

The Impact of Task Structure and Immediacy on International EAL Learners' Attention Orientation During Speech Production

Mohamadreza Jafary

Department of Education, Faculty of Education, LAP Program, McGill University, Canada

Sedigheh Shakib Kotamjani*

Department of Education, New Uzbekistan University, Uzbekistan

Azadeh Amoozegar

Faculty of Education and Liberal Arts, INTI International University, Negeri Sembilan, Malaysia

Correspondence

Email: s.shakibkotamjani@newuu.uz

Abstract

The paper examines the effects of task structure and task immediacy in attentional allocation in the process of engaging with (L2) speech production among international learners of English as an additional language (EAL) in Canada. In a series of four experimental conditions (structural support (structured vs. unstructured)) and immediacy (Here/Now), sixty B2 learners were asked to complete a narrative task. There/Then). Participants provided retrospective verbal reports to establish the sources of dysfluency in their speech. There are four processing stages: conceptualization, lexical encoding, syntactic encoding, and phonological encoding, which were based on the model of speech production proposed by Levelt to code the dysfluency markers. The findings indicated that unstructured tasks produced more pauses associated with conceptualization, suggesting greater requirements for message planning. Conversely, There/Then condition tasks, especially with structural support, had many more dysfluencies that were lexical, syntactic, and phonological encoding in nature, indicating increased focus on the linguistic form. The large effect size across conditions indicates that task design strongly influenced the allocation of attentional resources. Most notably, structured narratives in the past tense, without contextual support, significantly heightened attention to form, as evidenced by increased pauses during lexical, syntactic, and phonological encoding.

ARTICLE HISTORY

Received: 17 November 2025

Revised: 24 June 2026

Accepted: 16 July 2026

KEYWORDS

Task Complexity, Attention Orientation, Linguistic Form, Immediacy, Task Structure

How to cite this article (APA 7th Edition):

Jafary, M., Kotamjani, S. S., & Amoozegar, A. (2026). The impact of task structure and immediacy on international EAL learners' attention orientation during speech production. *Language Teaching Research Quarterly*, 55, 103–128. <https://doi.org/10.32038/ltrq.2026.55.06>

Introduction

In second language acquisition (SLA), noticing of linguistic form plays a central role in converting input to intake. According to the Noticing Hypothesis, put forward by Schmidt (1990, 2001), acquisition of language cannot take place unless the learner pays attention to certain features of linguistics. In task-based language teaching (TBLT), however, it is a continuous concern to maintain focus on form and still maintain meaningful communication (Dao et al., 2017; Ng et al., 2025). This issue has led scholars to explore the effects of task design variables, especially task complexity, on learners' allocation of attentional resources (Saeedi, 2023b).

Task complexity is perceived to influence the balance among fluency, accuracy, and complexity (CAF) (Wang & Jin, 2022). Skehan (2009) theorizes that attentional resources are limited: increasing cognitive demands can enhance one dimension, such as accuracy, at the expense of others, such as fluency. The Cognition Hypothesis introduced by Robinson (2011) proposes that learners have access to numerous pools of cognitive resources, and that resource-directing manipulations (e.g., $\pm$ Here-and-Now) can lead to simultaneous gains in both complexity and accuracy. The interpretation of these opposing claims has a psycholinguistic basis, drawing on Levelt's (1989) speech production model, which comprises four stages: conceptualization, formulation, articulation, and monitoring. Although there is a considerable amount of research in this field, previous studies have mostly looked at task dimensions, namely structure and immediacy, separately and have focused on the superficial CAF results. Their interactive effects or even psycholinguistic signs (markers of dysfluency) indicating real-time attentional distribution have received less attention. In response to this gap, the current research examines the interaction between the narrative task structure and immediacy, and their influence on the attentional orientation of speech by international EAL learners. This study provides a unique contribution to task-based SLA by operationalizing attentional shifts using psycholinguistic data, combining the Levelt model with the Trade-off and Cognition Hypotheses.

Although the area of task complexity research has developed over the years, less is known about the interaction of multiple dimensions of tasks to affect attentional distribution in real-time. Specifically, a few studies have operationalized it as psycholinguistic markers, including dysfluency markers. To fill this gap, this study analyses the joint effects of narrative task structure and immediacy on attentional orientation during oral performance among international students learning EAL. By situating the study within Levelt's (1990) speech production model and modeling the interaction between Skehan's Trade-off Hypothesis and Robinson's Cognition Hypothesis, this research makes a novel

contribution to task-based SLA. The given approach offers a more direct insight into the cognitive processes of EAL learners as they perform tasks, which is a gap in the current literature. The study is especially applicable to Canadian ESL settings in which learners belong to diverse language groups and communicate in real-world contexts.

The study is guided by the following research questions:

RQ₁: In what ways does task structure (structured versus unstructured) influence the attention allocation of international EAL learners during narrative speech production, as indicated by dysfluency markers associated with conceptualization, lexical encoding, syntactic encoding, and phonological encoding?

RQ₂: How does task immediacy (Here/Now versus There/Then) influence the attentional focus of international EAL learners across the same dimensions of speech processing?

RQ₃: What are the interactive effects of task structure and immediacy on learners' attention distribution during narrative speech tasks?

Literature Review

Task Complexity Effects on L2 Speaking Performance

Research on variables of task design has shown that task attributes influence the manner in which learners allocate cognitive resources when performing tasks (Vasylets et al., 2017; Chapman et al., 2025). Of these variables, task complexity has been given continued attention as the systematic manipulation of cognitive demands affects the way learners use their limited attentional capacity during task completion and learning (Kim, 2020). Empirical evidence repeatedly demonstrates that altering cognitive demands profoundly influences the allocation of attentional capacity among learners throughout task completion and acquisition. Research shows that higher task complexity, such as requiring learners to retell a story in the past tense without visual aids (i.e., a “There/Then” task), can enhance their focus on linguistic form, leading to greater accuracy in second language (L2) output (Gilabert, 2007; Rahimpour & Mehrang, 2010; Jung, 2020). Moreover, studies indicate that structured tasks are more likely to guide learners' attention towards the formal components of language, as reflected by improved accuracy scores (Skehan, 2021; Vasylets et al., 2017).

However, these findings are primarily grounded in observable linguistic variables—complexity, accuracy, and fluency and often lack psycholinguistic depth (Ahmadian et al., 2012; Javad Ahmadian et al., 2015; Najmi, 2021). Fukuta (2016) provided an analytical framework for dysfluency markers, categorizing them by their root causes, including lexical, syntactic, and phonological encoding, as well as conceptualization. The method is designed to offer a better understanding of learners' attention states while performing a task.

Empirical evidence suggests that task-complexity effects do not tend to be consistent in performance dimensions. Although in some studies accuracy improves as cognitive

demands increase, in other dimensions, especially fluency and lexical diversity, tend to decline, indicating that learners are forced to balance conflicting processing requirements within the limited attentional capacity (Skehan, 2009; Tavakoli & Foster, 2011). This conflict is consistent with the old argument between the Limited Attentional Capacity model of Skehan, a model that predicts trade-offs when cognitive load is high, and the Cognition Hypothesis of Robinson, a model that predicts that when the cognitive load is high, complexity may facilitate development by focusing attention on linguistic form (Robinson, 2011, 2021). According to Ortega (2005), manipulated structure and immediacy help boost form-based processing among learners and involve meaning. Another study conducted by Saeedi (2023a) also confirms this, finding that more complex tasks lead to changes in attention, and structured There/Then tasks are the most effective at focusing attention on a form. The model proposed by Levelt (1989) was used in Ahmadian et al. (2015), who found that under conditions of complex tasking, learners spent less effort on conceptualization and more on monitoring. This suggests that the acquisition of accuracy may be associated with shifts in attention across different processing stages.

Recent research demonstrated that task design influences the linguistic complexity and fluency of learners' speaking performance. Both Najmi (2021) and Chapman et al. (2025) found that task types significantly affected CAF, with decision-making tasks showing the highest scores, followed by narrative and personal information exchange tasks. Chapman et al. (2025) showed that the design of speaking tasks among adolescent learners influenced both the complexity and fluency effects: the length of the response during the speech exercise led to more linguistically elaborated speech but also to processing pressure, especially in the form of pausing. Similarly, Zhang and Bartlet (2025) have provided a partial inverted-U pattern in a technology-mediated environment using task-based tasks, in which moderately complex tasks tended to generate the most balanced Complexity, Accuracy, Lexical diversity, and Fluency (CALF) profile. In their analysis, the use of highly complex tasks tended to improve accuracy at the cost of lexical diversity, so that high demands can encourage careful monitoring at the cost of lexical access. Notably, Zhang and Bartlet also established that task sequencing had a major effect on fluency: ascending sequences facilitated fluency gains over time, whereas starting the speech with the hardest task in the sequence slowed the speaker's speed and then led to recovery. Taken together, these findings suggest that the complexity effects of the tasks may be dynamic rather than fixed, and that performance patterns are determined by the distribution of cognitive load over time (Zhang & Bartlet, 2025).

