

# A Study of Cloze Test in TEM-4 from the Perspective of Cohesion Theory

Wan Liu and Yi Zhang\*

School of Foreign Studies, Northwestern Polytechnical University, Xi'an, Shaanxi, China

Email: yizhang@nwpu.edu.cn (Y.Z.)

\*Corresponding author

Manuscript received August 14, 2023; revised September 30, 2023; accepted December 3, 2023; published March 26, 2024

**Abstract**—Cloze test has now become one of the core questions in various language tests. In Test for English Majors-Band 4 (TEM-4), cloze test accounts for 10% of the total score. However, cloze is also the most time-consuming and lowest-scoring test type for students. Therefore, this research aims to explore in-depth and discuss the cohesion devices that cloze tests of TEM-4 involve and strategies for coping with cloze. Results show that there are a total of 112 occurrences of cohesion devices in cloze. And the lexical cohesion device is the dominant one with 62 instances. Substitution occurs the least frequently (3 instances). In addition, specific examples are used to explain how to solve cloze questions based on cohesion theory. What's more, it is proposed that students should comprehend solving skills based on cohesion theory and consolidate the foundation of language learning to improve the capability of coping with cloze. It is expected that English major students can benefit from this study during the TEM-4 preparation.

**Keywords**—cloze test, cohesion theory, TEM-4

## I. INTRODUCTION

Cloze test is commonly used in the English tests such as CET-4, CET-6, TOEFL and so forth. Although there are many forms of cloze, they share the basic design principles. It is a test in which some of the words such as verbs, nouns, are removed from a cohesive and coherent article purposefully to form blanks. In order to make the article complete, test-takers are required to select an appropriate answer from the four corresponding ones given as alternatives or from a list of choices given in a word bank. Not only does it examine test-takers' capacity in grammar and vocabulary, but it also measures their flexibility in using logical reasoning, comprehensive judgment, and generalization ability to cope with the questions.

The correct answers for each of the blanks in the passage have close relations with texts in meaning and structure. Generally speaking, the selected answers must be consistent with the context and structure. Sometimes, from the point of view of one single sentence, the selected answer is feasible, but once viewed from the whole text, it is not appropriate. Therefore, we should select the answer from the overall structure and content of the passage. Moreover, the cloze test asks the test-takers to select among lots of choices to make the essay become complete again, which makes it, in test-takers' mind, one of the most difficult tests to fulfill. It takes the longest time but students always lose the most marks. In TEM-4, cloze is a text from which 10 words are removed, and it asks examinees to choose the correct answer from 15 choices given in a word bank based on the context. The difficulty of this question type is that the information in the article is no longer coherent after 10 words are deleted. Therefore, the examinee should focus on the contextual clues

and cohesion of the article instead of answering it in isolation.

In order to get a good score on the cloze test, examinees should have the ability of discourse awareness, prediction, logical inference, and the ability to grasp cohesion. This paper mainly studies the promoting effect of cohesion theory on the solution of cloze in TEM-4 and discusses how to improve the examinees' cloze ability from the perspective of cohesion theory. It is of practical significance for English major students during the TEM-4 preparation.

## II. LITERATURE REVIEW

### A. Studies on Cloze Test

The earliest form of cloze is known as Gestalt procedure or filling up the bare. It was invented and used by British scholar Taylor [1]. He defined the cloze as a complete essay composed of several paragraphs in which the originator deletes some words or phrases according to certain or any proportion. Examinees use their language knowledge and social background culture to fill up the words deleted in the article and to make the article restore its original complete appearance.

A study conducted by Kirkwood and Wolfe found that the cloze remains a very high sensitivity when testing participants' reading ability a test-taker has, since the score of the cloze, to some degree, is a reflection of reading comprehension [2]. They also revealed that the examinee's score in cloze and their ability to answer reading comprehension in the form of multiple choices have a great correlation, mostly between 0.6 and 0.7. Brown believed that the cloze test can be used in teaching English as a Second Language (ESL) or Foreign Language (EFL) [3]. It has certain effects and advantages to use cloze tests to measure students' academic performance. Gradually, it has been developed to be an indicator to test students' capacity to learn English. After reviewing related research both at home and abroad, we find that the research on cloze tests shows an increasing trend. In recent years, more and more scholars study cloze from different perspectives [4, 5].

