to read in English as a L2 underscores the importance of proficient L2 reading skills. Learning to read in L2 is complex, as learners must grasp literacy in an unfamiliar language [1]. Thus, there is an urgent need to support L2 reading from the early school years. However, teachers face challenges in providing interactive and adaptive learning experiences for a large number of students with limited time. Digital environments, such as ICALL systems, offer unique opportunities for new ways of learning and teaching [2]. These systems have been shown to enhance learning engagement [3] and achieve better language acquisition [4] through features such as automatic feedback [5], intelligent tutoring [6], and personalized support [7]. Despite these advancements, there remains a significant gap of the use of such tools in school settings, possibly because of the skepticism among practitioners [8] due to not only people's lack of knowledge of the field and its capabilities but also the fact that a lot of AI-based education applications do not meet educators' expectations of how effective language teaching and learning should be conducted [9]. Given the complexity of challenges in AI in education (AIED) and the field's traditional emphasis on technical aspects, many AI-driven educational tools and studies struggle to align with the most recent advancements in learning theories, empirical research findings, and pedagogical insights [9].

To address this gap between research on language education, foreign language teaching insights, and real-life classroom usage, an ICALL system that systematically and automatically provides various interactive support for L2 reading has been designed and developed, targeting learners of English as a foreign language (EFL). The design and development of the system is grounded in theories in Second Language Acquisition (SLA), educational sciences, and pedagogical insights from school practitioners, and leverages the affordance of the Natural Language Processing (NLP) tools and Large Language Model (LLM). In this article, we introduce the design rationale of the system and present the plans and status of studies assessing the effectiveness of the system in promoting L2 reading from various dimensions and examining learners' learning behaviors using interaction logs stored in the system. Specifically, the first goal is to design and develop an ICALL system that supports and enhances L2 reading comprehension based on the SLA theories and teachers' insights. The second goal is to examine the effectiveness of the ICALL system in promoting students' learning outcomes and motivation compared to traditional online reading practice. Lastly, the third goal is to investigate learners' self-regulated learning behavior from by combining interaction logs with self-report data. By exploring these dimensions, we aim to advance L2 learning research and refine educational tools to better support reading development in school contexts.

2. Background

2.1. Linguistic knowledge in reading comprehension

Reading is a complex cognitive task that necessitates the integration of textual information with prior knowledge. Effective comprehension relies on the reader's ability to efficiently process the visual information presented in the text

Proceedings of the Doctoral Consortium of the 19th European Conference on Technology Enhanced Learning, 16th September 2024, Krems an der Donau, Austria
mihwa.lee@uni-tuebingen.de (M. Lee)
https://uni-tuebingen.de/fakultaeten/wirtschafts-und-sozialwissenschaftliche-fakultaet/faecher/fachbereich-sozialwissenschaften/hector-institut-fuer-empirische-bildungsforschung/institut/personen/lee-mihwa/ (M. Lee)
0009-0000-0101-3549 (M. Lee)

[1]. Current theories of reading comprehension generally depict it as involving multiple interconnected layers of conceptual representation. These layers include a lower-level representation that draws on text-based elements such as vocabulary and grammar, and a higher-level representation where the textual content is incorporated into the reader's broader conceptual framework (e.g., combining information across sentences) [10, 11, 12]. The reader's vocabulary and grammar knowledge significantly shape the formation of these semantic structures throughout the reading process [10]. Specifically, vocabulary and grammar steer the parsing process, which builds meaning from local text segments. If the local-level representations are inaccurate or incomplete, overall text comprehension can be significantly hindered [10, 13]. Lexical-syntactic knowledge is essential for constructing the local-level representation, which forms the foundation for higher-level text comprehension [11, 14]. Thus, vocabulary and grammar facilitate the building of text-based propositions and contribute to deeper comprehension.

SLA researchers have also focused on the role of vocabulary and grammar knowledge in understanding L2 reading comprehension. Numerous studies have explored how these factors influence L2 reading comprehension, with findings consistently underscoring the importance of morphosyntactic knowledge [14]. Recent meta-analyses on L2 reading comprehension [14, 15] highlight that vocabulary and grammar knowledge are the two strongest predictors of L2 reading comprehension. Consequently, vocabulary and grammar knowledge have a significant impact—whether directly or indirectly—on reading comprehension.

From the instructional perspective, however, it is almost impossible for teachers to pinpoint vocabulary and grammatical knowledge that each learner does not understand while students are reading. One way to support teachers in this process is by utilizing supportive computer environments. Despite the importance of such fundamental linguistic skills in reading comprehension, however, there are a relatively small number of technological applications, though these (e.g., [16, 17]) involve only minimal use of technology. Therefore, it is crucial for researchers to account for both vocabulary and grammar when developing language learning applications. Additionally, further research is needed to explore whether and how support for these aspects can enhance learners' L2 reading processes, potentially shaping their learning behaviors and improving overall reading comprehension.

