

It is like a friend to me: Critical usage of automated feedback systems by self-regulating English learners in higher education

Long Li

The Chinese University of Hong Kong, Hong Kong SAR

Mira Kim

UNSW Sydney, Australia

This paper explores international students' engagement with educational technology for self-regulated English learning at an Australian university. Despite the increased use of automated feedback systems (AFSs) for language assessment, students' critical engagement with them for independent learning remains under-researched. The study primarily employed a qualitative approach to understand the students' preferred AFS tools and critical engagement throughout their personalised learning journeys but it also included a small-scale quantitative component. Data were gathered from seven students' e-portfolios, focus group interviews as well as a survey among 32 participants. Results highlight positive perceptions and successful use of AFSs, with students leveraging these tools to identify improvement areas, track progress and gain confidence. The study emphasises the importance of course structure, teacher guidance and a combination of human and automated feedback, in fostering learner autonomy and emotional self-regulation. The paper underscores the potential for sustained use of AFSs beyond the course, and the significance of guiding learners to critically use these tools for ongoing learning and growth rather than dependence. These findings have significant implications, as readily available artificial intelligence tools such as ChatGPT hold great pedagogical potential for self-regulated learning within and beyond the language learning field.

Implications for practice or policy

- Instructors can use AFSs as effective tools to help English learners in higher education when scaffolding critical engagement with automated feedback and emotional self-regulation and providing adaptability, as such scaffolding and flexibility are essential for mitigating the limitations of AFSs.
- Course leaders and universities should consider investing in AFSs as they can elevate the availability and sustainability of feedback for language enhancement and potentially any other type of learning.

Keywords: automated feedback systems (AFSs), computer- and mobile-assisted language learning, English language proficiency (ELP), self-regulated learning (SRL), engagement, ELSA, Grammarly

Introduction

A core goal of educational technology is to provide immediate, automated and ideally personalised feedback to students, increasingly through artificial intelligence (AI)-based assessment (Fang et al., 2023). Educational technology holds great potential in providing such feedback for English language learning. This paper explores international students' engagement with automated feedback systems (AFSs) in higher education, where the proficiency of English as the medium of instruction and assessment significantly impacts student learning. Its particular focus is on how learners critically use AFSs for English language enhancement when they are given autonomy to personalise their own learning, which is a form of self-regulated learning (SRL). Therefore, our findings and pedagogical discussion will be applicable beyond language learning in this era where the capacity of intelligent systems to support learning and even emotional aspects of learning is rapidly growing, as well captured in some chapters in du Boulay et al. (2023), with one of the biggest challenges being to create "intelligent learning environments" to

provide learners with the essential skills to engage in educational opportunities, as suggested by Lesgold (2023, p. xvii).

International students' need for enhancing English language proficiency (ELP) has been recognised by universities, and various models to address such a need have been developed over recent decades in Australia and other English-speaking countries worldwide (Arkoudis et al., 2012; Read, 2016). However, although international students continuously demand better support (Marginson et al., 2010), paradoxically, institutional supports attract limited student attention and engagement (Rochecoste & Oliver, 2014). There have thus been calls for a holistic and sustainable approach to the ELP needs of international students (Benzie, 2010). Against this background, a new personalised autonomous (PA) model has been adopted in a credit-bearing course, Personalised English Language Enhancement (PELE), with the course offered as a general education or elective course to students at all levels (undergraduate, postgraduate and HDR) across faculties, at UNSW Sydney, Australia, since 2017. The PA model promotes a cyclical learning process that consists of the following five stages (see Kim, 2014; Kim & Jing, 2019 for more details):

- (1) Awareness: Students identify their need or motivation to enhance language skills.
- (2) Needs analysis: They analyse their language needs in context and specify a project focus.
- (3) Planning: They design a personal project to enhance their communication in a specific context.
- (4) Implementation: They undertake the planned tasks and record their progress in an online portfolio over 5–6 weeks, attempting to form habits.
- (5) Reflection: They reflect on their progress, with quantitative and/or qualitative data, and share with their peers through oral presentation in the final week and a final essay.

Accordingly, the PA learning model aims to foster learners' ability to make informed decisions for their learning, take responsibility for their learning outcomes and engage in critical reflection to learn from the process (Kim, 2014). Data collected from 2016 to 2020 reveal that PELE's impacts on students are comprehensive and substantial, especially in terms of their "a) confidence in English skills in both academic and everyday contexts, b) self-efficacy for self-regulated English learning and academic and social engagement, and c) sense of belonging to PELE, faculty, and university" (Kim, 2023, p. 615). Self-efficacy refers to one's own feelings about one's ability to handle specific tasks and proves to be a reliable indicator in predicting students' performance (Bandura, 1997).

Despite these positive learning experiences, most PELE students experienced a challenge: how to identify their own areas to improve by themselves, which is an important skill in the PELE context. When students are introduced to the PA model, which is substantially different from conventional pedagogies, they initially experience mixed emotions, especially when they attempt to decide a specific goal for their project. The idea that they can improve their English skills in their preferred ways excites them. However, they often find it challenging to analyse their linguistic needs in context and narrow these down to a specific goal to achieve through their personal project. This challenge was clearly observed in class from the onset of PELE in 2016 and confirmed with findings from a set of paired-sample t tests conducted with identifiable data collected between 2018 and 2020. The improvement in students' self-efficacy to identify areas of improvement in their English language usage is relatively smaller than improvements in the other five self-efficacy variables for self-regulated English learning: confidence with locating resources, asking for English help, setting a realistic goal, using appropriate strategies as well as evaluating and reflecting (Kim, 2023, p. 608).

The present study was thus initiated to abate this challenge by investigating the benefits and limitations of using AFS tools. Subsequently, the research scope expanded to include the following research questions:

- (1) What kind of AFS tools do self-regulating English learners adopt?
- (2) How do learners perceive AFSs?
- (3) How do learners engage with AFSs across the PA learning cycle?
- (4) What are the pedagogical implications for scaffolding AFS tools in a SRL context?

A small body of studies, most of which focused on only one AFS tool, has reported mixed learner and teacher perceptions (Dikli & Bleyle, 2014), with commonly reported issues such as perceived inaccuracy (O'Neill & Russell, 2019); learner preference for teachers over automated feedback (Dikli & Bleyle, 2014); non-use, despite the positive perception, due to heavy cognitive load (Grimes & Warschauer, 2010) and an ongoing debate over whether an AFS tool is viewed more favourably by more advanced learners (Cavaleri & Dianati, 2016) or by less advanced learners (Chen & Cheng, 2008). Even though these studies have not all reported positive outcomes and some questions are still under debate, we hypothesise that AFS tools hold great potential in offering self-regulating learners useful and relevant feedback on their English communication, when they are introduced with explicit guidance about how they can be used in their learning context. A few studies have already highlighted the important role of the instructors in mediating AFS feedback (Hoang & Kunnan, 2016; O'Neill & Russell, 2019) and in regulating learner emotions and motivation (Zhang, 2020) to mitigate limitations and negative perceptions. Considering the growing but under-researched area of learner perception and engagement with automated feedback (particularly for speaking) (Evanini & Zechner, 2019; Xi, 2010; Zhang, 2020), the present study contributes new pedagogical insights to the field.

