

Human-Machine Collaboration in Contextualized Vocabulary Learning for Junior Secondary English: A Case Study of Generative AI-Assisted Thematic Discourse Creation

Haoxuan Fang, Hongkun Li*

Foreign Studies College, Hunan Normal University, Changsha, Hunan, China

*Corresponding author

Abstract: In response to the persistent challenge in junior secondary English vocabulary teaching, characterized by "high dictation scores but poor vocabulary application," this paper proposes a human-machine collaborative teaching model based on generative AI. Centered on the integrated framework of pronunciation, spelling, meaning, and use, the model follows a five-step procedure — word selection, constraint setting, dialogue, review, and input and output — and employs generative AI to support thematic discourse creation and multimodal input, with the aim of facilitating contextualized vocabulary learning. Scenario simulation and pilot application indicate that the model effectively enhances students' ability to actively use vocabulary, fosters AI literacy characterized by precise prompting and critical evaluation, and provides a feasible approach to contextualized vocabulary instruction.

Keywords: generative artificial intelligence; junior secondary English; vocabulary teaching; human-machine collaboration; contextualized learning

1. Introduction

The Guidelines for Teachers on the Application of Generative Artificial Intelligence explicitly state that teachers should make effective use of AI to facilitate learning transformation and explore emerging instructional approaches such as dialogic learning. Meanwhile, junior secondary English vocabulary teaching has long been confronted with the dilemma that students "memorize words without being able to use them" and "can write words but cannot pronounce them." Although students often achieve high accuracy in vocabulary dictation, they struggle to actively retrieve and use the words they have learned in productive tasks such as writing and speaking. The fundamental cause of this problem lies in the lack of contextualized vocabulary learning. Therefore, how to employ generative AI to create contextualized discourse for students and promote deep vocabulary learning has become an issue worthy of exploration. Taking the theme of "cross-cultural communication" as an example, this paper follows the integrated vocabulary learning framework of pronunciation, spelling, meaning, and use, presents the instructional design and operational procedure of a generative AI-based junior secondary English vocabulary teaching model, and summarizes replicable experiences and recommendations.

2. Problem Statement: Challenges in Junior Secondary English Vocabulary Teaching and AI Intervention

A common phenomenon in current junior secondary English vocabulary teaching is that students achieve "high dictation scores but demonstrate poor vocabulary application." Based on field visits to local junior secondary schools and interviews with teachers, it was found that students' vocabulary learning has long remained in a passive "memorization-dictation" mode, resulting in unsatisfactory vocabulary retention. Although they often perform well in vocabulary dictation, they struggle to use the words they have learned in authentic productive tasks (Ma, 2015, p. 7). In addition, a considerable proportion of students focus solely on spelling and Chinese meanings while neglecting pronunciation and contextual meaning, leading to a disconnect among the four dimensions of pronunciation, spelling, meaning, and use.

The fundamental cause of this dilemma is that vocabulary learning has long been detached from context (Malinowski, 1923). Students memorize words in isolation without developing a deep understanding, while teachers present vocabulary out of context through explicit instruction, making it difficult for students to develop the ability to use vocabulary appropriately (Xu, 2011, p. 81). However, implementing contextualized vocabulary teaching is far from straightforward. Creating contextualized materials independently requires considerable time and effort from teachers, while students also find it difficult to construct their own vocabulary learning models. Consequently, neither teachers nor students can, on their own,

bring about a substantive transformation of vocabulary learning.

Generative AI offers new technological possibilities for addressing this challenge. Characterized by rapid responsiveness and comprehensive information coverage, generative AI can be widely applied to instructional design and resource integration in language teaching (Wu, 2023, p. 75). With AI assistance, teachers can generate thematic discourse containing target vocabulary and further integrate AI-powered functions in video editing software to achieve the integrated learning of pronunciation, spelling, meaning, and use.

3. Practical Framework: Designing a Generative AI-Based Teaching Model for Junior Secondary English Vocabulary Instruction Based on the aforementioned technological possibilities, this study proposes the following instructional model.

