

# Cohesion and Coherence in Proofreading and Error

## Correction in TEM-8

Yucheng Mao<sup>1</sup>

<sup>1</sup> Foreign Languages College, Jiangxi Normal University, Nanchang, China

Correspondence: Yucheng Mao, Foreign Languages College, Jiangxi Normal University, Nanchang, China.  
E-mail: 1007815478@qq.com

Received: February 1, 2024

Accepted: February 20, 2024

Online Published: February 21, 2024

doi: 10.5539/elt.v17n3p54

URL: <https://doi.org/10.5539/elt.v17n3p54>

### Abstract

TEM-8 (Test for English Majors-Band 8) is an advanced English proficiency test specifically for English majors in China, and those who pass it are generally considered to have a good command of English. Of this kind of test, proofreading and error correction has long been a headache for teachers and students. The sample analysis in this study reveals that, however, many of the errors in this tricky question type are related to the knowledge of cohesion and coherence, yet few studies have linked the two. Therefore, the present study aims to investigate the distribution of errors in proofreading and error correction in TEM-8 using the cohesion and coherence theory. Errors in authentic TEM-8 test papers of the recent ten years are categorized according to a classification standard which is based on Halliday and Hasan's theory. The results show that 66% of the errors are related to cohesion and coherence with 56% of the former and 10% of the latter, which is followed by a tentative discussion about each error type from a psycholinguistic perspective. Despite limitations, this study exposes a close association between cohesion and coherence on the one hand, and proofreading and error correction on the other to some extent.

**Keywords:** cohesion, coherence, proofreading and error correction

### 1. Introduction

With high reliability and validity, the TEM-8 test papers are able to examine the English proficiency of English majors in institutions of higher learning in China (Liu, 2021). As an integral part of TEM-8 accounting for 10% of the total score, proofreading and error correction is usually a short essay of about 200 words which contains 10 grammatical, syntactical, and discourse errors. Involving abundant knowledge points, the proofreading and error correction section is such a thorny question type that requires careful treatment (Song & Bai, 2017), with its average score often being only 3 to 4 in recent years, and even an astonishing 1.79 in 2019. English learners generally find it a difficult task and many of them perform poorly in this section (Wu & Hu, 2021). Often times, when asked about how they deal with this section, most of the testees say they are not sure at all of their corrections. In most cases, what they do is simply guessing where the errors lie. Consequently, the average score of this section is always relatively low.

Students' performance in proofreading and error correction is not an optimistic picture at all, but instead a weakness in need of strengthening. In fact, a number of errors amid this part can be attributed to the knowledge about cohesion and coherence (Liu, 2007; Guan, 2010). It would be a completely different situation if students can be equipped with relevant skills when confronting these test questions. Therefore, by analyzing the distribution of errors related to cohesion and coherence in proofreading and error correction in TEM-8, the current study intends to (1) make more and more teachers realize the close relationship between the two, so as to pay adequate attention to them in their usual teaching practice; (2) inspire students to acquire targeted techniques during their process of learning, so as to achieve a more satisfactory performance in this section in the future.

## 2. Literature Review

### 2.1 Definitions of Key Terms

#### 2.1.1 Cohesion

A number of important works on the subject of cohesion were published as early as the early seventies. Beaugrande and Dressler (1981) argued that cohesion involves the way in which the components of a surface text, i.e., the actual words we hear or see, are interconnected in a sequence. According to Richards et al. (1992), cohesion refers to the grammatical and lexical relationships between different elements in a text. Thompson (1996) defined cohesion in this way: cohesion refers to the linguistic means by which a speaker can show that a text is empirically and interpersonally coherent, and thus cohesion is a textual phenomenon, i.e., we can point to features in a text that function as cohesion. Widdowson (1978) concluded that cohesion is best defined in terms of overt linguistic signaling relationships between propositions.

According to Halliday and Hasan's *Cohesion in English* (1976), the notion of cohesion is a semantic one. It refers to the relations of meaning that exist in a text which determine its nature. Cohesion occurs when the interpretation of certain elements of a discourse depends on the interpretation of another element. One element is a prerequisite for another, that is, it cannot be effectively decoded without the help of another element. Like other semantic relations, cohesion is expressed through the hierarchical organization of language. It is expressed partly through grammar and partly through vocabulary. Thus, we can refer to grammatical cohesion and lexical cohesion.

#### 2.1.2 Coherence

Following the release of *Cohesion in English*, coherence had become a popular topic of discussion, with influential theories including the following. Danes (1974) discussed the concept of coherence in terms of connexity, which is related to the concept of thematic progression. According to Beaugrande and Dressler (1981), coherence involves the mutual accessibility and relevance of the various components of the textual world, i.e., the configurations of concepts and relations that underlie the surface text.

