

Design on Corpus System for Classical Chinese Teaching in Middle School

Zhantao He

College of Literature, Bohai University, Jinzhou, 121013, China

hezhantao@126.com

Keywords: middle school; classical Chinese teaching; corpus system; system design; database design; function design

Abstract. Classical Chinese is an important carrier of ancient Chinese culture, has great cultural value and education value, classical Chinese teaching is an important part of Chinese teaching and learning. Build high school classical Chinese teaching corpus is promote the classical Chinese teaching information and is the effective means to solve the current predicament. In this paper, the research content mainly includes four aspects: one is corpus base, including text feature extraction and classification of corpus, optimization design and the rule base to establish; Second, analyzes the present situation and existing problems of high school classical Chinese teaching; Three is that based on SQL Server database management system, designs the corpus system constructed by six, the logic of a relational table structure; Four is designed by the corpus management, query statistics and user management and so on three big modules of the system function.

Introduction

Since the advent of computerized corpus, the fast development of corpus linguistics, the use of corpus for language contrast research, the study of language ontology has achieved fruitful results in [1, 2]. Becker has created the world's first English translation corpus, and puts forward the use of parallel corpus, multilingual corpus and comparable corpora, can find and make sure it's difficult to find the semantic characteristics of the conventional methods, the research of the text style, language habits, such as the redundancy of language, vocabulary co-occurrence, standardized degree and coherent form and syntactic pattern, and even the use of punctuation features, and help translators to choose appropriate translation strategies.

Classical Chinese is an important carrier of ancient Chinese culture, has great cultural value and education value. Classical Chinese teaching is an important part of Chinese teaching and learning, and is the window of the middle school students' understanding of traditional Chinese culture and the starting point, learning some classical Chinese is helpful to improve students' Chinese accomplishment and humanistic connotation. Classical Chinese teaching has always been a problem in the middle school language teacher. The quickening pace of curriculum reform, the implementation of new as well as the requirements for the college entrance examination, and the problem of classical Chinese teaching is also more highlights [3]. With the aid of English translation of corpus research results and related technologies, to build high school classical Chinese teaching corpus, can make full use of the excellent achievements of middle school Chinese teachers of classical Chinese teaching, giving full play to the advantages of education informatization, to provide a variety of auxiliary means for teachers and students, to improve the classical Chinese teaching effect.

Corpus Based

Corpus based includes the following contents:

(1) Text feature extraction [4]. The representation of a text and its feature selection is a basic text mining and information retrieval problem, to quantify the key of the extracted from the text to represent text information. From one without the structure of the original text into a structured computer can identify the processing of information, namely to scientific abstraction of text,

mathematical model is established, a term used to describe and replace text, and make the computer can through to the model calculation and operation to realize the recognition of the text. The purpose of feature selection is both to contain enough information to reflect the characteristics of the text, and not too complicated to make learning algorithm can't handle. The specific process includes word segmentation, pretreatment, word frequency statistics and feature selection.

(2) Classification of corpus [5]. For classification of corpus, according to different standards and reference, the researchers have different classification method. Donald E. Walke in the context of "The Ecology of Language", according to the content of the corpus is divided into the following four types: one is the heterogeneous type corpus. Widely collect all kinds of corpus, collect different types of text, not to determine the sampling principles and standards of corpus collected in the style of the original storage; Second is homogenous corpus. Collecting similar materials, all contain information, according to the content have the same property; Three is that corpus system type. Emphasizes the systematic, uniformity and rationality of select material, makes every effort to make the material has a wide range of representative, in order to reflect certain languages, the particular range of whole language facts; Four is, special corpus, collected to establish for a particular purpose.

(3) Optimal design. Index is the database table ordered by the value of one or more columns of a structure, using the index can be quick access to specific information in a database table. According to the need can create the common indexes, unique index, the primary key index and aggregation index. Among them, the common index, is the most basic type of index; The only index, don't allow any of these two lines with the same index value; The primary key index, defined for the table primary key will automatically create the primary key index, a specific type of primary key index is the only index; Clustered index, table the bank of China the physical order of the same as the logical sequence of key value, a relational table can contain only one clustered index. Corpus show "chapter number, clause number and word number" will create indexes.

