Available Online: 13 March 2026 DOI: 10.54254/2753-7080/2026.32246

A study on the impact of input modes on EFL learners' listening-reading comprehension

Shuyuan Yin

Xiangtan University, Xiangtan, China

yinshuyuan2024@163.com

Abstract. This study takes English learners in ordinary senior high schools in northern China as the research subjects, and compares the impact of unimodal reading and three types of bimodal input—reading after listening, reading before listening, and reading while listening—on their listening-reading comprehension. The results show that: 1) In the two tests, the effect of reading is significantly better than that of bimodal input; 2) Among the three bimodal input modalities, the reading after listening group performs significantly better than the reading before listening group in the free recall, and significantly outperforms the reading while listening group in the comprehension test; a significant difference is only found between the reading while listening group and the reading before listening group in the free recall. This study provides an empirical reference for the selection of input modalities in the teaching of listening-reading comprehension.

Keywords: input modality, listening-reading comprehension, input sequence, English learning

1. Introduction

Comprehension is the foundation and key of language acquisition, among which reading comprehension and listening comprehension are core skills in English learning, involving the process of constructing meaning from texts and sounds [1]. Listening-reading comprehension is indispensable in English teaching. Therefore, research on listening-reading comprehension in senior high school English is particularly important.

Input is the foundation of listening-reading comprehension, and input modality has become a key focus in the research of multimedia language teaching. Modality refers to the way people interact with the environment through their senses, and the combination of multiple sensory channels is called multi-modality [2]. For example, reading texts while listening to audio, watching videos or pictures, and reading textual materials with audio are all manifestations of input modalities [1]. Current research mainly focuses on comparing the effects of the listening only, the reading only, and the reading while listening. However, listening and reading involve different cognitive processes and strategies [3], and differences in their presentation sequence trigger different cognitive processing, thereby affecting comprehension outcomes. Thus, studying the sequence of listening and reading under bimodal input is crucial for analyzing the relationship between different cognitive mechanisms and learning effects, and it can also provide a basis for teaching plans that adapt to cognitive characteristics and improve learning efficiency.

Based on this, this study takes first-year senior high school EFL learners as the research subjects to explore the impact of four input modalities—reading only, reading after listening, reading before listening, and reading while listening—on listening-reading comprehension. The research is expected to screen out more optimal input modalities for senior high school English teaching and provide insights for improving students' listening-reading comprehension ability.

2. Literature review

Chang & Millett instructed 64 tenth-grade English learners to receive 90 minutes of reading or reading-while-listening training every week for 26 weeks, and found that both groups achieved improvements in reading speed and comprehension level, which could be maintained for three months without additional training, with the reading while listening group showing a more significant improvement in reading comprehension [4]. Song Beibei and He Yuan divided Chinese learners at elementary and intermediate levels into a reading only group and a reading while listening group; after receiving the corresponding input for the same duration, all participants completed a 10-minute text comprehension test. The results showed that the reading while listening group achieved significantly higher comprehension scores than the reading only group at both elementary and intermediate levels [5]. Li Shanshan and Meng Gaowang compared the discourse comprehension effects of English learners at different proficiency levels under the modalities of reading before listening, reading after listening, and reading while listening [6]. The findings indicated that regardless of proficiency level, the participants' comprehension rate under the reading while listening group was significantly higher than that under the reading after listening group, which was in turn higher than that under the reading before listening. The above studies all provide empirical support for the comprehension advantage of reading while listening.

Other studies, however, have found no such comprehension advantage in reading while listening. Gu Qiyi and Zang Chuanyun asked 99 English majors to complete multimedia listening-reading comprehension tasks under different input modes, and found that the reading only group performed the best in comprehension, followed by the reading while listening group, and the listening only group the worst, with significant differences between any two groups. The study also confirmed the existence of a redundancy effect in the simultaneous presentation of listening and reading input [7]. Serrano & Sánchez explored the impact of reading only and reading while listening on the processing of images and texts in graded readers and on reading comprehension among Spanish adolescent English learners. The results similarly showed no significant difference in reading comprehension scores between the two modalities, but it was also found that 67.6% of the learners preferred the reading while listening group, considering it easier for comprehension [8]. Sanchez et al. conducted a study on the impact of reading only and reading while listening on the reading of illustrated stories among adult native and second language learners, and their conclusions also supported that the two input modalities had similar effects on reading comprehension [1]. Overall, research on the impact of input modalities on listening-reading comprehension has focused on comparing the effects of reading while listening and reading only, and the research results remain controversial. Both the reading only modality and bimodal modalities with different sequences exert an impact on listening-reading comprehension, yet no definitive conclusion has been reached so far.