2.2. Feedback in reading comprehension

Feedback is information communicated to learners to modify their thinking or behavior to close the gap between their actual performance and target performance [18], thus aiming to improve learning [19], as well as enhance emotions and motivation during a learning situation [20]. In the field of both SLA and educational sciences, feedback is recognized as an important factor in supporting learning, particularly when it helps overcome insufficient or false hypotheses [18]. Feedback serves a cognitive function by informing readers of misunderstandings, filling gaps, and increasing awareness of their understanding [21]. This awareness of one's understanding level is crucial for teaching students to self-regulate their learning from texts, which involves both (meta)cognitive strategies, such as making inferences and monitoring comprehension, and motivational processes, like the desire to learn [18, 21, 22]. [22] emphasizes that effective feedback should guide students to consider both cognitive and motivational aspects in their learning process, particularly when using computer-assisted learning tools. Therefore, one can assume that providing adaptive and scaffolding feedback for learners potentially triggers changes of learners' attitudes (motivational component) and reading strategies (cognitive component), which consequently improves reading comprehension [21].

Despite the significant role feedback plays in reading comprehension, teachers are usually the only reliable source of feedback for learners in real-life classroom settings. However, their time and the amount they can spend with each student in class are very limited, resulting in few opportunities for learners to receive individual formative feedback. This is especially important given the substantial individual differences in aptitude and proficiency [23]. Another issue is the lack of research on how feedback enhances L2 reading comprehension. A recent meta-analysis [21] indicates that research on feedback has predominantly concentrated on learning outcomes related to reading comprehension, where few studies have explored the cognitive and affective processes triggered by feedback aimed at text comprehension. Furthermore, most of this research has focused on the effectiveness of feedback in reading comprehension in the first langauge (L1), with relatively little attention given to its impact on L2 learners' reading comprehension. Therefore, further empirical research is needed to investigate if and how such feedback, especially in the context of a computer-assisted learning environment, can enhance learners' learning processes and, consequently, influence their learning behavior and overall reading comprehension.

2.3. Self-regulated learning (SRL) in a computer-based learning environment

Self-regulated learning (SRL) broadly refers to an educational process in which learners proactively engage in academic tasks [24, 25]. [24] provides a widely accepted definition: "active, constructive process whereby students set goals for their learning and then attempt to monitor, regulate, and control their cognition, motivation, and behavior, guided and constrained by their goals and the contextual features of their environment" (p. 453). In academic literature, there is a consensus that SRL is essential for students' reading development. Proficient readers are typically highly motivated self-regulated learners who use various reading strategies effectively [26]. Motivation drives learners to use learning strategies, which helps regulate learning behaviors and improve outcomes [25, 27]. SRL strategies include planning, critical thinking, peer learning, effort regulation, and goal orientation. Classroom based research indicates that SRL strategies lead to higher learning performance [24, 25, 27]. Therefore, supporting students' motivation through, for instance, help options and feedback is crucial for promoting their use of learning strategies, which eventually enhances learning outcomes [28, 29]. Digital technologies have the potential to directly influence learners' motivation, strategy use, and outcomes by providing interactive and adaptive learning environments that cater to individual needs [30]. Hence, investigating the effects of these technologies on motivation, learning strategies, and outcomes is urgent.

Previous studies indicated that SRL is the crucial skill for success in computer-based learning environments as well [31]. However, learners cannot always regulate themselves successfully because of reasons such as lack of good strategy use, lack of metacognitive knowledge, failure to control of metacognitive processes, or lack of experience in learning environments with multiple representations. Thus, how to foster SRL ability has become a central issue in the field of education research and practice. In order to support SRL in the computer-mediated learning environments, instruments that capture students' self-regulation are critical.

2.4. Trace-based measurement of SRL

Offline instruments like self-reported questionnaires and semi-structured interviews have long been used to measure students' SRL processes in both educational sciences and SLA. However, these traditional methods face criticism due to their subjectivity, obtrusiveness, and limited ability to capture the dynamic nature of learning [32]. Such instruments are often unable to reflect all the elements learners attend to during their learning processes [33]. To address these limitations, researchers advocate for the integration of multiple data types, such as digital-trace data, which includes real-time interaction log data [34, 35]. This digital-trace data offers a more granular and continuous insight into SRL, allowing both researchers and practitioners to monitor students' learning behaviors and strategic decisions in online environments with remarkable detail and in real time [36]. These online measures are particularly valuable because they capture cognitive processes as they unfold during learning, offering a temporal perspective on cognitive change and presenting a moment-by-moment view of students' processing behaviors [37].

One challenge that is often addressed by researchers is importance of aligning the data collection with SRL model [32, 38]. To this end, researchers have often utilised theory-aligned coding schemes that define SRL processes at different levels of granularity by, for example, coding schemes [39]. Based on these schemes, previous research has primarily relied on clickstream data from Learning Management Systems (LMS) to measure SRL behaviors related to time management, a crucial sub-construct of SRL [40]. These studies consistently demonstrate that clickstream-based measures of time management predict student performance in online learning environments.

However, despite the promise of this microanalytic method and its availability due to recent technological advances, its application remains limited in the field of SLA and language learning studies [41]. Furthermore, most research has focused exclusively on time management, leaving other critical SRL sub-constructs largely unexplored in the context of digital-trace data collection [40]. This gap highlights the need for future studies to broaden their focus to include other dimensions of SRL to fully leverage the potential of interaction log data in understanding the complexities of student learning.

3. Research questions

Driven by the objective of advancing L2 learning research and refining educational tools to better support reading development in school contexts, this project centers the attention on the following research questions:

1. What are the characteristics that are considered beneficial for supporting and enhancing L2 reading comprehension? What are the students' and teachers' perceptions of our L2 reading system?
2. To what extent is the ICALL system for reading effective in promoting students' learning outcomes and motivation compared to traditional online reading practice?
3. What insights do the interaction logs reveal about the learners' usage of the system and their SRL behavior? How students with high self-regulation and low self-regulation behave differently?