3.1 Design Rationale

The proposed instructional design is grounded in the following principles. First, it adopts the integrated framework of pronunciation, spelling, meaning, and use, which emphasizes both the breadth and depth of vocabulary knowledge (Dakhi & Fitria, 2019, p. 15) and seeks the balanced development of four dimensions: understanding spoken words, spelling words accurately, comprehending meanings precisely, and using vocabulary appropriately. Second, it is built upon a human-machine collaborative mechanism, in which AI is responsible for content generation and information retrieval, while human users are responsible for review and judgment. Third, it draws on the Output Hypothesis and context theory, which hold that the meaning of vocabulary depends on discourse context and that productive language use is essential for promoting language internalization (Swain, 2005, p. 471).

3.2 Design Logic and Scope of Application

To address the persistent problem that students "memorize vocabulary without being able to use it," the proposed model is operationalized through a five-step procedure: word selection → constraint setting → dialogue → review → input and output. The model is applicable to instructional units organized around clearly defined themes in junior secondary English and is recommended to be implemented over 2-3 class periods. Two domestically compliant generative AI tools are employed: Tongyi Qianwen is used for vocabulary selection, thematic discourse generation, and content optimization, while iFLYTEK Spark is responsible for content review and linguistic refinement.

4. Simulated Teaching Scenario: Applying the Five-Step Approach to Vocabulary Instruction on the Theme of "Cross-Cultural Communication"

The student prompts, teacher-student interactions, and classroom responses presented below are derived from pedagogical reasoning and reconstructed on the basis of representative teaching scenarios collected through interviews with frontline teachers. They are intended to illustrate how the proposed instructional model may be implemented in authentic classroom settings.

4.1 Word Selection: AI Assistance with Human Review

The theme of "cross-cultural communication and interaction", as specified in the English Curriculum Standards for Compulsory Education (2022 Edition) (Ministry of Education, 2022, p. 16), is used as an example. At the initial stage, students are asked to independently identify relevant vocabulary from their junior secondary English textbooks and attempt to construct sentences using the selected words. The results indicate that most students are able to produce high-frequency words such as cultural and foreign, but perform less satisfactorily in identifying more sophisticated vocabulary and constructing meaningful sentences. One student remarked, "I know all these words, but I don't know how to tell this kind of story."

In the first round of vocabulary selection, the teacher guided students to prompt the AI as follows: "Please identify all the words related to the theme of cross-cultural communication from the People's Education Press (PEP) English textbooks for Grades 7 and 8." Most students provided relatively simple prompts. Based on these instructions, the AI-generated vocabulary list included words such as restaurant and hotel, which were clearly unrelated to the target theme. Students immediately questioned the results, and the teacher took the opportunity to ask, "Should we provide the AI with more specific constraints?" After discussion, the students revised their prompt: "Please provide all the vocabulary related to 'cross-cultural communication' in the PEP junior secondary English textbooks. Restrict the results strictly to this theme, exclude general-purpose words, and indicate the textbook source for each word." Based on the revised prompt, the AI generated a

vocabulary list of 17 words, including cultural and exchange.

Rather than allowing students to use the list directly, the teacher posed a key question: "Please verify each word individually. Does it genuinely belong to the theme of cross-cultural communication?" During the discussion, one student argued that exchange simply means "to exchange" and is therefore unrelated to cross-cultural communication. Another student immediately responded, "If it refers to the exchange of ideas and perspectives between different cultures, then it is indeed a cross-cultural activity." This exchange indicates that students had begun to develop contextual awareness and were using context to interpret the specific meanings of vocabulary items. Ultimately, the class removed three semantically ambiguous words and finalized an accurate vocabulary list for the theme of cross-cultural communication.