Therefore, this study examines the impact of four input modalities—reading while listening, reading before listening, reading after listening, and reading only—on listening-reading comprehension, aiming to provide more effective suggestions on input modes for English comprehension teaching in senior high school English classrooms.

3. Search methodology

3.1. Research questions

This paper mainly addresses the following questions: 1) How do the different input modality groups perform in free recall? 2) How do the different input modality groups perform in the detailed test?

3.2. Participants

The participants of this experiment were 87 first-year senior high school students from four parallel classes in an ordinary senior high school in northern China, with similar ages. Based on their mid-term exam scores of the semester, their overall English proficiency was slightly below average (Mean = 76.87, full score = 150). The four classes were randomly divided into four input condition groups: the Reading Only group (RO), the Reading After Listening group (RAL), the Reading Before Listening group (RBL), and the Reading While Listening group (RWL). One-way ANOVA showed no significant difference in English proficiency among all participants ($F = 0.575$, $p = 0.633$).

3.3. Input materials

The input material was three consecutive paragraphs of the article The Importance of Humor from the provincial textbook English (New Curriculum Standard) Selective Compulsory 1, with about 300 words. The difficulty of the text was slightly higher than the participants' English level, in line with the principles in Krashen's Input Hypothesis [9]. The audio recording of the text was the supporting audio of the textbook, with the speaking speed adjusted to about 110 words per minute, consistent with the listening speed of the senior high school entrance examination English listening test in this region.

3.4. Research instruments

This study adopted two testing instruments: free recall and detailed test.

Free recall: Referring to Lund's research [10], participants were required to write down what they had understood from listening and reading in Chinese to evaluate their overall comprehension of the text. The scoring was based on 23 scoring points determined by an experienced senior high school English teacher, with 1 point awarded for a completely correct answer to each point and 0.5 points for a partially correct answer.

Detailed test: The test items were selected from the supporting exercise book of the textbook, including 8 questions with 4 true/false questions and 4 multiple-choice questions. All questions were designed based on the key details of the text, examining participants' mastery of key words, literal information and inferential information in the text. 1 point was awarded for a correct answer to each question.

3.5. Research procedures

First, the four groups of participants received input processing in the corresponding modalities: the RO group read the material twice; the RAL group listened to the recording once and then read the text once; the RBL group read the text once and then listened to the recording once; the RWL group read the text while listening to the recording once. All groups received the input for 8 minutes. After the input, the participants first completed the free recall task, followed by the detailed test. All experimental instructions were given in the participants' native language to avoid language interference.

3.6. Data analysis

The data of this study included mid-term exam scores, free recall scores and detailed test scores. SPSS 26.0 was used for statistical analysis to test the differences in free recall and detailed test scores among the groups.

4. Results

4.1. Descriptive results

Table 1 presents the descriptive results of listening-reading comprehension. Among them, the RO group achieved the highest mean scores in both tests, which were 11.55 and 5.68 respectively; the RBL group had the lowest mean scores in free recall (7.07) and listening-reading comprehension (11.12), and also the largest standard deviations in both tests (2.42; 1.49), indicating a large fluctuation in the score distribution of this group in the two tests; the RWL group had the lowest mean score in multiple-choice questions (3.71) and the smallest standard deviation, meaning the multiple-choice scores of most students in this group were close to the mean.