According to van Dijk (1977), coherence is a semantic concept, which is not only linear and sequential but also hierarchical, and is not only a microstructure involving the connection of meanings within or between sentences but also a macrostructure referring to the overall semantic coherence of the text as a whole. Widdowson (1978) argued that cohesion is the overt relationship between propositions expressed through sentences, whereas coherence is the relationship between the illocutionary acts that propositions, which are not always overtly linked together, are used to perform. While according to Leech (1983), van Dijk and Widdowson were complementary, i.e., the former saw it as an internal semantic relation, while the latter saw it as a function of the text in social communication. Chinese scholar Zhu (1997) recognized coherence as a semantic concept, but it is semantic in the systemic-functional sense, a combination of van Dijk's text-internal semantic notion and Widdowson's notion of illocutionary development. Zhang (1999) accepted Zhu's position and considered coherence as a semantic concept in its broad sense.

In Halliday and Hasan's *Cohesion in English* (1976), Halliday presented his view of coherence from two perspectives: firstly, it hangs together internally; and secondly, it is related to situational and cultural contexts externally. Cohesion contributes to the coherence of a text because it makes the text hang together. Coherent texts are usually characterized by two features: first, they are coherent with the context and are therefore consistent in register; second, they are coherent with the text itself and are therefore cohesive. Hasan (1984) defined coherence as the uniformity of "hanging together".

### 2.2 Classification of Cohesion

Halliday and Hasan (1976) classified cohesion into grammatical and lexical cohesion.

Grammatical cohesion includes reference, substitution, ellipsis and conjunction. Reference deals with the links between words or phrases and other words or phrases in discourse. What reference does is to help the reader to keep track of the various participants in the text as he or she reads (Eggins, 1994). In substitution, one lexical item is replaced with another as an alternative to repeating the first. Ellipsis is a form of cohesion that is really a special case of substitution in which one phrase is substituted with nothing. In conjunctive cohesion, a relationship between phrases or sentences is expressed by using conjunctions such as "and", "or", "but", "yet", and "so".

Lexical cohesion includes lexical reiteration and collocation. Lexical reiteration involves the repetition of a lexical item, the use of a general word to refer back to a lexical item, and a number of cases in between, i.e., the

use of a synonym, near-synonym, or superordinate. Lexical collocation is accomplished by linking lexical items that frequently occur together. All lexical cohesion that is not covered by reiteration is treated under the general heading of collocation.

Collocation is the pairing of words arranged in a text, the relationship of meaning between a single lexical item and the lexical items with which it is customary to co-occur in a language. To some extent, the lexical items involved in a collocation are always predictable from each other (Crystal, 1992). McCarthy (1991) argued that collocational relationships are the basis of lexical research. Firth is often quoted as saying, “you shall know a word by the company it keeps” (Firth, 1957). Mastering appropriate collocations is part of native speakers’ competence. Therefore, collocation should be a central part of vocabulary learning.

Lexical collocations in a text can be understood through clues provided by the author and the reader’s knowledge of general discourse features and the world to which the discourse refers. Collocation includes two different language phenomena: the lexical collocational items belong to the same semantic field, such as “red” and “green” both belong to the semantic field of color; the lexical items do not belong to the same semantic field but they often co-occur, for example, “garden” and “dig”, “ill” and “doctor”. Halliday and Hasan (1976) set examples of this kind though they did not define it as collocation.

Other scholars also distinguished between cohesion and coherence (Grimes, 1971; van Dijk & Kintsch, 1983; Enkvist, 1978; Beaugrande & Dressler, 1981). They limited cohesion to explicit mechanisms in the text, both the types of cohesive ties described by Halliday and Hasan (1976) and other elements that bind the text together, such as tense, parallel structure, and subject-verb agreement.

### *2.3 Realization of Coherence*

As mentioned earlier, coherence is not a well-defined concept (van Dijk, 1977). The ambiguity of its definition may be related to the fact that coherence is an “interpretive process” that occurs when a reader reads a text (McCarthy, 1991). Coherence is generally seen as having two aspects: reader-based and text-based coherence (Johns, 1986). The former relates to the meaningful aspects of reader-writer interactions, while the latter refers to features related to the internal structure of the text itself. To summarize, coherence is in the mind of the reader and the writer, and both sides need to make every effort to make a text coherent. For the writer, the following should be done to achieve the goal of coherence.