The rest of this paper consists of an introduction to SRL and diagnostic tests as the theoretical framework and research background; an introduction to the use of AFSs in the present study; the presentation of data collection and analysis, and the results from the focus group interviews, the e-portfolios and the survey; a discussion of these results followed by a further discussion of contextual factors; and a conclusion on research and pedagogical implications for self-regulated English learning and for the wider field of educational technology.

SRL and diagnostic tests

The importance of helping students to identify their areas for improvement has been recognised not only in the PELE context but also in the SRL field in general. SRL is defined as “self-generated thoughts, feelings, and actions that are planned and cyclically adapted to the attainment of personal goals” (Zimmerman, 2000, p. 14). Since SRL strongly links feedback to learner motivation, affect and beliefs, scholars have called for further synthesis of SRL research with feedback research within diverse learning environments (e.g., Mory, 2004). The recursive nature of SRL means that feedback is a core impetus for SRL, contributing to learners’ regulation of subsequent cognitive engagement (Butler & Winne, 1995). Thus, our study is anticipated to fill a gap in the SRL field, and our findings can be applied to other SRL contexts.

To address the ongoing challenge experienced by PELE students in analysing their linguistic needs, the first resource we adopted in the PELE course was existing language testing. However, few models seemed suitable for our context, because most were developed for placement purposes rather than for SRL. In the field of language testing, there are four broad types of tests: proficiency, achievement, placement and diagnostic (Alderson et al., 2014). A true diagnostic test should provide feedback for three aspects (Hattie & Timperley, 2007): (a) the learner’s language goal, (b) the gap between a learner’s goal and their current level and (c) specific activities to achieve their goal. However, there remain very few truly diagnostic second and foreign language tests, and little research has been done on diagnostic English language testing (Alderson et al., 2014; Clark & Endres, 2021), especially for self-regulating English learners.

Nevertheless, we still tried out several models. For instance, the MASUS model (Bonanno & Jones, 2007) looked most promising, as it offers feedback enabling students to identify their strengths and weaknesses in discipline-specific academic writing. However, it is limited in our context due to its single focus on writing skills and the cost of teacher evaluation. In addition, although the traditional focus of university support for students’ ELP has been on academic writing, we have found that more than half of PELE students choose speaking as a goal when given autonomy (Kim, 2023, p. 607). Thus, few models were available for us to apply in our context in a sustainable way, both pedagogically and financially. We then introduced the Vocabulary Levels Test (Nation, 2022) and the Vocabulary Levels Test (Productive) (Laufer et al., 2022), as well as a diagnostic questionnaire specifically designed for PELE students. The questionnaire contains 20 situated questions: 10 for spoken communication and 10 for written

communication. Students rate their confidence or satisfaction on a Likert scale between 1 (*strongly disagree*) and 5 (*strongly agree*). The questions were organised, using Qualtrics software, in a sequence of areas where students can improve most effectively in a short period and experience the most tangible progress in communication. Questions for spoken communication are arranged in the following order: productive vocabulary (high-frequency words), academic vocabulary, English rhythm, pronunciation, spoken grammar, speaking confidence. Questions for written communication are arranged in the following order: reading strategies, academic vocabulary, academic writing structure, formal writing skills, nominalisation, grammar for academic writing. After answering a few questions, students are prompted with a suggestion. For example, if they have chosen “spoken”, they are first presented with statements related to vocabulary, such as:

I don't understand native speakers even though they don't speak too fast.
I am not able to have long conversations in English.
I feel very shy when speaking in English.

When students agree or strongly agree with any of these statements, they receive a suggestion such as “Hey, it looks like ‘expanding productive vocabulary (high-frequency words)’ would be a really good goal for your personal project. What do you think?”

Students can end the questionnaire here or proceed to the next question to obtain further suggestions. The questionnaire can help students to identify a broad missing piece (e.g., English pronunciation), but they cannot narrow it down further (e.g., the /θ/ sound). Therefore, we introduced some tools of AFSs in Term 3 (September – December 2021) to investigate whether these could be effective and sustainable tools for the goal-setting of PELE students.

Altogether, 93 students – 37 undergraduate, 5 postgraduate coursework students, 51 higher degree research (HDR) candidates – took the PELE course online (synchronously via Zoom and asynchronously via two learning management system platforms). Kim taught coursework students in a cohort and HDR candidates in a separate cohort.

AFSs

The tools in our project are necessarily wider in scope than in traditional research on automated writing evaluation and automated speaking evaluation. Thus, we adopted a general term – AFSs (e.g., Debusse et al., 2008). AFSs can be any mobile- or computer-assisted tool, such as Grammarly, which offers the potential to provide real-time information “with which a learner can confirm, add to, overwrite, tune, or restructure information in memory” (Butler & Winne, 1995, p. 275) about one's ELP. Language learning resources that do not offer feedback, such as YouTube videos, are not AFSs. AFSs are not necessarily designed for language learning: that is, Voice Analyst, an app for speech therapy, we adopted Voice Analyst, an app for speech therapy, for diagnosing intonation in the present study.

Although computer- and mobile-assisted language learning have always been promoted for students in PELE, we introduced additional AFS tools under this project, with a new pedagogical emphasis on using them for diagnostic purposes. We identified some promising new tools based on a literature review and Internet search. After trying out the AFS tools ourselves, we introduced several tools in the second lecture and encouraged students to consider adopting them to identify their “missing piece” in the Awareness and Planning stages and to use them for regular progress measurement in the Implementation stage and final progress measurement in the Reflection stage. We encouraged students to adopt AFS tools for their own language assessment. However, the scores generated by such tools did not directly impact the students' final grades. This use of AFS tools for formative feedback (Fang et al., 2023) aligns with recommendations for limiting automated assessment tools to low-stakes settings (Han & Lu, 2021; Ranalli et al., 2017).

