

Rethinking L2 academic speaking: Oral communication strategies and discourse markers in ChatGPT-mediated instruction

ALMONTASSAR BELLAH TAIEB^a  

DOUA KHENAT^b  

HANAE EZZAOUYA^c  

^{a,b} *Pázmány Péter Catholic University, Hungary*

^c *University of Innsbruck, Austria (current affiliation), Károli Gáspár University of the Reformed Church, Hungary*

Abstract

The introduction of generative artificial intelligence (GenAI) has created new opportunities to support L2 speaking instruction and development. However, empirical evidence is needed to establish its pedagogic value in higher education contexts, particularly for cognitively demanding second-language (L2) speaking tasks. This study investigates the effects of ChatGPT-mediated instruction on the use of oral communication strategies (OCSs) and discourse markers (DMs) during an academic debate session. Using a quasi-experimental design, a total of 56 Hungarian English as a Foreign Language (EFL) undergraduates were assigned to two instructional conditions: a GenAI-assisted experimental group and a non-GenAI-assisted control group. Data were collected through a posttest debate task and analysed using a combination of quantitative and qualitative techniques. The results indicated significant differences in the deployment of OCSs and DMs between the two instructional conditions. In particular, our analysis revealed an interesting contrast between the strong focus on fluency-maintenance and discourse-organisational devices in the control group as opposed to the preponderance of direct-interactive strategies alongside a wide range of DMs in the experimental group. The findings suggest a potentially beneficial role for ChatGPT-mediated instruction in the development of strategic competence and discourse-level awareness. These GenAI affordances, however, presuppose careful teacher mediation within a progressive instructional design that moves from AI-supported rehearsal to peer- and group-based interaction.

Keywords: ChatGPT, discourse markers, EFL learning, generative AI, L2 speaking, oral communication strategies

Introduction

One of the most enduring issues in the context of English as a Foreign Language (EFL) is the development of second-language (L2) speaking ability. As a productive skill, L2 speaking necessitates greater cognitive effort to plan and execute meaning while also attending to several background variables, including socio-pragmatic appropriateness (Geeslin et al., 2018; Lowie et al., 2018). These demands are particularly pronounced in L2 speaking tasks, such as academic debates, where learners are required to co-construct arguments, respond to criticism, and manage turn-taking effectively (Saidah et al., 2020; Želježič, 2017). Given the labour-intensive nature of oral tasks, L2 speaking ability might be prone to various communication breakdowns, despite the proper command of vocabulary and grammar.

In light of these communicative mishaps, two sets of interactional resources have become especially important for L2 speaking instruction and development. The first comprises oral communication strategies, which function primarily as compensatory mechanisms for resolving lexical and pragmatic gaps in spoken interactions, realised through approximation and stalling devices, among many other strategies (Seong, 2014). Interest in this line of inquiry has picked up pace in the last two decades, and recent research has underscored the relevance of exploring the interplay between learner-internal factors and instructional configurations in oral interaction (e.g., Cai & Zhang, 2023; Parcon & Reyes, 2021; Zhou et al., 2025). The second interactional resource involves discourse markers supporting, for example, the organisation of talk and speech management (Fraser, 2009). We contend that these resources should develop in parallel, and that their joint mobilisation is particularly important for L2 speaking tasks that place heavy cognitive pressure on L2 learners.

With the rapid development of artificial intelligence (AI), the English Language Teaching (ELT) enterprise has experienced a growing trend towards AI-mediated instruction, particularly through ChatGPT (e.g., Kohnke et al., 2023; Law, 2024). This AI system has attracted considerable attention from researchers working with L2 speaking, given its ability to simulate authentic interaction, provide instant feedback, and model language across diverse registers (e.g., Celik et al., 2025; Hsu et al., 2023). While existing studies have begun to document the pedagogical applications of AI in ELT (e.g., Rusmiyanto et al., 2023; Tai & Chen, 2025), research examining generative artificial intelligence (GenAI)-mediated L2 speaking tasks in higher education remains relatively limited, especially in relation to cognitively demanding forms of spoken interaction. To the best of our knowledge, the only study that has attempted to examine the use of AI in an academic debate is that of Bui et al. (2025). This gap remains particularly evident considering the role of AI in shaping learners' strategic and discourse-level competencies. The purpose of this paper is to investigate the extent to which ChatGPT-mediated instruction influences learners' deployment of oral communication strategies and discourse markers in an EFL academic debate setting. In doing so, we contribute to emerging discussions on AI-assisted language learning by exploring whether ChatGPT can support the development of learners' interactional resources for L2 speaking purposes.

Literature Review

Oral Communication Strategies

Oral communication strategies (OCSs) are real-life manifestations of strategic competence (SC), a subcomponent of the communicative competence framework proposed by Canale and Swain (1980). This underlying competence provides language users with a wide range of strategies for negotiating meaning, resolving communication problems, and compensating for limitations in linguistic resources (Celce-Murcia et al., 1995). As a behavioural manifestation of this strategic

capacity, OCSs have been conceptualised beyond a vocabulary-centred perspective (Bongaerts & Poulishse, 1989; Faerch & Kasper, 1984; Selinker, 1972; Tarone, 1980) to encompass core interactional functions involved in managing pragmatic, sociolinguistic, and discourse-level issues in spoken interaction (Dörnyei & Scott, 1997).

Within this evolving landscape, Dörnyei and Scott (1997) offered a more fine-grained pragmatic account of OCSs. Their granular model focuses on the specific *mishaps* that emerge in spoken interaction. This taxonomy is widely favoured in the literature for its ability to capture a broad spectrum of strategies, ranging from *direct* and *indirect* strategies to *interactional* strategies. Indirect strategies typically involve gap-fillers and hesitation devices (e.g., “uh”, “um”, “well”), enabling speakers to manage communication breakdowns by gaining additional processing time. However, the overuse of such fillers has often been associated with under-preparation, lack of interest, and/or lower levels of confidence. Direct strategies, by contrast, are considered self-initiated alternatives in communicative gaps (e.g., circumlocution, restructuring); however, interactional strategies represent collaborative attempts to resolve problems with other interlocutors (e.g., asking for clarification).

Beyond theoretical accounts, empirical research has increasingly viewed OCSs as context-sensitive and multidimensional, shaped by the interaction between cognitive, affective, and task-related variables. For example, Nash and Ramirez (2025) argued that the use of compensation strategies may serve as an inhibitory tool for the affective filter, thereby mitigating anxiety and boosting engagement. Similarly, Nakatani (2006) suggested that some language breakdowns often arise from the interplay between lexical deficits and cognitive-affective obstructions. This idea is further supported by research on strategy instruction, which has found that explicit training in OCSs may result in fluency gains and boost interactional confidence (Lai Kuen et al., 2017; Tavakoli et al., 2011). Another important dimension of OCSs is their relation to affective factors. There exist numerous studies that attribute the use of fillers and hesitation devices, as well as avoidance strategies, to high anxiety levels (Grzegorzewska, 2015; Tran, 2023); motivated students tend to exhibit better strategic behaviour, using a greater number of achievement strategies (Quadir, 2014; Soozandehfar, 2010; Su, 2021). These findings call for a holistic approach to prioritise the proceduralisation of this ability by integrating both cognitive and affective dimensions of SC.

Discourse Markers

Discourse markers (DMs) play a key role in spoken and written communication. Fung and Carter (2007) defined DMs as “intra-sentential and supra-sentential linguistic units which fulfil a largely non-propositional and connective function at the level of discourse” (p. 411). Although scholars have recognised their general linguistic properties, a multitude of labels have been used to connote the differing views on the functions of DMs. These include *semantic conjuncts* (Quirk et al., 1985), *discourse operators* (Redeker, 1991), and *pragmatic markers* (Fraser, 1990, 1999), to name a few.

Research on DMs is marked by the diversity of its theoretical foci. While some scholars recognise DMs as a broad class of expressions, including interjections and non-verbal expressions that contribute to discourse coherence (Schiffrin, 2005), others adopt a narrower conceptualisation of DMs by focusing on items that signal semantic or pragmatic relationships between utterances (Fraser, 1990), expressions carrying *procedural* meaning (Blakemore, 2002), or a specific subset of connectives used, for example, in argumentation (Carel, 1995). It is therefore not surprising that researchers have proposed numerous frameworks in an effort to provide a more fine-grained account of DMs. Fung and Carter (2007), for instance, identified a set of criteria to guide the identification of DM, namely *position*, *prosody*, *multigrammaticality*,

indexicality, and *optionality*. In a similar vein, Fraser (2009) considered DMs as “free morphemes”, capable of signalling “a specific message either about or in addition to the basic message” (p. 295).

