Castledown

OPEN ACCESS

Technology in Language Teaching & Learning
ISSN 2652-1687
https://www.castledown.com/journals/tltl/

Technology in Language Teaching & Learning, 7(2), 103217 (2025)
https://doi.org/10.29140/tltl.v7n2.103217

Exploring Generative Artificial Intelligence-Aided Language Learning Tools in Language Classrooms: Insights from Language Teachers of English in Hong Kong

KA LONG ROY CHAN

Marymount University, Arlington, VA, USA
George Mason University, Fairfax, VA, USA
kchan25@gmu.edu, drroyklchan@gmail.com

Abstract

The integration of technology has profoundly influenced language teaching and learning practices. Emerging technologies such as generative artificial intelligence (GenAI) are poised to shift educational paradigms even further. This research examined the use of AI-aided learning tools from the perspective of language teachers. The study engaged teachers from different tiers of schools in Hong Kong, both secondary and tertiary institutions, as participants. These participants evaluated and provided feedback on their overall views on their use of GenAI-aided language teaching and learning. Through a qualitative grounded theory approach to the dataset of semi-structured interviews, this study aimed to understand language teachers' views of the application of AI tools in classrooms, as well as the overall teaching and learning experience, with consideration of the tiers of schools in which the teachers were employed. Preliminary findings suggested varied opinions on the practicality and impact of AI tools in language education, highlighting both potential benefits and challenges. These insights are crucial for developing innovative teaching strategies that align with contemporary educational needs as well as the continual call for advancement in teacher education. This study contributes to the ongoing discourse on the development of GenAI in language education, offering valuable implications for curriculum design and instructional practices in secondary education.

Keywords: GenAI; GenAI-aided teaching; language teaching; teacher education; English language teaching.

Introduction

Over the past few decades, every introduction of new technology has been considered a catalyst to bringing substantial changes to higher education, like the use of online learning tools and, of course, generative artificial intelligence (GenAI), which has drastically transformed the landscapes of various domains, including the teaching and learning of languages (Chan & Chan, 2025; McDonald et al., 2025). At the beginning of the application of GenAI in 2023, O'Dea (2023) questioned whether GenAI would bring a critical change to higher education teaching and learning and concluded that "it is still too early to say whether GenAI is or will be the paradigm shift for higher education" (p. 814). However, with GenAI tools having been around for almost three years now (at the time of writing), it has widely been regarded by educators as one of the most discussed tools in education in decades because of its versatility in a number of tasks, including language inquiries, writing assistance, providing feedback, and material preparation (Choi et al., 2025; Naznin et al., 2025; Tang, 2024; Yu & Xie, 2025).

There have been reviews on how GenAI has singlehandedly impacted the educational landscapes across disciplines. Among these disciplines, one of the most discussed areas of application is language-related learning and teaching, especially writing. Interestingly, the results so far are mixed. Teng (2024) conducted a systematic review of 20 articles on how ChatGPT affected English as a foreign language (EFL) writing and concluded that despite the numerous advantages ChatGPT could bring to EFL classrooms, the integration of ChatGPT in writing classes posed issues related to academic integrity. Despite these negative findings, Su and Tran (2024) investigated the limitations of applying ChatGPT in Masters' level research writing courses. They summarized that ChatGPT might have a negative impact on students' writing achievements. Similarly, Aksakallı and Daşer (2025) surveyed undergraduates in Türkiye and suggested that students held positive perceptions of the use of ChatGPT in their English writing and believed that it would help. Huang and Mizumoto (2024) discussed how ChatGPT could affect student motivation and the English writing efficiency of Japanese undergraduate students. They concluded that ChatGPT positively enhanced both student motivation and writing efficiency.

Apart from writing, another frequently researched application of ChatGPT is its effect on general language teaching and learning, especially on the other skills (reading, listening, and speaking). Solak (2024) surveyed both English learners and teachers and concluded that the interviewees agreed that the use of ChatGPT might benefit the teaching and learning of not only reading skills but also vocabulary development. Wang and Feng's (2024) experiment using ChatGPT as an assistant for helping students to read proved the advantages of using ChatGPT as an aid in reading – their results were positive in the success rate in both reading comprehension and reading efficiency – and concluded that "the adoption of ChatGPT-assisted reading methodology exerts a catalytic effect, enabling students to navigate original novels with heightened efficiency and proficiency" (p. 26).

Speaking and listening, however, unlike writing and reading, which have been studied widely, has not drawn a lot of attention from scholars. One of the reasons might be the fact that ChatGPT focuses on textual input – which limits its application for skills like listening and speaking (Songsiengchai et al., 2023). Nonetheless, the emergence of more GenAI tools with the technology of large language models and advanced voice recognition has prompted more specialized GenAI products dealing with specific skills. Chan and Liu (2025) proposed a classification system for these GenAI speaking tools into three main categories based on their functionalities; the categories are listen–mimic (e.g., Murf.ai), analyze–feedback (e.g., ELSA Speech Analyzer), and practice–feedback (e.g., SpeechAce). These GenAI tools are said to benefit teachers' and students' teaching and learning experiences of teaching pronunciation with flexible GenAI tools (Chan and Liu, 2025). Karjagdi Çolak (2024) examined the

integration of GenAI tools in language-intensive classrooms and found a significant improvement in students' speaking tasks with immediate, tailored responses for the students. Regarding listening skills, Yuniarti et al. (2024) examined how GenAI tools could be applied in the teaching and learning of listening. The research team used a combination of ChatGPT and an AI-voice model (VoxBox AI) to train students' listening skills. The study suggested that the combination of GenAI tools resulted in a 23% increase in listening scores, which hints at the benefits of incorporating GenAI in listening training (Yuniarti et al., 2024).