### B. Studies on Cohesion Theory

The publication of Halliday and Hasan's *Cohesion in English* marked the establishment of the cohesion theory [6]. They defined "cohesion" as semantic relations that make a passage or a text a whole and a united one. If one linguistic component of a text plays a key role in the understanding of another component, then there is a cohesive relationship between these two. Besides, they also systematize their concepts and further classify cohesion into two categories. One is grammatical cohesion which includes reference,

substitution, ellipsis, and conjunction, the other is lexical cohesion which includes reiteration and collocation. The following are specific explanations of them. (1) Grammatical cohesion. Reference means that the meaning of one item is explained by another in the same text. It usually refers to the use of pronouns or deixis (I, you, he, she, they, we, this, that, here, there, etc.) to express semantic relations. Substitution refers to the replacement of a phrase, clause, or grammatical structure that has appeared in the text before. Ellipsis refers to the omission of certain elements of the language structure. Conjunction serves the function of linking sentences to express the logical relationship between two or more sentences, pointing out in what sense the sentences are related to each other. (2) Lexical cohesion. It is an approach to achieve cohesion in terms of vocabulary selection. Reiteration is a lexical device that means the repetition of lexical items. It includes repetition, synonyms, hyponyms, meronyms, antonyms, and general nouns. And collocation has to do with the relationship between words that often occur in the same context.

It is well acknowledged that this theory has been widely recognized and developed due to its strong explanatory power of cohesion. For example, Fu elaborated and illustrated its application in college English listening teaching in detail [7]. The conclusion can be drawn that cohesive devices provided students' better understanding of the listening material. As for the application of cohesion theory in cloze test, Keshavarz and Salimi conducted an empirical study on the relationship between the collocational ability of 100 Iranian English learners and their performance on multiple-choice cloze tests [8]. The results showed that there was a significant difference between collocational competence and cloze achievement. In his research, Flor described a simple method to conduct cloze test by combining sentiment-based and cohesion-based systems and demonstrated that cohesion is an important factor for the story cloze task [9].

From previous studies, it can be observed that cohesion theory is commonly used in the field of language teaching. Some researchers concentrated on how to apply cohesion theory to language teaching. Nowadays, most studies are concerning the application of cohesion theory to reading and listening. Although some scholars have started to study cloze from the perspective of discourse cohesion and coherence theory in recent years, most of them are related to CET-4 and other specialized tests. Besides, there are few studies on cloze tests in TEM-4 and even fewer studies on how students apply cohesion theory to cloze tests of TEM-4 and what strategies can be used to strengthen students' cloze ability. Therefore, it is necessary to carry out targeted research in TEM-4, which will be of great help to students' English learning.

### III. METHODOLOGY

#### A. Research Questions

- (1) What kinds of cohesion knowledge do cloze tests of TEM-4 mainly involve?
- (2) How do students apply cohesion theory to cloze tests of TEM-4?
- (3) What strategies can be used to strengthen students' cloze ability based on cohesion theory?

#### B. Sources of Data

The data for this study consists of 10 cloze tests of TEM-4 from 2012–2022, except for 2020 (Due to the impact of the COVID-19 pandemic, TEM-4 was not taken that year). The total number of words amounts to 3199, and the average is almost 320 in each passage. Every cloze test has 10 blanks. The examinees are required to select one word for each blank from a list of choices given in a word bank. There are 15 choices in the word bank containing prompts and distractors. They may not use any of the words in the bank more than once. Here the study only analyzes the cohesion devices which help test-takers choose the right answer and the whole passage will not be analyzed as the main aim of the current study is to help students solve cloze questions and improve their cloze capabilities.