(4) Rule base is established in this paper. In the process of building a corpus, language acquisition is very important, and language input is also very heavy work. In order to simplify the input process, improve the efficiency of the whole system running, the design rule library, as the auxiliary input processing according to the Chinese grammar rules of grammar feature extraction. At the same time, after corpus build into, can also be used as a set of rules of grammar allowed will not include in the rules of grammar phenomenon in the rule base. In addition, in the application of grammatical rules library corpus, to provide users with check content by the detailed grammatical knowledge information. Chinese grammar rule base structure mainly composed of two processes, a process is a process of word form phrases, another process is a process of words or phrases constitute parts of the sentence.

Current Situation on Classical Chinese Teaching in Middle School

Classical Chinese is a kind of ancient Chinese literary history, popular use of written language. Relatively modern text, since the words in ancient pre-Qin classics is a reflection of the language form as the foundation, to the Qin and Han dynasties after gradually formed a kind of standard of ancient Chinese written language [6]. Classical Chinese teaching is one of an important part of Chinese teaching and learning. Since the new curriculum reform, high school language teaching material to increase the content of classical Chinese, high examination has increased classical Chinese in weight. Visible, the status of classical Chinese in middle school Chinese teaching is more and more high, growing importance, so we must attach great importance to classical Chinese teaching [7]. But, at present middle school writings in classical Chinese teaching present situation unsatisfactory, the main problems are as follows:

(1) Students' interest in learning is not strong. Older has classical Chinese, with the modern social and cultural distance is larger, and the contemporary students' life. Classical Chinese with cultural background, ideology and the system of laws, for today's students are quite strange. Students generally believe that classical Chinese is a language that has no practical significance, and decoupling in modern society, in addition to the examination has no other quality but with

money. Relative to other content of classroom teaching, students in the classical Chinese teaching class depression, learning initiative is poor, the classroom atmosphere dull, difficult to inspire students' interest in learning.

(2) Ignore the subject status of students. Teachers and students is the "dominant" and "subject", the relationship between teachers is the organizer of teaching activities and the participants, students is the main body of teaching activity. In classical Chinese teaching value orientation of knowledge, teachers also appear less equality between teachers and students in the process of teaching the "dialogue", many students are perfunctory, have no chance of taking the initiative to search for knowledge, gradually formed the bad habit of not positive thinking, classical Chinese teaching is bound to get into trouble.

(3) The cultivation of students' humanities accomplishment is not enough. The new curriculum standard classical Chinese teaching goal is: "reading ancient poetry, and to read the simple classical Chinese can understand basic content with the help of reference books. Pay attention to accumulate, understand and apply, to improve their appreciation taste". Can be seen from this standard, classical Chinese teaching should focus on the influence of the students' spirit. But under the background of exam-oriented education, teachers and parents only focus on test scores, ignoring the quality education and comprehensive development.

(4) Too much emphasis on writings in classical style and modern teaching different [8]. High school classical Chinese teaching has its particularity. There are strong middle schools Chinese teaching reform, but rarely involved in the classical Chinese teaching. Some teachers will classical Chinese teaching is equivalent to foreign language teaching, too much emphasis on the different teaching methods and the modern culture, focusing on the analysis of the text content and the writing features of the article, but again not effective for writings in classical style teaching method is put forward, which affect the teaching effect.

Database Design

Usually database design includes the following three stages, the conceptual design stage, is the key to database design, through comprehensive, induction and abstraction of user requirements, to form an independent of the specific concept model of database management system; The logical design phase, the concept structure into a database management system supported by the data model, and carries on the optimization; Physical design phase, the logical model for selecting one of the most suitable for the physical structure of the application environment, mainly the design of the storage structure and access method.

Limited space, this article only carries on the logical structure design. Logical structure associated with specific database management system, this article is based on SQL Server database system to carry on the design. SQL Server is Microsoft's relational database management system. SQL Server is a comprehensive database platform, using the integrated business intelligence tools provides enterprise data management. The logical structure design results are shown in Table 1.