Several recent studies have examined how the integration of GenAI tools into language teaching affects teachers' professional identities, particularly through the lens of the ecological approach. As Van Lier (2004, cited in Satvati et al., 2025) suggested, the ecological approach provides a theoretical framework that views identity and individual development as shaped by multiple interconnected layers of social and environmental context. These layers include the microlevel (e.g., the immediate classroom environment), mesolevel (e.g., the institutional level, such as a school), and macrolevel (e.g., broader social and policy contexts). For example, Satvati et al. (2025) interviewed seven Iranian language teachers and found that GenAI integration positively influenced teachers' professional identities, particularly at the micro- and mesolevels. Similarly, Kamali et al. (2024) employed narrative interviews with 10 Iranian teachers to explore the challenges and benefits of using GenAI tools (especially ChatGPT) in language teaching. Their findings echoed those of Satvati et al. (2025), concluding that GenAI integration brought both opportunities and challenges across all three ecological levels. These insights provide valuable implications for teacher educators seeking to design responsive training and professional development in the age of GenAI.

The studies above suggested that after almost three years of having access to or using GenAI tools, researchers and teachers still have mixed reactions toward their integration into the teaching and learning process. As with the introduction of any new technology, the hesitation always comes less from the tools' utilities and more from the users' uncertainties and concerns about how they can apply the technologies – in the current case, GenAI tools. Despite growing interest in GenAI's integration in education, little is known about how language teachers will perceive and respond to these tools or the day-to-day challenge of attempting to implement them in real classrooms.

This qualitative study aimed to provide insights into the future application of GenAI in language teaching and to explore the challenges and frustrations experienced by language educators. These issues are particularly relevant for teacher educators hoping to understand how to better prepare teachers for GenAI-integrated classrooms. To this end, the study addressed the following central research question: What are the perceived benefits and challenges of using GenAI tools in language classrooms from the perspective of educators? Using a thematic analysis of interviews with 12 English teachers from various educational institutions in Hong Kong, this qualitative case study aimed to identify their attitudes toward GenAI tools and the significant concerns about their use in language classrooms. These insights are highly important for teacher educators in university programs to design courses that better prepare both preservice and in-service language teachers for the emerging challenges posed by GenAI technologies, which would also help teachers navigate the evolving GenAI landscape in education.

Methodology

The current qualitative case study employed semi-structured interviews to elicit data from 12 English teachers to discover language teachers' knowledge about the use of GenAI in language classrooms. The methodology is outlined in the following section.

Participants

Twelve English language teachers (five male, seven female) based in Hong Kong were recruited from secondary schools and universities across different tiers of the education system. The participant pool was intentionally diverse, representing a range of teaching experiences (ranging from novice teachers with less than 1 year of teaching experience to experienced teachers with over 10 years of experience) and institutional types (including different tiers from secondary schools and vocational tertiary institutes to research universities). By including educators from different educational contexts, it was hoped that the study would reveal how institutional settings might influence teachers' beliefs, practices, and professional identities. While this choice introduced a degree of heterogeneity into the sample, it also enriched the data by revealing both shared and divergent experiences across educational levels. Participants were selected via purposive sampling to ensure informed perspectives on technology-enhanced language learning, particularly the integration of GenAI tools in educational contexts. The use of purposive sampling assisted in distinguishing "key informants in the field who can help in identifying information-rich cases" (Suri, 2011). In this case, both teachers in secondary schools and universities have distinct experiences in using GenAI in teaching. Pseudonyms are used to protect the interviewee's identity. The demographics of the participants are listed in Table 1.

This study adhered to established research ethics protocols to ensure the rights and privacy of all participants. Prior to conducting the semi-structured interviews, informed consent was obtained from all 12 participating English teachers, with clear explanations of the study's purpose, procedures, and their right to withdraw at any time without consequence. Participants' identities were anonymized, and all data are securely stored and used solely for research purposes. Necessary ethical approval was obtained from the institutional review board, and care was taken throughout the research process to maintain transparency, confidentiality, and respect for the participants' professional and personal perspectives.