Among the various taxonomies proposed in the literature, Fung and Carter’s (2007) corpus-driven framework offers a useful account of DMs in classroom discourse. In their view, DMs can be grouped into four main categories. *Interpersonal* DMs are typically used to manage interactional relations, express stance, mark shared knowledge, and signal attitudes or responses (e.g., “you know”, “actually”). *Referential* DMs encode logical relations, such as contrast or addition (e.g., “so”, “because”). The category of *structural* DMs is particularly useful in organising discourse by managing topic shifts and turn-taking (e.g., “now”, “by the way”); however, *cognitive* DMs reflect speakers’ real-time processing, using hesitation, reformulation, and other DMs (e.g., “well”, “I mean”). These categories are summarised pictorially in Figure 1.

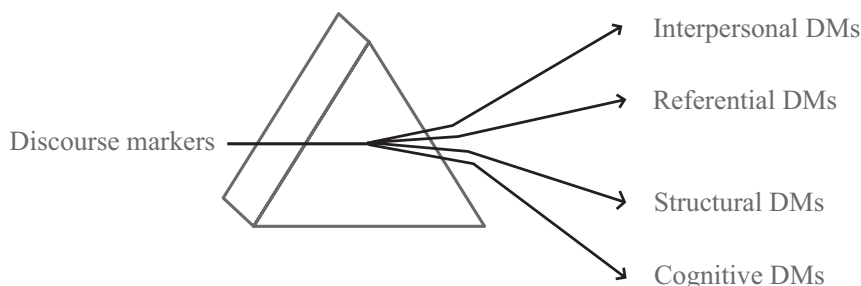


Figure 1. A framework of discourse markers in pedagogic discourse

There is a growing body of literature investigating learners’ use of DMs in spoken and written discourse. Fareh (2014) raised an interesting discussion regarding the limited attention given to DM instruction in the EFL classroom, which would inadvertently result in a less “communicatively competent language user” (p. 924). Studies have consistently shown that EFL speakers often rehash a limited set of markers (Alsaawi, 2022; Fareh et al., 2020; Fuller, 2003; Huneety et al., 2023; Özer & Okan, 2018). Thus, an explicit approach to teaching DMs may be warranted, given the strong empirical support for the effectiveness of direct instruction in promoting pragmatic development (Taguchi, 2015).

Generative Artificial Intelligence in Education

The emergence of GenAI has transformed the educational landscape. GenAI refers to a class of AI systems capable of producing novel content (e.g., texts, images, videos) and, most importantly, simulating characteristics akin to human creativity and problem-solving capacity. A key feature of GenAI technologies is their ability to perform a wide range of downstream tasks in response to a natural-language prompt (e.g., text summarisation, translation, information extraction). Unlike earlier AI technologies, GenAI outputs are probabilistic and may thus vary across users and contexts (Banh & Strobel, 2023).

GenAI technologies are increasingly integrated across many educational institutions and online platforms (e.g., Khan Academy, Duolingo, DeepL). While the use of such technologies has offered bountiful opportunities to support learning, their prospects are often viewed with some trepidation. Questions about whether humans and AI can live harmoniously (Haenlein & Kaplan, 2019) or the extent to which current teachers’ roles may risk becoming obsolete cannot

be ignored (Roll & Wylie, 2016; Stockwell, 2024). These concerns underscore the need for “new cognitive skills and literacy” (Karakaya et al., 2025, p. 9) and dynamic development across formal and informal learning pathways (Irshad, 2026; Ito, 2026).

Among the various GenAI systems, this study focuses on ChatGPT, a state-of-the-art large language model (LLM) chatbot developed by OpenAI and launched in late November 2022. At its core, ChatGPT is built on the generative pre-trained transformer (GPT) architecture, which is a type of neural network designed to process and generate contextually appropriate responses. Successive GPT iterations (e.g., GPT-3.5, GPT-4) have introduced sophisticated capabilities, and the latest architecture is found in GPT-5 (OpenAI, 2025).

In general, ChatGPT has been the subject of extensive debate. In education, while traditional classroom settings have often relied on scaffolding and peer interaction, ChatGPT-mediated instruction appears compatible with established pedagogical frameworks and mainstream educational practices (Karakaya et al., 2025) as well as with future workplace expectations (Khoumich & Nadif, 2026). As such, the increasing integration of ChatGPT, as with other AI systems, exemplifies how emergent technologies can be operationalised within the broader educational ecosystem (Deng et al., 2025; Kotsis, 2024). In Vygotsky’s sociocultural theory, for example, cognitive development is often mediated by various cultural tools, among which GenAI interactive technologies play an increasingly prominent role (Ciampa et al., 2025; Essien et al., 2024). ChatGPT-assisted learning environments can thus be understood as mediational artefacts that contribute to L2 proficiency, motivation, and engagement (Baidoo-Anu & Owusu Ansah, 2023; Nadif et al., 2026; Yan & Zhang, 2024). Educators can avail themselves of such a tool by anticipating applications that are both enriching and complementary to their teaching (Van Den Berg & Du Plessis, 2023), while simultaneously fostering conditions that promote greater satisfaction with the use of AI for professional development (El Anouar et al., 2026).

The Influence of GenAI on EFL Speaking

There is a growing consensus on the facilitative role of GenAI in EFL classrooms (Al-khresheh, 2024; Kaid et al., 2023). This paper proposes that it is timely to move beyond attitudinal accounts of perceptions while also charting new directions for current assessment practices to align with the technological prowess of GenAI. We characterise this transition as a shift beyond conventional pedagogical approaches towards instructional environments that capitalise on the affordances of GenAI systems. Earlier approaches to teaching L2 speaking have generally emphasised peer interaction and communicative tasks as key mechanisms to promote oral proficiency (Folse, 2006; Nation & Newton, 2009; Pakula, 2019; Thomas, 2019; Zhang et al., 2025). With the emergence of GenAI, however, these interactional dynamics have been drastically reconfigured, especially as conversational practices are no longer confined to traditional classroom settings.

GenAI has attracted considerable attention in L2 speaking research. At the elementary school level, studies have examined the impact of GenAI tools on oral English proficiency, willingness to communicate, and learners’ perception of AI-mediated speaking tasks (Rusmiyanto et al., 2023; Yuan, 2024). Working from this perspective can help redress the undue emphasis on developing receptive skills by directing more attention to the productive dimension of language (Yuan, 2024). Other studies have explored the suitability of AI-generated materials (Young & Shishido, 2023) and the advantage of incorporating various interactional configurations (Tai & Chen, 2024a). In the secondary education context, scholars have explored the effectiveness of GenAI in promoting adolescent EFL learners’ speaking proficiency, with particular focus on feedback delivery (Tai & Chen, 2024b) and interactional modalities (Tai & Chen, 2025). In higher education, research suggests that leveraging GenAI in L2 speaking classrooms may lead

to significant improvements in various cognitive and affective dimensions (Abdelhalim & Alsehibany, 2025; Celik et al., 2025; Huang & Kakham, 2026; Hsu et al., 2023; Wu et al., 2025) and can also facilitate error correction via automated feedback (Shadiev et al., 2024). Comparable gains have been observed in out-of-class contexts and non-formal language learning settings (Fathi et al., 2025; Tai, 2025; Yang et al., 2024). Across these studies, GenAI-assisted instruction has led to observable growth in several language-related outcomes (e.g., grammatical accuracy, lexical development).

The Present Study

As noted above, GenAI-mediated instruction has brought forth a wide range of pedagogical affordances in higher education. Although studies in this area have documented broader gains in L2 speaking proficiency, little attention has been paid to how GenAI can be harnessed to facilitate explicit instruction in OCSs and DMs. This gap is particularly notable given that research on OCSs and DMs has largely proceeded along separate lines, despite their seemingly complementary functions in spoken academic discourse.

By bringing together these two theoretically distinct interactional resources, the present study examines the extent to which ChatGPT-mediated instruction influences the frequency and range of OCSs and DMs in an EFL academic debate setting. Their joint mobilisation may, in turn, help EFL learners muster the interactional resources necessary for effective participation in spoken interaction. Addressing this gap, our study addresses the following question:

To what extent does ChatGPT-mediated instruction influence the frequency and range of oral communication strategies and discourse markers used by EFL undergraduates during an academic debate?