Beyond performance outcomes, studies have also indicated that task complexity and design shape learners' subjective experiences with tasks, which could help explain variation in performance patterns. Cho (2018) also explored the ways in which the complexity and modality of the tasks affect learners' perceptions of difficulty, the skills, and their interest in the task. Applying the repeated-measures design, Cho (2018)

established that both modality and complexity had a significant influence on perceived task difficulty. Whilst only modality had a significant impact on perceived skill and flow experience, writing had a better perceived task difficulty-to-skill balance and stronger flow experience. More importantly, flow experience was largely predicted by difficulty-skill balance, indicating that attentional focus and sense of control depend on the task demands relative to perceived abilities (Cho, 2018).

Task complexity has a significant impact on pronunciation and phonological results. Mora-Plaza et al. (2024) examined the hypothesis that increasing the cognitive demands of pronunciation-unrelated speaking activities would deteriorate L2 English pronunciation. They found that, with high task difficulty, segmental accuracy decreased, and speech was rated as more accented and slightly less comprehensible. Both Mora-Plaza et al. (2024) and Kim and Namkung (2024) claimed that increased cognitive load distracts from the focus on phonetic form and causes subsystem-specific costs in the face of language that may be saved in other dimensions. Such results support the trade-off stance of Skehan, but also clarify it: in the global context, there is no evidence of complexity decreasing the level of accuracy, but the linguistic subsystems could be affected differently, and phonological encoding seems to be especially susceptible to the high cognitive load (Mora-Plaza et al., 2024; Ngo et al., 2024; Skehan, 2009).

Recently, it has also been demonstrated that task complexity may influence both pragmatic and interpersonal aspects of L2 speaking; thereby complicating the perception that task complexity can be well elucidated by CALF measures. In an EAP achievement-testing setting, Chen (2022) used Robinson's (2007, 2021) resource-directing/resource-dispersing framework to pair decision-making tasks and investigated the effects of complexity on language use in rapport-building and on raters' ratings. According to Chen, the overall frequency of rapport-building elements did not always depend on task complexity, although task complexity did affect the association between these pragmatic features and raters' perceptions of collaboration, task completion, and style. Interestingly, the correlation between rapport language and ratings was more evident when the task was easier, suggesting that the greater the cognitive load, the less attention learners can devote to interpersonal and stylistic features of performance (Chen, 2022). Both Mora-Plaza et al. (2024) and Chen (2022) arrive at the same conclusion since they all show that the effects of task complexity are not only reflected in the outcomes of CALF but also in the underlying processes (product-level indicators, such as ratings, global judgments, and CALF measures), which are not fully specified.

Task Complexity Effect on Attentional Allocation and Speech Production Stages

Although research on task complexity has generated extensive findings on speaking performance, the current literature is limited in clarifying the effects of complexity on performance in real-time production. Most of the literature postulates indirect allocation of attention based on product-oriented indices such as complexity, accuracy, lexical

diversity, and fluency. The models of Limited Attentional Capacity by Skehan and the Cognition Hypothesis by Robinson presume that the effects of task complexity act by affecting the attentional distribution during speech production (Robinson, 2011, 2022; Skehan, 2009). The speech production model developed by Levelt (1989) is a principled framework that can be used to connect task demands and processing mechanisms. Speaking is seen as a staged process that includes conceptualization, formulation (lexical and syntactic encoding), phonological encoding, articulation, and continuous monitoring. When task demands are high, learners are supposed to redistribute the attentional resources during these stages, which may lead to visible failures in fluency, increased repair, or diminished phonological accuracy (Levelt, 1989).

To fill this explanatory gap, psycholinguistic methods treat dysfluency as an opening to the real-time processing. Fukuta (2016) proposed a framework for the analytical assessment of dysfluency signs, divided into underlying causes, such as conceptualization difficulty, as well as lexical, syntactic, and phonological encoding problems. This framework will allow researchers to see dysfluency not only as a loss of fluency worldwide but also as the manifestation of increased processing pressure when speaking (Fukuta, 2016; Fukuta & Yamashita, 2015). This would provide a more explicit connection between task complexity theories and empirical outcomes by mapping the distribution of attention onto specific production processes.

More recently, Ren et al. (2023) showed that as the task complexity increased, learners' focus on linguistic form increased as manifested in their increased monitoring during synchronous online interaction, leading to more self-repair. They also found that task sequence influenced the frequency of self-repair, with a simple-to-complex sequence facilitating more repair behaviors than a complex-to-simple sequence. The results are consistent with the suggestion that task features influence the allocation of attentional resources when learners produce language.

The recent empirical findings reinforce the usefulness of dysfluency markers for measuring task-induced cognitive strain and for monitoring individual differences. Albarqi and Tavakoli (2023) found that dual-task conditions led to more repetitions and hesitations. They found that the more proficient the learners, the fewer filled pauses they had and the better their self-monitoring ability. These results indicate that dysfluency patterns result from task-specific processing needs and learners' ability to control their attention when producing speech (Albarqi & Tavakoli, 2023). Similar process-based trends are observed in writing-based task complexity studies. According to Manchón et al. (2023), proficiency emerged as a better predictor of performance than working memory during complex manipulations, and thus, individual differences may mediate complexity effects that are not solely due to cognitive constraints. On the same note, in computer-assisted collaborative writing, Rong and Révész (2025) found that complex integrated tasks diverted cognitive resources toward planning and organization,

resulting in greater syntactic complexity but less accuracy in stimulated recall based on keystroke logging. Despite their focus on writing, these studies support the more general argument that task difficulty affects the allocation of cognitive resources between planning, formulation, and monitoring operations, with disproportionate impacts on performance metrics (Manchon et al., 2023; Rong & Révész, 2025).

Studies also show that task complexity interacts with task conditions that facilitate planning and decrease processing pressure. For example, Ryu and Lee (2018) found interaction effects between planning and task complexity in speaking: lexical complexity and accuracy were greater when complex tasks were supported by planning opportunities. Similarly, Xu and Zhang (2023) established that the syntactic complexity and functional adequacy of L2 written texts depended on the task complexity (with or without Here-and-Now), and Sánchez and Kalamakis (2023) showed that proficiency in written narrative performance depended on the interplay between task complexity (with or without Here-and-Now) and proficiency. Collectively, these results indicate that the effects of task complexity can be viewed as dynamic changes in the allocation of attention driven by task characteristics, learner variables, and the form of cognitive load (Ryu & Lee, 2018; Sánchez & Kalamakis, 2023; Xu & Zhang, 2023).

Recent evidence also indicates the emergence of an orientation to process in task-based language research. Recent studies on task-based speaking have shifted focus to the cognitive processes involved in task performance rather than assessing it solely in terms of complexity, accuracy, and fluency (CAF) (Hanzawa et al., 2026). The review highlighted the need for more psycholinguistically informed studies examining learners' attentional allocation during task performance, thereby supporting the present study's use of dysfluency markers and retrospective verbal reports to investigate real-time speech production processes.

While previous studies have examined structure OR immediacy, very few have examined their interaction using psycholinguistic markers within a between-groups design in multilingual EAL contexts. Recent research on task complexity has extended beyond documenting CALF trade-offs to studying broader performance domains, such as pronunciation, pragmatic-interactional competence, and technology-mediated task contexts (Chen, 2022; Kim & Namkung, 2024; Mora-Plaza et al., 2024). Recent task complexity research demonstrates a clear expansion from documenting CALF trade-offs toward examining broader performance domains, including pronunciation, pragmatic-interactional competence, and technology-mediated task contexts. Evidence suggests that complexity effects may be nonlinear, subsystem-specific, and sensitive to sequencing, planning, and learners' subjective experience of task difficulty and control (Cho, 2018; Zhang & Bartlett, 2025).