Table 1 *Demographics of Participants*

Participant*	Institute and banding	Age	Gender	Highest education level	Teaching experience
Anne	Secondary school	30–39	F	Master	10 years or more
Irene	Secondary school	30–39	F	Master	10 years or more
Leon	Secondary school	30–39	M	Bachelor	10 years or more
Mack	Secondary school	40–49	M	Master	10 years or more
Peter	Community college	40–49	M	Doctorate	10 years or more
Andy	Research university	40–49	M	Doctorate	10 years or more
Angel	Secondary school	20–29	F	Bachelor	Below 10 years
Christina	Secondary school community college	20–29	F	Master	Below 10 years
Cathy	Research university	30–39	F	Doctorate	Below 10 years
Erie	Vocational college	30–39	F	Master	10 years or more
Hilton	Secondary school	20–29	M	Master	Below 10 years
Kelly	Secondary school	20–29	F	Master	Below 10 years

*Pseudonyms are used to protect the interviewee's identity.

Data Collection

Data were collected through semi-structured interviews conducted in mid-2024. All the participants agreed and consented to the research before the start of the interview. Each of the participants answered

questions from the same prepared guide, which consisted of 10 open-ended questions designed to explore language teachers' perceptions and experiences regarding their use of GenAI-assisted teaching and learning. The questions focused on three core areas: (1) the participants' general ideas and attitudes toward GenAI-assisted language learning, (2) their personal experiences with using GenAI tools in teaching, and (3) their concerns, reservations, or perceived challenges related to the integration of such tools in the classroom. Semi-structured interviews provided a less rigid form of question protocol, which allowed more freedom for follow-up questions and interactions between the interviewer and the interviewee (Drever, 1995; Tyson, 1990). This freedom provides more information about an individual's beliefs and emotions, which is particularly beneficial for research in fields like education (Adamson, 2004). In order to prompt meaningful reflection and discussion, the interviewer showed the participants brief video demonstrations of selected GenAI-aided tools relevant to English language education. These videos, which included examples of using GenAI to support English listening and speaking instruction, served as stimulus materials to ground the participants' responses in concrete examples. The purpose of incorporating these materials was to ensure a shared understanding of the tools being discussed and to elicit more specific, experience-based reflections rather than generic opinions. The demonstrations were carefully selected according to their relevance to current pedagogical practices and their accessibility to teachers with varying degrees of technological familiarity. During each interview, the videos were shown early in the session, followed by targeted questions designed to prompt the participants' reactions, perceived usefulness, and concerns. The inclusion of this stimulus material is hoped to enhance the depth and focus of the data collected, encouraging participants to reflect critically on the pedagogical implications of GenAI and to draw connections to their own teaching contexts. Conducted by two qualified qualitative researchers, the interviews were conducted via Zoom and ranged from 30 to 45 minutes in length, resulting in over 400 minutes of recorded data. All the interview data were transcribed precisely by two other qualitative research assistants to ensure the accuracy of the transcription.

Data Analysis

Interview transcripts were analyzed by the researcher and two research assistants using a grounded theory approach to the dataset (Corbin & Strauss, 2015) using Nvivo 14.0. All of the transcriptions were skimmed for themes emerging from codes agreed upon by all the researchers. Researchers who searched for codes and themes emerging from the dataset were believed to be able to identify recurring patterns, concepts, and categories within the data (Patton, 2015). These themes were further refined through two rounds of comparison and memo-writing techniques, which are central to grounded theory. It was essential to begin directly with the datasets without having a predetermined set of codes of themes because, in grounded theory, "concepts are derived from data during analysis…[a]nalysis begins with the collection of those very first pieces of data...[c]oncepts derived from initial analysis guide collection of subsequent data" (Corbin and Strauss, 2015, p.15). The coding was conducted iteratively, with ongoing analytical memos used to track the themes' development and theoretical saturation across the 12 interviews. The analysis aimed not only to categorize teachers' perspectives but also to develop an interpretive understanding of how English language educators in Hong Kong perceived and responded to the emergence of GenAI in their professional contexts.

Findings

The datasets were collaboratively examined by the researcher and research assistants with the aim of uncovering the participants' perceptions of the use of GenAI in language classrooms. From the interview data, recurring codes emerged that informed the development of two overarching themes. Codes such as limited classroom integration, lack of clear pedagogical guidance, and trial-and-error use by teachers pointed to the immaturity of the applications of GenAI in language classrooms. Participants

frequently described their use of GenAI tools as experimental, noting a lack of institutional support or training. Meanwhile, codes like inaccurate grammar correction, overgeneralized feedback, and inconsistency across platforms contributed to the theme of concerns over the reliability of feedback and input from GenAI Tools. Teachers expressed hesitation in fully trusting these tools, citing frequent errors and the risk of misleading students. We will detail these two themes—the immaturity of the applications of GenAI in language classrooms and concerns over the reliability of feedback and input from GenAI tools—which are detailed below.

Immaturity of the applications of GenAI in language classrooms

When asked about their attitudes and experiences toward using GenAI in language teaching, all of the participants answered that they were still unsure about the position of GenAI tools in language education, especially on *when* and *how* to use GenAI in classes. Many of the participants who used GenAI in the classrooms were confused by the GenAI policies set by universities and, more importantly, were confused by the lack of a standardized use of GenAI across universities in Hong Kong. This confusion was indirectly evident in Dai et al.'s (2024) and Wang et al.'s (2024) summaries of how universities in the US and across Asia, at the beginning of the use of GenAI, adopted diverse approaches even within the same regions, including the acceptance in usage and the alignment to curriculum.