#### *2.3.1 Unity of Ideas*

Unity of ideas means that assertions in an essay should be related to all other elements, i.e., each idea should be related to the main idea of the paragraph in which it is made, as well as to other ideas in the same paragraph. In turn, the main idea of the paragraph is related to the central idea of the essay. Also, the conclusion summarizes the main ideas of the whole essay. There is a main thread running through the entire essay. In sum, unity of ideas is the singleness of purpose (Bander, 1978).

#### *2.3.2 Logical Organization of Points*

To achieve coherence, the writer needs to organize the discourse in a logical way, which is called logical organization of points. A coherent discourse usually has one topic and many subtopics, and points of a text need to be well organized to form a logical sequence from the beginning till the end of the essay. The general coherence of a discourse involves the overall coherence and the linear coherence, which are closely related to each other. The overall coherence is the main organization pattern, and the linear coherence connects all the parts on each level, including the connection between the microstructures. Whether the points of a text are organized in a logical way can be better judged from the level of a discourse.

### *2.4 Previous Studies on Cohesion and Coherence in Proofreading and Error Correction*

The concept of cohesion and coherence has been widely discussed in various subject areas since its debut, and it has also received considerable attention in the field of English teaching (Luo, 2019). However, compared to fields like reading and writing, its application in proofreading and error correction is far from adequate.

Ren (2017) analyzed how the discourse cohesion theory should be effectively used in English short text error correction in high school English teaching at this stage. Zeng (2016) further systematically illustrated the effective implementation of the discourse cohesion theory in the English short text error correction at the current stage of the college entrance examination on the basis of its authentic questions in the past ten years. Xie (2015) analyzed the effective application of the discourse cohesion theory to the specific questions of English short text error correction in the authentic questions of the college entrance examination. According to Liu (2015), at this stage, the teaching of English short text error correction in high school should focus on the understanding of the

discourse as a whole and its inner connection, thus requiring teachers to have the knowledge of discourse cohesion and coherence and the structural awareness of the discourse as a whole, as well as the ability to analyze the inner logic of thought and identify different discourse structures. Li (2015) believed that with the help of discourse cohesion, high school students can effectively develop their ability to analyze and solve problems in English. Yu (2013) argued that the main factors affecting the low correct rate of English short text error correction of most high school students at this stage are: difficulty in answering questions according to its standardized requirements leading to loss of marks, and the insufficient grasp of grammatical knowledge, which lead to low pragmaticity of discourse analysis. Ma & Zhang (2015) identified the main problems of current teachers in teaching English short text error correction as: neglecting the discourse cohesion work in the topic teaching, not paying attention to the feedback of the after-class exercises, etc.

Liu (2005) applied the knowledge about cohesion and coherence to analyze the errors in TEM-8 error correction questions. She adopted the theoretical frameworks of Halliday and Hasan (1976) on cohesion and coherence with some modifications to suit the requirements of her study. Before analyzing the entire corpus of 50 articles, she first analyzed two sample articles in detail then categorized the errors in the articles according to the established standard of error types, thus indicating the steps and methods of the whole study. Afterwards, all the corpus was analyzed by this method and then finally the findings of the whole study were summarized. The study found that among various types of errors in a total of 50 articles, 55.6% of them were related to the knowledge of cohesion and coherence.

Guan's (2008) study focused on the role of discourse cohesion and coherence in proofreading and error correction questions in the English graduation proficiency test. The research subjects were two natural classes of third-year English majors in the Open College of Guangdong University of Foreign Studies, divided into a control class and an experimental class. The experiment was conducted in three phases: pre-test, training, and post-test, and all data collected from the experiment were analyzed using SPSS. The results of the pre-test revealed no significant difference in proofreading and error correction scores between the two classes. After, the experimental class received specialized training on discourse cohesion and coherence, which lasted one month for a total of 10 class hours. The analysis of the post-test results revealed that the two classes produced significant differences in performance. The results of the experiment showed that learning about discourse cohesion and coherence helped to improve test takers' test scores in proofreading and error correction questions.

In summary, it can be found that although there is a large body of literature on cohesion and coherence in the field of English teaching, (1) the literature associated with proofreading and error correction is scarce or of low relevance; (2) few scholars have focused on the particular exam type of TEM-8; (3) coherence is far less compelling than and rarely studied along with cohesion. Therefore, inspired by and as an extension and supplement to Liu's (2005) study, the present study intends to investigate the role of cohesion and coherence in proofreading and error correction in TEM-8 and addresses the following research question:

Are there any errors related to cohesion and coherence in proofreading and error correction in TEM-8? If so, how are they distributed?