Simultaneously, the literature remains inadequate in two significant respects. First, it is assumed that the effects of complexity are mediated by the distribution of attention in speech production. The majority of studies continue to use product-based measures and ratings rather than psycholinguistic measures consistent with the production stages (Fukuta, 2016; Levelt, 1989). Second, few studies have explored the interaction between multiple task dimensions, specifically between task structure and immediacy (plus Here-and-Now), to determine attentional allocation at any given moment of speaking (Ortega, 2005; Skehan, 2021). The current paper fills these gaps by incorporating both the Skehan Limited Attentional Capacity model and the Cognition Hypothesis proposed by Robinson in the context of speech production by Levelt (1989), and operationalizes the idea of attentional allocation using the dysfluency indicators as related to conceptualization, lexical encoding, syntactic encoding, and phonological encoding. This process perspective provides an explanatory account of the effects of task complexity and shows that the different stages of L2 speech production are most vulnerable under higher cognitive demands.

Theoretical Framework

The theoretical framework underpinning the current work is based on three complementary schools of thought: Levelt's (1989) model of speech production, the Trade-off Hypothesis, and the Cognition Hypothesis developed by Skehan and Robinson, respectively. Collectively, these theories offer a psycholinguistic and task-related explanation of how attentional resources are distributed when performing in a second language (L2).

The model by Levelt (1989) is based on the conceptualization that speech production is a staged process that comprises conceptualization, formulation, articulation, and monitoring. In conceptualization, speakers produce preverbal messages; in formulation, there is lexical, syntactic, and phonological encoding; articulation is the speech plan implementation phase; and monitoring is the evaluation of output phase. Since these stages depend on limited attentional capacity, more processing at one stage can require more resources, limiting what is available to other stages. This model provides a cognitive structure in which the signs of dysfluency can be viewed as the signs of a redistribution of attention at processing levels.

Building on this psycholinguistic background, Skehan (2009) offers the Trade-off Hypothesis, which argues that learners operate with a finite number of attentional resources. The learners can direct their attention to one aspect of the performance, accuracy or difficulty, at the cost of another, fluency, when the task demands are elevated. In this light, the view suggests that task complexity can result in observable trade-offs among complexity, accuracy, and fluency (CAF), with competition for common cognitive resources.

On the contrary, Robinson's (2011) Cognition Hypothesis separates resource-dispersing and resource-directing variables. Resource-directing manipulations (e.g., $\pm$ Here-and-Now) are suggested to direct learners' attention to a particular domain of linguistics, which may encourage increased syntactic and lexical complexity but does not necessarily decrease accuracy. Robinson believes that greater processing and deeper cognitive complexity in certain aspects can encourage interlanguage development.

The current research synthesizes these schools of thought by considering task structure and task immediacy as cognitive factors that influence the distribution of attention in narrative speech. Task structure is considered to have an impact on the conceptualization stage by differentiating the level of organizational support, but not task immediacy (Here/Now vs.). There/Then) is a resource-guiding variable which heightens morphosyntactic and monitoring requirements. Attentional orientation is operationalized in terms of conceptualization, lexical encoding, syntactic encoding, and phonological encoding markers of disfluency, and, as a result, it would yield psycholinguistic results regarding how task design requires cognitive effort reallocation. By placing dysfluency analysis within these theoretical models, the study empirically tests the emergence of conflicting predictions from the Trade-off and Cognition Hypotheses in real-time L2 speech production.

Methodology

Design

This study employed a between-groups experimental design rather than a within-subjects design. This design has been employed to avoid potential carryover and practice effects when the same participants are involved in multiple task conditions. Considering the features of the narrative tasks, if learners are exposed to one condition, it may influence their performance in subsequent tasks, particularly in terms of lexical and structural recall. In this manner, each participant was exposed to only a single condition, thereby maintaining the manipulation intact and minimizing cross-condition contamination. Although within-subjects research methods are widely employed in the literature on task complexity (Robinson, 2011, 2022), the present study focused on minimizing priming and repetition effects by using a between-groups design. It should be noted that, despite avoiding inter-condition contamination, a sample size of 60 participants (15 per group) might limit the external validity of the findings. This is a trade-off between experimental control and statistical power, a limitation of the study.

The quantitative experimental design is complemented by retrospective verbal reports in order to gain psycholinguistic information about attentional focus on (EAL) learners' performance during narrative tasks in Canada. The independent variable was the complexity of the task, operationalized as various combinations of task structure and immediacy, yielding four different conditions. All respondents did a picture-based storytelling task in one of the following conditions (1) structured task in Here/Now

(present tense), (2) structured task in There/Then (present tense), (3) unstructured task in Here/Now, and (4) unstructured task in There/Then. The dependent variables were the conceptualization, syntactic formulation, lexical retrieval, and phonological encoding dysfluency episodes, which were compared to evaluate the attentional allocation at different stages of speech production.

Participants

The participants in this study were 60 international college students enrolled in ESL programs at a Canadian college. The college is a public post-secondary institution in Ontario, where the ESL curriculum focuses on English for Academic Purposes (EAP). Participants were recruited from "Level 4" intensive courses, which provide 20 hours of weekly instruction specifically designed to prepare international students for diploma-level academic integration. The students were between 18 and 32 years old and were non-native English speakers, with English as their additional language. The group included international students with different language backgrounds. There were 15 Ukrainians, 12 of Chinese origin, 12 of Iranian origin, 10 of Russian origin, and 12 Pakistanis. All participants were enrolled in upper-intermediate courses in ESL that are offered to college students at Canadian institutions. The respondents reported an average Canadian stay of 14 months ($SD = 6.3$), with individual stays ranging from 4 months to 3 years.

To have balanced English competence across the sample, all students were administered the Quick Oxford Placement Test (QOPT). Their findings indicated that they had the competence level that was in line with their placement in the course at the intermediate level. In the fall 2023 semester, recruiting was conducted in classes held twice a week. The decision to participate was per participant, and participants were counseled that their involvement would not influence their academic performance and that all confidential personal information would not be shared without their consent for research use.

Ethical Considerations

The Sommet College Ethics Approval No. IRB-2025-017 provided ethical approval for this study. Before participating, each participant provided their written consent. The consent process consisted of a complete description of the aims of the study, methods, possible risks and benefits, and guarantees of confidentiality and anonymity. They informed the participants that they can drop out at any time with no academic or personal consequences. To safeguard the privacy of the participants, all data were anonymized and stored safely. In line with journal guidelines, this article's Declarations section also includes a statement on ethics approval and consent procedures.

Instrumentation

A number of instruments were used to look at speech production and language skills:

Narrative tasks

Two different picture-based narrative tasks produced speech samples. Based on Tavakoli and Foster's (2011) framework, a single task included an unstructured task about a boy and girl conducting activities in a non-linear order (like going to the beach, cycling, going to the café, and staying in a hostel). The order of events could be changed without changing the story's flow. This task used a set of six individual, unnumbered sketches arranged in a randomized circular layout to prevent inherent chronological bias. The second task had a clear problem-solution narrative: kids playing football, a ball kicked into a hole, and getting it back by filling it with water. This contrast stresses the distinction between those tasks that involve producing free narrative and those that involve an underlying problem-solution structure. There is an obvious order of events and a solution to this task. These events were outlined in a series of six-strip, numbered comics, ensuring a logical, linear progression. These designs helped to investigate the effects of changes in narrative structure on fluency, accuracy, and complexity of language production (Robinson, 2011; Tavakoli & Skehan, 2009). These tasks were based on the work of Tavakoli and Foster (2011); they helped us explore how narrative structure influences attentional allocation during speech production.

Placement test

To maintain homogeneity among participants, all students had been administered the institution's homogeneity test before the study began. To determine the stability of the sample and qualify the intermediate level of competency, the Quick Oxford Placement Test (QOPT) was to be administered. This 60-item multiple-choice test, which was created jointly by Oxford University Press and Cambridge ESOL, has been tested to be reliable in gaining insight into the level of the learner (Allan, 2004) and serves as an effective way of ensuring that the participants are homogeneous and is administered in 30 minutes. The validation suggested by Allan (2004) confirms the international recognition of the QOPT as a valuable proficiency test that yields reliable outcomes across different learner levels and linguistic consistency within the whole group.

Procedure

The current study employed a between-groups design, which requires the participants to complete narrative tasks of varying difficulty. The data were gathered in a quiet language laboratory on the college campus to reduce external noise and distractions. The sessions were one-on-one, where a researcher was present to give instructions and operate the recording equipment. Each student was introduced to one set of pictures that matched one of the experimental conditions to be narrated.