A common theme expressed by the language teacher participants was their concern about immaturity when applying GenAI in language classrooms. One of the ideas was that GenAI tools did not seem to offer all-round training to language students in classrooms. Communication skills were one of them, as Christina, a lecturer at a tertiary institute, noted:

> *[W]ith communication skills—I think the use of GenAI … doesn't make it as applicable, perhaps because, I think, with communication skills, there seems to be a demonstration of…different ways of communicating… Because of the use of GenAI, the entire demonstration was a little bit mechanical. I don't think it was as convincing or as convincing as it could have been, in contrast to using real-life demonstrations or real speakers [in a language classroom].*

She suggested that the current GenAI tools might still be at the early stage of development, especially in teaching areas such as lifelike speaking and communication, which she believed signaled the immaturity of the application. Similarly, some of the participants argued that GenAI might be inflexible in accommodating students with different English backgrounds and proficiencies. For example, Erie, a language instructor in a vocational tertiary institute, stated:

> *I'm not sure whether the language levels in the materials are suitable for some of my students; they might find it difficult to understand, and thereby, they might not continue [to use it].*

Erie showed her concerns that even though the GenAI-aided tools seem to have the ability to adjust the level of difficulty of the materials, it would be hard to use several sets of vastly different materials in the same classroom. She also noted that the adjustment might not be as flexible as it seems, as it often involved knowledge of software skills, which concerned her about its application in a real classroom. This lack of strong software skills suggested another issue around GenAI applications in the classroom when the language teachers might not have the specific knowledge or skills to handle the GenAI tools.

As mentioned earlier, the alignment of GenAI with the existing curricula may also be a challenge to many of the language teachers, as they might not know when to use GenAI. While some participants assigned students to work with GenAI tools outside of class as a form of blended learning, others felt

the GenAI assignments might not be placed correctly in the curriculum. Hilton, a secondary school teacher, suggested that GenAI tools might not work with blended learning:

> *Since the whole thing is not conducted at school, why would they [the students] bother to create more homework for themselves unless it's put in the curriculum and unless it's actually on a desirable platform… More importantly, it [has to be] incorporated into their own curriculum… if this counted toward their continuous assessment marks, then they would be more conscientious and eager to perform [in the GenAI tasks].*

Moreover, Hilton expressed that one of the factors that could contribute to the success of integrating GenAI tools in language classrooms might be the curriculum design. He pointed out that if the GenAI tools were used in isolation from the curriculum, the students would have no reason to use them. He added later:

> *Whether this is counted as the marks [as] the continuous assessment… teacher[s] can choose to use it or not. Then, it's a voluntary pilot scheme. In the end, no one is going to use it. That's why whether it is incorporated into the curriculum, that means whether they're calculated as the continuous assessment marks, that's what parents and also students are most concerned.*

It was universal for other participants, regardless of their teaching level or their institution, to struggle to find the right place to apply GenAI tools to their existing curriculum. For example, Andy, a university lecturer, stated that involving GenAI tools might be a critical factor. For example, he said: "Let's say [using GenAI tools counts toward] one or two percent of the whole English curriculum. Then people would be serious [about the GenAI assignments]." Similarly, Mack, a secondary school teacher, pointed out that the key to using GenAI might be to design a curriculum that is "in line with [the students'] curriculum and raising some useful topics for them to practice with".

Almost all the participants admitted that they had tried to use GenAI in some way. However, they all expressed concern over the immaturity of GenAI tools currently because of the lack of coherent GenAI policies for teaching and learning, as well as being unfamiliar with the most suitable GenAI tools for their teaching. This problem hints at a lack of a GenAI-friendly curriculum for language education, which could provide teachers with an idea of how to integrate GenAI tools into their curriculum.

Concerns over the reliability of feedback and input from GenAI tools

In addition to their personal experience using GenAI, the semi-structured interview participants were asked to raise their concerns over the use of GenAI tools in light of demonstrations of some GenAI tools in language learning. One of the more prominent themes was the worry over the reliability of the GenAI tools, especially on the GenAI feedback form that had been given to students, as well as the GenAI inputs that were provided to the students.