### **3. Method**

#### *3.1 Research Procedures*

The present study consists of the following five steps:

Firstly, collect authentic TEM-8 test papers of the recent ten years (from 2022 to 2012 except 2020) to ensure timeliness, and extract texts from the proofreading and error correction part.

Secondly, establish the categorization standard of errors related to cohesion and coherence. The standard consists of Halliday and Hasan's (1976) theory and classification on cohesion and coherence as the primary basis, plus some modifications as the supplementation.

Thirdly, analyze a sample text by categorizing and demonstrating the errors that can be solved through using the knowledge of cohesion and coherence.

Fourthly, do the same for the remaining texts and calculate the respective percentage of each type of errors to reflect the role that cohesion and coherence play.

Eventually, according to the research findings, put forward a set of methods to improve students' performance in this part.

### 3.2 Establishment of Categorization Standard

In the classification of cohesion mentioned earlier, the current study focuses only on reference, conjunction, and lexical collocation because errors related to the rest rarely appear in proofreading and error correction, and expands Halliday and Hasan’s (1976) scope of cohesion to include set phrase, verb-object collocation, tense, parallel structure, and subject-verb agreement under the category of lexical collocation.

On the other hand, by analyzing how coherence is achieved, it can be seen that errors associated with coherence in proofreading and error correction are manifested mainly as incoherence in meaning, due to misuses of words similar in whether spelling or meaning without considering the context.

As a consequence, the categorization standard of errors related to cohesion and coherence for the present study is finalized as shown below.

Table 1. Categorization standard

Errors related to cohesion						Errors related to coherence
Reference	Conjunction	Lexical collocation				Incoherence in meaning
		Set phrase	Verb-object collocation	Tense	Parallel structure	

## 4. Results

### 4.1 Sample Analysis

Table 2. Sample analysis

Proofreading and error correction part of TEM-8, 2022	
<p>Many people believe that bilinguals have two different personalities, one for each of the languages they speak, and that switching between languages makes bilinguals act differently. <u>Whether</u> this may seem unbelievable to some, research actually supports this idea.</p>	(1) Whether→Although
<p>According to various studies, bilinguals, who are also bicultural <u>are</u> actively involved in both of their cultures, interpret situations differently depending on which language they speak. Although everyone, monolinguals and bilinguals alike, <u>are</u> able to change the way they feel and interpret events, biculturals do this without realizing when <u>switched</u> between languages.</p>	(2) ^ are→and (3) are→is (4) switched→switching
<p>In this study, bicultural Hispanic women in the United States viewed a series of TV commercials in both of their languages and were asked to judge the main characters. <u>Interesting</u> enough, when viewing the Spanish-speaking version the participants of the study saw the main character as an independent and extroverted female. On the other hand, when viewing the English-speaking version, they <u>looked</u> the same character as hopeless and <u>alone</u>. This study also noted, however, that this change in reasoning only happened with bicultural bilinguals.</p>	(5) Interesting→Interestingly (6) looked→considered (7) alone→lonely
<p>Depending on the languages spoken, biculturals will act differently. The changes are not only linguistic. As an English-Spanish bicultural, I do find <u>out</u> I act differently depending on which culture I am immersed in at the time. I am often aware of the fact <u>when</u> I speak to other Spanish speakers my voice is slightly louder and I gesticulate <u>more</u> when I talk to English speakers. Could we then say that bilinguals have two different personalities?</p>	(8) <del>out</del> (9) ^ when→that (10) more ^ →than

This sample text about bilinguals and biculturals contains 10 errors, which will be analyzed one by one in the following. (1) Misuse of conjunction. The statements of the first and the second half are in a clear adversative relation, so the conjunction “although” should be used. (2) Absence of conjunction. In the non-restrictive attributive clause led by “who”, there is a lack of conjunction between the two link verb-predicative structures which are in a parallel relation, so the coordinating conjunction “and” should be added before “are”. (3) Subject-verb disagreement. The subject of the sentence is the indefinite pronoun “everyone” and the predicate verb should be in the third-person singular form, so “are” should be changed to “is”. (4) Misuse of non-predicate verb. The subject of the main clause “biculturals” and the predicate of the subordinate clause “switch” are in an