Methods

Research Design

This study used a posttest-only quasi-experimental design to investigate the impact of GenAI-assisted tasks on the use of OCSs and DMs in an EFL academic debate. The sample consisted of 56 EFL undergraduate students enrolled at a Hungarian university during the autumn semester of 2025/26. The posttest-only structure was considered useful in minimising testing and sensitisation effects (Trochim, 2005). All participants were drawn from four intact first-year classes and were later assigned to conditions using a semi-random allocation procedure (Creswell, 2009). Specifically, two equally sized classes ($n = 28$) were assigned to the experimental group, which used ChatGPT-assisted preparation tools, while the remaining two classes ($n = 28$) served as the control group with teacher-led instruction.

The study was conducted in accordance with established ethical procedures for research involving human participants (British Association for Applied Linguistics [BAAL], 2021), namely voluntary participation, the disclosure of the study purpose, the use of audio recording, and the confidential and anonymised handling of data. Drawing on Robey and Taylor's (2018, p. 2) notion of "practitioner scholarship", in which inquiry is integrated with application to bridge the divide between research and practice, this paper adopted a practitioner-scholarship approach. Accordingly, one author occupied the dual role of teacher-researcher, while the other two researchers undertook the task of providing assistance and organising the classroom setup.

Procedure

The study was conducted over six 90-minute sessions, evenly divided between ChatGPT-assisted and non-assisted conditions. In both groups, the first two sessions served as practice sessions while the posttest debate task was administered in the third session. To maintain comparability, all sessions were designed to be as equivalent as possible in terms of length, instructional procedures, interactional format, debate topics, and preparation time. To ensure that participants' efforts were largely focused on deploying OCSs and DMs, we avoided tasks that required students to generate arguments from scratch. Instead, the tasks were built on arguments proposed in earlier activities, and, in this way, served as a scaffold for the training.

In both conditions, we used an interactive presentation to introduce OCSs and DMs. The starting point was the use of some relatable classroom scenarios, such as forgetting a word mid-sentence during a presentation or not knowing the exact English word for an object and demonstrating how communication can still be sustained through strategic language use. Several examples of effective OCSs were introduced (e.g., paraphrasing, circumlocution), which were later contrasted with less effective strategies (e.g., avoidance, language switching). Most activities were conducted in pairs to encourage collaborative reflection and negotiation of meaning.

GenAI-Mediated Tasks

In the ChatGPT condition, participants completed two OCS-focused tasks. First, participants engaged in simulated communication breakdowns, after which they were required to brainstorm suitable OCSs to resolve such fictitious situations. Using ChatGPT's conversation mode, participants engaged in a dialogue by negotiating their proposed OCSs in real time. These exchanges encouraged learners to reflect on their individual choices and to leverage ChatGPT's adeptness to provide immediate feedback and foster enthusiasm. This task was conducted in pairs, and students also took notes, which they shared with the whole class to support collective reflection. Participants used the following prompt:

"Give me situations where communication is problematic, and I will tell you how I would overcome the communication breakdowns; then you give me feedback about my proposed solution, and if I can improve it."

The second task was built on a previously introduced task in which students had worked in pairs to develop a tentative outline for a classroom presentation. Building on this task ensured topic familiarity and allowed participants to focus on strategic language use rather than argument generation. Participants interacted with ChatGPT by articulating their main points to receive targeted feedback on their use of OCSs and reformulations where necessary. The following prompt was used:

"This is a debate topic for which I had already prepared some arguments. It is about [TOPIC], and I will first present my arguments. Then, I want you to (1) evaluate whether my oral communication strategies are effective for the debate context, and (2) suggest how I can improve my strategies using, for example, paraphrasing, approximation, or circumlocution."

In the second session, participants completed two main tasks structured around the appropriate use of DMs. Participants were organised into pairs, then used ChatGPT to generate useful

DMs for an academic debate context. Each pair focused on a single DM category (e.g., hedging, rebuttal construction). To make the process even more engaging, students conducted a peer-to-peer exchange by moving around the classroom, eliciting each peer's assigned category, and taking note of ChatGPT-generated expressions.

"You are my debate coach. I am working with a partner to look up useful [CATEGORY] expressions I can use in an EFL academic debate context. Your task is to generate natural expressions for this category."

In the final task, participants used ChatGPT's generated DMs to practice debating skills using the platform's interactive feature to refine their previously prepared arguments. They also prompted the AI system to provide feedback about their DM use, followed by peer comparison and discussion of the AI-generated feedback.

"You are my debate coach and discussion partner. Engage with me in a short academic debate on the topic I provide. Afterwards, evaluate my use of DMs and provide feedback on their variety and naturalness in an academic debate context."

Non-GenAI-Mediated Tasks

In the non-GenAI condition, the participants completed parallel OCS-focused tasks. Each pair was provided with a handout containing a few breakdown situations and strategy options. The task required students to evaluate each strategy, justify their choice, and propose alternative responses if necessary. Towards the end, participants shared their decisions to support collective discussion and reflection. An example of a fictitious scenario is as follows:

"Your opponent uses a term you do not understand during a debate, but the discussion is moving quickly. You need to ask for clarification without interrupting too strongly."

Drawing on previously presented debate topics for which participants had already developed arguments, the second task required students to practice their argumentative skills and engage in peer feedback regarding the effectiveness of their OCSs. To enable this, participants relied on a checklist comprising examples of effective and ineffective OCSs in a debate context, which guided subsequent revision of their oral production skills.

In the next session, participants were required to complete two major tasks targeting the use of DMs. The opening task was to work in pairs using a worksheet containing a list of common DMs for academic debating. Each pair focused on a particular category (e.g., hedging, rebuttal construction). Students analysed the expressions, discussed their communicative function, and isolated appropriate DMs. The task was concluded by exchanging categories and taking notes on expressions identified by other pairs. The final task required students to revise previously generated arguments by integrating appropriate DMs to support their debating skills. Participants practised short debate exchanges while referring to the discourse marker handouts. The teacher's feedback focused on the coherence and pragmatic appropriateness of the selected markers within peer-to-peer interaction.

Following the training period, participants completed a structured debate task, which served as the posttest measuring their use of OCSs and DMs. The nature of this task mandated that students present arguments, respond to counterclaims, and negotiate meaning in real time. The debate topics were culturally aligned with the research setting. Students were randomly assigned to debating positions (i.e., proponents or opponents) and were allotted 30 minutes to prepare their arguments. All debate sessions were audio-recorded and transcribed verbatim.

Data Analysis

The data were analysed using a combination of quantitative and qualitative methods. At the beginning, a detailed transcription protocol and coding manual were developed. The codebook outlined the transcription conventions, the codes and their respective definitions to ensure systematic identification and classification of OCSs and DMs. Dörnyei and Scott’s (1997) framework was used to categorise OCSs, and strategies such as the use of L1 and gesturing were excluded from the current analysis. One reason was that students were repeatedly encouraged to avoid language switching since the teacher was not a native speaker of Hungarian. Because the posttest task was not video-recorded, non-verbal strategies were excluded. Other strategies, such as abandonment and asking for clarification, were also removed from the analysis because the classification did not yield any frequency counts. Drawing on Fung and Carter’s (2007) taxonomy, the present study employed a modified classification system that grouped DMs into two overarching categories: discourse-organisational markers (e.g., cause-effect relations) and interactional markers (e.g., hedging, turn-taking). Together, these categories comprised 24 individual DMs.

OCSs were coded using MAXQDA 24. An OCS was recorded whenever the speaker deployed a strategic device to overcome communication difficulty, maintain interactional flow, or compensate for linguistic constraints; multiple OCSs produced within a single turn were coded separately. Although most studies do not explicitly discuss their unit of analysis, we treated each strategic move as an individual instantiation of an OCS within a defined communicative event. The coding process was conducted independently by two researchers. Interrater reliability, calculated as percentage agreement, reached 92.5%. The transcripts were also coded for DMs using the same software. The classification was undertaken by two researchers, which resulted in an inter-rater reliability coefficient of 97%. All discrepancies were resolved through discussion until complete agreement was achieved. Using jamovi 2.7, descriptive statistics were first calculated, and chi-square analysis examined the association between instructional condition and the distribution of OCS and DM categories. Effect size was measured using Cramér’s *V*, which ranges from 0 to 1.0. Standardised residuals (*R*) allowed us to identify categories contributing most strongly to the difference between observed and expected frequencies. Values exceeding ± 1.96 indicated which of the main categories made a meaningful contribution to the chi-square statistic.