Table 1
AFSs used by learners in this study

Name of AFS	Targeted macro skill	Cost	Type of quantitative feedback	Type of qualitative feedback
PELE Diagnostic Questionnaire	All	Free	N/A	Suggested learning goals
Vocabulary Levels Test (Nation, 2022), Vocabulary Levels Test (Productive) (Laufer et al., 2022)	None	Free	Vocabulary size and level	N/A
Grammarly (https://www.grammarly.com)	Writing	Free/Paid	Task score; originality score; word count metrics; readability metrics; vocabulary metrics	Error detection; Error analysis; Suggested corrections
PaperRater website (https://www.paperrater.com)	Writing	Free/Paid	Task score; sub-scores; originality score	Error detection; error analysis; suggested corrections
ELSA (https://elsaspeak.com/en/)	Speaking	Free/Paid	Diagnostic test: Task score and an equivalent IELTS speaking score; sub-scores Regular exercises: Score for the proximity to a native speaker	“Try again” feedback or verification feedback; emojis; error detection; error analysis; self-recording; model answer; speaking tips; video lessons
Speech Rate Meter (https://play.google.com/store/apps/details?id=by.intoncore.SpeechRateMeter&hl=en&gl=US)	Speaking	Free	Calculated speech rate per minute	N/A
Speechace website (or Englace app) (https://www.speechace.com)	Speaking	Free	Task score and an equivalent IELTS speaking score; Speech rate per minute; Number of bad pauses	Error detection; error analysis; self-recording; model answers
IELTSAce app (https://ieltsace.com/s/ielts/)	Speaking	Free/Paid	An IELTS speaking score; IELTS speaking score by sentence	Error detection; error analysis; self-recording; model answers

Liulishuo (流利说) app (https://m.liulishuo.com/liulishuo.html)	Speaking	Free/Paid	Diagnostic test: Overall band; estimated IELTS speaking score; sub-scores Regular exercises: Task score and sub-scores	Generic band descriptions; Video lessons
MaiMemo (墨墨背单词) app (https://www.maimemo.com)	N/A (Vocabulary)	Free	Vocabulary size; tracking of learning behaviours & achievements	N/A
Duolingo English Test (https://englishtest.duolingo.com/applicants)	All	US\$49	Overall score; Four sub-scores	Generic score descriptions
Versant English Placement Test (https://versantstore.pearson.com/p/VERSPT-EPT)	All	US\$15	Overall score; sub-scores; typing speed & typing accuracy	Generic band descriptions; tips on improvement
Voice Analyst app (https://apps.apple.com/us/app/voice-analyst-pitch-volume/id732245213)	Speaking	US\$18	Pitch and volume	N/A
SpeakingPal app (https://apps.apple.com/us/app/speakingpal-speak-english/id437219346)	Speaking	Free/Paid	Task score out of three stars	Error analysis; self-recording; model answers; video lessons
Write and Improve with Cambridge (https://writeandimprove.com)	Writing	Free	Band from A0 to C2; task validity score; sentence quality band; tracking of learning behaviours; completion time	Generic band description; error detection and labelling

We also created an AI Club as one of the voluntary PELE social clubs that promote out-of-class language learning and bonded groups (Senior, 1997). We ran this four times, in Weeks 4–8, for those who used any AFS tool to exchange experiences. We also shared our experiences as AFS users. Students shared their positive experiences and challenges with automated feedback tools that we recommended but also inspired us with new tools and their critical insights. Consequently, the pool of AFS tools expanded, as summarised in Table 1.

As Table 1 shows, the tools differ in their targeted macro skills and cost to users. Three tools target all macro skills, three for writing, and eight for speaking. We did not include some well-researched automated writing feedback systems such as Criterion (e.g., Dikli & Bleyle, 2014), MY Access! (e.g., Chen & Cheng, 2008) and Pigaiwang (e.g., Zhang, 2020), because they appear more suitable for academic writing classes with prescribed topics. Furthermore, by including tools recommended by learners, we echo Chen and Cheng (2008) in studying tools that are already widely used. ChatGPT did not exist during the data collection though it has been introduced in PELE more recently with its great potential in offering language feedback (Lodge et al., 2023).

We introduced the tools for diagnostic purposes in the PELE context while considering the principles of diagnostic test feedback. Firstly, the PELE diagnostic questionnaire, ELSA, Grammarly and PaperRater provide feedback for learners' goal selection. The PELE diagnostic questionnaire suggests a macro goal, such as pronunciation or academic writing structure. On the other hand, ELSA, Grammarly and PaperRater guide learners to narrow down the macro skill to a sub-area, by providing sub-scores or sub-feedback. Secondly, some tools help learners to measure the gap between their goal and their current level. Thus, when learners already set a learning goal but wish to understand the gap between that and their current level, they can incorporate AFSs such as vocabulary tests, Speech Rate Meter (for speed) and VoiceAnalyst (for intonation). Thirdly, several tools provide exercises to help learners achieve their goals; In particular, ELSA and SpeakingPal provide targeted lessons to fill gaps in domain knowledge (i.e., pronouncing the /θ/ sound). Furthermore, tools such as ELSA and MaiMemo track progress, which may enhance learner motivation for self-regulated activities. Duolingo English Test and Versant English Placement Test provide the least diagnostic feedback. Although no single tool is a complete diagnostic test, we scaffolded learners to critically synthesise tools to obtain sufficient feedback across the three key aspects of diagnostic tests.

Data collection and analysis

We collected two sets of student self-reported data to answer our questions. We collected our primary data from seven students who participated in the study for a term, including their e-portfolios and focus group interviews (research questions 1–3). We also collected data from a survey ($n = 32$) to understand students' perceptions from larger samples (research questions 1 & 2). We explored pedagogical implications (research question 4) by triangulating the two data sets in relation to existing literature. Ethical clearance was given by the Human Research Ethics Committee at UNSW Sydney.

Focus groups interviews

We conducted two focus groups: one face-to-face and the other online due to logistical reasons. Seven international students from diverse backgrounds participated, as shown in Table 2. The focus groups fall within the recommended size of five to eight participants for an ideal focus group (Krueger & Casey, 2014, p. 82). We designed semi-structured questions to elicit learner reflection on their cognitive, behavioural and affective engagement (Ellis, 2010) and pedagogical suggestions. We employed thematic analysis (Braun & Clarke, 2021) to analyse interview transcripts in NVivo to identify common experiences bearing stronger pedagogical implications. Subsequently, we visited their e-portfolios to identify the stages of AFS incorporation in the phases of the PA model.

Table 2
Focus group participants

Participant's pseudonym	Country of origin	Gender	Level of education	Focused macro skill
Ali	Iran	Male	HDR	Speaking
Tangmu	China	Male	HDR	Writing
Aaron	Germany	Male	HDR	Speaking
Yuvani	Sri Lanka	Female	HDR	Speaking
Haiyue	China	Female	Undergraduate	Writing
Yuxi	China	Female	HDR	Speaking
Rashida	Egypt	Female	HDR	Writing

Survey

We invited students to take the survey to see whether they used any AFS tools for their project. The survey was included in the Week 9 module of PELE's main asynchronous learning management system platform for students to complete voluntarily, outside of class time. We made the survey voluntary because using AFS tools was not compulsory and therefore not everyone used them. In addition, we wanted to give students an opportunity to opt out of the study, as they were already invited to participate in a different study in the term. Out of the cohort of 93 students, 32 valid survey responses were recorded. Due to this small sample size, the survey findings cannot be generalised but will be used to complement the focus group findings only. The survey asked students to indicate, on a 5-point Likert scale, the level of usefulness of the Table 1 tools they have used. We analysed the survey results using SPSS to run a one-way ANOVA to compare mean ratings for each AI tool.