The structured task in here/now (ST/HN)

The Structured Task/Here-Now (ST/HN) condition reduced cognitive load in two ways: a more basic narrative structure and continued contextual perceptivity in terms of visual accessibility. Students were shown a structured picture series for 30 seconds before they

produced present-tense descriptions under the Here/Now construct. This arrangement kept the images before the participant during the entire speaking period, eliminating the need to process them intensively since it was easy to make conclusions about visually sequenced events (Robinson, 2011). Immediacy also enhanced narrative construction by frequently referring to images and reducing cognitive complexity (Ahmadian et al., 2012).

Performing the structured task in there/then (ST/TT)

The Structured Task There-Then (ST/TT) condition demanded more cognitive load, as it involved past-tense narration and no subsequent visual representation of the images. Participants watched the orderly sequence, were then left out of sight for 30 seconds, and then started speaking in the past tense. This setup required participants to rely heavily on their memories while imposing heightened cognitive engagement (Xu et al., 2023). The intentional use of the There/Then framework through past-tense storytelling removed immediate visual scaffolding, which made the task as a whole more challenging.

Performing the unstructured task in here/now (UST/HN)

The Unstructured Task/Here-Now (UST/HN) condition exhibited heightened structural complexity due to ambiguous pictorial sequences that lacked definitive narrative initiation or resolution (Xu & Zhang, 2023). Participants were given 30 seconds to look at the unnumbered, non-linear images, which remained in front of them while they narrated in the present tense. Participants recounted stories in the present tense while observing visuals, maintaining immediacy but requiring enhanced cognitive processing to navigate the narrative's intrinsic structural disorganization.

Performing the unstructured task in there/then (UST/TT)

The Unstructured Task/There-Then (UST/TT) condition exemplified maximal complexity by concurrently altering task structure and immediacy. Participants viewed the unstructured images for 30 seconds; the images were then removed, and participants were required to narrate the events from memory using the past tense. Participants recounted unstructured narratives in the past tense without visual aids, significantly increasing cognitive burden (Robinson, 2011). This setup required complete reliance on memory recall and linguistic formulation, which put more strain on the cognitive resources needed for both conceptual planning and linguistic encoding (Skehan, 2021). These four conditions systematically determined levels of task complexity. Subsequent speech analysis assessed fluency, accuracy, and complexity characteristics to ascertain the impact of increasing complexity on narrative performance among international college students using English as an additional language. The results provided essential insights into the cognitive processing requirements across various task settings and their implications for L2 learners' ability to produce fluent, accurate narratives. Right after that, participants had a retrospective interview to identify the cognitive factors that led to their dysfluencies.

Data Coding Scheme

Speech performance analysis utilized dysfluency markers such as pauses, hesitations, false starts, repeats, and self-repairs to assess attention allocation during language output (Bui & Skehan, 2018). Dysfluency frequencies were calculated in relation to discourse units. A bilingual researcher performed retrospective interviews. Participants were encouraged to utilize their native language if desired. A second bilingual assistant recorded, transcribed, and translated all of the interviews into English. When interviews were conducted in participants' L1, translations were subsequently back-translated by an independent bilingual to ensure accuracy and reduce bias.

To verify accuracy, a third researcher examined a selection of the original and translated transcripts; any discrepancies were discussed until an agreement was achieved. The interview technique adhered to Fukuta (2016), employing playback of recorded narrations to prompt participant reflection on challenging circumstances. This strategy helped make it clear if the pauses were caused by conceptual, lexical, syntactic, or phonological processes. This triangulation of real-time dysfluencies with learners' reflections reinforces the connection between observed pauses and Levelt's (Levelt, 1989) stages of speech production.

Speech data were coded categorically to illustrate the underlying cognitive processes, including conceptualization (content planning and message formulation), syntactic encoding (word order and morphosyntactic processing), lexical choice (vocabulary selection), and phonological encoding (pronunciation and prosodic features). These four categories were used to code transcripts and speech data. A fifth of the data was coded by two trained raters, who achieved high inter-rater reliability (0.87). The discrepancies were resolved by conversation. To ensure accuracy, researchers analyzed pauses using GoldWave software.

The inter-rater reliability was also achieved with two raters independently coding 25 percent of the data (Cohen 0.87), and the differences were solved through discussion. GoldWave software was used to accurately measure pauses. Silent pauses were coded as 300 milliseconds or more of silent speech, which was acoustically identified in GoldWave. The filtering was done to avoid pauses shorter than 300 ms since the micro-hesitations, which do not imply any substantive planning processes, were excluded. We also provided effect sizes (partial η^2 with ANOVAs and Cohen d with pairwise comparisons). The examples of retrospective comments demonstrated how instances of dysfluency in learners reflected real-time difficulties in cognitive processes and in the strategic use of attention across various speech production processes.

The following are the retrospective comments (RCs) of some participants:

Conceptualization

Ex: In the picture, several students are playing soccer. Participant RC: In the first place, I thought that they were children, and then I realized that they were on a school field. I figured they must be students.

Syntactic encoding: Example

I see some kids who are playing soccer in the park. RC of the participant: I was not certain whether I would say they are playing or just play. I used the present continuous tense because I felt it was the best option.

Phonological encoding

Sample: There is a scene in a cafe where the boy and the girl are having something to drink. Participant, RC: I was not sure how the word cafe was pronounced. I was thinking about whether it was /kæ'fei/ or /kafi/.

Lexical choice: Example

Afterward, they go to the beach to swim. Participant's RC: I wasn't sure if "beach" was the best word. I was thinking maybe "seaside" would be more fitting.

Quantitative Data Analysis

Quantitative analysis of the protocol-coded data was performed using SPSS software. To calculate the frequency of episodes involving conceptualizing, syntactic encoding, phonological encoding, and lexical choice, descriptive statistics were used. After this, the means of various episodes were compared between groups through one-way ANOVA. The Scheffe post hoc test has been used to determine specific differences between the groups.

Results

The two research questions of the study concerned the effects of task demands and the variables of structure and immediacy on the attention orientation of EAL students, as indicated by the number of dysfluency markers across conceptualization, syntactic encoding, lexical choice, and phonological encoding. The results are presented for the four experimental conditions. Table 1 provides the means and standard deviations for the participants' performance across these conditions.

Descriptive Statistics

Table 1 presents the means and standard deviations for the four dependent variables: lexical encoding, syntactic encoding, phonological encoding, and conceptualization across the four task conditions (ST/HN, UST/HN, ST/TT, UST/TT).

Table 1*Mean Frequencies of Dysfluency Markers Across Task Conditions*

Dependent Variable	ST/HN	UST/HN	ST/TT	UST/TT
Lexical Encoding	5.73	2.33	9.00	5.60
Syntactic Encoding	2.86	0.86	4.73	2.80
Phonological Encoding	0.73	0.13	1.33	0.80
Conceptualization	3.06	5.86	0.72	3.60

Note: The values in parentheses are standard deviations.

Table 2*Mean Frequency of Dysfluency Markers by Processing Stage and Task Type*

Dependent Variable	Source	Sum of Squares	df	Mean Square	F	Sig.
Lexical Encoding	Between Groups	333.467	3	111.156	12.66	.000
	Within Groups	491.867	56	8.783		
	Total	825.333	59			
Syntactic Encoding	Between Groups	112.183	3	37.394	12.71	.000
	Within Groups	164.800	56	2.943		
	Total	276.983	59			
Phonological Encoding	Between Groups	10.850	3	3.617	14.07	.000
	Within Groups	14.400	56	0.257		
	Total	25.250	59			
Conceptualization	Between Groups	199.783	3	66.594	14.50	.000
	Within Groups	257.200	56	4.593		
	Total	456.983	59			

*Note. Values represent the mean number of episodes. *p < .05.*

As displayed in Table 2, manipulating task complexity through the variables of structure and immediacy significantly affected the number of dysfluency episodes related to lexical encoding, $F(3, 56) = 12.66$, $p < .001$, partial $\eta^2 = .40$. Regarding syntactic encoding, the statistical analyses indicate that the group mean differences caused by varying task demands were significant, $F(3, 56) = 12.71$, $p < .001$, partial $\eta^2 = .41$. Similarly, phonological encoding results showed statistically significant mean differences, $F(3, 56) = 14.07$, $p < .001$, partial $\eta^2 = .43$. Finally, for conceptualization, the group mean differences also reached statistical significance, $F(3, 56) = 14.50$, $p < .001$, partial $\eta^2 = .44$. According to Cohen's (Cohen et al., 2018) guidelines, all partial η^2 values represent large effect sizes, indicating that task condition explained a substantial proportion of variance for each dependent variable.

