Analyzing the micro coherence in English writing and implications for the teaching of English writing

SUI Dan-ni[1], CHEN Zheng[2]

(1. School of Foreign Languages, Shenyang University, Shenyang 110044, China;

2. Department of Basic Studies, China Criminal Police University, Shenyang 110035, China)

Abstract: Based on LSA's (Latent Semantic Analysis) measuring textual coherence and Halliday and Hasan's cohesive concepts and their categorization of cohesive devices, this paper analyzes coherent relatedness and cohesive devices of 100 Chinese students' English writings. The results indicate that (1) LSA is proved to be an objective and appropriate method in measuring the textual coherence, and (2) close attention should be paid to the cultivation of textual coherence, which will be of significance to the teaching of English writing. Finally, the paper provides some suggestions for college English writing.

Key words: coherence; cohesion; teaching of writing; LSA (Latent Semantic Analysis)

1. Introduction

Coherence and cohesion are 2 main features of a text. Coherence and cohesion theory can be applied to English writing, because it plays an important role in writing. In terms of discourse structure, Carroll (2000) mentioned 2 levels of discourse structure: local structure (microstructure), that is, in the relationships between individual sentences in the discourse and global structure (macrostructure), that is, relations between individual paragraphs meaning the unity of the article. Local structure is an intrinsic component part while global structure serves as macro-adjudgement and control. In Chinese students' writings, it is found that the students keep their points of view shifting from one to another. Because of different modes of thinking, a topic is not directly developed but rather explained with examples. The passage looks like a text but has no semantic connections between one part of the text and another. Therefore, the appropriate cohesive devices can facilitate the coherence in an article, even paragraphs, and the lack of coherent ability can exert negative influence upon advanced English learners.

LSA is a fully automatic statistical technique for extracting and inferring relations of expected contextual usage of words in passages of discourse. Word, sentence and passage meeting representation derived by LSA have been found capable of simulating a variety of human cognitive phenomena, ranging from developmental acquisition of recognition vocabulary to word-categorization, sentence-word semantic priming, judgments of essay quality and discourse comprehension (Landauer, 1998). The application of LSA in discourse comprehension plays an important role in the paper. According to the original research for the application of LSA, it is an effective method for text-based research, especially for measuring textual coherence in reading comprehension, because the comprehension of text depends heavily on its coherence (Foltz, 1996). The method is proved to be automatic and fast, permitting quick measurements of the semantic similarity between pieces of textual

---

SUI Dan-ni, female, lecturer of School of Foreign Languages, Shenyang University; research field: psycholinguistics.
CHEN Zheng, lecturer of Department of Basic Studies, China Criminal Police University; research field: applied linguistics.

information. This paper tentatively studies the measurement of textual coherence with LSA in Chinese English majors' writings and provides some implications for college English writing through analysis.

2. Research methods

2.1 Participants

The investigation was conducted at a university in China. The subjects under study were 100 juniors from various departments. They were randomly sampled.

2.2 Material

All the participants were doing classroom-writing assignments. All subjects were required to write an English composition of 250-300 words on a given topic.

The writing task was as follows:

> Nowadays with the development of economy, existing cities are growing bigger and new cities are appearing. What do you think is ONE of the major problems that may result from this process of urbanization? Write an essay of about 300 words on the topic given below.
>
> One Major Problem In The Process Of Urbanization.

All these compositions were graded by CET-4 standards.

2.3 Procedure

First is to establish a word-documents matrix called the original matrix, because this matrix contains too much information. Secondly, the Singular Value Decomposition (SVD) is used to reduce the semantic dimension. Thirdly, a reconstructed matrix which is similar to the original one but much smaller is reached. It keeps the essence of the semantic relationships in the texts but discards the incidental and irrelevant details.

3. Data analysis and results

3.1 A general result and analysis of micro coherence calculated by LSA

The following items are analyzed by LSA and linguistic statistical methods. In order to make the experimental data reliable, 2 groups of key words are input: $word_{(0)}$ and $word_{(1)}$. The former is gained from a small corpus of English compositions of college students and retrieved from Wordsmith; the latter is retrieved in the same way but gained from *China Daily*. So $word_{(0)}$ is apt to students' use, while the other one is apt to native speakers.

Table 1   The comparison between $r_{(1)}$, $r_{(2)}$ and actual score

	Higher level	Middle level	Low level
$r_{(1)}$	0.61	0.45	0.21
$r_{(2)}$	0.43	0.22	0.12
Actual score	14	9	4

Note: r: relatedness.

Here are 3 samples with different levels shown in Table 1: lower level, medium level and higher level. LSA calculates the coherence between sentences and then the mean of the sentence coherence is the coherence of micro coherence. From Table 1, people can see that there exists obvious distance between $r_{(1)}$ and $r_{(2)}$ that $r_{(1)}$ is inclined to the ideal data, because 0.61 is close to 1.00, which means the coherence of this composition is higher. But here the $r_{(2)}$ is taken as the analytic standard, because it is inclined to the use of native speakers. The comparison of measurements of coherence between the students' actual scores and data carried out by LSA from $word_{(0)}$ and $word_{(1)}$ is made and

the results represent that the application of LSA in measuring textual coherence is probably an objective method in measuring sentence coherence and the result is almost similar to the sort order made by teachers.