Findings

In this section, we present key findings from the focus group interviews, an e-portfolio analysis with an example of how AFSs were incorporated into the personal projects and the survey results.

Focus group interviews

Overall, the interview participants found the chosen tools valuable and satisfactory. However, interview participants also critically evaluated the limitations of various tools and highlighted the continuous need for human feedback and critical thinking. To maximise reliability in identifying common themes, we revisited the common nodes (identified in the transcripts of 5 or more out of the 7 participants) after coding to ensure that the agreement is supported with clear textual evidence. Based on common nodes, three themes emerged as follows:

- benefits from accuracy, improved competence, positive emotion and safety
- limitations in validity, reliability and emotion
- critical engagement with AFS tools.

Theme 1: Benefits from accuracy, improved competence, positive emotion and safety

Having collectively incorporated a wide range of AFS tools, each of the focus group participants showed satisfaction with at least one AFS tool. Participants improved their targeted aspect of English communication through the tool(s), with ELSA (5/7) and Grammarly (4/7) being the most discussed tools overall.

There appears to be three main reasons for their high level of satisfaction: perceived accuracy, helpfulness in improving competence and contribution to positive emotions. Firstly, most participants (6/7) explicitly evaluated their AFS feedback and, especially, numerical scores as accurate, consistent and helpful in reflecting their English skills:

When you took this free test and got 6.5 Or B2, you felt that, oh yeah, I think that's an accurate reflection of my writing. (Haiyue, on Write and Improve with Cambridge)

Those comments, mostly, I felt that it's actually matching with humans. (Yuvani, on ELSA)

This strong agreement on the accuracy of AFS tools does not mean that participants accepted AFS feedback uncritically; in fact, all questioned the accuracy, as explored under Theme 2.

Participants also appreciated numerical scores as quantitative evidence of progress that satisfied their intrinsic motivation to know their progress. For example, Yuvani commented on ELSA:

For others, it may be different. Because we are from computer science and engineering backgrounds, we want that specific number for everything. So, it satisfied my internal voice.

Secondly, all participants (7/7) indicated that the tools helped them to improve competence, whether in pronunciation, IELTS, presentation skills, or daily communication. For example, Aaron commented:

I plan to continue with ELSA because of the overall feedback in my real life and especially that I can now pronounce my own name. That's just something, before the PELE, really deemed to be impossible, and I kind of gave up on myself. And yet, the outcome is just so surprisingly positive.

By practising the /r/ sound on ELSA, Aaron learnt to pronounce his name intelligibly to English speakers so that, he noted, baristas could recognise his name. This led to improved confidence and self-efficacy in overcoming a task previously deemed impossible.

Similarly, Haiyue, an undergraduate engineering student, improved her IELTS writing from 6.5 (or B2) to 7 (or C1) by using the website Write and Improve with Cambridge:

I only got a like, B2. And that's equivalent to 6.5 on IELTS. And I was not really satisfied with this result. That's why I decided to start an academic writing task in PELE project ... After the five weeks of PELE project, I got a C1, which is equivalent to 7 in the IELTS exam. So I'm pretty satisfied with this result; at least 7 is something. And I think I can keep using this website to improve my writing.

It should be noted that AFS tools helped Aaron and Haiyue not only to make progress but also to specifically measure it quantitatively. Both learners demonstrate satisfaction with their progress, achieved with the assistance of AFS tools, and high motivation to continue their tool use, as they believe they can make further improvements, suggesting improved self-efficacy.

The final reason is the positive emotions experienced by most participants (6/7) in engaging with AFS tools. Positive affect can be attributed to progress and improved competence, as in the case of Haiyue and Aaron. It can also be attributed to inherent features of certain AFS tools, such as gamification. For example, Ali, who explored a wide range of AFS tools, commented:

And maybe using the apps and being happy to learn has a little contribution to my emotion and psychological status.

It is also worth mentioning that most participants (5/7) considered AFSs as a safe place to seek feedback and improve their communication skills so that they could become more confident in communicating with humans. For example, Rashida commented:

Despite having AI to be your examiner, it's difficult somehow; it's stressful somehow. However, this level of stress made me able to interact with people more comfortably.

Rashida argued that the stress induced by AI evaluation is tolerable and beneficial, enabling her subsequently to communicate with humans with less stress. This perception of AI tools as a safer place to

prepare for human interaction is strongly echoed by Yuvani, who said, “I feel it’s [ELSA] like a friend to me”.

In short, all participants rated their experience with AFS tools as satisfactory due to perceived accuracy, helpfulness in improving competence and positive emotions thanks to safety, progress and gamification.

Theme 2: Limitations in validity, reliability and emotion

Although focus group participants perceived AFSs as accurate and reliable, all participants identified the limitations of AFSs and stressed the danger of blindly accepting all feedback. They also raised concerns about the validity and reliability of automated feedback. Validity issues include incorrect feedback and scope and insufficient feedback. Ali, a PhD candidate in computer science, questioned the size of data that powers PA:

We know that the apps were written in computer languages using machine learning and deep learning. And we know that this processing power that we have at this time, and the amount of data for training at this stage of our life is not very intelligent. But it's a good point to start and have as supplementary to other ways that we learn languages.

Ali’s caution about the intelligence of AI was widely shared. One reason is that AFS feedback can be invalid. For instance, Haiyue criticised Write and Improve with Cambridge and Grammarly for providing nonsensical detections of incorrect words in her writing. In addition, the scope of feedback can be incorrect when AFS tools disregard users’ intended English variety or context. For instance, participants expressed frustration with ELSA Speak and Liulishuo for penalising features of Australian English against standard American English. Likewise, participants were unsatisfied with inappropriate correction for the intended context of communication, as some learners expressed the wish to teach the machine discipline-specific norms, an area for which generative AI holds great potential.

Furthermore, even correct feedback can be limited in its scope. Some participants cautioned that the narrow scope of feedback can be misleading. For example, in a pronunciation drill on ELSA, the absence of feedback on other aspects of spoken English, such as intonation, may mislead learners to neglect aspects that work in tandem with articulatory pronunciation in authentic communication. This means that a lack of automated feedback should not be treated as meaning there is no room for improvement. Likewise, Tongmu highlighted some limitations in Grammarly feedback for his email writing by comparing Grammarly and a piece of teacher feedback: because Tongmu’s email asked a peer mentor for a last-minute favour, the teacher suggested adding an apology to soften his demand, “I’m sorry for the last-minute request ... I totally understand it if you are not available”. Tongmu appreciated the teacher’s comment and argued that Grammarly does not offer helpful feedback on politeness and intercultural communication, even though Grammarly Premium offers eight target tones, including confidence and friendliness.

In addition to invalidity, participants also noted inconsistency. For instance, Tongmu lamented that Grammarly does not consistently identify the same error depending on the text length and that Grammarly’s feedback for structure is less consistently reliable than that for grammar.