The following sections present the plans and status of the studies addressing those research questions in detail. At the time of submission, the first research question is being addressed in a study that is currently taking place.

4. Methodology

Involving teachers or stakeholders in education research whose results will be used in schools is considered very important because schools and teachers should not only be treated as consumers of the research results [42]: successful research that has a practical impact in schools is always the outcome of bi-directional efforts. This the bi-directional effort will not be a one-off process, but a process that will involve multiple iterations of interactions between the research team and the teachers. Consequently, a multi-cycle Action Research paradigm was chosen to guide the research process. The Action Research Model (see Figure 1) is a systematic, collective, collaborative, and self-reflective scientific inquiry aimed at improving educational practices and ad-dressing the practical concerns of teachers [43, 44], where a key characteristic of action research is the involvement of stakeholders, including teachers, students, and researchers. Throughout the project, we adhere to this approach in the process of development, testing, and implementation of the system. Figure 2 illustrates the overview of the research design based on the Action Research Model.

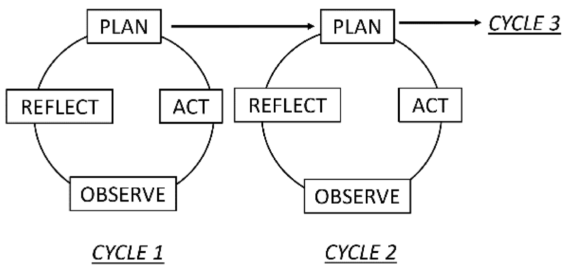

Figure 1: Action Research Model, adpated from [43]

RQ1: To answer the first research question, we first conducted an intensive literature review about L2 reading comprehension in order to decide on the characteristics to be implemented in an ICALL system for L2 reading. The aim was to understand the key factors that contribute to effective L2 reading comprehension and how these can be supported in a digital learning environment. As discussed in the Background section (see section 2.1 and section 2.2), reading comprehension involves integrating text information, heavily relying on vocabulary and grammar. Feedback plays a crucial role in improving comprehension by helping learners bridge understanding gaps and enhance self-regulation.

Action Research Phase	PLAN	ACT	OBSERVE	REFLECT
Target RQ	RQ1	RQ1	RQ1 and RQ3	RQ2 and RQ3
Outcome	• System design	• System prototype	• Perceptions of the system • Learning behavior	• Learning outcomes • Learning motivation • Learning behavior
Method	• Literature review • Regular consultations with teachers	• System development	• Mixed method (survey + semi-structured interview) • Log analysis	• Mixed method (survey + semi-structured interview) • Log analysis
Stakeholders involved	• Researchers • Teachers	• Researchers	• Researchers • Teachers • Students	• Researchers • Teachers • Students

Figure 2: Overviview of methodology based on Action Research Model

The first prototype of our ICALL system, called *ARES*, includes features to support these aspects (more discussion of the technical side of the system development can be found in Lee et al. (2024)). Following the Action Research Model, multiple consultations with English practitioners and teachers from German secondary schools (”Gymnasiums”) were conducted. This collaborative co-design approach ensured that the system's features met not only the SLA theories, but also practical classroom needs and pedagogical insights. Figures 3–8 illustrate some features of the first prototype of the system, developed upon after the initial “Plan” phase of the first iteration cycle of the Action Research Model. Using the NLP tools, key features on the learner side include:

- On-demand interactive lookup on language means: learners can access detailed explanations and examples of language means directly within the reading text, adaptively helping them understand grammar rules in context according to their need (see Figure 3).
- On-demand interactive vocabulary lookup: learners can access detailed explanations and examples of vocabulary in terms of its form, meaning, and use directly within the reading text, adaptively helping them understand vocabulary in context according to their need (see Figure 4).
- Elaborated feedback: learners receive detailed, personalized feedback on their reading and comprehension activities, highlighting areas of strength and providing targeted suggestions for improvement (see Figure 5).

In addition to the features that support learners, educational systems should also support teachers so that they can be used in real-life classroom contexts. At the same time, however, it should not replace the teachers. Rather, it should help teachers. Therefore, with the LLM (ChatGPT4o[1]), the system includes features and resources that empower teachers to effectively support L2 reading development in their classrooms, while at the same time it is designed in a way that teachers' expertise is always involved in the process (more discussion of the technical side of the system development can be found in Lee et al. (2024)). They can post-edit suggestions by the LLM, confirm them, or add their own questions manually. In this way, teachers make the ultimate decision about what to show the students. Key features on the teacher side include:

- Customization of annotations on language means: teachers can customize which annotations on language means are shown to students to align with their instructional goals and the specific needs of their students (see Figure 6).