3.2 A detailed result and analysis from cohesive devices

Sentences in a paragraph are coherent, when there is a clear transition from one to another. There are 3 main methods of transition to achieve coherence. They are transitional words and phrases, repetition of key words and pronoun reference, which are identical respectively with the cohesion methods of connective, lexical cohesion and reference. Connective is one familiar type of marked connective relationship in text. According to Halliday and Hasan (1976), there are 4 types of connective relations. But sometimes it is difficult to group some connective elements. So, the authors analyze the students' writings by Halliday's categories of the connectives: elaboration, extension and enhancement. Lexical cohesion has to do with repetition, general word, synonym/antonym, hyponymy and collocation. Reference mainly refers to the analysis of personal reference and demonstrative reference.

3.2.1 Connectives

According to the data in Table 2, here are mainly 2 connectives mistakes in students' writing: firstly, the improper use of some phrases or words, for instance, the confusion between "on the other hand" and "otherwise", "at last" and "later", "on the contrary" and "in contrast", etc.; secondly, the overuse of connectives makes the text tend to be redundant and causes the logical disorder between adjoining sentences.

Table 2   Failure rate of connectives in students' writing

Connectives	Elaboration	Extension	Enhancement
Number of connectives	474	969	515
Failure rate (%)	12.8	40.4	25.4

3.2.2 Lexical cohesion

The statistical results (see Table 3) indicate that the students do not do well in general word and collocation. As to general word, the students usually choose the wrong words to express some concepts. The habitual wordings may influence the collocation in students' writing, such as "do efforts", "fast speed", "living level", and so on.

Table 3   Failure rate of lexical cohesion in students' writing

Lexical cohesion	Repetition	General word	Synonym/antonym	Hyponymy	Collocation
Number of lexical cohesion	992	402	698	392	1006
Failure rate (%)	5.9	26.1	1.2	7.4	15.2

3.2.3 Reference

Reference includes personal reference, demonstrative reference and comparative reference. The former 2 types are frequently used in students' writing and will be focused in the present study. Personal reference includes personal noun with subjective case or objective case, adjective possessive and noun possessive; demonstrative reference includes definite article, demonstrative pronouns (this, that, these, those) and demonstrative adverbs (here, there, now, then, etc.).

The results (see Table 4) show that most students use first person and third person frequently, because they usually use these 2 persons to set examples. The frequent mistake in using the third person is the misuse across genders; while the second person is used as general reference and the usually mistakes are improper reference. In addition, the inconsistency of personal reference is always observed in students' writing.

Besides, the students could not discriminate the exact context of definite article "the", for example,

"Nowadays, world changes quickly", there should be "the" before "world"; "The urbanization is the tendency of the whole world", the word "urbanization" appears firstly, so it is unnecessary to add "the" before "urbanization".

Table 4   Failure rate of personal reference and demonstrative reference in students' writing

Reference	Personal reference							Demonstrative reference			
	I, my, etc.	you etc.	we, us etc.	he, etc.	they, etc.	it, its, etc.	one, etc.	the	this, that	these, those	here, there
Number of reference	814	185	305	984	206	195	121	826	107	49	16
Failure rate (%)	0.6	7.4	1.6	2.1	2.9	1.6	4.7	15	1.5	1.5	0

4. Discussion

Writing is using language to make a passage or text for communication. If the writer-reader communication is intended to be successful, both cohesion and coherence need to be maintained in a piece of writing. So in college English teaching, teachers need to give proper lectures on cohesion and coherence in connection with English writing. English learners should learn to employ clear ways of thinking and proper cohesive devices on which coherence is based. Since clear ways of thinking can create a proper semantic space which makes the writing smooth and keep the relevance to the subject. At the same time, cohesive devices play an important role in English writing, suggesting that it will be meaningful and useful to introduce textual cohesion to the teaching of English.

Besides, teachers should offer students some model texts. They can ask students to analyze them, paying attention to the writers' choice of words, such as connecting words, personal pronouns, demonstrative pronouns and adjectives, etc.. Within a passage, there are logical and linguistic links between sentences. At the same time, students are required to examine a writer's organization of ideas, in other words, the writer's arrangements of text. Which sentence best expresses the main idea of the text? What supporting details are given to develop the main idea? How is the text organized? After they have mastered adequate vocabulary and fully comprehended the text, students can practice consciously in their own writing. Before writing, the students had better write down an outline of their way of thinking. In this way, they can develop their writing according to the main line. LSA can be employed as a statistical method to estimate the students' writing coherence. They can make proper corrective measures towards the weak points in the sentences or paragraphs. Practice makes perfect. The more they practice, the more they get familiar with the coherent techniques.

5. Conclusion

In all, the coherence and cohesion serve as an effective tool to improve students' writing. What is more, the approach can be applied in other areas of English study, such as listening, speaking, translating, especially reading. As a channel of input, reading can act on writing. If students' attention is drew to cohesion and coherence in a text, and to how coherence is achieved through cohesion in a reading or listening class, students will have a better understanding of coherence and cohesion and adopt them in their writing.

References:
Carroll, D. W. (2000). Thomson learning Asia. *Psychology of language*. Beijing: Foreign Language Teaching and Research Press, 99-103.
Foltz, P. W. (1996). Latent semantic analysis for text-based research. *Behavior research methods, instruments and computers*, 28(2), 197-202.
Halliday, M. A. K. & Hassan, R. (1976). *Cohesion in English*. London: Longman.
Landauer, T. K., Foltz, P. W. & Laham, D. (1998). Introduction to latent semantic analysis. *Discourse Processes*, 25, 259-284.