#### C. Research Procedures

This study combines quantitative and qualitative analysis. As for quantitative analysis, cohesion ties that are involved in the cloze test are first identified, connotated, and then counted to see the general distribution. Secondly, in terms of qualitative analysis, specific examples of each kind of cohesion device are given to show how to apply cohesion theory to cloze questions. Thirdly, based on cohesion theory, strategies for students to cope with cloze tests are put forward.

## IV. RESULTS AND DISCUSSION

#### A. Cohesion Devices in Cloze Tests of TEM-4

The result of the study shows that every cohesion device is used in cloze tests of TEM-4. Moreover, the proportions of each device in the 10 cloze tests are different, as shown in the Table 1 below.

Table 1. Frequency of cohesive ties in cloze tests of TEM-4

Cohesive Ties	Frequency	Percentage
Reference	12	10.71%
Substitution	3	2.68%
Ellipsis	5	4.46%
Conjunction	30	26.79%
Lexical	62	55.36%
Total	112	100%

As can be observed in the above table, there are a total of 112 occurrences of cohesion devices in cloze. And it is apparent that the lexical cohesion device is the dominant one in cloze tests with 62 instances, making up 55.36%; the type of conjunction is the second most common one employed in the cloze tests with 30 cases, equal to 26.79%, followed by the reference type (10.71%). Next is the ellipsis (4.46%) and the smallest number belongs to substitution (2.68%). And it is also worth mentioning that some cloze questions involve more than one cohesion device. Take the following sentence for example.

Example 1: *Religious traditions were passed on through (34) generations and spread to others, in writing. Scientific and technological progress was achieved and (35) disseminated through writing. (2016)*

In the instance above, to fill in the correct answer, we need to use the cohesion devices of conjunction and reiteration of lexical cohesion. Through this conjunction “and”, we can predict that the word that needs to fill in is in a parallel or

progressive relationship with “achieved” both in grammatical and semantic structure. As for lexical cohesion, the correct answer “disseminate” is in effect the synonym of “spread”. Candidates can take advantage of the preceding sentence as a clue in the process of answering this question. Students who have discourse awareness will know that this sentence is closely related to the former information. Therefore, examinees can make the choice “disseminated” based on the additive conjunction and synonyms reiteration in lexical cohesion. As lexical cohesion devices account for a large proportion of the whole cohesion ties, it is necessary to count the detailed classification below it in order to better explore the various specific cohesion knowledge involved in the cloze tasks. Table 2 presents the frequency of lexical cohesion and Table 3 shows the frequency of different types of reiteration in cloze tests of TEM-4.

Table 2 Frequency of lexical cohesion in cloze tests of TEM-4

Lexical Cohesion	Frequency	Percentage
reiteration	24	21.43%
collocation	38	33.93%

Table 3 Frequency of reiteration in cloze tests of TEM-4

Reiteration	Frequency	Percentage
repetition	0	0
synonym	14	12.50%
hyponym	3	2.68%
meronym	2	1.79%
antonym	1	0.89%
general-nouns	3	2.68%

From Table 2, it can be seen that collocation type is more frequent than reiteration (38>24), indicating that the task stresses students’ competence in jumping out of the scope of the sentence of the target word and making a judgment at the discourse level based on the context. From Table 3, we find that synonym is the dominant one among reiteration types with 14 instances, and repetition doesn’t occur in the cloze tests. One of the reasons may be the questioner aims to examine the meaning and synonymous substitution of words. If there are just repeated words, it’s too simple for English majors. Overall, the cloze questions are rarely designed for just one single sentence, its majority is related to discourse cohesion. It proves that the cloze test in TEM-4 has higher requirements on the discourse level, which strengthens the inspection of students’ capability to use the cohesion theory to solve cloze questions.

*B. Analysis of the Application of Cohesion Theory in Answering Cloze*

*1) Application of grammatical cohesion*

**Reference:** there are some items that make reference to something else for interpretation instead of being interpreted semantically in their own right. We characterize this particular type of cohesion as reference. Look at the following example.