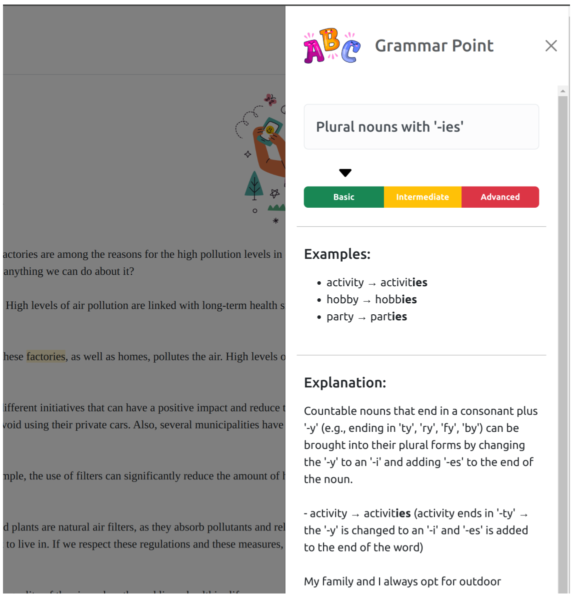

Figure 3: Lookup of language means

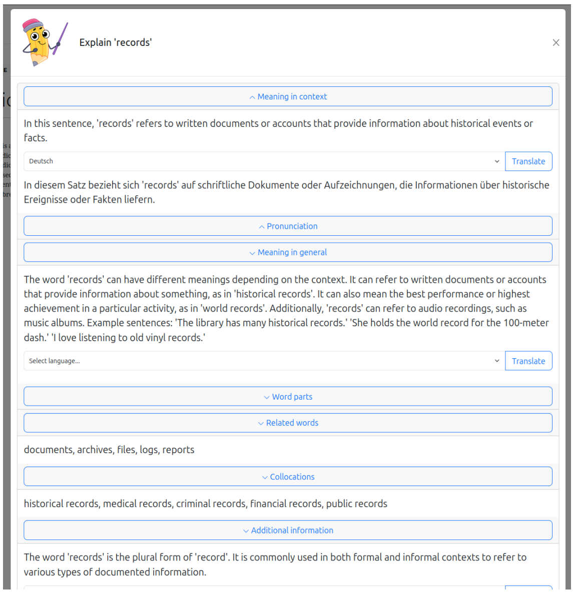

Figure 4: Vocabulary lookup

[1] https://chatgpt.com/

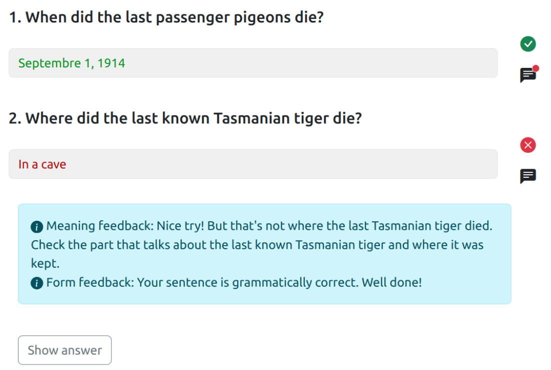

Figure 5: Feedback for student's response

- Question generation: the system generates reading comprehension questions (factual and inferential) based on the reading text, helping teachers provide questions tailored to the text (see Figure 7).
- Feedback generation: the system creates personalized feedback for students based on their performance, helping teachers provide individualized support (see Figure 8).
- Evaluation: the system evaluates student responses to comprehension questions, providing immediate grading, which reduces the grading burden on teachers.
- Minimalistic analytics: the system provides simple analytics on student performance and engagement, offering teachers quick insights without overwhelming them with data.
- Text bank and uploading: the system not only includes a library of reading texts of a variety of topics but also lets the teachers upload texts, allowing them to tailor the reading materials to their curriculum and students' interests.

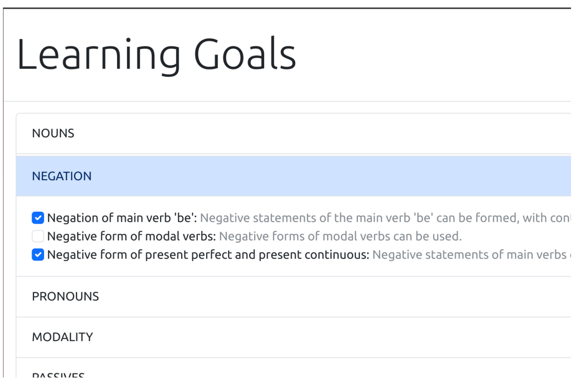

Figure 6: Selection of annotations of language means

In terms of the technical aspect, ARES is built with Java at the backend, with a Jetty[2] server and a Docker[3] container. The database is PostgreSQL[4], and the frontend is based on a popular JavaScript framework, HTML, and Bootstrap[5] that provides a highly extensible component-based