4.2 Constraint Setting: Learning to Prompt AI Precisely

The teacher asked students to use the selected vocabulary to instruct the AI to integrate the words into a coherent discourse. The students' initial prompt was: "Write an essay using all the words above." In response, the AI generated a text that was excessively long, mixed multiple genres, included vocabulary beyond the junior secondary curriculum, and contained overly complex and lengthy sentences.

To address these issues, the teacher guided students by posing four questions concerning genre, thematic dimensions, vocabulary range, and grammatical complexity. After discussion and reflection, the students revised their prompt as follows: "Using all the thematic vocabulary listed in the table above, write a narrative telling a complete story. The length and difficulty should be appropriate for Grade 9 English. All grammatical structures should be limited to the core sentence patterns taught in junior secondary school, and individual sentences should not be overly long. In addition, no vocabulary beyond the junior secondary curriculum should be used, and all words in the passage should fall within the junior secondary English syllabus." Following these revisions, the AI generated a narrative of appropriate length with a complete and coherent structure.

The second step further develops the meaning dimension of vocabulary learning, but with a different emphasis from the first step. While the first step focuses on identifying which words belong to the target theme, the second step emphasizes the discourse context in which these words are to be used. The introduction of explicit constraints encourages students to consider the relationship between vocabulary and its contextual appropriateness.

4.3 Multi-Round Dialogue: Human-Machine Collaborative Discourse Generation

Although the first AI-generated version of the discourse incorporated all the target vocabulary, its content lacked coherence. Students raised new questions: "Why does the story suddenly move to Chinatown?" and "Why does the protagonist suddenly feel shocked here?"

The teacher then guided students into the stage of iterative dialogue for discourse refinement, encouraging them to engage in progressive prompting. The following is an example of a series of interactions.

In the first round, a student pointed out: "At the beginning of the story, the protagonist has just arrived in Canada, but in the next sentence he suddenly appears in someone else's home. The transition is too abrupt." In response, the AI added details such as "a long flight", as well as scenes involving asking for directions and taking a taxi, thereby improving the continuity of the plot.

In the second round, a student commented: "The ending is too abrupt. It does not convey any broader cultural significance. Please revise the ending." The AI then expanded the conclusion by adding the protagonist's reflections on cultural differences and intercultural understanding.

In the third round, the students requested: "The passage is too long. Shorten it to no more than 200 words while retaining the core storyline and all the thematic vocabulary." The AI subsequently generated a 198-word narrative that preserved both the complete storyline and all the target vocabulary.

Throughout this process, students shifted from being passive recipients of AI-generated content to becoming reviewers of the AI's output. They also learned to anticipate the characteristics of AI-generated responses, such as specifying word limits in advance, thereby developing a synergistic effect in human-machine collaboration (Chen & Ostolaza, 2025, p. 58).

This stage jointly emphasizes the dimensions of meaning and use. Regarding meaning, students evaluate whether vocabulary is used logically by examining the coherence and plausibility of the storyline. Regarding use, students learn how to organize vocabulary into appropriate narrative contexts while revising and refining the plot.

4.4 Three-Dimensional Review: Optimizing AI-Generated Discourse from Multiple Perspectives

The teacher introduced a three-dimensional review framework encompassing vocabulary, discourse, and values, guiding students to achieve deeper processing of the target vocabulary during the review process.

During the review, students identified the sentence "People on the street are all with more colourful and different clothes. I have never seen this in China." as reflecting a cultural stereotype. One student objected, "But this situation does

exist. Foreign styles are indeed more open than ours." The teacher then prompted further discussion by asking, "Do you think this sentence is appropriate in this context?"

Through the discussion, students came to realize that AI does not possess genuine consciousness. Its outputs are generated through the retrieval and organization of online information and therefore lack the ability to determine whether information is appropriate or contextually suitable. Value judgment must ultimately be made by humans. After careful deliberation, the class decided to delete the sentence and revise the surrounding transitions accordingly.

This step primarily focuses on deepening the dimension of meaning, guiding students to engage in more profound reflection on vocabulary use by asking: "Should this word or expression be used in this context?"