Example 2: *If a (35) couple invites us to a party, we should be sure to invite them to one of ours. (2018)*

In the example above, the pronoun “them” in the main sentence implies that the words filled in the blank should have plural meaning. However, there is an indefinite article “a” before the blank, which usually represents the singular

form. Combining these two conditions and observing the 15 options in the word bank, only the word “couple” matches well, so the correct answer can be quickly obtained according to the cohesion device of reference.

**Substitution:** it is the replacement of one item by another. The parts replaced usually can be found from the context. In general, there are a few questions designed for substitution in the cloze of TEM-4 with only 3 instances. Take the cloze test in 2018 for instance.

Example 3: *If a woman does us a favor, we should do her (34) one in return. (2018)*

In Example 3, according to the adverbial “in return”, the answer to question 34 should be the same as “favor” in the “if” clause. Since there is no word similar to it in the choices given in the word bank, we can consider using the means of substitution cohesion to express the same meaning, so “one” refers to the word “favour” in the preceding clause.

**Ellipsis:** it is the omission of one item. It can be used to highlight the main information and avoid repetition. In the past ten years, results show that there are only 5 questions that involve ellipsis cohesion devices. It is possible because it’s easy to answer, while the questioners want to improve the difficulty of the test. We take the cloze in 2019 as an example to see how students can apply the ellipsis cohesion device to answer questions.

Example 4: *Certainly there can be no doubt that machines contributed to the Protestant Reformation and the decline of the Catholic Church’s power in Europe or (31) that television has changed the way family members interact. (2019)*

The syntactic structure in the example shows that what comes after the blank has already formed an independent clause, and the conjunction “or” indicates that a functional word should be filled in here. From the cloze context, the last independent clause can be inferred that it is the appositive of “no doubt”, so the word “that” can only be filled in here. After filling it out, we will read the sentence again to check whether it is grammatically correct and semantically reasonable. It can be observed that “there can be no doubt” is omitted before “that television has changed the way family members interact”. With the help of ellipsis cohesion devices, we can further make sure whether the words we fill in are correct.

**Conjunction:** it refers to a cohesive relationship among clause sentences and a means which uses connective components to reflect the logical relationship between discourse units. Due to the use of conjunction words, readers can understand the semantic relations between sentences, and even the meanings of subsequent sentences can be logically predicted by the former. From table 1, we can observe that it accounts for a relatively high frequency in the examination. Take the following question as an example to see the conjunction devices of cohesion.

Example 5: *Every English speaker understands what the word “food” refers to, even (35) though the particular kind of food that comes to mind may vary to each individual. (2021)*

Both the former and latter sentences are complete in the question of examining conjunction relation. The subject of

the former sentence refers to that “every English speaker understands food”, while the latter refers to that “individuals understand food in a different way”, and between the two there is the word “even” indicating that the two sentences are in adversative relation. Examinees can make the right choice based on the logic relation and the indication of “even”.

## 2) Application of lexical cohesion

**Repetition:** it’s worth noting that repetition is the most basic and necessary means of cohesion. However, it does not appear in specific questions in our data source. Maybe because the expressions of repetition are too simple for English major students, the questioner has never examined the repetition of vocabulary in the past decade, so we will not give examples here.

**Synonyms:** it refers to the appearance of two words with similar meanings in the same text. This type is more difficult than repetition, which requires candidates to have a certain lexical and syntactic basis to identify the synonyms. From this study, it is easy to find that synonyms type is quite important compared with other types of reiterations. The 38th question in 2014 examined synonyms in lexical reiteration.

Example 6: *For example, Americans were out of work for an average joblessness (38) rose to about 33 weeks in 2010 and 39 weeks in 2011 as the share seeking benefits soared to 95%. (2014)*

The correct answer “rose” in 38 is the synonym of “soar” in the following clause. “As” is a conjunction in grammatical cohesion, it is a signal of causal relation. And there are lots of numbers mentioned here. Based on these two clues, we can think of the word “rise” to describe the number’s increase and it is logical right after reading the whole sentence.