The lack (or absence) of emotions was also identified as a limitation. For instance, Yuxi, the only participant to limit AFS tools to the initial phase, highlighted the lack of emotion in such tools as a key reason for her preference for human interaction:

After the first two weeks, I just stopped using the tools and used more time to practice with real people ... the software ... they don't have any emotion. And also, when I imitate the TED talker speech, I can't copy their emotion. So after using the tools, I found that my imitation sounds a little bit artificial. And that's also why I stopped using that tool.

In short, the participants demonstrated a keen awareness of the validity, reliability and emotional limitations of automated feedback.

Theme 3: Critical engagement with AFS tools

Participants demonstrated critical engagement with AFS tools' feedback on cognition, behaviour and emotion. Participants not only questioned the validity and reliability of automated feedback but also exercised autonomy and agency in selectively accepting feedback. A key strategy to verify automated feedback is integrating it with human feedback. This synthesis was explicitly encouraged and scaffolded in the PELE course. Almost all participants (6/7) sought human feedback, with some citing a deliberate effort to seek feedback from a PELE teacher, a PhD supervisor or friends, and others pointing to serendipitous feedback. This synthesis empowered participants to make an informed decision to embrace an AFS tool despite recognising its shortcomings. For instance, Aaron synthesised ELSA with human feedback as he realised that optimising scores on AFS tools does not equate to improved communication with humans:

I know that I can do very good in the [ELSA] exercises now with the sounds ... However, I can't really tell myself how much are implicit in daily life, and that's where I need the human input, because the AI doesn't observe me in daily life.

Another strategy to mitigate the limited scope of automated feedback is the restricted use of AFS tools. Most learners (5/7) demonstrated sensitivity about the context of communication, perceiving certain AFS tools as being more helpful or valid in either an academic or everyday context. This awareness of contextualising AFS tools was enhanced in the PELE course, which guides students to analyse their linguistic needs within context (Halliday, 1978).

Once participants' self-efficacy for self-assessment improved, some exercised agency to compare themselves to native English speakers with AFS tools. An excellent example of this creative usage of AFS tools is Tongmu, who incorporated Grammarly as the main feedback source for email writing. Before PELE, Tongmu would uncritically accept Grammarly suggestions but noticed a deterioration in his writing and spelling ability. Encouraged by PELE teachers to question feedback, Tongmu transformed himself into a critical user of Grammarly. Curious about the gap between himself and native speakers, he applied Grammarly to evaluate the emails by native English speakers and institutions and found errors in them, leading to this comment:

And [I] also analysed some emails from other people, like different organisations and other people's writing ... Everybody needs it [Grammarly].

The knowledge that even native speakers err altered Tongmu's belief. Previously, he faced pressure to write perfect English like a native speaker with the belief that only international students such as him need to improve their English communication. Now, he felt more confident with improved skills after freeing himself from the illusion of perfection.

Another important critical engagement is emotional self-regulation. Although tools with gamification features contributed to more positive emotions towards learning, some participants experienced frustration. Yuvani offered an excellent example of emotional self-regulation with a "mix-and-match" strategy of using different speaking AFS tools:

"I was struggling with the J sound. But ELSA Speak was only proposing exercises on that sound again and again; I felt so frustrated at the moment, because I was already struggling, and it was making me struggle more. And if I felt so down or emotionally weak by some kind of sound, I usually changed to Speechace app, because compared to the ELSA app, SpeechAce app will give a better score, because the Speechace app focuses on sentence level, not just one word. That's how I manage the emotional imbalance in that case."

Emotional self-regulation appears to be a long-term impact of PELE on students (Jing, 2024). PELE guides students to reflect on their emotions throughout their learning process and record these in their Logbook & Journal and encourages them to use these as data to understand and accept themselves. Although self-regulation is a significant predictor of student engagement (J. C. Y. Sun & Rueda, 2012), paradoxically, emotional self-regulation becomes more challenging when students engage with automated feedback than with teacher feedback (Zhang, 2020). Yuvani combined, on the one hand, AFSs that provide more micro feedback with potentially lower scores, and on the other hand, AFSs that provide more macro feedback with encouraging scores. This innovative synthesis shows successful emotional self-regulation and adaptability based on knowledge of software affordances.

E-portfolios

We analysed the focus group participants' project e-portfolios to gain insight into their use of AFSs in various stages of the PA cycle. The e-portfolio is an assessment task that includes (a) Personal Project Design, where students detail the project goal and plans for language enhancement, including self-assessment; (b) Logbook & Journal, where they document their learning activities and ongoing reflections; and (c) Final Reflection, where they provide self-generated evidence of language enhancement. The review of the e-portfolios reveals that most focus group participants, except one, incorporated AFS feedback in all phases of the PA model.

Now, we present an example case to show how AFS tools were used in all phases of the PA model. Aaron, a PhD candidate from Germany, is one of those who utilised AFS feedback in all phases of the PA model. Firstly, he found his "missing piece" by synthesising feedback from his PELE tutor and ELSA, both of which point to the /r/ sound as the most significant potential missing piece in his pronunciation. Aaron's Logbook & Journal entries show active incorporation of ELSA both as a source of explicit knowledge (i.e., a detailed explanation about how to make the /r/ sound in English) and a major tool for practising and evaluating pronunciation. Aaron again adopted ELSA for his final progress measurement, which showed progress in pronouncing the /r/ sound, echoing the positive human feedback he sought. Aaron's active engagement suggests that AFS feedback can be incorporated as a key component in all phases of SRL as a form of technology-assisted self-assessment.

Survey results

Survey results also show that AFS tools were perceived as useful, with 13 out of 14 of these tools being rated above 3/5, as shown in Figure 1; although only one tool received an average rating below 3/5 (Versant English Placement Test). Together with the major findings from the qualitative data, these positive results suggest that AFS tools have a great potential to be used as effective learning tools in SRL environments.