Several participants were bothered about whether the feedback provided was comparable with the standards in real life or with the feedback given by human raters. Peter, a lecturer at a community college, expressed his belief that GenAI could be deficient compared with human raters:

> *I think the human rater could also provide a little bit more contextual clues based on students' performance. A human rater could add some more examples of what the students have performed that will help students understand more about their real performances. So, I think the AI model could help students to a great extent. But in a perfect scenario, with the supplement of a human rater, providing more useful, relevant examples for illustration of AI feedback will help students, maybe a little bit more.*

Peter also pointed out that GenAI tools could provide decontextualized feedback, which would not be as helpful as the feedback provided by human raters. He suggested a system combining both a GenAI rater and a human rater, which could be an ideal solution. Interestingly, Cathy suggested a similar idea. In her interview, she emphasized that "AI is not that comparable to real people for some reasons", but "time is never on the [teachers'] side…I guess [GenAI] has an advantage because it can analyze students' areas for improvement… since it can analyze students more comprehensively for *us* [language teachers]". What Cathy suggested for language classrooms was to have teachers first use GenAI tools for a labor-intensive analysis of students' material. Then teachers could make use of the analyses to provide further in-depth feedback. The suggestion of utilizing GenAI tools for the more tedious jobs was common among all the participants in this study. The teachers supported the use of GenAI tools because "they are fast and could be repeated" (Leon, a secondary school teacher) and "GenAI could provide [a] reliable and quick diagnosis" (Andy, university lecturer). Another place where teachers found GenAI could be useful is that it is easily accessible to students on a 24/7 basis, which could provide training to students when they are not at school. Hilton, a secondary school teacher, stressed that "the best thing is we [teachers] have limited time, but they [GenAI] don't, so they can keep using it outside of school". Kelly, another secondary school teacher, also agreed that "a strong advantage of AI models is that they're accessible at any time…so my students who can't reach me after work....This app [GenAI tool] offers a good alternative for them". All these comments suggest that even though a teacher might have concerns over the reliability, especially the feedback, of GenAI, collaboration between humans and GenAI tools could be feasible, since GenAI offers efficiency and accessibility, which humans cannot.

Apart from the reliability issues, teachers also expressed concerns about the inputs that GenAI tools bring to language classrooms. When asked about their experience of listening to GenAI-produced speeches that were used as inputs for students to train their listening skills, almost all of the participants worried that it might not be beneficial input for students. Christina, a community college lecturer, explained how GenAI tools may not be the most suitable kind of listening input in her mind:

> *The voice of the AI was telling. I could tell right out of the gate that the voice was AI… it's toned up and sped up a little bit. Still, I could tell that it was a little bit mechanical, especially in the communication skills part. [GenAI] would sound very mechanical, and it's absolutely fine to sound like that at the learning or the beginner stage. But when it comes to communication, I think it would be better to use real-life speakers."*

Christina suggested that, at the moment, the GenAI tools still sound different from daily conversation, and she was afraid that it might set a wrong expectation for students if teachers decided to use GenAI voice clips as listening inputs. Similar descriptions of GenAI voices being "mechanical" were also found among other teachers, like "monotonous, like a robot" (Kelly, secondary school teacher) and "not natural" (Anne, secondary school teacher). Teachers were worried that these inputs might affect students' learning in the long run, causing them problems when they use English in real life for communication.

Another theme about the inputs from GenAI tools from teachers was a concern about the GenAI's voice, but focusing this time on the familiarity with the various varieties of English that the GenAI tools use. Prior to the interview, when asked to watch the demonstration, most of the participants noticed the GenAI voices switched among several varieties of English (for example, from General American English to Received Pronunciation). Several teachers commented that this feature of incorporating different varieties of English from GenAI could potentially be beneficial to students, especially when "they [students] will have more exposure [to different varieties of English]" (Irene, secondary school teacher). "[The different varieties of English are things] teachers are unable to provide in this speaking

context" (Leon, secondary school teacher). While teachers seem to be positive about the inclusion of different English varieties, they also expressed worries about whether these inputs would affect students' learning. Cathy, a lecturer at the university, questioned the practicality of including multiple varieties of English in the GenAI voices:

> *I think it really depends on the purpose of including different accents… I'm not sure whether including particularly heavy accented speeches would be helpful for examination purposes. It really depends on who is using the [GenAI] app and the purpose of including different accents.*

Cathy continued to explain that even though she could understand the rationale behind incorporating different varieties of English, she worried about whether students should start with the varieties that they are more familiar with; in this case, the Hong Kong English variety. Several other participants supported this idea. For example, Peter (a university lecturer) said, "[i]t's not realistic to foresee those students in Hong Kong…could be [speaking like] native speakers". Christina (a lecturer at a community college) also noted, "I don't think [using] a native accent is really necessary in Hong Kong", and she further stated her suspicions about using native varieties instead of the local variety in Hong Kong:

> *I think in local schools, or for local students, it would be more comforting and a little bit more approachable for them to hear English in a local accent. Hearing someone who teaches them in a native accent, or a more Americanized or British accent, might be intimidating to them—a local accent might be an advantage."*

The participants above noted that the inputs of the GenAI tools given to students were mechanical and nonlocal, and worried about the effect on their student's learning of the language. Together with their feedback that the GenAI was potentially inaccurate, the suggestion was that participants were also doubtful about GenAI providing optimal inputs and feedback for students in language classrooms.

Discussion

The present study has provided numerous insights into how language teachers in Hong Kong view the incorporation of GenAI tools in language classrooms. The two themes identified using the ground theory approach of coding clearly indicated that the 12 English teachers in Hong Kong were uncertain about the use of GenAI tools in language classrooms. Interestingly, teachers, regardless of their teaching experience and type of institution, all showed similar concerns over the reliability of GenAI and the lack of understanding as to when and what method of GenAI should be used. The following sections detail how the two themes can offer further insights into language teacher education and GenAI curriculum design in language classrooms.