active relation, so “switch” should be in the form of a present participle and changed to “switching”. (5) Error of part of speech. The adjective “Interesting” should be replaced by the adverb “Interestingly” to indicate that the author found the matter interesting. (6) Misuse of verb. The word “looked” can not be combined with “as” to convey the meaning of consideration, so the word “looked” should be replaced by “considered”. (7) Misuse of near-synonym. “Alone” means “no other person is with you”, while “lonely” has an additional meaning “and that you are sad and unhappy as a result”, so the word “alone” should be replaced by “lonely”. (8) Redundancy of adverb. The verb “find” means “discover”, while the phrase “find out” means “figure out (by research, reasoning, etc.)”, so the word “out” should be deleted. (9) Absence of leading word. When explaining specific content, the appositive clause should be led by “that” which can not be omitted, so the leading word “that” should be supplemented before “when”. (10) Absence of conjunction. The sentence lacks a conjunction of the comparative adverbial clause, so the conjunction “that” should be added after “more”. Among the errors above, 6 are related to the knowledge of cohesion and coherence, which are distributed in the table shown below.

Table 3. Error distribution

Reference	Errors related to cohesion						Errors related to coherence
	Conjunction	Lexical collocation				Subject-verb agreement	Incoherence in meaning
		Set phrase	Verb-object collocation	Tense	Parallel structure		
	(1)	(6) (8)			(2)	(3)	(7)

4.2 Data summary and Interpretation

Table 4. Data summary

Year	Errors related to cohesion							Errors related to coherence
	Reference	Conjunction	Set phrase	Verb-object collocation	Tense	Parallel structure	Subject-verb agreement	Incoherence in meaning
2022		(1)	(6) (8)			(2)	(3)	(7)
2021			(3)	(2)				(8)
2019	(1) (10)	(9)	(4) (7) (8)		(5)	(2)		(6)
2018		(4) (6)	(1) (5) (7)		(8)			
2017	(8)	(2) (4)	(1) (10)	(7)	(9)			(5)
2016	(4)	(10)	(1) (8)	(9)				(5)
2015		(2)	(4) (5) (8) (9)	(10)				(7)
2014	(5)	(8)	(3)	(2)				(6) (10)
2013	(2) (7) (8)		(3) (5)		(6) (10)	(4)		(9)
2012	(6)	(10)	(1) (5)		(4)			(2)
Total	9	10	22	5	6	3	1	10
%	9%	10%			37%			10%

The results of the analysis of all texts according to the practice used in the sample analysis are summarized in the table above, which are broadly similar to Liu’s (2005) findings while differ slightly. Out of a total of 100 errors in these 10 texts, 66% are related to cohesion and coherence, with many more errors related to cohesion than to

coherence, as many as 56% of the former compared to only 10% of the latter. Furthermore, of errors related to cohesion, those under the category of lexical collocation accounts for the highest percentage at 37%, much higher than 9% of reference errors and 10% of conjunction errors. At this point, an answer to the research question can be given: there are indeed a certain number of errors related to cohesion and coherence in proofreading and error correction in TEM-8, which is worth noting to some extent, and most of the errors are centered on lexical collocations.

## 5. Discussion

### 5.1 About Reference

Anaphoric reference, the reference in which some current expression (anaphor) is related to one encountered earlier (antecedent), has commanded a great interest among psychologists. One reason is its ability to explore the role of working memory, “the temporary storage of information that is being processed in any range of cognitive tasks” (Baddeley, 1986), in discourse comprehension. To understand a simple pair of sentences, the antecedent must be held in working memory long enough to be linked with the anaphor. In general, the antecedent and the anaphor will not be too far apart. Chances are that, though, the distance between antecedent and anaphor is much longer, which generally imposes a burden on working memory and ultimately disrupts comprehension (McKoon, Gerrig, & Greene, 1996).

Question (10) of 2019 is such an example. In the sentence “As it only introduces the basic notions, its main purpose is to illustrate the framework’s descriptive and explanatory potential through extensive discussions of their application to diverse facets of language structure”, “their” should be replaced by “its”, as it refers to the antecedent “the framework”, which is a little bit far away from the anaphor. Candidates may be distracted by the recent word “discussions” and fail to spot the error. Therefore, when encountering such kind of errors, testees are advised to review the preceding text a bit further carefully to locate the corresponding antecedent accurately.

Another more general problem about reference relates to the use of the definite article “the”. According to Halliday and Hasan (1976), like the demonstratives, “the” is essentially a specifying agent, serving to identify a particular individual or subclass within the class designated by the noun. However, “the” can only do so through dependence on something else, since it contains no specifying element of its own.

In other words, “the” can be regarded as the sign of reference. For example, in question (1) of 2019, the lacking “the” functions as an antecedent of the postpositive attributive in the expression “the vast literature that now exists in CG and in cognitive linguistics more generally”. According to Zhao & Zhang (2010), English learners in China have considerable problems with the use of articles, often omitting necessary definite articles and sometimes using redundant ones. Since the use of the definite article “the” depends to a large extent on the discourse context (Chen, 2016), it can be argued that looking beyond the sentence level to determine whether the modified word is used in a specific or general sense corresponds to whether “the” is used or not.