4.5 Multimodal Input and Contextual Output: Constructing the Four-in-One Framework through Three-Level Tasks

After the previous four steps, students completed the discourse generation process and became relatively familiar with the thematic vocabulary. However, there remains a substantial gap between "being able to recognize" words and "being able to use" them. More importantly, the previous four steps did not involve auditory input or pronunciation practice, making it necessary to incorporate perception and training of word pronunciation at this stage.

The teacher used the "AI text-to-video" function of video editing software to transform the generated discourse into an approximately two-minute video with standard voice-over and subtitles. The speech rate was natural, and no subtitle prompts were provided for the target vocabulary.

(1) Multimodal input and auditory training.

Students held printed copies of the text while the teacher played the video. In the first viewing, students checked the thematic vocabulary list and marked a check each time they heard a target word. High-frequency words could be quickly recognized through listening, whereas low-frequency words revealed deficiencies in the "pronunciation" dimension. In the second viewing, students circled the thematic vocabulary in the text and wrote down the contextual meanings in English. Although most students' interpretations differed from the textbook definitions, they were appropriate within the given contexts. In the third viewing, students conducted shadowing practice, focusing on the pronunciation of thematic vocabulary as well as phenomena such as connected speech and sound reduction. After these three rounds of input, "pronunciation" and "form" and "meaning" were preliminarily integrated.

(2) Independent output and awareness of human-machine division of responsibilities.

The teacher asked students to expand any plot from the text by using at least five thematic words and explicitly required that the expansion must be completed independently. Any use of AI had to be indicated with the specific stage of assistance, and direct copying was considered invalid. The following is an example of one student's expanded writing:

"Our group made a play about ethnic diversity at the talent show. Emily spoke English, and I used Chinese. This mixed our mother tongue with foreign languages. Luca did an Italian dance, and I played Chinese music. Everyone clapped happily. Mrs. Parker said, 'This is good cross-cultural teamwork!'"

The student used the thematic vocabulary appropriately, gradually achieving the four-in-one goal of "understanding spoken words, writing words accurately, comprehending meanings precisely, and using words effectively."

(3) Reflection and transfer.

The teacher posed two reflective questions: "At which step can AI help you?" and "Which step must rely on yourself?" By making the implicit thinking and judgment processes in human-machine interaction explicit, students developed transferable principles for using AI effectively.

5. Analysis of Design Advantages: Triple Enhancement in Efficiency, Quality, and Competence

5.1 Efficiency Dimension: From "Teacher-Created Materials" to "Human-Machine Collaboration"

With the support of optimized prompts and generative AI dialogue, vocabulary selection, discourse generation, and revision can be conducted more efficiently. Compared with traditional approaches relying entirely on manual work, this method significantly reduces the time required while improving the quality and relevance of the generated output (Weng, 2025, p. 10). According to the practical implementation, traditional methods require approximately 2-3 hours of concentrated effort to generate a discourse, whereas the entire process of discourse generation with AI assistance takes only about 35 minutes. The substantial improvement in efficiency enables teachers to devote more classroom time to monitoring and supporting students' mastery of the four dimensions of vocabulary learning: pronunciation, spelling, meaning, and use.

5.2 Quality Dimension: A Four-Dimensional Evaluation System Based on "Pronunciation, Spelling, Meaning, and Use"

This design evaluates students' vocabulary mastery from the four dimensions of pronunciation, spelling, meaning, and use, rather than simply quantifying vocabulary acquisition through "dictation accuracy."

(1) Pronunciation: auditory recognition and pronunciation accuracy.

After three rounds of listening discrimination practice, students' sensitivity to recognizing the thematic vocabulary through listening improved significantly. Many students gradually shifted from being reluctant to speak to actively engaging in shadowing practice.

(2) Spelling: spelling accuracy.

During students' review of the AI-generated discourse, they identified spelling errors caused by excessive data processing in AI generation, such as custom being incorrectly spelled as custum. This process reflects students' transition from passive memorization to active review and evaluation.