As mentioned above, while most of the language teachers were frustrated about the integration of GenAI tools into their classes in terms of *when* and *what*, it seems the issue for the majority was *how* to use GenAI tools. While it is true that language education in Hong Kong lacks a coherent policy on the use of GenAI, efforts to draft such policies have been made by international education organizations, such as through the guidebook published by the United Nations Educational, Scientific and Cultural Organization (2023).The frustration of language teachers in using GenAI tools echoes the results of Temiz and Kafadar's (2024) study, in which 15 Turkish EFL teachers were interviewed, indicating that many of the teachers feared the integration of GenAI because of inadequate training in the technology. Temiz and Kafadar (2024), however, suggested that including GenAI tools and other new technologies in teacher education is the key to preparing language teachers for more effective use of GenAI tools in language classrooms. This perception aligns with that of many other scholars, including Moorhouse

and Kohnke (2024), Tang (2024), and Chan and Liu (2025), who argue that teacher training courses focusing on GenAI tools—both continuing education for in-service teachers and newly developed courses for preservice teachers—are essential for preparing language educators to teach in increasingly GenAI-integrated classrooms. Providing up-to-date teacher training to preservice teachers and in-service teachers would likely have more impact on teachers' teaching practices (Chan, 2023).

Apart from their concerns over the integration of GenAI, teachers were also concerned about the varieties of English that GenAI tools produce as inputs for students in language classrooms. While existing GenAI tools can produce audio using different familiar accents (Wang and Zou, 2025), it is uncertain how listeners would react differently to GenAI-generated speech compared with natural human speech (Hwang et al., 2024). While some relevant studies have provided some insights into GenAI, there is little literature that directly examines the accuracy of accented GenAI-generated speech. Hu et al. (2024) carried out an experiment focused on the prosodic performance of the speech produced by ChatGPT. They concluded that even though ChatGPT could provide highly realistic sounds, the contours of the generated speech might be different from human speakers. Graham and Roll (2024) examined the accent accuracy of OpenAI's large language model (LLM) in speech recognition. Their results showed that the LLM may show a higher accuracy in response to native varieties of English than their non-native counterparts because of the scarcity of non-native varieties in the database (Graham and Roll, 2024). These studies show that GenAI tools, from the production and recognition sides, might still face the challenges of producing speech with accurate accents that resemble human speech. This conclusion echoed the participants' fear of the accuracy of GenAI tools' inputs, especially in the Hong Kong context, where Hong Kong English conveys a significant language identity to local speakers (Chan, 2024; Hansen Edwards, 2019).

It is noteworthy that the most common theme of the 12 participants was that they were not sure how to integrate GenAI into their current language curriculum. Erstad and Voogt (2018) argued that whenever new technologies with the potential to benefit students emerge, some form of adjustment is always necessary. However, Chan and Liu (2025) also noted that teachers may be reluctant to use new technologies. Furthermore, Niepes (2025) urged the sustainable integration of GenAI into language education. He also proposed the Triadic GenAI Integration Theory, stressing the fact that sustainable integration of GenAI in language education requires three elements, including (1) value-driven sustainable implementation, (2) AI literacy-embedded curriculum, and (3) English proficiency and AI integration alignment.

While the first point focuses on how the integration should align with core education values with a long-term assessment of the impact of GenAI tools, the second and the third points emphasize the need to align the strengths of GenAI tools into the existing curriculum and match the specific GenAI tools to the teaching different language skills. This model summarizes the needs for a meaningful GenAI tools curriculum. Similarly, Chan and Liu (2025) proposed a flowchart for language teachers to use with multiple GenAI tools to train students' English-speaking skills. He also concluded that this flowchart should be for other language skills like listening and reading. Choi et al. (2025) also suggested that, given the potential benefits for language teaching, continuous learning opportunities for GenAI use, such as online courses or webinars, should be provided by schools to keep teachers updated on the latest advancements in GenAI. There is an urgent need for GenAI-related teacher education courses, GenAI-integrated curricula, and skill-specific GenAI curricula.

Conclusion

This qualitative study investigated the perception of language teachers toward the use of GenAI tools in Hong Kong language classrooms. As the participants suggested in their interviews, even though

they acknowledged the powers of GenAI tools, they expressed their concerns in two significant areas: the integration of GenAI tools into the existing curriculum and the accuracy of the feedback and input provided by GenAI tools to students. The results indicated that it is now time to call for a proposal for better integration of GenAI tools into the curriculum as suggested by Niepes' (2025) Triadic GenAI Integration Theory. More up-to-date teacher training for both pre- and in-service teachers is needed to help teachers transition to the GenAI era of teaching.