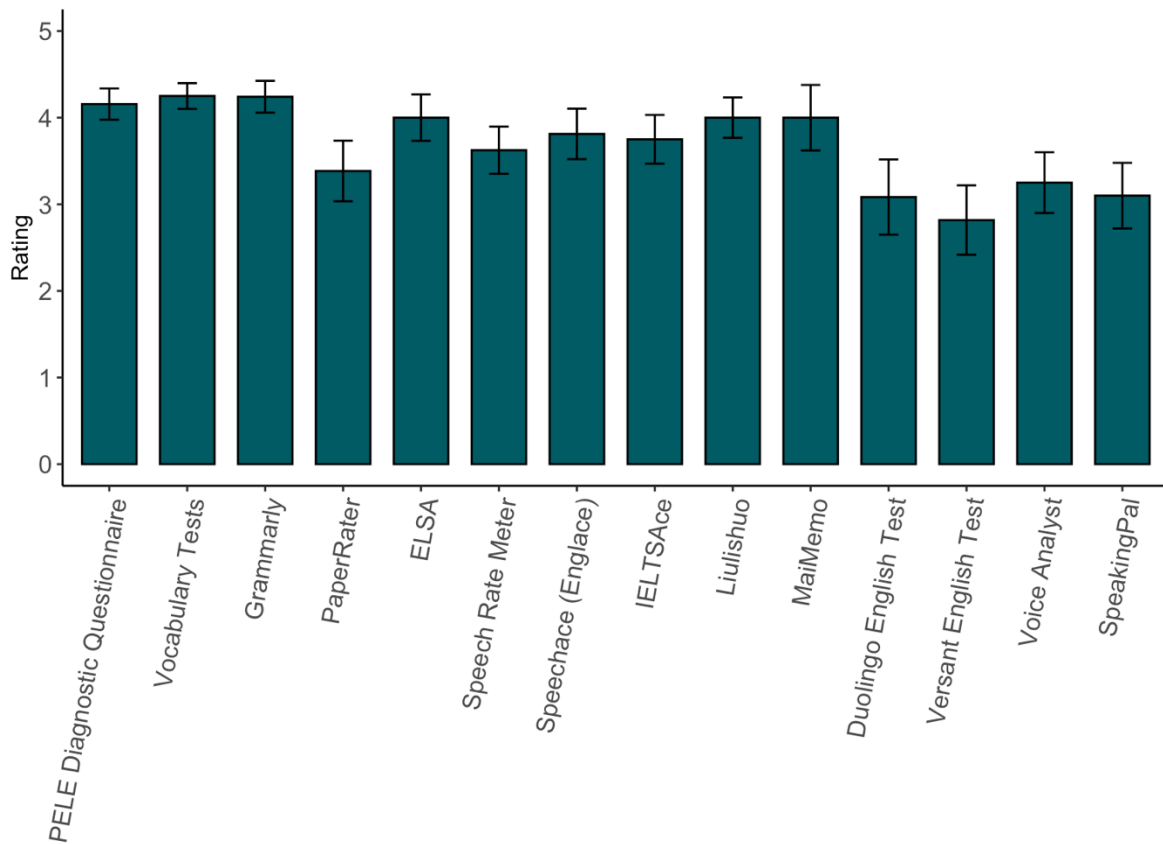


Figure 1. Mean score and standard error of 14 AFSs

Table 3 lists the top six best-rated tools. The two AFSs that PELE students used in class – Vocabulary Levels Tests and the PELE Diagnostic Questionnaire – were highly rated. Among the other tools, Grammarly and ELSA were the most popular and highly rated, warranting more teacher and researcher attention. These findings cannot be generalised in any way but can be useful information for those who want to introduce AFS tools for self-regulating English learners.

Table 3
Top six best-rated AFSs

Name of AFS	N (max = 32)	Rating	SE
Vocabulary Levels Test & Vocabulary Levels Test (Productive)	32	4.2500	0.14892
Grammarly	29	4.2414	0.18339
PELE Diagnostic Questionnaire	32	4.1563	0.18022
ELSA app	18	4.0000	0.26813
Liulishuo (流利说) app	11	4.0000	0.23355
MaiMemo (墨墨背单词) app	8	4.0000	0.37796

Discussion

Our data clearly show that those who participated in this case study experienced constructive engagement with AFS tools for their self-regulated English learning. These students utilised these tools throughout the PA learning cycle, as an effective guide to identify areas of improvement and measure their ongoing progress. They perceived feedback as a reasonably accurate language assessment, beneficial not just for their English skill development but also for self-efficacy. Learners' perceived self-improvement is critical because self-efficacy significantly contributes to a sense of learner agency

(Zimmerman, 2000) and sustains learners' interest, motivation and performance (Panadero et al., 2017). Learners also appreciated the instant and unlimited access to quantifiable feedback, as they used this to indicate their progress as frequently as needed. They found this to be safer and less intimidating than seeking human feedback. Having improved their communication using the automated feedback, students then reported confidence in approaching humans for authentic communication and seeking additional feedback, to mitigate the shortcomings of AI feedback. This seems to be a virtuous cycle in which learners are more motivated to seek human feedback once they have engaged with automated feedback. Most importantly, students effectively navigated through potential challenges and weaknesses of AFS tools with critical cognitive, behavioural and emotional strategies.

What we found most encouraging in introducing AFS tools is that all focus group participants were willing to continue to use them for their ongoing English language enhancement beyond the course. Considering that the AFSs we recommended are publicly available (albeit at a cost to purchase some tools), it is worth noting that the pedagogical impact of AFS tools introduced in one term can extend beyond the constraint of the course length, promoting sustainability, autonomy and self-efficacy for post-PELE language enhancement.

A caveat should be given that the positive perception of and engagement with AFS tools we found are contextualised within the PELE course. Firstly, and most importantly, PELE students choose their own communicative goals and AFS tools, which differs from most studies of automated writing evaluation. Secondly, although PELE students use AFSs in the low-risk setting of language self-assessment, they are encouraged to use these tools to enhance the validity, reliability and availability of such self-assessment as required for various assessment tasks in this credit-bearing course. This assists autonomous learners with extrinsic motivation to adopt AFSs. Thirdly, focus group participants' competence in critically accepting AFS feedback has been encouraged by PELE's scaffolding of obtaining and critically synthesising feedback from multiple sources (i.e., self, friends and teachers). Finally, PELE scaffolds emotional self-regulation, emphasising reflection and lifelong learning. Thus, the course structure and teacher's guidance are both crucial in helping students to develop critical engagement strategies such as the context-sensitive and goal-oriented choice of AFS tools, "shopping around", questioning and selective acceptance, complementing AFS feedback with human feedback and emotional self-regulation and adaptability. These pedagogical practices in PELE, aiming to enhance autonomy, competence, relatedness and emotional self-regulation, may have contributed to the more satisfactory and critical cognitive, affective and behavioural engagement with AFS tools, as reported by our participants. Furthermore, a sense of isolation and reduced availability of in-person human contact, brought about by the COVID-19 pandemic restrictions in Sydney at the time, may have resulted in the availability of AFS feedback being evaluated more positively.

We propose the following suggestions for contexts beyond PELE where educators wish to support students' ELP with more effective and sustainable means such as AFSs. Firstly, when scaffolding the use of automated feedback, educators should encourage learners to diversify feedback by complementing AFSs with formal (i.e., from teachers) and/or informal (i.e., from friends) human feedback. This synthesis is beneficial for developing students' critical thinking, because AFS feedback can be incorrect (O'Neill & Russell, 2019) and correct AFS feedback may be incorrectly dismissed (Koltovskaia, 2020). Secondly, instead of assuming that learners know how to interpret automated feedback and regulate their emotional engagement, we strongly echo Zhang (2020) on the importance of teacher scaffolding of how to respond to AFS feedback emotionally and cognitively. For instance, teachers should scaffold and model reflective learning where one constantly modifies one's learning behaviours through measuring regular progress and analysing one's emotional reaction to AFSs. Given the prevalence of automated tools for learning and practising spoken English, we encourage instructors and researchers to go beyond automated essay scoring and conduct further research into the burgeoning field of speaking AFS tools. Finally, it is paramount that educators adopt a positive outlook towards AFSs for pedagogical efficacy, given the impact of instructor perception on learner perception (Chen & Cheng, 2008; P. C. Sun et al., 2008). In our context, we consider that Kim's transformation of her personal experimentation with AFS tools into pedagogical materials contributes to more positive learner perception.