Thus, based on the one-way ANOVA results, it can be concluded that modifying task demands, through different combinations of structure and immediacy, significantly influenced learners' shifts in attention. This is reflected in the varying numbers of dysfluency markers related to lexical encoding, syntactic encoding, phonological encoding, and conceptualization.

The Effects of Task Structure on Attention Allocation

To address Research Question 1 (In what ways does task structure influence the attention allocation of international EAL learners during narrative speech production?), we compared the structured (ST) and unstructured (UST) conditions within both immediacy levels (Here/Now and There/Then). The analysis focused on four categories of dysfluency markers: conceptualization, lexical encoding, syntactic encoding, and phonological encoding. The results are summarized in Table 3 and detailed below. As shown in Table 3, in general, performing the more demanding unstructured narrative resulted in fewer dysfluency episodes related to form (i.e., lexical, syntactic, and phonological encodings). However, it led to more instances of dysfluency markers related to conceptualization. The statistical significance of these differences was confirmed through the Scheffe test results, which are reported below.

Table 3

Post-hoc Scheffe Test Results on the Effects of Task Structure on Learners' Attention Orientation

Cross-Group Comparisons	Dependent Variable	ST/HN vs. UST/HN	Mean Difference	Sig.	ST/TT vs. UST/TT
	Lexical Encoding	3.40	0.027*	3.40	0.027*
	Syntactic Encoding	2.00	0.024*	1.93	0.031*
	Phonological Encoding	0.60	0.021*	0.53	0.050*
	Conceptualization	-2.80	0.009*	-2.86	0.007*

Note: Mean difference is significant at the .05 level.

As presented in Tables 1 and 3, the mean difference for episodes related to lexical encoding between the ST/HN ($M = 5.73$, $SD = 3.34$) and UST/HN ($M = 2.33$, $SD = 2.71$) groups was statistically significant, $p = .027$, indicating that the ST/HN group produced more instances of episodes related to lexical encoding. In terms of syntactic encoding, the ST/HN group ($M = 2.86$, $SD = 1.92$) showed a higher mean compared to the UST/HN group ($M = 0.86$, $SD = 1.35$), with a significant difference, $p = .024$. Likewise, a significant difference was observed for phonological encoding, with the ST/HN group ($M = 0.73$, $SD = 0.45$) producing more dysfluency markers than the UST/HN group ($M = 0.13$, $SD = 0.35$), $p = .021$. Regarding conceptualization, the UST/HN group ($M = 5.86$, $SD = 2.44$) produced significantly more episodes related to conceptualization than the ST/HN group ($M = 3.06$, $SD = 2.46$), $p = .009$.

For the ST/TT ($M = 9.00$, $SD = 2.03$) and UST/TT ($M = 5.60$, $SD = 3.52$) groups, a significant mean difference was found for lexical encoding episodes, $p = .027$, indicating that the ST/TT group produced more instances of these episodes. The ST/TT group ($M = 4.73$, $SD = 1.86$) also showed significantly more syntactic encoding episodes compared to the UST/TT group ($M = 2.80$, $SD = 1.65$), $p = .031$. For phonological encoding, the ST/TT group ($M = 1.33$, $SD = 0.61$) outperformed the UST/TT group ($M = 0.80$, $SD = 0.56$), with a significant mean difference, $p = .05$. Regarding conceptualization, the UST/TT group ($M =$

3.60, SD = 2.29) had significantly more episodes related to conceptualization than the ST/TT group ($M = 0.72$, $SD = 1.03$), $p = .007$. These results suggest that increasing task demands, particularly in terms of structure, leads learners to focus on message conveyance, emphasizing conceptualization primarily.

The Effects of Task Immediacy on Attentional Focus

To address Research Question 2, how does task immediacy (Here/Now vs. There/Then) influence the attentional focus of international EAL learners, we compared the Here/Now (HN) and There/Then (TT) conditions across both structured and unstructured tasks. This analysis also focused on the four categories of dysfluency markers: conceptualization, lexical encoding, syntactic encoding, and phonological encoding. The relevant comparisons are presented in Table 4 and elaborated below.

Table 4

Post-hoc Scheffe Test Results on the Effects of Immediacy on Learners' Attention Orientation

Cross-Group Comparisons	Dependent Variable	ST/HN vs. ST/TT	UST/HN vs. UST/TT
		Mean Difference	Sig.
	Lexical Encoding	-3.10	.040*
	Syntactic Encoding	-1.80	.035*
	Phonological Encoding	-0.55	.022*
	Conceptualization	2.20	.042*

Note: Mean differences are significant at the .05 level.

As displayed in Tables 1 and 4, the mean difference for episodes related to lexical encoding between ST/HN ($M = 5.73$, $SD = 3.34$) and ST/TT ($M = 9.00$, $SD = 2.03$) groups was significant, $p = .036$, $d = 1.07$ (large), suggesting that the ST/TT group produced more episodes related to lexical encoding. Regarding syntactic encoding, the participants in the ST/HN group ($M = 2.86$, $SD = 1.92$) produced fewer dysfluency markers than those in the ST/TT group ($M = 4.73$, $SD = 1.86$), $p = .040$, $d = 0.97$ (large). Similarly, the mean difference between the ST/HN ($M = 0.73$, $SD = 0.45$) and the ST/TT group ($M = 1.33$, $SD = 0.61$) was significant for episodes related to phonological encoding, $p = .021$, $d = 1.26$ (very large). As for conceptualization, a comparison between the ST/HN ($M = 3.06$, $SD = 2.46$) and ST/TT ($M = 0.72$, $SD = 1.03$) groups revealed that the ST/HN group produced more episodes related to conceptualization, $p = .040$, $d = -1.04$ (large).

Comparing the mean difference between the UST/HN ($M = 2.33$, $SD = 2.71$) and UST/TT ($M = 5.60$, $SD = 3.52$) groups showed a statistically significant difference in terms of episodes related to lexical encoding, $p = .036$, suggesting that the UST/TT group produced more episodes related to lexical encoding. Similarly, the mean difference between the UST/HN ($M = 0.86$, $SD = 1.35$) and UST/TT ($M = 2.80$, $SD = 1.65$) groups was significant for episodes related to syntactic encoding, $p = .031$, with the UST/TT group showing a higher mean. For phonological encoding, the mean for the UST/HN group ($M = 0.13$, $SD = 0.35$) was significantly lower than the UST/TT group ($M = 0.80$, $SD = 0.56$), $p = .008$.

Regarding conceptualization, the mean for the UST/HN group ($M = 5.86$, $SD = 2.44$) was significantly higher than the UST/TT group ($M = 3.60$, $SD = 2.29$), $p = .048$.

These findings confirm the hypothesis that as task-immediacy difficulty increases, the frequency of episodes of lexical, syntactic, and phonological encoding increases, whereas the incidence of pauses due to conceptualization decreases.

The Interactive Effects of Task Structure and Immediacy on Attention Distribution

Finally, we investigated the interactive effects of task structure and immediacy on learners' attention distribution during narrative speech tasks to answer Research Question 3 (What are the interactive effects of task structure and immediacy on learners' attention distribution during narrative speech tasks?). We examined how the combination of task structure (structured vs. unstructured) and task immediacy (Here/Now vs. There/Then) influenced learners' attention allocation. Specifically, we compared the Structured/There-Then (ST/TT) condition with the other three task conditions to identify whether any compounded or synergistic effects occurred. The results are presented in Table 5 and discussed below. As shown in Table 5, learners in the ST/TT condition produced the highest number of dysfluency markers related to lexical, syntactic, and phonological encodings. Additionally, these learners showed the lowest number of episodes related to conceptualization, indicating that their focus was more on the form of speech rather than message conveyance.