### 5.2 About Conjunction

In psycholinguistics, investigators have developed propositional models of sentence representation (Anderson, 1976; Kintsch, 1974, Norman, Rumelhart, & the LNR Research Group, 1975). All of these proposals assume that a sentence can be represented as a proposition consisting of two or more concepts and some form of relation between them. When a sentence is first encountered, its meaning is extracted and a proposition that represents this meaning is constructed. For instance, in question (1) of 2022, the compound sentence “Whether this may seem unbelievable to some, research actually supports this idea” can be represented as two propositions “Be unbelievable (this)” and “Support (research, idea)”.

It is conjunction, then, that makes these two independent propositions coherent in a single sentence. Based on Halliday and Hasan’s *Cohesion in English* (1976), conjunction is divided into four types: additive, adversative, causal and temporal. Additives, the so-called “and” links such as “and”, “furthermore”, “or”, and “neither”, indicate a relationship of coordination, drawing on the notion of addition in both a positive and a contrastive sense. Adversatives, the so-called “but” links whose basic meaning is “contrary to expectation” (Halliday & Hasan, 1976), suggest what follows is a relation turning by using “yet”, “but”, etc. Causals, the so-called “so” links, mark a cause-and-effect or reason-and-consequence relationship, which is expressed by “because of”, “for”, etc. Temporals, the so-called “then” links, mark the sequence or order, in which there are expressions like “then”, “and then”, “first”, “at last”, and so on. In the example, the two propositions are clearly in an adversative relation, therefore the conjunction “whether” should be changed into “although”. It follows that, by simplifying a compound sentence into propositions, students can more easily judge their semantic relationships if they are familiar with the above categorization.

### 5.3 About Lexical Collocation

Regarding how vocabulary is stored in memory, psycholinguists have proposed the spreading activation model. Collins and Loftus (1975) assumed that words are represented in the internal lexicon in a network organized as a web of interconnecting nodes, with the distance between the nodes determined by both structural characteristics such as taxonomic relations and considerations such as typicality and degree of association between related concepts. The process of spreading activation was likened to the effect of dropping a rock into a pool of still water (Wessells, 1982). For example, in question (2) of 2021, when the word “status” is accessed, the modifying word can be “approach” rather than “become”, as the former is semantically much closer and more typical than the latter.

A more recent spreading activation model that incorporates lexical as well as conceptual aspects was presented by Bock and Levelt (1994), who assumed that knowledge of words exists at three different levels as follows. The conceptual level consists of nodes that represent concepts, which are connected to other nodes by various relations. The lemma level refers to syntactic aspects of word knowledge (Bock & Levelt, 1994; Levelt, 1989). The lexeme level captures a word’s phonological properties, or how a word sounds. Among the three levels, the lemma level contributes to explaining the way set phrases are stored in memory, for the usage of a word in syntax would be memorized simultaneously with its meaning by an ideal learner. For example, the word “find” would be acquired along with its various collocations such as “find out”, as in the case of question (8) of 2022. In this regard, when acquiring a new word, learners should not limit themselves to memorizing its meaning, but also grasp its syntactic usage.

### 5.4 About Incoherence in Meaning

In Liu’s (2005) study, errors related to coherence were embodied in contradiction in meaning resulting from misuses of semantically opposite words, which accounted for 5.66% of the total number of errors. Interestingly, the present study fails to find this type of error. Instead, the subtle difference between words that are similar in whether spelling or meaning have been examined in recent years, which place a higher demand on examinees as the easily-confused nuance between similar words is much less obvious than the disparity between antonyms. It is often the case that students are familiar with both words but just can not tell them apart when put together. Taking question (7) of 2022 as an instance, “lonely” has an extra negative emotion than “alone” besides the meaning “someone is by himself”, yet which one is more coherent can only be judged by examining the context.

This process can be viewed as involving lexical access. In Morton’s (1969) logogen model, each word (or morpheme) in the lexicon is represented as a logogen, which specifies the word’s various attributes (semantic, orthographic, phonological, and so on). The logogen is activated in either of two ways: by sensory input or by contextual information. In the sensory route, as orthographic or phonological features of the input stimulus are detected, they are matched to the logogen, which functions as a scoreboard or counter. When the counter rises above a predesignated threshold, the item is recognized. With regard to contextual information, the semantic and syntactic structure of a sentence may influence the activation of the logogen for a given word. A word can be anticipated due to the expectations created by earlier words in the sentence. The activation of the earlier words influences the logogen for the final word, temporarily lowering its threshold. Thus, it is easier to recognize a word in context than if presented in isolation (Tulving, Mandler, & Bauml, 1964), which inspires learners to seize lexical meaning contextually rather than isolatedly.