Conclusion

We conclude this paper with additional suggestions from students for PELE teachers, which might also be helpful for others who consider using AI-based assessment in wider educational contexts. All respondents agreed that there is no one-size-fits-all tool: the most suitable tool depends on a myriad of factors, including the learner's goal, the current level of communication and even the learner's mobile phone operating system. Respondents unanimously agreed that teachers should continue introducing AFS tools at the start of the course. The introduction should be clear and brief, systematising all such tools based on the targeted linguistic goal, benefits and limitations and by drawing on previous learner experience and research findings. Teachers can also cultivate learner agency by scaffolding past students' critical engagement with AI tools. Students also advised teachers to encourage more students to try AFS tools, with three participants suggesting more "obligatory" activities, such as in-class activities for all learners to try these tools. Teachers can co-construct the learning experiences with their current student cohort by inviting them to share their experiences with the tools covered in the present study and beyond.

Most respondents also cautioned that low digital literacy (i.e., in "their parents' generation"), older mobile phones and slow Internet might cause struggles navigating AFSs, even though they perceive most AFSs as being user-friendly. In addition, not all learners can access paid tools. To navigate these hurdles, we suggest that institutions and teachers provide additional support to self-directed learners who require technical support or cannot afford it. Furthermore, one participant cautioned that teachers could set the right expectations by scaffolding the potentially different levels of availability and features of AFS tools on Android versus iOS devices.

With the advancement of educational technology, AFS tools can be effective for learners to use in a particular course or as ongoing, lifelong learning tools. This is particularly important considering the major issues with the availability and sustainability of professional feedback, such as the cost of and waiting time for human evaluation. Instructors should provide a safe and supportive environment where even students with lower digital literacy feel competent to use these tools. It is also important for instructors to critically introduce the AFS tools in terms of both their strengths and weaknesses, create opportunities and incentives for learners to try them, encourage peer sharing, monitor learner engagement to provide timely support and scaffold emotional self-regulation, to balance potential weaknesses of these tools. Moreover, instructors should encourage self-regulating learners to complement AFS feedback with human feedback when limitations are identified. Further research is required to investigate sustainable ways of complementarity. This also includes investigating teachers' responses to learners' perceptions of the (in-)accuracy of feedback from AFS tools. We strongly encourage more research into learner and teacher perception of and engagement with the tools, beyond the prolific works exploring the validity and reliability of AI assessment.

A limitation of this study is the relatively small sample size of the voluntary survey, which potentially over-represents motivated learners. Further research is needed to understand learners who do not, or are less motivated to, engage with AFSs, which is particularly important in contexts where the use of AFSs is compulsory.

Author contributions

Long Li: Conceptualisation, Literature review, Data collection and analysis, Writing – original draft, Writing – review and editing; **Mira Kim:** Conceptualisation, Writing – original draft, Writing – review and editing.

References

- Alderson, J. C., Haapakangas, E.-L., Huhta, A., Nieminen, L., & Ullakonoja, R. (2014). *The diagnosis of reading in a second or foreign language*. Routledge.
- Arkoudis, S., Baik, C., & Richardson, S. (2012). *English language standards in higher education: From entry to exit*. ACER Press.

- Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review*, 84(2), 191–215. <https://doi.org/10.1037/0033-295X.84.2.191>
- Benzie, H. J. (2010). Graduating as a 'native speaker': International students and English language proficiency in higher education. *Higher Education Research & Development*, 29(4), 447–459. <https://doi.org/10.1080/07294361003598824>
- Bonanno, H., & Jones, J. (2007). *The MASUS procedure: Measuring the academic skills of university students: A diagnostic assessment*. <https://www.sydney.edu.au/content/dam/corporate/documents/engage/industry-and-business/consultancy-services/masus-procedure.pdf>
- Braun, V., & Clarke, V. (2021). *Thematic analysis: A practical guide*. Sage.
- Butler, D. L., & Winne, P. H. (1995). Feedback and self-regulated learning: A theoretical synthesis. *Review of Educational Research*, 65(3), 245–281. <https://doi.org/10.3102/00346543065003245>
- Cavaleri, M. R., & Dianati, S. (2016). You want me to check your grammar again? The usefulness of an online grammar checker as perceived by students. *Journal of Academic Language and Learning*, 10(1), A223–A236. <https://journal.aall.org.au/index.php/jall/article/view/393>
- Chen, C.-F. E., & Cheng, W.-Y. E. C. (2008). Beyond the design of automated writing evaluation: Pedagogical practices and perceived learning effectiveness in EFL writing classes. *Language Learning & Technology*, 12(2), 94–112. <https://doi.org/10.125/44145>
- Clark, T., & Endres, H. (2021). Computer-based diagnostic assessment of high school students' grammar skills with automated feedback—an international trial. *Assessment in Education: Principles, Policy & Practice*, 28(5-6), 602–632. <https://doi.org/10.1080/0969594X.2021.1970513>
- Debusse, J. C., Lawley, M., & Shibl, R. (2008). Educators' perceptions of automated feedback systems. *Australasian Journal of Educational Technology*, 24(4), 374–386. <https://doi.org/10.14742/ajet.1198>
- Dikli, S., & Bley, S. (2014). Automated essay scoring feedback for second language writers: How does it compare to instructor feedback? *Assessing Writing*, 22, 1–17. <https://doi.org/10.1016/j.asw.2014.03.006>
- du Boulay, B., Mitrovic, A., & Yacef, K. (Eds.). (2023). *Handbook of artificial intelligence in education*. Elgar. <https://doi.org/10.4337/9781800375413>
- Ellis, R. (2010). Second language acquisition, teacher education and language pedagogy. *Language Teaching*, 43(2), 182–201. <https://doi.org/10.1017/S0261444809990139>
- Evanini, K., & Zechner, K. (2019). Overview of automated speech scoring. In K. Zechner & K. Evanini (Eds.), *Automated speaking assessment* (pp. 3–20). Routledge. <https://doi.org/10.4324/9781315165103>
- Fang, Y., Roscoe, R. D., & McNamara, D. S. (2023). Artificial intelligence-based assessment in education. In B. du Boulay, A. Mitrovic, & K. Yacef (Eds.), *Handbook of artificial intelligence in education* (pp. 485–504). Elgar.
- Grimes, D., & Warschauer, M. (2010). Utility in a fallible tool: A multi-site case study of automated writing evaluation. *The Journal of Technology, Learning and Assessment*, 8(6), 1–43. <https://ejournals.bc.edu/index.php/jtla/article/view/1625/1469>
- Halliday, M. A. K. (1978). *Language as social semiotic: The social interpretation of language and meaning* (Vol. 42). Edward Arnold.
- Han, C., & Lu, X. (2021). Can automated machine translation evaluation metrics be used to assess students' interpretation in the language learning classroom? *Computer Assisted Language Learning*, 36(5-6), 1–24. <https://doi.org/10.1080/09588221.2021.1968915>
- Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, 77(1), 81–112. <https://doi.org/10.3102/003465430298487>
- Hoang, G. T. L., & Kunnan, A. J. (2016). Automated essay evaluation for English language learners: A case study of MY Access. *Language Assessment Quarterly*, 13(4), 359–376. <https://doi.org/10.1080/15434303.2016.1230121>
- Jing, B. (2024). *A longitudinal study on the long-term impact of the Personalised English Language Enhancement course* [Unpublished doctoral dissertation]. The University of New South Wales.
- Kim, M. (2014). Action research on advanced bilingual enhancement in translator education. In K. Kunz, E. Teich, S. Hansen-Schirra, S. Neumann, & P. Daut (Eds.), *Caught in the middle: Language use and translation* (pp. 195–213). Saarland University Press.