Table 5

Post-hoc Scheffe Test Results on the Combined Effects of Structure and Immediacy on Learners' Attention Orientation

Cross-Group Comparisons	Dependent Variable	ST/TT vs. ST/HN	ST/TT vs. UST/HN	ST/TT vs. UST/TT
	Lexical Encoding	3.26 (.036*)	6.66 (.000*)	3.4 (.027*)
	Syntactic Encoding	1.86 (.040*)	3.86 (.000*)	1.93 (.031*)
	Phonological Encoding	.6 (.021*)	1.2 (.000*)	.53 (.050*)
	Conceptualization	-2.33 (.040*)	-5.13 (.000*)	-2.86 (.007*)

Note: Mean difference is significant at the .05 level

As displayed in Tables 1 and 5, the mean difference for the occurrence of episodes related to lexical encoding between ST/TT ($M = 9.00$, $SD = 2.03$) and ST/HN ($M = 5.73$, $SD = 3.34$) groups was significant, $p = .036$, $d = 1.05$ (large). Regarding syntactic encoding, the participants in the ST/TT group ($M = 4.73$, $SD = 1.86$) produced more instances of dysfluency markers than their counterparts in the ST/HN group ($M = 2.86$, $SD = 1.92$), $p = .040$, $d = 0.97$ (large). Similarly, the mean difference between the ST/TT ($M = 1.33$, $SD = 0.61$) and the ST/HN group ($M = 0.73$, $SD = 0.45$) was significant for episodes related to phonological encoding, $p = .021$, $d = 1.26$ (very large). Finally, concerning conceptualization, a comparison between the ST/TT ($M = 0.72$, $SD = 1.03$) and ST/HN (M

= 3.06, SD = 2.46) groups revealed that the former group generated fewer instances of dysfluency markers caused by conceptualization, $p = .040$, $d = -1.04$ (large).

Comparing the mean difference between the ST/TT ($M = 9$, $SD = 2.03$) and UST/HN ($M = 2.33$, $SD = 2.71$) groups showed a statistically significant mean difference in terms of episodes related to lexical encoding, $p < .001$, $d = 2.42$ (very large), showing that the former group produced more instances of this type of episode. In the same vein, the mean differences between the ST/TT ($M = 4.73$, $SD = 1.86$) and the UST/HN group ($M = .86$, $SD = 1.35$) $p < .001$, $d = 2.28$ (very large) were significant with respect to episodes related to syntactic encoding, $p = .000$, pointing to the higher mean of the former group. As for episodes related to phonological encoding, the mean for the ST/TT group ($M = 1.33$, $SD = .61$) was significantly higher than the UST/HN group ($M = .13$, $SD = .35$), $p < .001$, $d = 2.48$ (very large). On the other hand, the results for episodes related to conceptualization showed that the mean of the ST/TT group ($M = .72$, $SD = 1.03$) was significantly lower than the UST/HN group ($M = 5.86$, $SD = 2.44$), $p < .001$, $d = -2.22$ (very large).

The mean difference for episodes related to lexical encoding between the ST/TT ($M = 9$, $SD = 2.03$) and UST/TT ($M = 5.6$, $SD = 3.52$) groups was significant, $p = .027$, $d = 1.08$ (large), suggesting that the former group produced more instances. Regarding syntactic encoding, the participants in the ST/TT group ($M = 4.73$, $SD = 1.86$) also yielded more dysfluency markers than those in the UST/TT group ($M = 2.8$, $SD = 1.65$), $p = .031$, $d = 1.07$ (large). Likewise, the mean difference between the ST/TT ($M = 1.33$, $SD = .61$) and the UST/TT group ($M = .8$, $SD = .56$) was significant with regard to episodes associated with phonological encodings, $p = .05$. Finally, concerning episodes connected with conceptualization, the comparison between the ST/TT ($M = .72$, $SD = 1.03$) and UST/TT ($M = 3.6$, $SD = 2.29$) $p = .007$, $d = -1.18$ (large) revealed that the latter group produced more dysfluency markers caused by conceptualizing their message.

Altogether, the above results indicate that reducing task demands in terms of structure while simultaneously increasing them with respect to immediacy results in a maximum focus on form at the expense of focus on meaning. This significant piece of evidence, which constitutes the major contribution of the present research, will be expounded further below.

Discussion

The current research analyzed the effect of task structure and task immediacy on the allocation of attention during L2 speech production, with dysfluency markers acting as predictors of the processing load at different stages of speech production. The results show that task complexity does not necessarily increase difficulty; it only reallocates attentional resources toward conceptualization and formulation in a systematic and meaningful manner.

In terms of task structure, unstructured tasks caused much more conceptualization-linked dysfluency, whereas structured tasks caused more lexical, syntactic, and phonological encoding-linked dysfluency. This trend implies that when learners lack organizational support, they allocate more cognitive resources to message planning, leaving fewer for linguistic encoding. On the contrary, structured tasks minimize conceptual demands and enable learners to pay more attention to forms. This observation is consistent with previous research conducted by Ahmadian et al. (2012) and Tavakoli & Foster (2011), which indicates that task design influences how learners allocate cognitive effort through performance dimensions. For example, Chapman et al. (2025) and Zhang and Bartlet (2025) show that higher task demands may also facilitate elaboration and increase processing pressure, particularly through pausing behavior.

Task immediacy also contributes to understanding how task complexity influences the attentional allocation that occurs during speech production. The There/Then condition produced a greater number of dysfluencies related to lexical, syntactic, and phonological encoding but a smaller number of ones related to conceptualization, indicating increased reliance on internal representations and grammatical processing in the absence of pictures and temporal displacement. The results suggest that while a higher level of task demands caused learners to shift their focus from conceptualization to linguistic encoding, it was found that task complexity also affected the underlying cognitive processes involved in speech production (Hanzawa et al., 2026). This pattern is consistent with the Cognition Hypothesis developed by Robinson (2007, 2011), which suggests that a shift in attention to linguistic form occurs when resource-directing variables are encouraged. Meanwhile, the rise in form-related dysfluencies could be considered as a result of the greater loads involved in monitoring and encoding language, which aligns with Skehan's (2009) assumption that learners operate under limited attentional capacity. Thus, an increase in immediacy-related demands does not expand learners' cognitive resources but rather redistributes them across different stages of speech production. This interpretation aligns with recent studies, which showed that when tasks are made more difficult, attention is paid to linguistic form, but there are also processing demands associated with the task, in this case, in pronunciation, fluency, and grammatical encoding (Mora-Plaza et al., 2024; Ngo et al., 2024).

The relationship between task structure/immediacy provides the clearest demonstration of the link between task design and attentional distribution. The Structured/There-Then condition gave rise to the greatest number of dysfluencies in linguistic encoding and the fewest in conceptualization. This shows that in the case of a reduction of conceptual demands by structural support and an increase in cognitive demands by temporal displacement, the attentional resources are focused on form-related processing. This observation is consistent with recent research showing that the impact of task complexity is not linear but depends on how cognitive demands are allocated across task dimensions (Zhang & Bartlet, 2025). It is also in line with process-based studies that revealed that

complex tasks can shift attention toward linguistic organization and monitoring, namely, when planning demands are facilitated externally (Rong & Révész, 2025).

The present results not only confirm earlier research but also provide more detailed insight into the effect of immediacy on learners' cognitive processing during speech production. The results not only confirm that there/then conditions encourage greater attention to linguistic form, but also argue that, under increased task demands, attention shifts from conceptualization to linguistic encoding. This result suggests that task complexity not only affects the quality of the linguistic products but also the cognitive processes of language production. The present study includes dysfluency analysis elements that complement retrospective verbal reports to provide process-oriented evidence for Robinson's (2007, 2011) claim that resource-directing variables promote greater attention to form and also endorse Skehan's (2009) perspective on limited processing capacity. That is, the results indicate that greater demands for immediacy do not increase learners' attentional resources, but rather shift them across stages of speech production.

Notably, these findings help reconcile the apparent theoretical differences between Robinson's Cognition Hypothesis and Skehan's Trade-off Hypothesis. Rather than treating these views as opposing, the findings show that the two operate simultaneously. Students make trade-offs because of constrained attentional capacity, as postulated by Skehan (2009), although the pattern of trade-offs depends on how a task's variables direct attention, as advanced by Robinson (2011). Recent research tends to support this integrative approach, since the effects of task complexity are dynamic, context-dependent, and contingent on factors such as sequencing, modality, and the learner's perception (Cho, 2018; Kim & Namkung, 2024).

One of the main contributions of this research is the conceptualization of dysfluency as a marker of attentional distribution rather than of lower proficiency. Greater dysfluency during lexical, syntactic, or phonological encoding reflects increased monitoring and processing demands rather than lower language proficiency. This reading aligns with psycholinguistic models that view pauses, repetitions, and repairs as indicators of the demands of real-time processing (Fukuta, 2016). The study offers a more detailed description of how attentional resources are allocated to execute tasks by associating dysfluency markers with the stages of speech production, as outlined by Levelt (1989), a factor that has been under-researched in the past due to reliance on product-based indicators.