## 6. Conclusion

In this study, Halliday and Hasan’s (1976) theory of cohesion and coherence is used as the basis for setting up the error classification standard to categorize the errors in proofreading and error correction in TEM-8, and the results show that 66% of the errors are related to the knowledge of cohesion and coherence, which are noteworthy for frontline teachers and educators and call for deeper future studies to verify the feasibility and effectiveness. However, there also exist the following limitations: (1) It is difficult to establish an absolutely reasonable classification standard for cohesion and coherence; (2) It is impossible to be completely accurate in categorizing errors. Nevertheless, this study may still shed some light on the significance of the knowledge of cohesion and coherence in proofreading and error correction questions to some extent.

## Acknowledgments

The study has been supported by the Innovation Fund for Postgraduates of Jiangxi Normal University—The Effects of Neural Machine Translation on Chinese College Students’ English Writing Performance (YJS2023087).

## References

- Anderson, J. R. (1976). *Language, Memory, and Thought*. Hillsdale, NJ: Erlbaum.
- Baddeley, A. (1986). *Working Memory*. New York: Oxford University Press.
- Bander, R. G. (1978). *American English Rhetoric: A Writing Program in English as a Second Language*. New York: Holt, Rinehart and Winston.
- Beaugrande, R. de, & Dressler, W. U. (1981). *Introduction to Text Linguistics*. London: Longman.
- Bock, K., & Levelt, W. (1994). Language production: Grammatical encoding. In M. A. Gernsbacher (Ed.), *Handbook of Psycholinguistics* (pp. 945-984). San Diego, CA: Academic Press.
- Chen, Y. (2016). *A Study of Misuses of the Definite Article by Chinese English Majors*. Master's Thesis, Huaqiao University.
- Collins, A. M., & Loftus, E. F. (1975). A spreading activation theory of semantic processing. *Psychological Review*, 82, 407-428. <https://doi.org/10.1016/B978-1-4832-1446-7.50015-7>
- Crystal, D. (1992). *An Encyclopedic Dictionary of Language and Languages*. Oxford: Blackwell.
- Danes, F. (1974). Functional sentence perspective and the organization of the text. In F. Danes (Ed.), *Papers on Functional Sentence Perspective* (pp. 106-128). Prague: Academia.
- Eggins, S. (1994). *An Introduction to Systemic Functional Linguistics*. London: Pinter Pub.
- Enkvist, N. E. (1978). Coherence, pseudo-coherence, and non-coherence. In J. O. Ostman (Ed.), *Cohesion and Semantics*. Abo: Akademi Foundation.
- Firth, J. R. (1957). A synopsis of linguistic theory. In *Studies in Linguistic Analysis* (pp. 1-32). Oxford: Blackwell.
- Grimes, J. E. (1971). Kinds of information in discourse. *Kivung*, 4, 64-74.
- Guan, L. (2008). *Functions of Cohesion and Coherence in PEC*. Master's Thesis, Guangdong University of Foreign Studies.
- Guan, L. (2010). Cohesion and coherence in proofreading and error correction. *Journal of Guangdong University of Foreign Studies*, 21(03), 95-98.
- Halliday, M. A. K., & Hasan, R. (1976). *Cohesion in English*. London: Longman.
- Hasan, R. (1984). Coherence and cohesive harmony. In J. Flood (Ed.), *Understanding Reading Comprehension: Cognition, Language and the Structure of Prose* (pp. 181-219). Newark: International Reading Association.
- Johns, A. M. (1986). Coherence and academic writing: Some definitions and suggestions for teaching. *TESOL Quarterly*, 20(2), 227-265. <https://doi.org/10.2307/3586543>
- Kintsch, W. (1974). *The Representation of Meaning in Memory*. Hillsdale, NJ: Erlbaum.
- Leech, G. N. (1983). *Principles of Pragmatics*. London: Longman.
- Levelt, W. J. M. (1989). *Speaking: From Intention to Articulation*. Cambridge, MA: MIT Press.
- Li, H. (2015). Study on the evaluation model of the formation of portfolio based on English writing. *Journal of Hubei Open Vocational College*, 28(13), 148-149.
- Liu, K. (2021). Evaluation of reliability and validity of TEM-8 test papers. *Overseas English*, (05), 88-89.
- Liu, L. (2005). *Functions of Cohesion and Coherence in Proofreading and Error Correction Part in TEM8*. Master's Thesis, Guangdong University of Foreign Studies.
- Liu, L. (2007). On the role of cohesion and coherence in TEM8 proofreading and error correction questions. *Journal of Mudanjiang Normal University*, (03), 43-45. [https://doi.org/10.2307/358654310.13815/j.cnki.jmte\(pss\).2007.03.015](https://doi.org/10.2307/358654310.13815/j.cnki.jmte(pss).2007.03.015)
- Liu, X. (2015). Construction of English portfolio at vocational level based on the philosophy of learner development. *Journal of Beijing City University*, (04), 96-99+104. <https://doi.org/10.16132/j.cnki.cn11-5388/z.2015.04.018>
- Luo, W. (2019). Overview of domestic cohesion and coherence research in the last 20 years. *Journal of Kaifeng Institute of Education*, 39(10), 59-60.
- Ma, Z., & Zhang, J. (2015). Investigation and analysis of the learning strategies of primary and secondary school