Results

Analysis of OCS Categories

Table 1 presents the frequencies of OCS categories across instructional conditions. Three main categories were observed: indirect strategies (fluency-maintenance and overcorrection), direct strategies (circumlocution and paraphrase), and interactional strategies (asking for help). The GenAI-assisted group exhibited a higher frequency of OCS use (*n* = 209) than did the non-GenAI-assisted group (*n* = 144). Across both groups, direct and indirect strategies occurred more frequently, whereas interactional strategies remained comparatively low.

Table 1. Descriptive statistics of OCS categories in the debate posttest

	Category			
	Indirect	Direct	Interactional	Total
Non-GenAI-assisted	84.02% ₍₁₂₁₎	11.11% ₍₁₆₎	4.86% ₍₇₎	100% ₍₁₄₄₎
GenAI-assisted	53.11% ₍₁₁₁₎	29.66% ₍₆₂₎	17.22% ₍₃₆₎	100% ₍₂₀₉₎

Note. Raw frequencies are reported in parentheses.

A closer inspection of Table 2 indicates important differences in the frequency distribution of individual OCS strategies. In the non-GenAI-assisted group, students relied predominantly on fluency-maintenance strategies (76.4%), with the remaining strategy types occurring at marginal levels. In contrast, the GenAI-assisted group displayed a diversified OCS repertoire, characterised by reduced use of fluency-maintenance devices (35.9%) and a substantially higher proportion of circumlocution (18.2%), asking for help (17.2%), and overcorrection devices (17.2%). The latter pattern suggests awareness of form- and meaning-oriented repair and negotiation strategies in response to ChatGPT-mediated tasks.

Table 2. Specific tendencies in the use of OCSs in the debate posttest

	Group	
	Non-GenAI-assisted	GenAI-assisted
Fluency-maintenance	76.4% ₍₁₁₀₎	35.9% ₍₇₅₎
Circumlocution	4.2% ₍₆₎	18.2% ₍₃₈₎
Paraphrase	6.9% ₍₁₀₎	11.5% ₍₂₄₎
Asking for help	4.9% ₍₇₎	17.2% ₍₃₆₎
Overcorrection	7.6% ₍₁₁₎	17.2% ₍₃₆₎
Total	100% ₍₁₄₄₎	100% ₍₂₀₉₎

Note. Raw frequencies are reported in parentheses.

A chi-square test was performed to examine whether the distribution of OCS categories differed between the two conditions (see Table 3). A statistically significant association was found between instructional condition and OCS category, $\chi^2(2, N = 353) = 36.4, p < .001$, with Cramér’s *V* indicating a near-significant effect size (Cramér’s *V* = 0.321). The post-hoc analysis showed that instructional conditions were associated with differences in strategy use. A cursory examination of residuals revealed marked differences in the use of indirect strategies, which occurred significantly more frequently in the non-GenAI-assisted group (*R* = 6.01) and less frequently in the GenAI-assisted group (*R* = −6.01). The opposite pattern was also observed for direct strategies (*R* = −4.13 and 4.13, respectively). Interactional strategies, however, were the least used strategies, particularly in the non-GenAI-assisted group (*R* = −3.49). Together, these

findings indicate that differences in OCS categories largely emerged from systematic instructional differences rather than mere variation in individual strategy profiles.

Table 3. *Chi-square and post-hoc test analysis of OCS categories in the debate posttest*

$\chi^2(2, N = 353) = 36.4, p < .001;$ Cramér's $V = .321$		Category		
		IND	DT	INT
Non-GenAI-assisted	Observed counts	121	16	7
	Expected counts	94.6	31.8	17.6
	R	6.01	−4.13	−3.49
GenAI-assisted	Observed counts	111	62	36
	Expected counts	137.4	46.2	25.5
	R	−6.01	4.14	3.49

Analysis of DM Categories

In Table 4, two main categories were analysed: discourse-organisational markers (DOM) and interactional markers (IM). The GenAI-assisted group exhibited a higher frequency of DMs ($n = 298$) than its non-GenAI-assisted counterpart ($n = 159$). When we consider these categories, a complex picture emerged: a reversed distributional pattern becomes evident across both student cohorts. Discourse-organisational markers seem to play a markedly important role in the non-GenAI-assisted group as opposed to interactional markers in the other group.

Table 4. *Descriptive statistics of discourse-marker categories in the debate posttest*

	Category		
	DOM	IM	Total
Non-GenAI-assisted	61% ₍₉₇₎	38.99% ₍₆₂₎	100% ₍₁₅₉₎
GenAI-assisted	38.59 % ₍₁₁₅₎	61.40% ₍₁₈₃₎	100% ₍₂₉₈₎

Note. Raw frequencies are reported in parentheses.

As shown in Table 5, the current dataset revealed an interesting contrast in the deployment of DMs. The sparse use of DMs in the non-GenAI-assisted group was particularly evident, with comparatively higher proportions in categories related to expressing contrast (10.1%), opinions (10.7%), and signalling cause-and-effect relationships (11.9%). On the other hand, the GenAI-assisted student group demonstrated a broader functional profile through increased use of interactionally and epistemically oriented markers, mainly by embedding more hedging devices (16.8%), cause-and-effect (14.8%), and opinion markers (15.1%). This pattern suggests a shift towards more evaluative and cautious debating practice. At the same time, the limited

use of agreement- and data-oriented markers somewhat suggests a preferential attachment towards stance-taking rather than using purely informational structuring devices (e.g., referring to data).

Table 5. *Specific tendencies in the use of DMs in the debate posttest*

	Group	
	Non-GenAI-assisted	GenAI-assisted
ADD	—	3.4% ₍₁₀₎
CONT	10.1% ₍₁₆₎	8.4% ₍₂₅₎
CAUS	11.9% ₍₁₉₎	14.8% ₍₄₄₎
SIMI	3.1% ₍₅₎	—
EMPH	3.8% ₍₆₎	4.4% ₍₁₃₎
SUMM	4.4% ₍₇₎	—
STRT	1.9% ₍₃₎	0.7% ₍₂₎
MOVE	6.3% ₍₁₀₎	2.0% ₍₆₎
ADEV	—	—
SHFT	—	1.0% ₍₃₎
RESTA	3.1% ₍₅₎	4.0% ₍₁₂₎
HEDG	6.3% ₍₁₀₎	16.8% ₍₅₀₎
COPP	1.3% ₍₂₎	0.7% ₍₂₎
AGRE	9.4% ₍₁₅₎	2.0% ₍₆₎
PAGR	5.0% ₍₈₎	0.3% ₍₁₎
DISA	1.9% ₍₃₎	3.4% ₍₁₀₎
VASK	2.5% ₍₄₎	4.7% ₍₁₄₎
VOWN	—	2.0% ₍₆₎
REBU	3.1% ₍₅₎	2.3% ₍₇₎
EXAM	5.7% ₍₉₎	7.0% ₍₂₁₎
DATA	9.4% ₍₁₅₎	4.7% ₍₁₄₎
ANAL	0.0% ₍₀₎	0.7% ₍₂₎
OPIN	10.7% ₍₁₇₎	15.1% ₍₄₅₎
TURN	0.0% ₍₀₎	1.7% ₍₅₎
Total	100% ₍₁₅₉₎	100% ₍₂₉₈₎

Note. Raw frequencies are reported in parentheses; ADD = Addition / Reinforcement; CONT = Contrast / Opposition; CAUS = Cause-Effect; SIMI = Comparison and similarity; EMPH = Emphasis; SUMM = Summarizing; STRT = Starting; MOVE = Continuing; ADEV = Adding and developing; SHFT = Shifting; RESTA = Restating; HEDG = Hedging; COPP = Careful opposition; AGRE = Agreement; PAGR = Partial agreement; DISA = Disagreement; VASK = Asking for clarification; VOWN = Clarifying own point; REBU = Counterarguments; EXAM = Giving examples; DATA = Referring to data / Evidence; ANAL = Using analogies; OPIN = Expressing opinions; TURN = Turn-taking.