**Hyponyms:** superordinate has the function of generalization. And the relationship between superordinate and hyponyms is inclusion. Sometimes in order to avoid using the same word again and again, we use this kind of relation to make diversified vocabulary. And of course, attention should be paid to this technique in order to answer cloze effectively.

Example 7: *Whole grains are very popular in many dietary guidelines because they contain high levels of (40) nutrients like zinc, copper, manganese and iron. (2013)*

The correct answer “nutrients” is the superordinate of the words “zinc, copper, manganese and iron” and these words “zinc, copper, manganese and iron” are all hyponyms of “nutrients”. The word “like” is the clue which means that the following words are explanations for the former word.

**Meronyms:** it denotes a constituent part of or a member of something. That is, it’s in a part-whole relationship. Similar to hyponymy, this type can be used to make a summary or generalize something. Look at Example 8.

Example 8: *There would be no ball-points, no typewriters, no computers, no Internet, no magazines, no movie credits, no shopping lists, no newspapers, no tax returns. But such (31) lists of subjects almost miss the point. (2016)*

“Ball-points, typewriters, computers, Internet and so on” are meronyms of the lists as they are parts of the lists. Therefore, the blank here should be “lists”.

**Antonyms:** sometimes antonyms are adopted to develop

the semantic cohesion in the text. The examinee should pay attention to analyze these phenomena from the perspective of antisense judgment to make the correct choice.

Example 9: *Language is clearly (34) cultural and not natural: it is human made and accepted with some variations throughout a particular culture. (2021)*

The correct answer “cultural” is an obvious comparison with the word “natural” following it. The word “and” and “not” is the clue words of this question as “and” indicates additive relation and “not” suggests there should be an antonym.

**General nouns:** this type includes “things, person, do, moves and so on”, which are used cohesively when they have the same referent. It’s similar to hyponymy and meronymy. Let’s look at the following example.

Example 10: *To (38) leaf through a box of old paper correspondence is to know what been (39) lost in this shift: the pretty stamps, the varying look and feel of handwritten and typed correspondence, the tangible (40) object that was once in the sender’s hands. (2017)*

Here “object” refers to “the pretty stamps, the varying look and feel of handwritten and typed correspondence”. It is the general noun that can be used to refer to any one of them.

**Collocation:** lexical collocational cohesion is achieved through the association of lexical items that regularly co-occur. In order to ensure the harmony and integrity of the discourse, it is necessary to comprehend the linguistic knowledge of certain specific topics and understand the vocabulary chain composed of related words. Take the following sentence as an example.

Example 11: *Usually it is left to (31) employers to deduce the potential from a list of extracurricular adventures on a graduate’s resume, but now the University of Bristol has launched an award to formalise the achievements of students who devote time to activities outside their courses. (2012)*

The correct answer “employers” and the latter word “resume” are co-occurrence words. These vocabularies usually appear in the field of job hunting. And considering the whole passage is about graduates’ outdoor activities to increase the content of the resume and improve competitiveness, examinees can easily think of the word “employer”.

## C. Strategies for Coping with Cloze

As an important means to test students’ comprehensive use of language and reading comprehension ability, cloze plays an important role in English tests. Due to its large proportion of English tests, this requires English major students must strive to improve the cloze solving ability. Here two strategies are proposed. Firstly, students need to comprehend solving skills based on cohesion theory. An isolated sentence or paragraph in a cloze may be difficult to be understood for students, while grasping the whole passage’s internal cohesion relationships, to a large extent, can eliminate interference and provide some clues to make the final choice. Results show that cohesion devices are widely used in the cloze. Therefore, learning cohesion devices systematically is needed when answering cloze. Secondly, students should

consolidate the foundation of language learning. The cloze test is to examine the examinee's ability in comprehensive usage of language, however, flexible use of language on cloze does need to consolidate the language base, and this is the top priority in any type of English learning. Furthermore, the cloze examines candidates' inference ability based on discourse cohesion, if examinees do not have the ability to do lexical differentiation and identify grammatical structures of sentences first, they cannot apply appropriate cohesion devices to answer any questions. Therefore, paying close attention to the language foundation cannot be relaxed.