This study had some limitations. For example, the study had limitations in terms of participants' demographics and institutional representation. The age range of the participants was relatively narrow, and all were affiliated with either secondary schools or universities, potentially limiting the generalizability of the findings across other educational contexts. Moreover, the study focused exclusively on English language teachers, excluding educators of other languages whose experiences with GenAI tools may differ. The current study should be seen as a preliminary study on how language teachers view GenAI tools. In turn, it provides significant insights into fields like language teacher education and language education. The limitations suggest that future research could benefit from a more diverse participant pool in terms of age, institutional background, and subject specialization. Since GenAI tools have become an inevitable presence in education, we, as language educators, must step ahead and prepare for the rise of the GenAI era.

References

Adamson, J. (2004). Talking about lesson planning: The use of semi-structured interviews in teacher education. *The Journal of Language and Learning, 2*(2), 114–128.

Aksakallı, C., & Daşer, Z. (2025) Unlocking EFL learners' insights into ChatGPT use for L2 writing: The impacts of usage frequency and gender variations. *Current Psychology, 44*, 7957–7977. https://doi.org/10.1007/s12144-025-07437-3

Chan, K. L. R. (2023). Trilingualism in Hong Kong: A World Englishes framework for EMI English teachers in university. *LEARN Journal: Language Education and Acquisition Research Network, 16*(1), 1–17. https://so04.tci-thaijo.org/index.php/LEARN/article/view/263428

Chan, K. L. R. (2024). The construct of English native speaker in Hong Kong. *Journal of English and Applied Linguistics, 3*(1), 1–7. https://doi.org/10.59588/2961-3094.1074

Chan, K. L. R., & Chan, N. C. L. (2025). Using the internet as a corpus for English teacher training under multilingualism: Challenges and potentials. In S. Karpava (Ed.), *Multilingualism and multiculturalism in language education* (pp. 175–190). Springer. https://doi.org/10.1007/978-3-031-76043-3_9

Chan, K. L. R., & Liu, J. (2025). Using multiple GenAI Tools in pronunciation teaching: An AI-assisted teaching flowchart in blended learning and flipped learning. *Journal for Online Educators*. Advance online publication.

Choi, K. Y., Wu, C., & Moorhouse, B. (2025). Exploring the use of generative artificial intelligence (GenAI) in English language teaching: Voices from in-service teachers at an early-adopting Hong Kong secondary school. *Technology in Language Teaching & Learning, 7*(3), 102516. https://doi.org/10.29140/tltl.v7n3.102516

Corbin, J. M., & Strauss, A. L. (2015). *Basics of qualitative research: Techniques and procedures for developing grounded theory*. SAGE.

Dai, Y., Lai, S., Lim, C. P., & Liu, A. (2024). University policies on generative AI in Asia: Promising practices, gaps, and future directions. *Journal of Asian Public Policy*, 18(2), 260–281. https://doi.org/10.1080/17516234.2024.2379070

Drever, E. (1995) *Using semi-structured interviews in small-scale research. A teacher's guide*. Scottish Council for Research in Education.

Erstad, O., & Voogt, J. (2018). The twenty-first century curriculum: Issues and challenges. In J. Voogt, G. Knezek, R. Christensen, K. W. Lai (eds), *Second handbook of information technology in primary and secondary education.* Springer international handbooks of education. Springer. https://doi.org/10.1007/978-3-319-71054-9_1

Graham, C., & Roll, N. (2024). Evaluating OpenAI's Whisper ASR: Performance analysis across diverse accents and speaker traits. *JASA Express Letters, 4*(2), 025206. https://doi.org/10.1121/10.0024876

Hansen Edwards, J. G. (2019). *The politics of English in Hong Kong: Attitudes, identity and use.* Routledge. https://doi.org/10.4324/9781315178547

Hu, N., Kim, J., Orrico, R., Gryllia, S., & Arvaniti, A. (2024). *Can OpenAI's TTS model convey information status using intonation like humans?* [Conference presentation] Proceedings in Speech Prosody (2–5 July 2024). Leiden, The Netherlands. https://doi.org/10.21437/SpeechProsody.2024-7

Huang, J., & Mizumoto, A. (2024). Examining the effect of generative AI on students' motivation and writing self-efficacy. *Digital Applied Linguistics, 1*, 102324. https://doi.org/10.29140/dal.v1.102324

Hwang, A, H., Siy, J. O., Shelby, R., & Lentz, A. (2024). *In whose voice? Examining AI agent representation of people in social interaction through generative speech* [Conference presentation]. Proceedings in Designing Interactive Systems Conference (1–5 July 2024). IT University of Copenhagen, Denmark. https://doi.org/10.1145/3643834.3661555

Kamali, J., Paknejad, A., & Poorghorban, A. (2024). Exploring the challenges and affordances of integrating ChatGPT into language classrooms from teachers' points of view: An ecological perspective. *Journal of Applied Learning & Teaching, 7*(2), 168–180. https://doi.org/10.37074/jalt.2024.7.2.8

Karjagdi Çolak, M. (2024). Enhancing speaking skills through task repetition and ChatGPT integration in remedial EFL lessons: An action research approach. *Focus on ELT Journal, 6*(4), 1–16. https://doi.org/10.14744/felt.6.4.1

McDonald, N., Johri, A. Ali, A., & Collier, A. H. (2025). Generative artificial intelligence in higher education: Evidence from an analysis of institutional policies and guidelines. *Computers in Human Behavior: Artificial Humans, 3*,100121 https://doi.org/10.1016/j.chbah.2025.100121.