Based on chi-square analysis, there was a statistically significant association between instructional condition and DM category, $\chi^2(1, N = 457) = 20.9, p < .001$, indicating that the frequency distribution of DMs varied significantly across the student groups (see Table 6). The effect size demonstrated a small-to-moderate association between instructional condition and DM usage (Cramér’s $V = .214$). Considering the residuals, the post-hoc analysis indicated that the control group produced significantly more discourse-organisational markers ($R = 4.58$) and fewer interactional markers ($R = -4.58$) than expected; the opposite was true in the experimental group ($R = -4.58$ and 4.58). The results suggest that distinct DM usage patterns have principally emerged from differences in instruction, with the GenAI-assisted group favouring interactionally oriented discourse practices.

Table 6. *Chi-square and post-hoc analysis of discourse-marker categories in the debate posttest*

$\chi^2(1, N = 457) 20.9, p < .001;$ Cramér’s $V = .214$		Category	
		DOM	IM
Non-GenAI-assisted	Observed counts	97	62
	Expected counts	73.8	85.2
	R	4.58	−4.58
GenAI-assisted	Observed counts	115	183
	Expected counts	138.2	159.8
	R	−4.58	4.58

Discussion

The purpose of this article was to examine how ChatGPT-mediated tasks influenced the use of OCSs and DMs in debate interactions. The findings revealed notable differences in the distribution of OCS and DM categories across instructional conditions. We outline the specific tendencies below.

From an OCS perspective, students in the non-GenAI-assisted group relied heavily on indirect strategies, whereas the GenAI-assisted group used more direct and interactional strategies. Based on these findings, it can be deduced that Hungarian EFL undergraduates managed communication breakdowns by resorting to fluency-maintaining strategies (e.g., fillers and hesitation devices, overcorrection) without directly repairing or negotiating meaning. This behaviour is consistent with contexts where learners lack access to strategy-training programs and tend to prioritise task completion or avoid cognitively demanding strategies, rather than troubleshoot communication breakdowns themselves (Pratama & Zainil, 2020). Another interesting finding was the infrequent use of interactional strategies, which may be culturally bound to the nature of the Hungarian educational system. Although direct evidence from the present study is limited, the current pattern may reflect the broader socialisation patterns in which opportunities for dialogic negotiation and spontaneous interaction are somewhat limited. This may indirectly contribute to a reluctance to engage in conversations, coupled with a strong preference for more teacher-led instructional configurations. The severity of the situation is most apparent at the university level, where students are expected to share conversational grounds with their peers and teachers, yet they find themselves unable to do so (Win & Kálmán, 2023). We further

speculate that this pattern may be related to learners' personality traits (Kövi et al., 2023), although this warrants further investigation.

In contrast, the GenAI-assisted group used direct and interactional strategies more frequently, suggesting a greater willingness to engage in solving communication problems. Direct strategies usually reflect increased effort to participate in spoken interactions despite limited linguistic resources (e.g., compensating for missing vocabulary, restructuring an utterance), while interactional strategies were used to coordinate understanding (e.g., checking comprehension, requesting clarification, otherwise mutually managing meaning). As a result, the strategic patterns outlined here were indicative of their problem-solving and meaning-negotiation approach, using, for instance, clarification requests. One emergent situation in the experimental group was the high frequency of overcorrection despite the availability of GenAI. We attributed these findings to how GenAI scaffolding may have increased overcorrection on the part of the student. In other words, learners were more likely to be conscious of their mistakes through repeated GenAI feedback cycles. Such *optimisation* behaviour may inadvertently generate overcorrections in oral production. Overall, our findings are consistent with previous research showing persistent reliance on fluency-maintenance and message-abandonment strategies despite ample strategy training. This pattern may also reflect underdeveloped SC, as evidenced by the heavy use of low-cost strategies. Longer training programmes may therefore be warranted (Ghout-Khenoune, 2012). Towards the end of the analysis, we examined the distributional patterns of DMs across both conditions. Our findings indicated that GenAI-assisted instruction significantly influenced both the frequency and functional orientation of discourse-marker use. In particular, the GenAI-assisted group drew on a variety of DMs and, more importantly, demonstrated a stronger preference for interactional devices (e.g., opinion markers, asking for clarification, restating). On the other hand, students in the non-GenAI-assisted group used more discourse-organisational markers (e.g., contrastive, agreement, data-oriented markers). The preponderance of these markers created a discourse that is more text-structuring in nature. This style, however, aligns more closely with written academic discourse than with spoken interactions (Huneety et al., 2023).

Previous studies have shown that L2 learners' spoken discourse is often confined to highly frequent organisational devices such as *well*, *but*, or *because* (Fuller, 2003; Kapranov, 2020). Interestingly, our analysis partly diverged from this pattern, with the experimental group demonstrating a more diversified interactional profile beyond basic connective functions. It seems thus that GenAI-assisted instruction had influenced not only the frequency of DMs but also learners' rhetorical orientation and debating skills. Such a tendency seems to be a concomitant of academic debate tasks, in which persuasive engagement competes with evidence-driven reasoning (Želježič, 2017). Similar to the previous discussion, this finding may also raise an important pedagogical question regarding the limited use of interactional markers in the control group. Overall, our results corroborate previous findings on the teachability of DMs to promote learners' pragmatic-interactional competence (Taguchi, 2015). This is particularly evident in our debate setting, which naturally encourages the use of a wider range of DMs to maintain interactional coherence and negotiate interpersonal meanings.

Although our findings corroborate previous research on the potential benefits of GenAI for L2 speaking, several limitations should be acknowledged. One major limitation concerns the unpredictability of ChatGPT output, particularly in how it interprets and generates corrective feedback. This underscores the need for careful monitoring of AI tools to ensure their use is compliant with learning goals and does not introduce inappropriate or misleading feedback. As a result, caution must be taken when considering the pedagogical potential of GenAI by reinforcing the role of teachers in filtering and evaluating GenAI outputs throughout the

instructional process (Karakaya et al., 2025). Additionally, simultaneous prompting generated some interactional disruption, which at times made it difficult for some students to interact with ChatGPT effectively. Future replications should consider smaller treatment groups to minimise such interference. Longitudinal and mixed-method designs would also be valuable with the aim of providing more comprehensive accounts of how these interactional resources develop and function under more controlled instructional conditions.

Implications

The findings discussed above have important implications for ELT practitioners. ChatGPT-assisted L2 speaking instruction can serve as a pedagogical scaffold for rehearsing communication breakdowns, experimenting with alternative formulations, and receiving targeted feedback on strategy use in L2 speaking practice. ELT teachers, working with AI tools, should design prompts that direct learners beyond fluency-maintenance devices towards more interactional strategies, such as circumlocution, paraphrasing, and meaning-negotiation techniques. A substantial portion of classroom instruction should also be devoted to discourse-pragmatic markers (Fareh, 2014), with particular focus on hedging, stance-taking, and rebuttal strategies to support pragmatic appropriateness in EFL spoken interaction.

Across the various stages of L2 speaking instruction, ChatGPT should be viewed as an *adaptive* pedagogical tool that complements rather than substitutes for ELT practitioners. This requires teachers to evaluate GenAI output and guide L2 learners towards developing context-sensitive strategic repertoires instead of mechanically reproducing AI-modelled language. Particular attention should also be paid to GenAI feedback on learners' oral production, especially since learners may prefer concise forms of corrective feedback over extended metalinguistic explanations. We strongly caution that excessive exposure to AI oral correction feedback may encourage heightened monitoring of linguistic forms, potentially resulting in excessive self-correction and disrupted fluency in L2 spoken interaction. For some learners, frequent correction may prove disheartening if feedback is disproportionately focused on linguistic accuracy. ELT teachers should thus guide learners in distinguishing, for example, between productive self-repair and unnecessary modification of communicatively effective utterances. Supporting this argument, Soyoo et al. (2026) found that learners in the teacher-led group outperformed those in the ChatGPT-only group, an advantage the authors attributed to the cognitive, affective, and technical scaffolding provided by the language teacher. Although the present study did not evaluate the relative effectiveness of AI and teacher feedback, the emerging evidence suggests that these feedback modalities should not be positioned as competing alternatives. Instead, ELT practitioners should adopt a dual-feedback approach that harnesses the complementary affordances of both modalities within their instructional practices.