Table 1. Results of free recall and detailed test

Test	Modality	N	Min	Max	Mean	SD
Free Recall	RO	22	7	16	11.55	2.36
	RAL	22	5	13	8.89	1.97
	RWL	21	5	12	8.67	1.79
	RBL	22	4	14	7.07	2.42
Detailed Test	RO	22	4	7	5.68	1.04
	RAL	22	2	7	4.50	1.30
	RBL	22	2	8	4.05	1.49
	RWL	21	2	5	3.71	1.00

4.2. Intergroup comparison results

One-way ANOVA on free recall scores showed a significant main effect of input modality ($F = 16.330$, $p = 0.000$, $\eta^2 = 0.371$).

Table 2 shows the post-hoc multiple comparison results of free recall among the four groups. The RO group performed significantly better than all bimodal groups; the RAL group and the RWL group achieved significantly higher scores than the RBL group, while there was no significant difference between the RAL group and the RWL group ($p = 0.739$).

Table 2. Post-hoc multiple comparison results of free recall test

Group (I)	Group (J)	Mean Difference (I-J)	Significance	Standard Error
RO	RAL	2.66	.000	.65
	RBL	4.48	.000	.65
	RWL	2.88	.000	.66
RAL	RBL	1.82	.006	.65
	RWL	.22	.739	.66
RBL	RWL	-1.59	.017	.66

One-way ANOVA on detailed test scores revealed a significant main effect of input modality (F = 10.632, $p = 0.000$, $\eta^2 = 0.278$).

Table 3 presents the intergroup comparison results of the detailed test among the four groups. The RO group performed significantly better than all bimodal groups; the RAL group scored significantly higher than the RWL group; there were no significant differences between the RAL group and the RBL group, nor between the RBL group and the RWL group ($p = 0.224$; $p = 0.380$).

Table 3. Post-hoc multiple comparison results of detailed test

Group (I)	Group (J)	Mean Difference (I-J)	Significance	Standard Error
RO	RAL	1.18	.002	.37
	RBL	1.64	.000	.37
	RWL	1.97	.000	.38
RAL	RBL	.46	.224	.37
	RWL	.79	.039	.38
RBL	RWL	.33	.380	.38

5. Discussion

This study shows that reading only has a significant advantage over other modalities in both free recall and the detailed test, with the RO group achieving much higher listening-reading comprehension scores than the bimodal groups, which is consistent with the conclusions of many previous studies [11, 12]. The RO group's advantages in the two tests are most likely due to the participants' greater familiarity with the reading-only modality [13] and the single cognitive channel it involves. For Chinese EFL learners, this type of input reduces intrinsic and extraneous cognitive load, allowing more cognitive resources to be allocated to text comprehension, thus achieving better comprehension effects [14].

Secondly, the lower scores of the bimodal groups are probably due to the participants' poor listening ability and the high cognitive load of the bimodal input process. On the one hand, the cultivation of listening ability for Chinese EFL learners is mainly conducted in classrooms with a lack of effective practice methods, resulting in their listening proficiency being far inferior to their reading ability [15]; in addition, during listening, participants passively receive information in a continuous linear manner, and it is difficult to remedy misunderstandings, leading to a higher cognitive load [13]. Given that the participants in this study had a weak English foundation, the bimodal input groups needed to process auditory input in addition to visual processing, and this dual cognitive load was likely to cause cognitive overload [12]. On the other hand, due to their insufficient listening ability, the participants in the bimodal input groups may have subconsciously relied on reading—their stronger skill that consumes fewer cognitive resources—to understand the text. However, the presence of listening input may still interfere with reading, resulting in poorer comprehension effects of the bimodal input groups compared with the RO group.

The intergroup comparison of free recall scores among the three bimodal groups shows that the effect of the reading after listening is better than that of the reading while listening without a significant difference, and both groups perform significantly better than the reading before listening group. This is partially inconsistent with the findings of Li Shanshan and Meng Gaowang [6], who found that the reading while listening group performed the best (especially the low-proficiency subgroup), with the other two groups showing similar effects. This difference is most likely due to the different input processing mechanisms and repetition opportunities. In Li's study, the bimodal groups received input for the same duration with no additional learning tasks; in contrast, the RWL group in this study conducted a single simultaneous processing of input,

and the intensified chunking of words distracted their attention, thus the RWL group did not perform as well as that in Li's study.