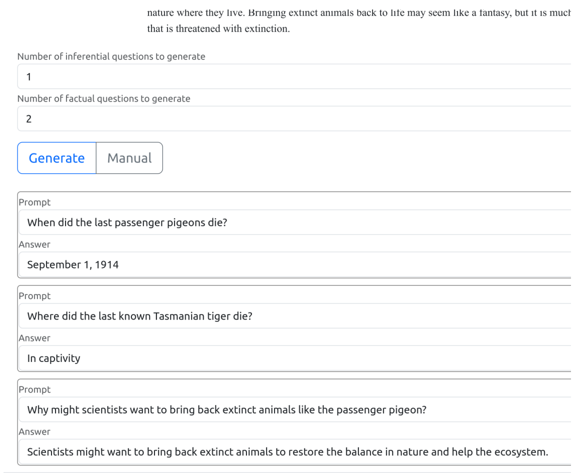

Figure 7: Question generation for a reading text

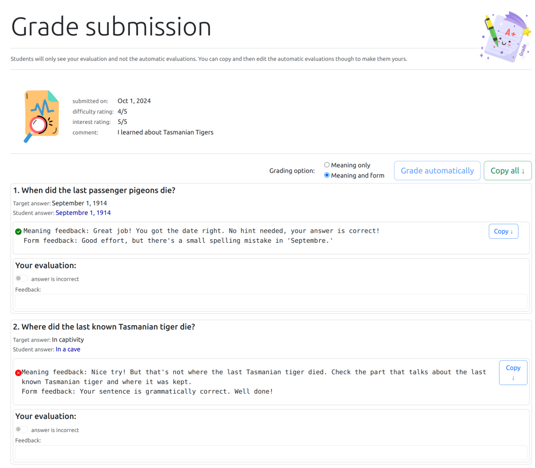

Figure 8: Grading of individual submission

design. In order to enable Learning Analytics, all user activities such as button clicks, lookups of language means, reading comprehension question attempts, assignment submissions, viewing of specific feedback messages, and any other relevant user actions are logged through xAPI[6], an interoperability specification for learning technology, and stored in a Learning Record Store (LRS) in the database.

Since the first version of the system is deployed, a study investigating teachers' and students' perceptions of the system is currently taking place in two intact English classes at secondary schools (students around age 13-14) in southwest Germany with the purpose of evaluating the system's usability and overall task and system design. These mixed-gender classes are part of the academic track of the German education system. The curriculum at this grade level is equivalent to A2-B1 levels on the Common European Framework of Reference for Languages (CEFR) [46], representing the students' fourth year of EFL instruction in school. Over an eight-week period, students read two texts weekly using ARES as part of their homework assigned by teachers. A mixed-method approach with quantitative data from self-reports and qualitative data from semi-structured interviews is employed. System

[2]https://jetty.org/index.html
[3]https://www.docker.com/
[4]https://www.postgresql.org/
[5]https://getbootstrap.com/

[6]https://xapi.com/

perceptions are assessed through a self-report questionnaire of comprehensive evaluation of educational technology adapted from [47], which contains closed-ended items in eight evaluation categories such as Usability, Design, and Learning Motivation with a 7-point Likert scale. Additional open-ended items asking what students and teachers liked or disliked, and what they wish for the system are included as well. For the analysis of the learning behavior from logs, students' self-reported SRL skills in online learning (*Online Self-Regulated Learning Questionnaire, OSLQ*, adapted from [48]) are also collected from students. After filling in the questionnaires, teachers and several students will be invited for a follow-up semi-structured interview to gain their perceptions of the system in-depth, which will follow the guideline suggested by [49]. For quantitative data analysis of the self-reports, the mean and standard deviation of of each close-ended item and category will be calculated. For quantitative analysis from the semi-structures interview, a reflexive thematic analysis [50] will be conducted. The results will be discussed with the English teachers at the participating schools to refine the system's usability and task design.

RQ2: To answer the second research question, the study investigating the effectiveness of the system is planned to take place this school year in English classes (students around age 13-14) in secondary schools in southwest Germany. The study will be administered via the ARES system, and use a posttest/pretest design consisting of a battery of tests and questionnaires. After providing parental consent, participants will be introduced to the ARES system and complete the pre-tests and pre-questionnaires. The teachers will be asked to assign at least two reading assignments per week over an eight-week period via the ARES interface. Based on the methodology of [23], rotational within-class randomization will be employed based on the condition. In the first four weeks, half of each class will serve as the intervention group, using the system with lookup and feedback on comprehension questions features, while the other half will read plain texts without such aids. In the second four weeks, this will be reversed. After eight weeks, participants will be instructed to complete the post-tests, post-questionnaires, and background questionnaire.

Learning outcomes will be measured by pre- and post-tests that measure their English vocabulary knowledge (*Updated Vocabulary Levels Test*, [51]), English reading comprehension (Reading section of *TOEFL® Primary™ Step 2*), general English proficiency (*Elicited Imitation test*, [52]), and L2 reading motivation (*Reading Motivation Questionnaire*, adapted from [53]). OSLQ [48] will be also used to measure the students' self-reported SRL skills. During the eight-week period, participants' interaction with the system will be tracked. We plan to conduct a pretest-posttest and pre questionnaire-post questionnaire comparison across groups, in which we expect improvements in the measurements on which participants had access to the aids while learning.

RQ3: To answer the third research questions, the subset of log data that is stored as students interact with the system from the aforementioned studies addressing RQ1 and RQ2 will be used. Student's behavioral data will be firstly collected as learning logs from the ARES system by extracting students' interactions in the LRS in form of xAPI statements stored in the system database. As noted in the Background section (see section 2.4), it is critical to align the data collection with SRL model [32, 38]. Consequently, the analysis will be guided by the SRL processes proposed by [32] that defines the three macro level [54] of SRL processes: Planning, Engagement, and Evaluation and Reflection. Each process phase is further divided into several micro-level SRL processes in order to define fine-grained SRL processes. Details about this theoretical framework and the SRL processes it encompasses are provided in Table 1. Next, to extract the SRL behavior implied by the actions, the actions will be aggregated into a common xAPI statement structure with the theoretical framework of SRL processes proposed by [32]. Among seven micro-level SRL processes proposed by [32], five processes are identified according to the functionality of the system where the actions are taken. Table 2 summarizes the actions in the systems in ARES mapped to each proposed macro-level and micro-level SRL processes.