- Kim, M. (2023). A personalised autonomous model to resolve a prolonged dilemma in international students' English language needs in higher education. *Higher Education Research & Development*, 42(3), 603–618. <https://doi.org/10.1080/07294360.2022.2105823>
- Kim, M., & Jing, B. (2019). A personalised autonomous model for enhancing translation students' linguistic competence. In M. Koletnik & N. Froeliger (Eds.), *Translation and language teaching: Continuing the dialogue* (pp. 127–146). Cambridge Scholars Publishing.
- Koltovskaia, S. (2020). Student engagement with automated written corrective feedback (AWCF) provided by *Grammarly*: A multiple case study. *Assessing Writing*, 44, Article 100450. <https://doi.org/10.1016/j.asw.2020.100450>
- Krueger, R. A., & Casey, M. A. (2014). *Focus groups: A practical guide for applied research*. Sage Publications.
- Laufer, B., Nation, P., & Cobb, T. (2021). *Vocabulary Levels Test (Productive)*. Retrieved January 4, 2024, from <https://www.lexutor.ca/tests/levels/productive/>
- Lesgold, A. (2023). Foreword. In B. du Boulay, A. Mitrovic, & K. Yacef (Eds.), *Handbook of artificial intelligence in education* (pp. xvi–xviii). Elgar.
- Lodge, J. M., Thompson, K., & Corrin, L. (2023). Mapping out a research agenda for generative artificial intelligence in tertiary education. *Australasian Journal of Educational Technology*, 39(1), 1–8. <https://doi.org/10.14742/ajet.8695>
- Marginson, S., Nyland, C., Sawir, E., & Forbes-Mewett, H. (2010). *International student security*. Cambridge University Press. <https://doi.org/10.1017/CBO9780511751011>
- Mory, E. (2004). Feedback research revisited. In D. H. Jonassen (Ed.), *Handbook of research on educational communications and technology* (2nd ed., pp. 745–783).
- Nation, P. (2022). *Vocabulary Levels Test*. Retrieved January 4, 2024, from <https://www.lexutor.ca/tests/vlt/?mode=test>
- O'Neill, R., & Russell, A. (2019). Stop! Grammar time: University students' perceptions of the automated feedback program Grammarly. *Australasian Journal of Educational Technology*, 35(1), 42–56. <https://doi.org/10.14742/ajet.3795>
- Panadero, E., Jonsson, A., & Botella, J. (2017). Effects of self-assessment on self-regulated learning and self-efficacy: Four meta-analyses. *Educational Research Review*, 22, 74–98. <https://doi.org/10.1016/j.edurev.2017.08.004>
- Ranalli, J., Link, S., & Chukharev-Hudilainen, E. (2017). Automated writing evaluation for formative assessment of second language writing: Investigating the accuracy and usefulness of feedback as part of argument-based validation. *Educational Psychology*, 37(1), 8–25. <https://doi.org/10.1080/01443410.2015.1136407>
- Read, J. (Ed.). (2016). *Post-admission language assessment of university students*. Springer. <https://doi.org/10.1007/978-3-319-39192-2>
- Rochecouste, J., & Oliver, R. (2014). English language growth and the international student. *HERDSA Review of Higher Education*, 1, 63–81. <https://www.herdsa.org.au/herdsa-review-higher-education-vol-1/63-81>
- Senior, R. M. (1997). Transforming language classes into bonded groups. *ELT Journal*, 51(1), 3–11. <https://doi.org/10.1093/elt/51.1.3>
- Sun, J. C. Y., & Rueda, R. (2012). Situational interest, computer self-efficacy and self-regulation: Their impact on student engagement in distance education. *British Journal of Educational Technology*, 43(2), 191–204. <https://doi.org/10.1111/j.1467-8535.2010.01157.x>
- Sun, P. C., Tsai, R. J., Finger, G., Chen, Y. Y., & Yeh, D. (2008). What drives a successful e-learning? An empirical investigation of the critical factors influencing learner satisfaction. *Computers & Education*, 50(4), 1183–1202. <https://doi.org/10.1016/j.compedu.2006.11.007>
- Xi, X. (2010). Automated scoring and feedback systems: Where are we and where are we heading? *Language Testing*, 27(3), 291–300. <https://doi.org/10.1177/0265532210364643>
- Zhang, Z. V. (2020). Engaging with automated writing evaluation (AWE) feedback on L2 writing: Student perceptions and revisions. *Assessing Writing*, 43, Article 100439. <https://doi.org/10.1016/j.asw.2019.100439>

Zimmerman, B. J. (2000). Attaining self-regulation: A social cognitive perspective. In M. Boekaerts, P. R. Pintrich, & M. Zeidner (Eds.), *Handbook of self-regulation* (pp. 13–39). Academic Press.
<https://doi.org/10.1016/B978-012109890-2/50031-7>

Corresponding author: Long Li, klinterpreting@gmail.com

Copyright: Articles published in the *Australasian Journal of Educational Technology* (AJET) are available under Creative Commons Attribution Non-Commercial No Derivatives Licence ([CC BY-NC-ND 4.0](https://creativecommons.org/licenses/by-nc-nd/4.0/)). Authors retain copyright in their work and grant AJET right of first publication under CC BY-NC-ND 4.0.

Please cite as: Li, L., & Kim, M. (2024). It is like a friend to me: Critical usage of automated feedback systems by self-regulating English learners in higher education. *Australasian Journal of Educational Technology*, 40(1), 1-18. <https://doi.org/10.14742/ajet.8821>