Moorhouse, B., & Kohnke, L. (2024). The effects of generative AI on initial language teacher education: The perceptions of teacher educators. *System, 122*, 103290. https://doi.org/10.1016/j.system.2024.103290

Naznin, K., Al Mahmud, A., Nguyen, M.T., & Chua, C. (2025). ChatGPT integration in higher education for personalized learning, academic writing, and coding tasks: A systematic review. *Computers, 14*, 53. https://doi.org/10.3390/computers14020053

Niepes, G. L. (2025). Developing theory on sustainable integration of GenAI in tertiary English language teaching: The triadic GenAI integration theory. ELT Forum, 14(1), 29–40. https://doi.org/10.15294/elt.v14i1.11033

O'Dea, X. (2024). Generative AI: is it a paradigm shift for higher education? *Studies in Higher Education, 49*(5), 811–816. https://doi.org/10.1080/03075079.2024.2332944

Patton, M. Q. (2015). *Qualitative research and evaluation methods* (4th edn.). SAGE.

Satvati, N., Kamali, J., Safian Boldaji, F., Khodadadi, M., & Akhondi, S. (2025). AI Integration into language education and teacher identity: An ecological perspective. *Language Teaching Research Quarterly, 47* , 1–19. https://doi.org/10.32038/ltrq.2025.47.01

Solak, E. (2024). Revolutionizing language learning: How ChatGPT and AI are changing the way we learn languages. *International Journal of Technology in Education, 7*(2), 353–372. https://doi.org/10.46328/ijte.732

Songsiengchai, S., Sereerat, B., & Watananimitgul, W. (2023). Leveraging artificial intelligence (AI): ChatGPT for effective English language learning among Thai Students. *English Language Teaching, 16*(11), 68–79. https://doi.org/10.5539/elt.v16n11p68

Su, A. A. T., & Tran, T. H. N. (2024). The benefits and weaknesses of ChatGPT on students' learning writing. *International Journal of AI in Language Education, 1*(1), 20–28. https://doi.org/10.54855/ijaile.24112

Suri, H. (2011). Purposeful sampling in qualitative research synthesis. *Qualitative Research Journal, 11*(2), 63–75. https://doi.org/10.3316/QRJ1102063

Tang, K. H. D. (2024). Implications of artificial intelligence for teaching and learning. *Acta Pedagogia Asiana, 3*(2), 65–79. https://doi.org/10.53623/apga.v3i2.404

Temiz, G., & Kafadar, E. N. (2024). An investigation of in-service EFL teachers' perception of AI and experiences with its integration in ELT: A phenomenological approach. In Y. Özkan, M. Kara, A. Dinçer, & D. Sucak (Eds), *English studies: A multifaceted lens* (pp. 140–157). Black Swan Publishing.

Teng, M. F. (2024). A systematic review of ChatGPT for English as a foreign language writing: Opportunities, challenges, and recommendations. *International Journal of TESOL Studies, 6*(3), 36–57. https://doi.org/10.58304/ijts.20240304

Tyson, P. (1990). Talking about lesson planning: The use of semi-structured interviews in teacher education. *Teacher Education Quarterly, 18*(3), 87–96. http://www.jstor.org/stable/23475113

United Nations Educational, Scientific and Cultural Organization. (2023). *Guidance for generative AI in education and research*. United Nations Educational, Scientific and Cultural Organization. https://www.unesco.org/en/articles/guidance-generative-ai-education-and-research

Wang, C., & Zou, B. (2025). D-ID Studio: Empowering language teaching with AI avatars. *TESOL Journal, 16*, e70034. https://doi.org/10.1002/tesj.70034

Wang, H., Dang, A., Wu, Z., & Mac, S. (2024). Generative AI in higher education: Seeing ChatGPT through universities' policies, resources, and guidelines. *Computers and Education: Artificial Intelligence, 7*, 100326. https://doi.org/10.1016/j.caeai.2024.100326

Wang, X., & Feng, Y. (2024). *An experimental study of ChatGPT-assisted improvement of Chinese college students' English reading skills: A case study of dear life* [Conference presentation]. Proceedings of the 15th International Conference on Education Technology and Computers (ICETC) 2023, Barcelona, Spain (September 26–28). https://doi.org/10.1145/3629296.3629300

Yu, X., & Xie, Q. (2025) Generative AI vs. teachers: Feedback quality, feedback uptake, and revision. *Language Teaching Research Quarterly, 47*, 113–137. https://doi.org/10.32038/ltrq.2025.47.07

Yuniarti, F., Wulandari, F., & Rakhmawati, D. (2024). Personalized listening practice: integrating AI (ChatGPT) and (Voxbox) native speaker input in language learning. *Inovish Journal, 9*(2), 137–148. https://jurnal.polbeng.ac.id/index.php/IJ/issue/view/18