To explore how students with varying levels of self-regulation approach their learning, we will compare their self-reported SRL skills with behavioral patterns recorded in the system. Behavioral variables will be tracked for each student and assignment to provide a detailed profile of their learning behaviors. K-means cluster analysis will be employed to group students based on (1) their self-reported SRL skills and (2) their behavioral patterns as reflected in the trace data in order to identify patterns that highlight how well their self-perceptions align with their actual learning behaviors. The resulting clusters will then be compared to examine correlations between subjective and objective measures of SRL, which will help reveal whether students with strong self-reported SRL skills also demonstrate strong behavioral evidence of self-regulation, or whether there are discrepancies between the two, providing valuable insights into the alignment (or misalignment) between students' perceived and actual learning strategies.

5. Conclusion and contribution

Driven by the need to fill the gap between research on language education, foreign language teaching insights, and real-life classroom usage, this article presented the development of the pedagogically grounded ICALL system that provides various learning supports for L2 reading comprehension and an overview of ongoing and planned studies, which focus on different aspects of the ICALL system, examining learners' behaviors through interaction logs. The results of this project will provide AIED researchers and language educators with an interdisciplinary perspective and further insights on the feasibility and capabilities of using the current NLP and AI (LLM) tools in language learning applications and inform system and task design decisions for enhancing learning outcomes. Apart from the research plans and studies outlined in this article, future directions include examining the feasibility of leveraging the LLM to generate short answer questions and feedback, the classification accuracy of annotations on language means, and the efficacy of different feedback types for students with different levels of SRL skills.

Macro-level SRL process	Micro-level SRL process	Description
Planning	Task Analysis	To get familiar with the learning context and the definition and requirements of a (learning) task at hand
	Goal Setting	To explicitly set, define, or update learning goals
	Making Personal Plans	To create plans and select strategies for achieving a set learning goal
Engagement	Working on the Task	To consistently engage with a learning task, using tactics and strategies
	Applying Strategy Changes	To revise learning strategies, or apply a change in tactics
Evaluation & Reflection	Evaluation	Evaluating one's learning process and comparing one's work with the goal
	Applying Strategy Changes	Reflecting on individual learning and sharing learning experiences

Table 1
Theoretical framework guiding trace-based measurement of SRL processes, articulated in [32]

Macro-level SRL process	Micro-level SRL process	Behavioral variables
Planning	Task Analysis	Total number of visits to an assignment overview; Total number of visits to a tutorial video; Sum of total time spent on a tutorial video
Engagement	Working on the Task	Total number of visits to assignments before deadline; Sum of total time spent on assignments before deadline; Total number of visits to reading questions before deadline; Total number of assignments completed; Total number of correct responses; Total number of incorrect responses
	Applying Strategy Changes	Total number of access to vocab help; Total number of access to language means help
Evaluation & Reflection	Evaluation	Total number of visits to assignments after deadline; Sum of total time spent on assignments after deadline; Total number of visits to feedback to correct responses after deadline; Total number of visits to feedback to incorrect responses after deadline; Total number of visits to target answers after deadline
	Applying Strategy Changes	Total number of visits to class average score after deadline; Total number of visits to own score after deadline

Table 2
Matching map between the SRL processes and learning behavior data in the system

Acknowledgments

This PhD project is funded by the German Ministry of Education and Science (BMBF) under the funding number 01IS22076.

References

[1] L. Verhoeven, Second language reading acquisition, in: M. L. Kamil, P. D. Pearson, E. B. Moje, P. Afflerbach (Eds.), Handbook of reading research, volume 4, Routledge, New York, NY, 2011, pp. 661–683. doi:10.4324/9780203840412.ch27.

[2] L. A. Amaral, D. Meurers, On using intelligent computer-assisted language learning in real-life foreign language teaching and learning, ReCALL 23 (2011) 4–24. doi:10.1017/S0958344010000261.

[3] C.-C. Liu, P.-C. Wang, S.-J. D. Tai, An analysis of student engagement patterns in language learning facilitated by Web 2.0 technologies, ReCALL 28 (2016) 104–122. doi:10.1017/s095834401600001x.

[4] A. Oberg, P. Daniels, Analysis of the effect a student-centred mobile learning instructional method has on language acquisition, Computer Assisted Language Learning 26 (2013) 177–196. doi:10.1080/09588221.2011.649484.

[5] H. Ai, Providing graduated corrective feedback in an intelligent computerassisted language learning environment, ReCALL 29 (2017) 313–334. doi:10.1017/s095834401700012x.

[6] I.-C. Choi, Efficacy of an ICALL tutoring system and process-oriented corrective feedback, Computer Assisted Language Learning 29 (2016) 334–364. doi:10.1080/09588221.2014.960941.

[7] M. Heilman, K. Collins-Thompson, J. Callan, M. Eskenazi, A. Juffs, L. Wilson, Personalization of reading passages improves vocabulary acquisition, International Journal of Artificial Intelligence in Education 20 (2010) 73–98. doi:10.3233/JAI-2010-0003.

[8] A. Kate, Three reasons to be skeptical of artificial intelligence in schools, 2020. URL: https://www.edweek.org/technology/three-reasons-to-be-skeptical-of-artificial-intelligence-in-schools/2020/02.