As noted previously, EFL teachers should remember that, on the whole, L2 learners differ in the extent to which they can benefit from AI-mediated instruction. A primary concern is that L2 learners do not hold stable characteristics equally amenable to pedagogical intervention. ELT teachers should nevertheless take individual traits into account when designing interactive tasks. In line with the pedagogical approach proposed in the present study, we believe that AI-mediated L2 speaking tasks should be embedded within a scalable pedagogical sequence that moves from individual ChatGPT-assisted rehearsals towards increasingly authentic peer-based interactional spaces, so that strategic resources developed with AI transfer reliably to real-life communication. On a final note, L2 learners are expected to engage with GenAI tools as *active* and *critical users* rather than passive recipients of modelled language. Such engagement may promote greater learner autonomy and communicative self-efficacy, thereby enabling learners to use AI as a tool for developing their strategic resources.

Conclusion

The present study set out to examine the extent to which ChatGPT-mediated instruction influenced the strategic deployment of OCSs and DMs in an EFL academic debate context. The findings point to a promising application of ChatGPT in fostering L2 speaking instruction in higher education. The key contribution of our research lies in extending the emerging literature on GenAI-mediated L2 speaking instruction by shifting attention towards interactional processes in the context of L2 academic debate. At the same time, the study bridges two distinct strands of research on interactional resources and demonstrates the value of examining their instantiation *in situ*. We thus consider learners to occupy a *quasi-apprenticeship* position within the AI-mediated instructional environment, according to which ChatGPT functions as an on-the-spot pedagogical resource that extends traditional classroom dynamics.

Although the present study did not directly assess the effectiveness of AI-generated feedback, we contend that EFL students are likely to achieve better learning outcomes when both feedback modalities are closely integrated into the instructional space. Accordingly, ELT practitioners working with GenAI tools are encouraged to critically anticipate applications in which GenAI tools are seamlessly embedded within the instructional design, in ways that enhance existing practices rather than being construed as “external” or “disruptive” to L2 speaking (Bhar, 2026).

References

- Abdelhalim, S. M., & Alsehibany, R. A. (2025). Integrating AI-powered tools in EFL pronunciation instruction: Effects on accuracy and L2 motivation. *Computer Assisted Language Learning*, 1–25. <https://doi.org/10.1080/09588221.2025.2534015>
- Al-khresheh, M. H. (2024). Bridging technology and pedagogy from a global lens: Teachers’ perspectives on integrating ChatGPT in English language teaching. *Computers and Education: Artificial Intelligence*, 6, 100218. <https://doi.org/10.1016/j.caeai.2024.100218>
- Alsaawi, A. (2022). Use of discourse markers among senior university students. *Arab World English Journal*, 13(1), 161–172. <https://doi.org/10.24093/awej/vol13no1.11>
- British Association for Applied Linguistics [BAAL]. (2021). *Recommendations on good practice in applied linguistics* (4th ed.). <https://www.baal.org.uk>
- Baidoo-Anu, D., & Owusu Ansah, L. (2023). Education in the era of generative artificial intelligence (AI): Understanding the potential benefits of ChatGPT in promoting teaching and learning. *Journal of AI*, 7(1), 52–62. <https://doi.org/10.61969/jai.1337500>
- Banh, L., & Strobel, G. (2023). Generative artificial intelligence. *Electronic Markets*, 33(1), 1–17. <https://doi.org/10.1007/s12525-023-00680-1>
- Bhar, S. K. (2026). Artificial intelligence in EFL speaking instruction: A systematic review of pedagogical design, affective conditions and instructional input. *Encyclopedia*, 6(4), 74. <https://doi.org/10.3390/encyclopedia6040074>
- Blakemore, D. (2002). *Relevance and linguistic meaning: The semantics and pragmatics of discourse markers*. Cambridge University Press. <https://doi.org/10.1017/CBO9780511486456>
- Bongaerts, T., & Poulishse, N. (1989). Communication strategies in L1 and L2: Same or different? *Applied Linguistics*, 10(3), 253–268. <https://doi.org/10.1093/applin/10.3.253>
- Bui, T. T. G., Ngo, N. V. A., & Mai, D. H. (2025). Applying AI tools to enhance argumentation skills in debate activities for EFL students. *International Journal of AI in Language Education*, 2(3), 83–100. <https://doi.org/10.54855/ijaile.25236>
- Cai, Y., & Zhang, L. J. (2023). Effects of mobile-supervised question-driven collaborative dialogues on EFL learners’ communication strategy use and academic oral English performance. *Frontiers in Psychology*, 14, 1142651. <https://doi.org/10.3389/fpsyg.2023.1142651>
- Canale, M., & Swain, M. (1980). Theoretical bases of communicative approaches to second language teaching and testing. *Applied Linguistics*, 1(1), 1–47. <https://doi.org/10.1093/applin/I.1.1>
- Carel, M. (1995). Pourtant: Argumentation by exception. *Journal of Pragmatics*, 24(1–2), 167–188. [https://doi.org/10.1016/0378-2166\(94\)00106-0](https://doi.org/10.1016/0378-2166(94)00106-0)

- Celce-Murcia, M., Dörnyei, Z., & Thurrell, S. (1995). Communicative competence: A pedagogically motivated model with content specifications. *Issues in Applied Linguistics*, 6(2), 6–35. <https://doi.org/10.5070/1462005216>
- Celik, B., Yildiz, Y., & Kara, S. (2025). Using ChatGPT as a virtual speaking tutor to boost EFL learners' speaking self-efficacy. *Australian Journal of Applied Linguistics*, 8(1), 102418. <https://doi.org/10.29140/ajal.v8n1.102418>
- Ciampa, K., Wolfe, Z., & MacDonald, E. C. (2025). AI as a cultural tool: Sociocultural scaffolding and ethical integration in upper elementary instruction. *Computers in the Schools*, 1–26. <https://doi.org/10.1080/07380569.2025.2602506>
- Creswell, J. W. (2009). *Research design: Qualitative, quantitative, and mixed methods approaches*. Sage Publications, Inc.
- Deng, R., Jiang, M., Yu, X., Lu, Y., & Liu, S. (2025). Does ChatGPT enhance student learning? A systematic review and meta-analysis of experimental studies. *Computers & Education*, 227, 1–35. <https://doi.org/10.1016/j.compedu.2024.105224>
- Dörnyei, Z., & Scott, M. L. (1997). Communication strategies in a second language: Definitions and taxonomies. *Language Learning*, 47(1), 173–210. <https://doi.org/10.1111/0023-8333.51997005>
- El Anouar, E. M., El Adnani, J. M., El Azzabi, M., Laghribi, S., Es-seqally, M., Nadif, B., & Fekkak, A. (2026). Artificial intelligence impact on teachers' professionalism: Case of ChatGPT. In N. Idrissi, A. Hair, M. Lazaar, Y. Saadi, H. Chakib, M. Erritali, & S. El Kafhali (Eds.), *Artificial intelligence and green computing* (Vol. 1589, pp. 347–359). Springer Nature Switzerland. https://doi.org/10.1007/978-3-032-02312-4_28
- Essien, A., Salami, A., Ajala, O., Adebisi, B., Shodiya, A., & Essien, G. (2024). Exploring socio-cultural influences on generative AI engagement in Nigerian higher education: An activity theory analysis. *Smart Learning Environments*, 11(1), 1–17. <https://doi.org/10.1186/s40561-024-00352-3>
- Faerch, C., & Kasper, G. (1984). Two ways of defining communication strategies. *Language Learning*, 34(1), 45–63. <https://doi.org/10.1111/j.1467-1770.1984.tb00995.x>
- Fareh, S. (2014). Macrolinguistic errors in Arab EFL learners' essays. *Procedia - Social and Behavioral Sciences*, 141, 923–933. <https://doi.org/10.1016/j.sbspro.2014.05.161>
- Fareh, S., Jarad, N., & Yagi, S. (2020). How well can Arab EFL learners adequately use discourse markers? *International Journal of Arabic-English Studies*, 20(2), 85–98. <https://doi.org/10.33806/ijaes2000.20.2.4>
- Fathi, J., Rahimi, M., & Teo, T. (2025). Applying intelligent personal assistants to develop fluency and comprehensibility, and reduce accentedness in EFL learners: An empirical study of Google Assistant. *Language Teaching Research*, 1–36. <https://doi.org/10.1177/13621688251317786>
- Folse, K. S. (2006). *The art of teaching speaking: Research and pedagogy for the ESL/EFL classroom*. University of Michigan Press.
- Fuller, J. M. (2003). Discourse marker use across speech contexts: A comparison of native and non-native speaker performance. *Multilingua - Journal of Cross-Cultural and Interlanguage Communication*, 22(2), 185–208. <https://doi.org/10.1515/mult.2003.010>
- Fung, L., & Carter, R. (2007). Discourse markers and spoken English: Native and learner use in pedagogic settings. *Applied Linguistics*, 28(3), 410–439. <https://doi.org/10.1093/applin/amm030>
- Fraser, B. (1990). An approach to discourse markers. *Journal of Pragmatics*, 14(3), 383–398. [https://doi.org/10.1016/0378-2166\(90\)90096-V](https://doi.org/10.1016/0378-2166(90)90096-V)
- Fraser, B. (1999). What are discourse markers? *Journal of Pragmatics*, 31, 931–952. [https://doi.org/10.1016/S0378-2166\(98\)00101-5](https://doi.org/10.1016/S0378-2166(98)00101-5)
- Fraser, B. (2009). An account of discourse markers. *International Review of Pragmatics*, 1(2), 293–320. <https://doi.org/10.1163/187730909X12538045489818>
- Geeslin, K., Gudmestad, A., Kanwit, M., Linford, B., Long, A., Schmidt, L., & Solon, M. (2018). Sociolinguistic competence and the acquisition of speaking. In R. Alonso (Ed.), *Speaking in a second language* (pp. 1–24). John Benjamins Publishing Company. <https://doi.org/10.1075/aals.17.01gee>
- Ghout-Khenoune, L. (2012). The effects of task type on learners' use of communication strategies. *Procedia - Social and Behavioral Sciences*, 69, 770–779. <https://doi.org/10.1016/j.sbspro.2012.11.472>