The comparison of free recall scores among the bimodal groups in this study can be explained from the perspectives of modal matching and modal characteristics in memory retrieval. When the modality used in the testing phase is consistent with that in the memory encoding phase, the efficiency of knowledge retrieval is higher; meanwhile, written input allows learners to conduct in-depth processing, encoding and storage of information, which is more conducive to the long-term retention of memory compared with transient auditory input [16]. The free recall task in this study was conducted in a visual mode, and the three bimodal groups showed differences in memory effects due to their different combinations of input modalities: after receiving auditory information, the RAL group could supplement the information that may have been missed or ambiguous during listening through the subsequent reading session, thus forming a more complete and stable memory representation. In addition, the shorter time interval between the reading session and the visual test enables a faster modal matching between input and testing, further improving the efficiency of memory retrieval. According to the theoretical mechanisms of Amer and Mayer, the RWL group constructs mental models of language and images by integrating lexical information and auditory input simultaneously, and establishes connections between dual representations with the help of prior knowledge [17, 18]. This information processing method oriented to holistic comprehension, although not significantly surpassing the RAL group in memory performance, can still effectively grasp the overall information of the text, supporting its good performance in the free recall task. In contrast, in the reading before listening modality, the subsequent auditory input not only has a weaker effect on memory retention than written input, but may also interfere with or even destroy the text comprehension established through the previous reading due to the learners' insufficient listening ability and discontinuous information reception, ultimately leading to the RBL group's significantly lower free recall scores than the other two bimodal groups.

The RWL group scored significantly lower than the RAL group in the detailed test, with no significant differences between any other two bimodal groups. This can be explained from the perspectives of cognitive resource allocation and information processing mechanisms. Compared with other bimodal groups, the RWL group conducted a single simultaneous processing of input with half the input time, and needed to process dual-channel auditory and visual information synchronously, which places higher demands on the cognitive system [19]. In addition, the reading while listening prompts participants to focus more on the holistic comprehension of the text rather than decomposing it into scattered parts [17]. This holistic orientation is likely to compress the cognitive space for detailed processing, making it difficult for learners to dig deep into the specific information in the text, thus affecting their performance in the detailed comprehension task. In contrast, the RAL group does not need to process bimodal information synchronously; it first constructs an overall framework through listening to form a basic verbal representation, avoiding the cognitive distraction caused by the simultaneous processing of bimodal input. The subsequent reading session allows learners to review details autonomously and verify the information obtained from listening [15], strengthening the correspondence between speech and words. This in turn reduces extraneous processing and allows more resources to be concentrated on understanding detailed information, thus giving the RAL group an advantage in the detailed comprehension task.

Overall, the research results support that all four input modes can promote listening-reading comprehension, and for the learners in this study, the reading only is the most effective comprehension method at present. Although bimodal input has shown advantages in many previous studies, such advantages cannot be realized when learners' cognitive resources and listening ability are insufficient to process the input information.

6. Conclusion

This study explores the impact of four input modalities on the listening-reading comprehension of Chinese EFL learners. The results indicate that all four input modalities are conducive to listening-reading comprehension, and for English learners in ordinary senior high schools, the reading only modality is more effective than bimodal input modalities, among which the reading after listening is the most effective one, followed by the other two bimodal modalities. In short, each of the four modalities has its own strengths.

Since different input modalities have significantly different effects on different learning objectives—for example, bimodal input can assist in improving listening-reading comprehension ability in specific scenarios—English teachers need to dynamically select appropriate input modalities based on specific teaching objectives and students' proficiency levels. For students with a weak English foundation and for the teaching objective of improving comprehension efficiency and reducing cognitive load, the traditional reading only modality can be prioritized, and its advantages such as text enhancement (e.g., bold font) and self-paced reading can be utilized to help students improve their detailed comprehension effects. If the teaching objective is to cultivate students' listening-reading coordination ability and the students have a certain English foundation, bimodal input teaching can be introduced. Teachers do not need to adopt the complex reading while listening modality directly; instead, they can start with transitional forms such as reading before listening and reading after listening to enable learners to gradually adapt to the rhythm of bimodal input. This can reduce the risk of cognitive overload while giving full play to the positive role of bimodal input in listening-reading comprehension, ultimately optimizing learning effects under different teaching objectives.