(3) Meaning: accuracy of contextual interpretation.

At the initial stage, students tended to first recall the Chinese meanings of words. After completing the five-step process, most students were able to describe the contextual meanings of vocabulary in English. Although their expressions differed from direct Chinese translations, they were fully appropriate within the given contexts. This indicates that students had begun to shift from mechanically memorizing Chinese definitions to actively interpreting vocabulary meanings in context.

(4) Use: Active retrieval and application of thematic vocabulary.

Based on students' performance in the initial sentence-construction task and the final plot-expansion task, the five-step process significantly enhanced students' ability to actively apply the target vocabulary. This design selected five target words — cultural, exchange, global, foreign language, and mother tongue — as the testing content and conducted a preliminary trial test on vocabulary use. The results are shown in Table 1.

Table 1. Results of the Trial Test

Test Time	Sample Size	Average Number of Correctly Used Words per Student (/5)	Accuracy Rate
Pre-test	15 students	1.8	36%
Post-test	15 students	3.9	78%

The results indicate that students' authentic vocabulary application ability improved significantly. They no longer used individual words mechanically and in isolation but were able to apply vocabulary flexibly and appropriately within specific contexts.

5.3 Competence Dimension: Cultivating Awareness of AI Use

This design is expected to enable students to develop an awareness of AI use characterized by "clearly articulating requirements, critically reviewing outputs, and independently completing final products." Students demonstrate greater precision in prompting, more critical evaluation of generated outputs, and stronger awareness of documenting AI assistance.

6. Teaching Recommendations: Transferable Experiences and Operational Guidelines

6.1 Constraint-Based Prompt Template

Through multiple rounds of practice, a constraint-based prompt template was developed as follows:

"Please use these vocabulary items [list] and write a discourse on the topic of [theme] in the form of [genre], with an approximate length of [word count] words and a difficulty level equivalent to [level]. The discourse should meet the following requirements: [specific requirements such as coherence and logical consistency]."

Taking the theme of cross-cultural communication as an example: "Please use all the vocabulary items in the table above. Focusing on the theme of cross-cultural communication, write a narrative/story of approximately 200 words at a Grade 9 junior secondary English level. The story should have a complete plot, coherent development, and rigorous logic."

6.2 Five-Step Operational Process

This design extracts a five-step operational procedure of "word selection → constraint setting → dialogue → review → input and output." This process aligns with students' cognitive development patterns in English vocabulary learning and can effectively facilitate the construction and mastery of the integrated framework of pronunciation, spelling, meaning, and

use. Table 2 presents the core tasks of each step and its corresponding role within the framework of pronunciation, spelling, meaning, and use.

Table 2. Five-Step Operational Process

Step	Core Task	Main Dimension	Teacher's Role
Word Selection	AI-assisted selection and student-based human review	Spelling + Meaning	Guiding the boundaries of inquiry
Constraint Setting	Precise prompting of AI and specification of genre/difficulty/word count	Spelling + Meaning	Demonstrating constraint conditions
Dialogue	Multi-round human-machine collaboration to generate corresponding discourse	Meaning + Use	Guiding progressive questioning
Review	Three-dimensional review and discourse optimization (vocabulary-discourse-values)	Spelling + Meaning	Cultivating critical judgment
Input and Output	Multimodal input → expansion writing → reflection	Pronunciation + Meaning + Use	Providing multimodal materials and establishing boundaries for appropriate AI use

6.3 Dynamic Discourse and Plot Expansion

Teachers can utilize AI text-to-video functions to generate dynamic discourse and construct students' perception of the "pronunciation" dimension of vocabulary through a three-level progressive input process: listening discrimination → contextual interpretation → shadowing practice.

After completing the auditory input, teachers guide students to independently expand the discourse plot by using the thematic vocabulary. This process enables students to repeatedly retrieve and apply thematic vocabulary within the target thematic context, integrating fragmented learning into a theme-oriented vocabulary network and achieving a transition from "recognizing words" to "using words."