- Grzegorzewska, L. (2015). The relationship between anxiety and the use of communication strategies. *Konin Language Studies*, 3(3), 295–311. <https://www.ksj.konin.edu.pl/wp-content/uploads/2016/02/KSJ-33-295-311.pdf>
- Haenlein, M., & Kaplan, A. (2019). A brief history of artificial intelligence: On the past, present, and future of artificial intelligence. *California Management Review*, 61(4), 5–14. <https://doi.org/10.1177/0008125619864925>
- Huang, Y., & Kakhram, P. (2026). The impacts of AI conversational agents on EFL learners' oral proficiency and foreign language speaking anxiety. *Frontiers in Education*, 11, 1799269. <https://doi.org/10.3389/feduc.2026.1799269>
- Hsu, H.-L., Chen, H. H.-J., & Todd, A. G. (2023). Investigating the impact of the Amazon Alexa on the development of L2 listening and speaking skills. *Interactive Learning Environments*, 31(9), 5732–5745. <https://doi.org/10.1080/10494820.2021.2016864>
- Huneety, A., Alkhalwaldeh, A., Mashaqba, B., Zaidan, Z., & Alshdaifat, A. (2023). The use of discourse markers in argumentative compositions by Jordanian EFL learners. *Humanities and Social Sciences Communications*, 10(1), 41. <https://doi.org/10.1057/s41599-023-01525-0>
- Irshad, M. (2026). Beyond formal training: An account of informal, self-directed professional development. *Australian Journal of Applied Linguistics*, 8(4), 103119. <https://doi.org/10.29140/ajal.v8n4.103119>
- Ito, Y. (2026). Navigating AI in language education: Exploring the role of online language teacher communities. *Australian Journal of Applied Linguistics*, 8(4), 103133. <https://doi.org/10.29140/ajal.v8n4.103133>
- Kaid, J., Ali, M., & Mohammed, A. A. Q. (2023). Impact of ChatGPT on learning motivation: Teachers and students' voices. *Journal of English Studies in Arabia Felix*, 2(1), 41–49. <https://doi.org/10.56540/jesaf.v2i1.51>
- Karakaya, K., Alpat, M. F., Uçar, H., Karakaya, Ö., & Bozkurt, A. (2025). Preparing teachers for the algorithmic educational landscape: A critical mapping of generative AI integration in language teacher education. *Technology in Language Teaching & Learning*, 7(2), 102841. <https://doi.org/10.29140/ttl.v7n2.102841>
- Kapranov, O. (2020). The use of discourse markers in oral discourse in EFL classroom by pre-service EFL teachers. *Onomázein*, (6), 126–146. <https://doi.org/10.7764/onomazein.ne6.07>
- Khouchich, A., & Nadif, B. (2026). Revolutionizing business English instruction: Adopting tailored artificial intelligence solutions for teaching English to business students. *American Journal of STEM Education*, 18, 125–140. <https://doi.org/10.32674/2mt3qb76>
- Kohnke, L., Moorhouse, B. L., & Zou, D. (2023). ChatGPT for language teaching and learning. *RELC Journal*, 54(2), 537–550. <https://doi.org/10.1177/0033688223116286>
- Kövi, Z., Wongpakaran, T., Wongpakaran, N., Zábó, V., Birkás, B., & Mirnics, Z. (2023). Comparison of typical Thai and Hungarian personality profiles using the Zuckerman–Kuhlman–Aluja Personality Questionnaire. *Scientific Reports*, 13, 13508. <https://doi.org/10.1038/s41598-023-40654-z>
- Kotsis, K. T. (2024). Integrating ChatGPT into the inquiry-based science curriculum for primary education. *European Journal of Education and Pedagogy*, 5(6), 28–33. <https://doi.org/10.24018/ejedu.2024.5.6.891>
- Lai Kuen, G., Rafik-Galea, S., & Swee Heng, C. (2017). Effect of oral communication strategies training on the development of Malaysian English as a second language learners' strategic competence. *International Journal of Education and Literacy Studies*, 5(4), 57. <https://doi.org/10.7575/aiac.ijels.v5n4.p57>
- Law, L. (2024). Application of generative artificial intelligence (GenAI) in language teaching and learning: A scoping literature review. *Computers and Education Open*, 6, 100174. <https://doi.org/10.1016/j.caeo.2024.100174>
- Lowie, W., Verspoor, M., & van Dijk, M. (2018). The acquisition of L2 speaking: A dynamiperspective. In R. Alonso (Ed.), *Speaking in a second language* (pp. 105–125). John Benjamins Publishing Company. <https://doi.org/10.1075/aals.17.05low>
- Nadif, B., Zemrani, M., Ait Hammou, B., Mansour, S.-E., Fakhrou, A., & Bouih, A. (2026). Employing generative AI applications in learning: A factor analysis of motivations and effects among Moroccan