An important limitation of this study is the lack of qualitative data. Relying solely on quantitative data makes it difficult to capture "hidden information" such as students' real thoughts and thinking processes when receiving bimodal input, their strategies for dealing with incomprehensible content, or their intuitive feelings about teaching materials, leading to an insufficiently in-depth and thorough analysis of the research results. In addition, the RWL group in this study reads while listening only once, resulting in greater time pressure than the other groups, and the design of this link needs to be optimized. Future research can further explore the impact of the reading while listening modality and other modalities on listening-reading comprehension with the same input time or different frequencies of input.

References

[1] Sanchez, P. A., Conklin, K., Rodgers, M., & Parente, F. (2021). The effect of auditory input on multimodal reading comprehension: An examination of adult readers' eye movement. *Modern Language Journal, 105*(4), 936-956.

[2] Gu, Y. D. (2007). An analysis of multimedia and multimodal learning. *Computer-Assisted Foreign Language Education*, (02), 3-12.

[3] Dai, J. (2007). Input modes, input frequency and discourse comprehension. *Foreign Language Teaching and Research*, (04), 285-293.

[4] Chang, C.-S., & Millett, S. (2015). Improving reading rates and comprehension through audio-assisted extensive reading for beginner learners. *System, 52*, 91-102.

[5] Song, B. B., & He, Y. (2024). The impact of text enhancement and input modes on the learning of verb-noun collocation chunks and text reading comprehension in Chinese as a second language. *Chinese Teaching in the World*, (01), 126-140.

[6] Li, S. S., & Meng, G. W. (2022). The impact of presentation modes of visual and auditory input on second language discourse comprehension. *Journal of Heihe University*, (07), 120-122.

[7] Gu, Q. Y., & Zang, C. Y. (2011). The impact of input modes on second language comprehension and incidental vocabulary acquisition. *Journal of PLA University of Foreign Languages*, (03), 55-59.

[8] Serrano, R., & Sánchez, P. A. (2019). Young L2 learners' online processing of information in a graded reader during reading-only and reading-while-listening conditions: A study of eye-movements. *Applied Linguistics Review, 13*(1), 49-70.

[9] Krashen, S. (1985). *The Input Hypothesis: Issues and Implications.* London: Longman.

[10] Lund, R. J. (1991). A comparison of second language listening and reading comprehension. *Modern Language Journal, 75*(2), 196-204.

[11] Diao, Y. L., & Sweller, J. (2007). Redundancy in foreign language reading comprehension instruction: Concurrent written and spoken presentations. *Learning and Instruction, 17*(1), 78-88.

[12] Gu, Q. Y., & Yin, N. (2017). Cognitive load in the discourse comprehension process of Chinese EFL learners: A comparative study of three input modes. *Foreign Language Teaching and Research*, (05), 754-766.

[13] Dai, J. (2005). Film and television subtitles and foreign language teaching. *Computer-Assisted Foreign Language Education*, (03), 18-22.

[14] Sweller, J. (2005). *Implications of cognitive load theory for multimedia learning. In R. E. Mayer (Ed.), The Cambridge Handbook of Multimedia Learning (pp.19-30).* New York: Cambridge University Press.

[15] Dai, J. (2014). The gap between listening and reading comprehension abilities of Chinese EFL learners: An investigation and reflection based on reading and listening tests. *Foreign Language Teaching and Practice*, (04), 66-74.

[16] Nelson, R. J., Balass, M., & Perfetti, A. C. (2005). Differences between written and spoken input in learning new words. *Written Language and Literacy, 8*(2), 25-44.

[17] Amer, A. A. (1997). The effect of the teacher's reading aloud on the reading comprehension of EFL students. *ELT Journal, 51*(1), 43-47.

[18] Mayer, R. E. (2024). The Past, Present, and Future of the Cognitive Theory of Multimedia Learning. *Educational Psychology Review, 36*(1), 8.

[19] Mayer, R. E. (2014). *Cognitive Theory of Multimedia Learning. In R. E. Mayer (Ed.), The Cambridge Handbook of Multimedia Learning (pp.31-48)*. New York: Cambridge University Press.