7. Reflection and Prospects

7.1 Limitations

First, the design is primarily based on a single theme and a short implementation period. This study takes the theme of cross-cultural communication as an example, with each thematic teaching cycle completed within one to two weeks. This model demonstrates clear effectiveness for intensive learning of thematic vocabulary. However, in the long term, whether students can transfer this awareness to future language use remains to be further examined.

Second, students' AI application abilities vary. Some students encounter difficulties in the "precise prompting" stage and do not know how to implement specific operations, requiring individualized guidance from teachers. How to meet the needs of students at different levels within limited instructional time remains an issue requiring further investigation.

Third, the production of video-based learning materials involves certain technical requirements. For teachers with relatively limited information technology skills, this may still constitute a barrier.

7.2 Future Prospects

Based on the above limitations, this design can be further explored in the following directions:

First, cross-unit and cross-grade transfer of lexical fields. Vocabulary related to the same theme can be systematically organized longitudinally to form a "thematic lexical field," enabling students to repeatedly retrieve and consolidate vocabulary over a longer learning period.

Second, the development of differentiated support strategies. A "prompt template library" can be provided for students with weaker foundations, while "independent prompting challenges" can be designed for students with stronger foundations.

Third, expansion of multimodal output. At present, the output tasks in this design mainly focus on written expansion. Future studies may introduce more oral output activities to further strengthen the connection between "pronunciation" and "use."

This design constructs a five-step operational process of "word selection → constraint setting → dialogue → review → input and output" based on generative AI, which can promote students' balanced development across the four dimensions of pronunciation, spelling, meaning, and use, while cultivating their awareness of AI use characterized by precise prompting, critical evaluation, and conscious documentation. Although this design still has limitations in terms of thematic transfer, responses to student differences, and technological barriers, it provides a feasible pathway for the routine implementation of contextualized vocabulary teaching.

Acknowledgments

This paper was supported by the following fund project: A Study on Product-Oriented English Teaching in Rural Senior High Schools (Z2023123).

References

- [1] Chen, Z. Z., & Ostolaza, L. M. 2025. It Takes Two to Tango: A Longitudinal Mixed-Methods Investigation of Human-AI Co-Adaptation Across Iterative Dialogues[J]. *Studies in Applied Linguistics and TESOL*, 25(2): 51-60.
- [2] Dakhi, S., & Fitria, T. N. 2019. The Principles and the Teaching of English Vocabulary: A Review[J]. *Journal of English Teaching*, 5(1): 15-25.
- [3] Malinowski, B. 1923. The Problems of Meaning in Primitive Languages[A]. In J. Maybin, *Language and Literacy in Social Practice*[C]. Clevedon: Multilingual Matters Ltd. 1-10.
- [4] Swain, M. 2005. The Output Hypothesis: Theory and Research[A]. In E. Hinkel, *Handbook of Research on Second Language Teaching and Learning*[C]. New York: Routledge. 471-483.
- [5] Ministry of Education. 2022. *English Curriculum Standards for Compulsory Education (2022 Edition)*[M]. Beijing: People's Education Press.
- [6] Ma, Y. 2015. A Survey Study on Vocabulary Teaching Strategies of Junior Secondary English Teachers[J]. *Basic English Education*, (2): 3-8.
- [7] Weng, S. X. 2025. Research on Prompt Design Strategies for Generative Artificial Intelligence Empowering Junior Secondary English Reading Teaching[J]. *Digital Teaching in Primary and Secondary Schools*, (1): 10-16.
- [8] Wu, H. P. 2023. How Can Language Teachers "Chat" with Generative Artificial Intelligence: A Case Study of English Teaching[J]. *Digital Teaching in Primary and Secondary Schools*, (8): 75-79.
- [9] Xu, X. O. 2011. A Study on English Vocabulary Teaching in Junior Secondary English Instruction[J]. *Overseas English*, (4): 81, 88.