- university students. *American Journal of STEM Education*, 24, 403–424. <https://doi.org/10.32674/vt55r634>
- Nakatani, Y. (2006). Developing an oral communication strategy inventory. *Modern Language Journal*, 90(2), 151–168. <https://doi.org/10.1111/j.1540-4781.2006.00390.x>
- Nash, A., & Ramirez, O. (2025). Perceptions of how English foreign language instructional strategies may enhance adult learners' speaking skills. *Australian Journal of Applied Linguistics*, 8(3), 102574. <https://doi.org/10.29140/ajal.v8n3.102574>
- Nation, I. S. P., & Newton, J. (2009). *Teaching ESL/EFL listening and speaking*. Routledge.
- OpenAI. (2025). *OpenAI GPT-5 system card*. <http://arxiv.org/abs/2601.03267>
- Özer, H. Z., & Okan, Z. (2018). Discourse markers in EFL classrooms: A corpus-driven research. *Journal of Language and Linguistic Studies*, 14(1), 50–66. <https://www.jlls.org/index.php/jlls/article/view/759>
- Pakula, H. M. (2019). Teaching speaking. *Apples - Journal of Applied Language Studies*, 13(1), 95–111. <https://doi.org/10.17011/apples/urn.201903011691>
- Parcon, R. E., & Reyes, Z. Q. (2021). Exploring the oral communication strategies used in online classroom discussion. *ELT Forum: Journal of English Language Teaching*, 10(3), 198–211. <https://doi.org/10.15294/elt.v10i3.47152>
- Pratama, V. M., & Zainil, Y. (2020). EFL learners' communication strategy on speaking performance of interpersonal conversation in classroom discussion presentation. In *Proceedings of the 7th International Conference on English Language and Teaching (ICOELT 2019)* (pp. 29–36). Atlantis Press. <https://doi.org/10.2991/assehr.k.200306.006>
- Quadir, M. (2014). Relationships between learners' motivation factors and speaking strategy factors to learn oral communication in English. *The English Teacher*, 43(3), 113–135. <https://meltajournals.com/index.php/TET/article/view/132>
- Quirk, R., Greenbaum, S., Svartvik, J., & Leech, G. (1985). *A comprehensive grammar of the English language*. Longman.
- Redeker, G. (1991). Linguistic markers of discourse structure. *Linguistics*, 29(6), 1139–1172. <https://doi.org/10.1515/ling.1991.29.6.1139>
- Robey, D., & Taylor, W. T. (2018). Engaged participant observation: An integrative approach to qualitative field research for practitioner-scholars. *Engaged Management ReView*, 2(1), 1–15. <https://doi.org/10.28953/2375-8643.1028>
- Roll, I., & Wylie, R. (2016). Evolution and revolution in artificial intelligence in education. *International Journal of Artificial Intelligence in Education*, 26(2), 582–599. <https://doi.org/10.1007/s40593-016-0110-3>
- Rusmiyanto, R., Huriati, N., Fitriani, N., Tyas, N. K., Rofi'i, A., & Sari, M. N. (2023). The role of artificial intelligence (AI) in developing English language learner's communication skills. *Journal on Education*, 6(1), 750–757. <https://doi.org/10.31004/joe.v6i1.2990>
- Saidah, N. R., Munir, A., & Anam, S. U. (2020). Communication strategies used by EFL learners in task-based English debate. *Linguistic, English Education and Art (LEEA) Journal*, 3(2), 414–425. <https://doi.org/10.31539/leea.v3i2.1254>
- Selinker, L. (1972). Interlanguage. *International Review of Applied Linguistics*, 10(1–4), 209–232. <https://doi.org/10.1515/iral.1972.10.1-4.209>
- Seong, Y. (2014). Strategic competence and L2 speaking assessment. *Studies in Applied Linguistics and TESOL*, 14(1), 13–24. <https://doi.org/10.7916/SALT.V14I1.1313>
- Shadiev, R., Feng, Y., Zhussupova, R., & Altinay, F. (2024). Effects of speech-enabled corrective feedback technology on EFL speaking skills, anxiety and confidence. *Computer Assisted Language Learning*, 1–37. <https://doi.org/10.1080/09588221.2024.2430761>
- Soozandehfar, S. M. A. (2010). Is oral performance affected by motivation? *Journal of Pan-Pacific Association of Applied Linguistics*, 14(2), 105–119. https://www.kci.go.kr/kciportal/landing/article.kci?arti_id=ART002404230
- Soyoof, A., Reynolds, B. L., Rassaei, E., Kao, C., & Van Ha, X. (2026). From teachers to chatbots: Scaffolded corrective feedback and student trust in online L2 English classrooms. *Computers and Education: Artificial Intelligence*, 10, 100530. <https://doi.org/10.1016/j.caeai.2025.100530>

- Stockwell, G. (2024). ChatGPT in language teaching and learning: Exploring the road we're travelling. *Technology in Language Teaching & Learning*, 6(1), 2273. <https://doi.org/10.29140/ttl.v6n1.2273>
- Su, Y. C. (2021). College students' oral communication strategy use, self-perceived English proficiency and confidence, and communication anxiety in Taiwan's EFL learning. *Educational Studies*, 57(6), 650–669. <https://doi.org/10.1080/00131946.2021.1919677>
- Tai, T.-Y. (2025). Exploring the effects of intelligent personal assistants on bilingual education teachers' L2 speaking proficiency. *Interactive Learning Environments*, 33(1), 440–451. <https://doi.org/10.1080/10494820.2024.2350646>
- Tai, T.-Y., & Chen, H. H. (2024a). Improving elementary EFL speaking skills with generative AI chatbots: Exploring individual and paired interactions. *Computers & Education*, 220, 1–14. <https://doi.org/10.1016/j.compedu.2024.105112>
- Tai, T.-Y., & Chen, H. H. (2024b). The impact of intelligent personal assistants on adolescent EFL learners' speaking proficiency. *Computer Assisted Language Learning*, 37(5–6), 1224–1251. <https://doi.org/10.1080/09588221.2022.2070219>
- Tai, T.-Y., & Chen, H. H. (2025). Impact of generative AI chatbots and interaction modes on the speaking proficiency of adolescent EFL learners. *Computer Assisted Language Learning*, 1–30. <https://doi.org/10.1080/09588221.2025.2572999>
- Taguchi, N. (2015). Instructed pragmatics at a glance: Where instructional studies were, are, and should be going. *Language Teaching*, 48(1), 1–50. <https://doi.org/10.1017/S0261444814000263>
- Tarone, E. (1980). Communication strategies, foreigner talk, and repair in interlanguage. *Language Learning*, 30(2), 417–428. <https://doi.org/10.1111/j.1467-1770.1980.tb00326.x>
- Tavakoli, M., Dastjerdi, H. V., & Esteki, M. (2011). The effect of explicit strategy instruction on L2 oral production of Iranian intermediate EFL learners: Focusing on accuracy, fluency and complexity. *Journal of Language Teaching and Research*, 2(5), 989–997. <https://doi.org/10.4304/jltr.2.5.989-997>
- The jamovi project. (2025). *jamovi* (Version 2.7) [Computer software]. <https://www.jamovi.org>
- Thomas, N. (2019). Teaching L2 speaking: Recommending a holistic approach. *rEFLections*, 26(1), 134–145. <https://doi.org/10.61508/refl.v26i1.204022>
- Tran, T. T. L. (2023). The affective factors that influence oral English communication in non-major EFL classrooms. *International Journal of Language and Literary Studies*, 5(1), 127–155. <https://doi.org/10.36892/ijlls.v5i1.1161>
- Trochim, W. M. K. (2005). *Research methods: The concise knowledge base*. Atomic Dog Publishing.
- Van Den Berg, G., & Du Plessis, E. (2023). ChatGPT and generative AI: Possibilities for its contribution to lesson planning, critical thinking and openness in teacher education. *Education Sciences*, 13(10), 998. <https://doi.org/10.3390/educsci13100998>
- VERBI Software. (2025). *MAXQDA* (version 26) [computer software]. VERBI Software. <https://www.maxqda.com>
- Win, C. C., & Kálmán, C. (2023). Hungarian higher education EFL students' perceptions of learner autonomy. *Glottodidactica*, 50(2), 29–47. <https://doi.org/10.14746/gl.2023.50.2.2>
- Wu, T.T., Hapsari, I. P., & Huang, Y. M. (2025). Effects of incorporating AI chatbots into think–pair–share activities on EFL speaking anxiety, language enjoyment, and speaking performance. *Computer Assisted Language Learning*, 1–39. <https://doi.org/10.1080/09588221.2025.2478271>
- Yan, D., & Zhang, S. (2024). L2 writer engagement with automated written corrective feedback provided by ChatGPT: A mixed-method multiple case study. *Humanities and Social Sciences Communications*, 11, 1–14. <https://doi.org/10.1057/s41599-024-03543-y>
- Yang, C. T. Y., Lai, S. L., & Chen, H. H. J. (2024). The impact of intelligent personal assistants on learners' autonomous learning of second language listening and speaking. *Interactive Learning Environments*, 32(5), 2175–2195. <https://doi.org/10.1080/10494820.2022.2141266>
- Young, J. C., & Shishido, M. (2023). Investigating OpenAI's ChatGPT potentials in generating chatbot's dialogue for English as a foreign language learning. *International Journal of Advanced Computer Science and Applications*, 14(6), 65–72. <https://doi.org/10.14569/IJACSA.2023.0140607>
- Yuan, Y. (2024). An empirical study of the efficacy of AI chatbots for English as a foreign language learning in primary education. *Interactive Learning Environments*, 32(10), 6774–6789. <https://doi.org/10.1080/10494820.2023.2282112>

- Želježič, M. (2017). Debate in EFL classroom. *ELOPE: English Language Overseas Perspectives and Enquiries*, 14(1), 39–54. <https://doi.org/10.4312/elope.14.1.39-54>
- Zhang, Y., Fauzi, W. J., & Kassim, H. (2025). Input-based instruction as a teaching strategy for enhancing EFL speaking skills: A systematic review. *Discover Education*, 4(1), 534. <https://doi.org/10.1007/s44217-025-01008-1>
- Zhou, S., Tso, W. S., & Aubrey, S. (2025). Enhancing L2 learners' self-regulated speaking through digital oral dialogue journaling: An EFL classroom-based study. *RELC Journal*, 1–17. <https://doi.org/10.1177/00336882251313704>