Pedagogically, the results emphasize the role of principled task sequencing in task-based language teaching. Here/Now conditions for tasks seem to assist with fluency and message building by alleviating conceptual load and are therefore appropriate in the early instructional stages. Structured There/Then tasks, on the other hand, have been

shown to promote more attention to linguistic form and can work better to promote morphosyntactic development. Unstructured activities can be communicatively rich, but will overload learners' conceptual resources unless aided by planning or scaffolding. These results support the recent call to design tasks that balance cognitive load and attentional focus rather than merely making them more complex (Kim & Namkung, 2024). These findings suggest that task complexity should be manipulated strategically rather than progressively, with task characteristics selected according to instructional objectives rather than assuming that more complex tasks always produce superior learning outcomes.

Overall, the results suggest that task complexity is not only a measure of cognitive difficulty, but also a way of allocating attentional effort. In the present study, the authors show how task structure and immediacy interact to influence learners' attentional allocation during real-time speech production by integrating Levelt's speech production model with Robinson's Cognition Hypothesis and Skehan's Trade-off Hypothesis. More significantly, this study brings a process-oriented approach, which complements traditional performance-oriented approaches, and serves to advance the psycholinguistic understanding of task complexity in multilingual contexts of EAL.

Conclusion

This study investigated the role of task structure and task immediacy in determining attentional allocation in L2 speech production. The study did not conceptualize dysfluency as a performance deficiency, but instead as a psycholinguistic marker of the distribution of cognitive resources in Levelt's (1989) model of speech production. The results indicate that the task design distributes attention systematically across the conceptualization and formulation levels, suggesting that the impact of complexity is not homogeneous but rather stage-dependent.

The findings provide robust evidence; however, the cross-sectional design and national focus on Canadian EAL intermediate learners limit generalizability. Longitudinal designs should be adopted to monitor long-term effects on learning, examine the interplay between task structure and immediacy over proficiency levels, and consider multimodal task design and AI-based support such as ASR or virtual agents. Following these patterns will further explain how task complexity and design characteristics influence learners' focus of attention and impact language development over time. This study reinforces the theoretical need to combine Levelt's (1989) and Robinson's (2011) views, providing clear pedagogical implications for incorporating narrative tasks. It adds to the current discussion on task-based SLA and provides a basis for task design in L2 teaching.

The findings empirically support both the Limited Attentional Capacity model of Skehan (2009) and the Cognition Hypothesis of Robinson (2011, 2022), indicating that the two theories are not mutually exclusive. Rather, attentional trade-offs and resource-

allocation effects are contingent on the distribution of task demand across production processes. In cases of lower conceptual loading and increased linguistic encoding requirements, learners devote more attention to morphosyntactic and phonological formulation, demonstrating that task difficulty can strategically bias form-formulation.

Pedagogically, the results highlight the significance of task sequencing in TBLT, which is guided by principles. Depending on the instructional objectives, learners' attention to tasks can be manipulated. Creating meaning-oriented tasks can support the formation of messages and fluency, and carefully designed decontextualized tasks can help focus on linguistic form. In multilingual ESL/EAL, matching task structure to learners' cognitive preparedness would help balance fluency and accuracy acquisition.

There are a number of limitations that need to be taken into account. A cross-sectional design cannot be used to draw conclusions about long-term acquisition, and the sample was confined to intermediate-level learners in a single instructional context. Longitudinal designs, different levels of proficiency, and technology-mediated task environments should be embraced in future research to advance the studies of the influence of task characteristics on attentional allocation characteristics and developmental consequences. The proposed research can help understand the role of attentional redistribution during speech production, as it provides a more accurate explanation of the effects of complexity than product-based indicators alone. By combining evidence from psycholinguistics and task complexity theory, the study offers a sophisticated explanation of the design of narrative tasks that can be optimal for L2 language teaching.

ORCID

 <https://orcid.org/0009-0003-7816-1330>

 <https://orcid.org/0000-0001-6249-0293>

 <https://orcid.org/0000-0003-4988-5113>

Publisher's Note

The claims, arguments, and counter-arguments made in this article are exclusively those of the contributing authors. Hence, they do not necessarily represent the viewpoints of the authors' affiliated institutions, or EUROKD as the publisher, the editors and the reviewers of the article.

Acknowledgements

The authors gratefully acknowledge the cooperation of the international EAL students who participated in this study and the Canadian ESL program that facilitated data collection.

Funding

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

CRedit Authorship Contribution Statement

Mohamadreza Jafary: Conceptualization, Methodology, Data collection, Investigation, Formal Analysis, Writing- Original Draft

Sedigheh Sahkib Kotamjani: Methodology, Supervision, Writing – Review & Editing, Project Administration, Funding

Azadeh Amoozegar: Writing – Review & Editing, Supervision- Funding Acquisition

Generative AI Use Disclosure Statement

Generative AI tools (Grok-4 by xAI) were used solely for basic text editing, minor copyediting, and formatting assistance, consistent with standard word-processing functions (e.g., Microsoft Word grammar/style checks). No AI was involved in generating content, analyzing data, creating code, drafting text, refining research design, or producing figures/tables/images. All content reflects the original authors' work.

Ethics Declarations

World Medical Association (WMA) Declaration of Helsinki–Ethical Principles for Medical Research Involving Human Participants

This study was conducted in accordance with the ethical standards outlined in the World Medical Association Declaration of Helsinki (2013). Ethical approval was obtained from the Institutional Review Board (IRB), IRB-2025-017. All participants were informed about the purpose of the study, and informed consent was obtained before participation. Participation was voluntary, and all data were anonymized to ensure confidentiality. The study involved no physical or psychological risk to participants.

Competing Interests

The authors declare that there are no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Data Availability

The datasets generated and/or analyzed during the current study are not publicly available due to ethical and privacy considerations, but are available from the corresponding author on reasonable request. All data have been anonymized to protect participants' confidentiality.

References

- Ahmadian, M. J., Abdolrezapour, P., & Ketabi, S. (2012). Task difficulty and self-repair behavior in second language oral production. *Applied Linguistics Review*, 3(2), 297–324. <https://doi.org/10.1515/applirev-2012-0014>