#### V. CONCLUSION

This paper is about the application analysis of cohesion theory to the cloze test of TEM-4. Based on the cohesion theory, we not only identified the cohesion devices and analyzed the application of cohesion theory to the cloze, but also proposed strategies for improving students' cloze ability. Results show that there are a total of 112 occurrences of cohesion devices in cloze. And the lexical cohesion device is the dominant one with 62 instances. Substitution occurs the least frequently (3 instances). In addition, specific examples are used to explain how to solve related questions based on cohesion theory. It is suggested that students should comprehend solving skills based on cohesion theory and consolidate the foundation of language learning to improve the capability of coping with cloze.

Cloze is a common test type in almost all language tests. It is said that the English major students' score is lower than other question types. Therefore, the current study on cloze is valuable and meaningful. However, there are also some limitations. Firstly, only ten tests are involved in my study, which cannot sufficiently conclude the general distribution of TEM-4. Secondly, the application analysis part of cohesion theory is only analyzed from the perspective of students' problem-solving, and it would be better to analyze it from the perspective of the teacher such as how to teach students to master a certain type of question. Overall, it is hopeful that English major students can benefit from this study during the TEM-4 preparation.

#### CONFLICT OF INTEREST

The authors declare no conflict of interest.

#### AUTHOR CONTRIBUTIONS

Wan Liu conducted the research, analyzed the data, and wrote the paper; Yi Zhang reviewed the paper; all authors had approved the final version.

#### FUNDING

This research was sponsored by the Practice and Innovation Funds for Graduate Students of Northwestern Polytechnical University: PF2023024.

#### REFERENCES

- [1] W. L. Taylor, "Cloze procedure: A new tool for measuring readability," *Journalism Quarterly*, vol. 30, no. 4, pp. 415–433, 1953.
- [2] K. J. Kirkwood and R. G. Wolfe, *Matching Students and Reading Materials: A Cloze-Procedure Method for Assessing the Reading Ability of Students and the Readability of Textual Materials*, Ontario Government Bookstore, 800 Bay St., Toronto, Ontario M7A 1L2, Canada, 1980.
- [3] H. D. Brown, *Language Assessment: Principles and Practice*, London: Longman, 2004.
- [4] R. S. Williams, O. Ari, and C. N. Santamaria, "Measuring college students' reading comprehension ability using cloze tests," *Journal of Research in Reading*, vol. 34, no. 2, pp. 215–231, 2011.
- [5] S. Kleijn, H. P. Maat, and T. Sanders, "Cloze testing for comprehension assessment: The HyTeC-cloze," *Language Testing*, vol. 36, no. 4, pp. 553–572, 2019.
- [6] M. A. K. Halliday and R. Hasan, *Cohesion in English*, Routledge, 1976.
- [7] C. Fu, "Application of cohesion theory in college listening EFL teaching," *Studies in Literature and Language*, vol. 14, no. 6, pp. 38–43, 2017.
- [8] M. H. Keshavarz and H. Salimi, "Collocational competence and cloze test performance: A study of Iranian EFL learners," *International Journal of Applied Linguistics*, vol. 17, pp. 81–92, 2007.
- [9] M. Flor, "Sentiment analysis and lexical cohesion for the story cloze task," in *Proc. the 2nd Workshop on Linking Models of Lexical, Sentential and Discourse-level Semantics*, 2017, pp. 62–67.

Copyright © 2024 by the authors. This is an open access article distributed under the Creative Commons Attribution License which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited ([CC BY 4.0](https://creativecommons.org/licenses/by/4.0/)).