- students. *Educational Research*, 36(06), 85-95.
- McCarthy, M. (1991). *Discourse Analysis for Language Teachers*. Cambridge: Cambridge University Press.
- McKoon, G., Gerrig, R. J., & Greene, S. B. (1996). Pronoun resolution without pronouns: Some consequences of memory-based text processing. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 22, 919-932. <https://doi.org/10.1037/0278-7393.22.4.919>
- Morton, J. (1969). Interaction of information in word recognition. *Psychological Review*, 76, 165-178. <https://doi.org/10.1037/h0027366>
- Norman, D. A., Rumelhart, D. E., & the LNR Research Group. (1975). *Explorations in Cognition*. San Francisco: W. H. Freeman.
- Ren, J. (2017). A review of research on evaluation of college English writing portfolios at home and abroad. *Journal of Kaifeng Institute of Education*, 37(10), 68-69.
- Richards, J. C., Platt, J., & Platt, H. (1992). *Longman Dictionary of Language Teaching & Applied Linguistics*. London: Longman Group Limited.
- Song, J., & Bai, Q. (2017). Exploring the trend of proposing questions in the short-text correction section of the Test for English Majors-Band 8 (TEM-8). *Data of Culture and Education*, (18), 200-201.
- Thompson, G. (1996). *Introducing Functional Grammar*. London: Arnold.
- Tulving, E., Mandler, G., & Baumal, R. (1964). Interaction of two sources of information in tachistoscopic word recognition. *Canadian Journal of Psychology*, 18, 62-71. <https://doi.org/10.1037/h0083296>
- van Dijk, T. A. (1977). *Text and Context: Explorations in the Semantics and Pragmatics of Discourse*. London: Longman.
- van Dijk, T. A., & Kintsch, W. (1983). *Strategies of Discourse Comprehension*. Orlando: Academic Press.
- Wessells, M. G. (1982). *Cognitive Psychology*. New York: Harper & Row.
- Widdowson, H. G. (1978). *Teaching Language as Communication*. Oxford: Oxford University Press.
- Wu, K., & Hu, S. (2021). An empirical study on TEM8 proofreading performance by English majors at different levels. *Journal of University of Shanghai for Science and Technology*, 43(02), 127-134. <https://doi.org/10.13256/j.cnki.jusst.sse.2021.02.005>
- Xie, X. (2015). Research and countermeasures on short-text correction of new questions in College Entrance Examination English. *Journal of Fujian Institute of Education*, 16(11), 83-84.
- Yu, W. (2013). Suggesting how to effectively combine written expression and short-text correction in English teaching. *English Journal for Middle School Students*, (35), 96.
- Zeng, J. (2016). Testing viewpoints of cloze & proof-reading and evaluation of pragmatic competence in ESP. *TEFLE*, (03), 86-90.
- Zhang, D. (1999). A survey of studies on text coherence. *Journal of Foreign Languages*, (06), 24-31+80.
- Zhao, Z., & Zhang, S. (2010). Use of the definite article in English by non-English major college students: An association theory perspective. *Foreign Language Education*, 31(01), 40-44. <https://doi.org/10.16362/j.cnki.cn61-1023/h.2010.01.017>
- Zhu, Y. (1997). Halliday's standard of discourse coherence - External misconceptions and its own shortcomings. *Foreign Language Teaching and Research*, (01), 23-27.

### Copyrights

Copyright for this article is retained by the author(s), with first publication rights granted to the journal.

This is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/4.0/>).