- Ahmadian, M. J., Tavakoli, M., & Vahid Dastjerdi, H. (2015). The combined effects of online planning and task structure on complexity, accuracy and fluency of L2 speech. *Language Learning Journal*, 43(1), 41–56. <https://doi.org/10.1080/09571736.2012.681795>
- Albarqi, G., & Tavakoli, P. (2023). The effects of proficiency level and dual-task condition on L2 self-monitoring behavior. *Studies in Second Language Acquisition*, 45(1), 212–233. <https://doi.org/10.1017/S0272263122000146>
- Allan, D. (2004). *The Oxford placement test*. Oxford University Press.
- Bui, G., & Skehan, P. (2018). Complexity, accuracy, and fluency. In J. I. Lontos (Ed.), *The TESOL encyclopedia of English language teaching* (pp. 1–7). Wiley. <https://doi.org/10.1002/9781118784235.eelt0046>
- Chapman, M., Montee, M., Wang, T., West, G. B., Kemp, J., & Kim, A. (2025). Speaking task design and complexity and fluency features of adolescent English learners. *Language Testing*. Advance online publication. <https://doi.org/10.1177/02655322251328665>
- Chen, H. J. (2022). Task complexity and second language performance in EAP achievement testing: Implications for the Cognition Hypothesis. *Applied Linguistics Review*, 13(4), 693–717. <https://doi.org/10.1515/applirev-2021-0199>
- Cho, M. (2018). Task complexity, modality, and flow in task-based instruction. *The Modern Language Journal*, 102(1), 224–243. <https://doi.org/10.1111/modl.12460>
- Cohen, L., Manion, L., & Morrison, K. (2018). Approaches to quantitative data analysis. In L. Cohen, L. Manion, & K. Morrison (Eds.), *Research methods in education* (8th ed., pp. 716–754). Routledge. <https://doi.org/10.4324/9781315456539-38>
- Dao, P., Iwashita, N., & Gatabonton, E. (2017). Learner attention to form in ACCESS task-based interaction. *Language Teaching Research*, 21(4), 454–479. <https://doi.org/10.1177/1362168816651462>
- Fukuta, J., & Yamashita, J. (2015). Effects of task demands on attention orientation in L2 oral production. *System*, 53, 1–12. <https://doi.org/10.1016/j.system.2015.06.01>
- Fukuta, J. (2016). Effects of task repetition on learners' attention orientation in L2 oral production. *Language Teaching Research*, 20(3), 321–340. <https://doi.org/10.1177/1362168815570142>
- Gilabert, R. (2007). Effects of manipulating task complexity on self-repairs during L2 oral production. *International Review of Applied Linguistics in Language Teaching*, 45(2), 215–240.
- Hanzawa, K., Suzuki, Y., Yanagisawa, A., & Fukuta, J. (2026). A scoping review of oral task repetition: Evolving concepts for speaking skills development. *Research Methods in Applied Linguistics*, 5(1), Article 100292. <https://doi.org/10.1016/j.rmal.2025.100292>
- Jung, J. Y. (2020). The impact of task complexity on the development of L2 grammar. *English Teaching (Seoul)*, 75(1), 93–117. <https://doi.org/10.15858/engtea.75.1.202003.93>
- Kim, N. (2020). Conditions and tasks: The effects of planning and task complexity on L2 speaking. *Journal of Asia TEFL*, 17(1), 34–52. <https://doi.org/10.18823/asiatefl.2020.17.1.3.34>
- Kim, Y. J., & Namkung, Y. (2024). Methodological characteristics in technology-mediated task-based language teaching research: Current practices and future directions. *Annual Review of Applied Linguistics*, 44, 56–78. <https://doi.org/10.1017/S0267190524000096>
- Manchón, R. M., McBride, S., Mellado Martínez, M. D., & Vasylets, O. (2023). Working memory, L2 proficiency, and task complexity: Independent and interactive effects on L2 written performance. *Studies in Second Language Acquisition*, 45(3), 737–764. <https://doi.org/10.1017/S0272263123000141>
- Mora, J. C., Mora-Plaza, I., & Bermejo Miranda, G. (2024). Speaking anxiety and task complexity effects on second language speech. *International Journal of Applied Linguistics*, 34(1), 292–315. <https://doi.org/10.1111/ijal.12494>
- Mora-Plaza, I., Mora, J. C., Ortega, M., & Aliaga-García, C. (2024). Is L2 pronunciation affected by increased task complexity in pronunciation-unfocused speaking tasks? *Studies in Second Language Acquisition*, 46(4), 1117–1149. <https://doi.org/10.1017/S0272263124000470>
- Najmi, K. (2021). The role of task types with different task complexity in EFL learners' speaking performance complexity, accuracy, and fluency across language proficiency. *Journal of New Advances in English Language Teaching and Applied Linguistics*, 3(2), 629–651. <https://doi.org/10.22034/jeltal.2021.3.2.5>
- Ng, M. L., Rahman, M. M., Rahman, M. H., Lai, T. T., & Tan, R. S. E. (2025). Exploring obstacles and effectiveness of task-based language teaching (TBLT) approach in EFL speaking instruction: Insights from an adult EFL classroom. *World Journal of English Language*, 15(6), 151–162. <https://doi.org/10.5430/wjel.v15n6p151>
- Ngo, T. T. N., Chen, H. H. J., & Lai, K. K. W. (2024). The effectiveness of automatic speech recognition in ESL/EFL pronunciation: A meta-analysis. *ReCALL*, 36(1), 4–21. <https://doi.org/10.1017/S0958344023000113>

- Ortega, L. (2005). What do learners plan? Learner-driven attention to form during pre-task planning. In R. Ellis (Ed.), *Planning and task performance in a second language* (pp. 77–109). John Benjamins.
- Rahimpour, M., & Mehrang, F. (2010). Investigating effects of task structure on EFL learners' oral performance. *English Language Teaching*, 3(4), 10–16. <https://doi.org/10.5539/elt.v3n4p10>
- Ren, W., Wu, Y., & Peng, Y. (2026). Effects of task complexity, task sequence, and interlocutor familiarity on Chinese EFL learners' self-repair in synchronous online interaction. *Language Teaching Research*, 30(3), 1544–1561. <https://doi.org/10.1177/13621688231176066>
- Robinson, P. (2007). Task complexity, theory of mind, and intentional reasoning: Effects on L2 speech production, interaction, uptake and perceptions of task difficulty. *International Review of Applied Linguistics in Language Teaching*, 45(3), 193–213. <https://doi.org/10.1515/iral.2007.009>
- Robinson, P. (2011). Second language task complexity, the cognition hypothesis, language learning, and performance. In P. Robinson (Ed.), *Second language task complexity: Researching the cognition hypothesis of language learning and performance* (pp. 3–39). John Benjamins.
- Robinson, P. (2021). The cognition hypothesis, the triadic componential framework and the SSARC model: An instructional design theory of pedagogic task sequencing. In M. J. Ahmadian & M. H. Long (Eds.), *The Cambridge handbook of task-based language teaching* (pp. 205–225). Cambridge University Press.
- Rong, X., & Révész, A. (2025). The effects of task complexity on second language writing processes and products in computer-assisted collaborative writing. *Journal of Second Language Writing*, 69, Article 101216. <https://doi.org/10.1016/j.jslw.2025.101216>
- Ryu, J., & Lee, B. (2018). Effects of task complexity and planning on L2 speaking performances and focus on form. *Korean Journal of Applied Linguistics*, 34(4), 255–286. <https://doi.org/10.17154/kjal.2018.12.34.4.255>
- Saeedi, M. (2023a). The effect of manipulating narrative task demands on EFL learners' attention orientation. *Journal of Language Horizons*, 6(4), 125–148. <https://doi.org/10.22051/lghor.2022.36735.1519>
- Saeedi, M. (2023b). Effects of repeating tasks of different complexity levels on EFL learners' oral production. *Teaching English Language*, 17(2), 391–415. <https://doi.org/10.22132/tel.2024.377408.1418>
- Sánchez, L., & Kalamakis, S. Z. (2023). The interplay of task complexity ($\pm$ Here-and-Now) and proficiency in shaping written narratives in Spanish as a foreign language. *System*, 112, Article 102949. <https://doi.org/10.1016/j.system.2022.102949>
- Schmidt, R. (1990). The role of consciousness in second language learning. *Applied Linguistics*, 11(2), 129–158. <https://doi.org/10.1093/applin/11.2.129>
- Schmidt, R. (2001). Attention and awareness in SLA. In P. Robinson (Ed.), *Cognition and second language instruction* (pp. 3–32). Cambridge University Press.
- Skehan, P. (2009). Modeling second language performance: Integrating complexity, accuracy, fluency, and lexis. *Applied Linguistics*, 30(4), 510–532. <https://doi.org/10.1093/applin/amp047>
- Skehan, P. (2021). The psycholinguistics of task-based performance. In M. J. Ahmadian & M. H. Long (Eds.), *The Cambridge handbook of task-based language teaching* (pp. 3–26). Cambridge University Press. <https://doi.org/10.1017/9781108868327.002>
- Tavakoli, P., & Foster, P. (2011). Task design and second language performance: The effect of narrative type on learner output. *Language Learning*, 61(1), 37–72. <https://doi.org/10.1111/j.1467-9922.2011.00642.x>
- Tavakoli, P., & Skehan, P. (2005). Strategic planning, task structure and performance testing. In R. Ellis (Ed.), *Planning and task performance in a second language* (pp. 239–277). John Benjamins. <https://doi.org/10.1075/llt.11.15tav>
- Vasylets, O., Gilabert, R., & Manchón, R. M. (2017). The effects of mode and task complexity on second language production. *Language Learning*, 67(2), 394–430. <https://doi.org/10.1111/lang.12228>
- Xu, T. S., & Zhang, L. J. (2023a). Effects of cognitive task complexity and online planning on second language learners' argumentative writing. *Frontiers in Psychology*, 14, Article 1121994. <https://doi.org/10.3389/fpsyg.2023.1121994>
- Wang, L., & Jin, C. (2022). Effects of task complexity on linguistic complexity for sustainable EFL writing skills development. *Sustainability*, 14(8), Article 4791. <https://doi.org/10.3390/su14084791>
- Zhang, W., & Bartlett, J. J. (2025). The impact of cognitive task complexity and task sequence on L2 speaking performance: A technology-mediated TBLT study. *Forum for Linguistic Studies*, 7(7), 918–933. <https://doi.org/10.30564/fls.v7i7.9735>