[9] X. Chen, E. Bear, B. Hui, H. Santhi-Ponnusamy, D. Meurers, Education theories and ai affordances: Design and implementation of an intelligent computer-assisted language learning system, in: M. Rodrigo, N. Matsuda, A. Cristea, V. Dimitrova (Eds.), Artificial Intelligence in Education. Posters and Late Breaking Results, Workshops and Tutorials, Industry and Innovation Tracks, Practitioners' and Doctoral Consortium. AIED 2022., Lecture Notes in Computer Science, Springer, Cham, 2022, pp. 582–585. doi:10.1007/978-3-031-11647-6_120.

[10] J. Jung, Second language reading and the role of grammar, Working Papers in TESOL and Applied Linguistics 9 (2009) 29–48. doi:10.7916/D88915FW.

[11] W. Kintsch, The role of knowledge in discourse comprehension: A construction-integration model, Psychological Review 95 (1988) 163–182. doi:10.1037/0033-295x.95.2.163.

[12] W. Kintsch, T. A. van Dijk, Toward a model of text comprehension and production, Psychological Review 85 (1978) 363–394. doi:10.1037/0033-295X.85.5.363.

[13] K. Koda, Reading and language learning: Crosslinguistic constraints on second language reading development, Language Learning 57 (2007) 1–44. doi:10.1111/j.1467-9922.2007.00411.x.

[14] Y. Choi, D. Zhang, The relative role of vocabulary and grammatical knowledge in L2 reading comprehension: A systematic review of literature, International Review of Applied Linguistics in Language Teaching 59 (2021) 1–30. doi:10.1515/iral-2017-0033.

[15] H. Chen, H. Mei, How vocabulary knowledge and grammar knowledge influence L2 reading comprehension: a finer-grained perspective, European Journal of Psychology of Education (2024) 1–23. doi:10.1007/s10212-024-00793-x.

[16] D. Meurers, R. Ziai, L. Amaral, A. Boyd, A. Dimitrov, V. Metcalf, N. Ott, Enhancing authentic web pages for language learners, in: Proceedings of the NAACL HLT 2010 Fifth Workshop on Innovative Use of NLP for Building Educational Applications, BEA 2010, Association for Computational Linguistics, Los Angeles, California, 2010, pp. 10–18.

[17] S. Walker, P. Schloss, C. R. Fletcher, C. A. Vogel, R. C. Walker, Visual-syntactic text formatting: A new method to enhance online reading, Reading Online 8 (2005).

[18] J. Hattie, H. Timperley, The power of feedback, Review of Educational Research 77 (2007) 81–112. doi:10.3102/003465430298487.

[19] V. J. Shute, Focus on formative feedback, Review of Educational Research 78 (2008) 153–189. doi:10.3102/0034654307313795.

[20] C. J. Fong, E. A. Patall, A. C. Vasquez, S. Stautberg, A meta-analysis of negative feedback on intrinsic motivation, Educational Psychology Review 31 (2019) 121–162. doi:10.1007/s10648-018-9446-6.

[21] E. K. Swart, T. M. Nielen, M. T. S. de Jong, Does feedback targeting text comprehension trigger the use of reading strategies or changes in readers' attitudes? a meta-analysis, Journal of Research in Reading 45 (2022) 171–188. doi:10.1111/1467-9817.12389.

[22] M. ter Beek, L. Brummer, A. S. Donker, M.-C. J. Opdenakker, Supporting secondary school students' reading comprehension in computer environments: A systematic review, Journal of Computer Assisted Learning 34 (2018) 557–566. doi:10.1111/jcal.12260.

[23] D. Meurers, K. D. Kuthy, F. Nuxoll, B. Rudzewitz, R. Ziai, Scaling up intervention studies to investigate real-life foreign language learning in school, Annual Review of Applied Linguistics 39 (2019) 161–188. doi:10.1017/S0267190519000126.

[24] P. R. Pintrich, The role of goal orientation in self-regulated learning, in: M. Boekaerts, P. R. Pintrich, M. Zeidner (Eds.), Handbook of self-regulation, volume 4, Academic Press, San Diego, 2000, pp. 451–502. doi:10.1016/B978-012109890-2/50043-3.

[25] B. J. Zimmerman, Becoming a self-regulated learner: Which are the key subprocesses?, Contemporary Educational Psychology 11 (1986) 307–313. doi:10.1016/0361-476X(86)90027-5.

[26] H. Li, Z. Gan, Reading motivation, self-regulated reading strategies and english vocabulary knowledge: Which most predicted students' english reading comprehension?, Frontiers in Psychology 13 (2022) 1041870. doi:10.3389/fpsyg.2022.1041870.

[27] D. H. Schunk, B. J. Zimmerman, Motivation and self-regulated learning: Theory, research, and applications, Lawrence Erlbaum Associates Publishers, Mahwah, 2008.

[28] R. S. Jansen, A. van Leeuwen, J. Janssen, R. Conijn, L. Kester, Supporting learners' self-regulated learning in massive open online courses, Computers Education 146 (2020) 103771. doi:10.1016/j.compedu.2019.103771.

[29] S.-H. Jin, K. Im, M. Yoo, I. Roll, K. Seo, Supporting students' self-regulated learning in online learning using artificial intelligence applications, International Journal of Educational Technology in Higher Education 20 (2023) 37. doi:10.1186/s41239-023-00406-5.

[30] M. L. Bernacki, J. A. Greene, H. Crompton, Mobile technology, learning, and achievement: Advances in understanding and measuring the role of mobile technology in education, Contemporary Educational Psychology 60 (2019) 101827. doi:10.1016/j.cedpsych.2019.101827.

[31] T. Adeyinka, S. Mutula, A proposed model for evaluating the success of webct course content management system, Computers in Human Behavior 26 (2010) 1795–1805.

[32] M. Siadaty, D. Gašević, M. Hatala, Trace-based micro-analytic measurement of self-regulated learning processes, Journal of Learning Analytics 3 (2016) 183-–214. doi:10.18608/jla.2016.31.11.

[33] P. Robinson, A. Mackey, S. Gass, R. Schmidt, Attention and awareness in second language acquisition, in: S. Gass, A. Mackey (Eds.), The Routledge handbook of second language acquisition, Routledge, New York, NY, 2012, pp. 247–267.

[34] D. Gašević, S. Dawson, G. Siemens, Let's not forget: Learning analytics are about learning, TechTrends 59 (2015) 64–71.

[35] G. Siemens, R. Baker, Learning analytics and educational data mining: Towards communication and collaboration, in: Proceedings of the 2nd international conference on learning analytics and knowledge, Association for Computing Machinery, New York, NY, 2012, pp. 252–254. doi:10.1145/2330601.2330661.

[36] G. Siemens, R. Baker, Student learning in higher education: A commentary, Educational Psychology Review 29 (2017) 353–362. doi:10.1007/s10648-017-9410-x.

[37] I. Roll, P. H. Winne, Understanding, evaluating, and supporting self-regulated learning using learning analytics, Journal of Learning Analytics 2 (2015) 7–12. doi:10.1007/s10648-017-9410-x.

[38] Q. Yu, Y. Zhao, The value and practice of learning analytics in computer assisted language learning, Studies in Literature and Language 10 (2015) 90.

[39] M. Raković, S. Iqbal, T. Li, Y. Fan, S. Singh, S. Surendrannair, J. Kilgour, J. van der Graaf, L. Lim, I. Molenaar, M. Bannert, J. Moore, D. Gašević, Harnessing the potential of trace data and linguistic analysis to predict learner performance in a multi-text writing

task, Journal of Computer Assisted Learning 39 (2023) 703–718.

[40] Q. Lia, R. Baker, M. Warschauer, Using clickstream data to measure, understand, and support self-regulated learning in online courses, The Internet and Higher Education 45 (2020) 100727.

[41] Q. Yu, Learning analytics: The next frontier for computer assisted language learning in big data age, in: SHS Web of Conferences, volume 17, EDP Sciences, 2015, p. 02013. doi:10.1051/shsconf/20151702013.

[42] E. Farley-Ripple, H. May, A. Karpyn, K. Tilley, K. McDonough, Rethinking connections between research and practice in education: A conceptual framework, Educational Researcher 47 (2018) 235–245. doi:10.3102/0013189X18761042.

[43] S. Kemmis, R. McTaggart, The Action Research Planner, 3rd ed., Deakin University Press, Geelong, 1988.

[44] R. N. Rapoport, Three dilemmas of action research, Human Relations 23 (1970) 499–513. doi:10.1177/001872677002300601.

[45] M. Lee, B. Rudzewitz, X. Chen, Developing a pedagogically oriented interactive reading tool with teachers in the loop, in: Proceedings of the 13th Workshop on Natural Language Processing for Computer Assisted Language Learning (NLP4CALL 2024), 2024. In press.

[46] Council of Europe, Common European Framework of Reference for Languages: Learning, teaching, assessment – Companion Volume, Cambridge University Press, Cambridge, 2020.

[47] J. W. M. Lai, J. D. Nobile, M. Bower, Y. Breyer, Comprehensive evaluation of the use of technology in education – validation with a cohort of global open online learners, Education and Information Technologies 27 (2022) 9877-–9911. doi:10.1007/s10639-022-10986-w.

[48] L. Barnard, Y. W. Lan, M. Y. To, V. O. Paton, S.-L. Lai, Measuring self-regulation in online and blended learning environments, The Internet and Higher Education 12 (2009) 1–6.

[49] E. Rosalind, J. Holland, What is qualitative interviewing?, Bloomsbury Academic, London, 2013.

[50] V. Braun, V. Clarke, Thematic analysis: A practical guide, Sage Publications, Los Angeles, 2022.

[51] A. Godfroid, K. M. Kim, The updated vocabulary levels test: Developing and validating two new forms of the vlt, International Journal of Applied Linguistics 168 (2017) 34-–70. doi:10.1075/itl.168.1.02web.

[52] A. Godfroid, K. M. Kim, The contributions of implicit-statistical learning aptitude to implicit second-language knowledge, Studies in Second Language Acquisition 43 (2021) 606-–634. doi:10.1017/S0272263121000085.

[53] W. Wang, Z. Gan, Development and validation of the reading motivation questionnaire in an english as a foreign language context, Psychology in the Schools 58 (2021) 1151-–1168. doi:10.1002/pits.22494.

[54] U. Kroehne, F. Goldhammer, How to conceptualize, represent, and analyze log data from technology-based assessments? a generic framework and an application to questionnaire items, Behaviormetrika 45 (2018) 527–563. doi:10.1007/s41237-018-0063-y.