Table 1. Database desing on corpus system for classical Chinese teaching in middle school

Chapter information table (Ancient)				Chapter information table (Modern)			
No	FieldName	DataType	Width	No	FieldName	DataType	Width
1	AnChapterNumber	varchar	10	1	MoChapterNumber	varchar	10
2	AnChapterTitle	varchar	100	2	MoChapterTitle	varchar	100
3	AnChapterAuthor	varchar	30	3	MoChapterAuthor	varchar	30
4	AnWritingYear	datetime	4	4	MoTranslateYear	datetime	4
5	AnChapterKeyWords	varchar	100	5	MoChapterKeyWords	varchar	100
6	AnStorageDocument	varchar	max	6	MoStorageDocument	varchar	max
7	AnDocumentSize	decimal	10,4	7	MoDocumentSize	decimal	10,4

Clause information table (Ancient)				Clause information table (Modern)			
No	FieldName	DataType	Width	No	FieldName	DataType	Width
11	AnChapterNumber	varchar	10	11	MoChapterNumber	varchar	10
12	AnClauseNumber	varchar	15	12	MoClauseNumber	varchar	15
13	AnClauseContent	varchar	200	13	MoClauseContent	varchar	200
14	AnClauseIntroduction	varchar	500	14	MoClauseIntroduction	varchar	500
Word information table (Ancient)				Word information table (Modern)			
No	FieldName	DataType	Width	No	FieldName	DataType	Width
21	AnChapterNumber	varchar	10	11	MoChapterNumber	varchar	10
22	AnClauseNumber	varchar	15	12	MoClauseNumber	varchar	15
23	AnWordNumber	varchar	20	23	MoWordNumber	decimal	6,2
24	AnWordContent	varchar	100	24	MoWordContent	decimal	6,2
25	AnWordIntroduction	varchar	200	25	MoWordIntroduction	decimal	6,2

System Function Design

Modularity is a complex system is decomposed into a better management module, system function general modular design method. Modularization is used to break up, organization and packaged software. Each module needs to complete a specific function, all modules assembled by some method, as a whole, required to complete the whole system function. In the structure of the system, modules can be combined, decomposition and replacement of the unit. Modularity is a complex system is decomposed to become better to management module. Can set different functions in different components, dividing a problem into many small independent and interact with each other component, to deal with complex, large software.

Abstraction is often used in humans to know the world and describe the process of thinking method, is put these similar aspects of concentration and in summary, ignored the differences between them. Modular design of the software can have different level of abstraction. At the highest levels of abstraction, you can use the problems of environment language description method. While on the lower level of abstraction, we can use procedural methods. Based on the abstract modular system function design results are shown in Fig. 1.

The function of the system design embodies the following features:

(1) Easy to use, good sharing. Due to the system based on B/S structure, the application server run data load lighter, more transparent data storage management function. Corpus and its related applications are stored in the server side, users can connect to the Internet, and you can pass the client browser to use this system.

(2) Retrieval flexible way. Users can be single keyword, or keywords retrieval, still can add a wildcard fuzzy retrieval. You can also specify when retrieving span, limited the keywords of the context. Show centered on keywords retrieval results, analyze and summarize the retrieval word language.

(3) Management convenient and good security. The system administrator in the management of the corpus of the client can and maintenance work. At the same time, the only authorized users can use all functions of this system, the check to the corpus of all resources, conducive to the protection of intellectual property rights. At the same time, based on the role of the management mechanism, the user authorization way is flexible.

Conclusion

The latest Chinese teaching outline stipulates that proportion of classical Chinese in middle school Chinese teaching and learning ability to achieve the level. This fully shows the importance of

classical Chinese teaching in middle school Chinese teaching. At the same time as one of Chinese subject knowledge, the classical Chinese is essential. Classical Chinese learning and teaching is to teach students that knowledge of classical Chinese and the Chinese traditional culture the main means and effective way. As a language teacher, should be in classical Chinese teaching practice. To construct a system of classical Chinese corpus, this paper fills the blank of classical Chinese corpus, can promote high school classical Chinese teaching reform, and guide classical Chinese teaching to deepen, break through the limitations of traditional learning methods and expand classical Chinese to learn. To promote the high school classical Chinese teaching is of great significance.

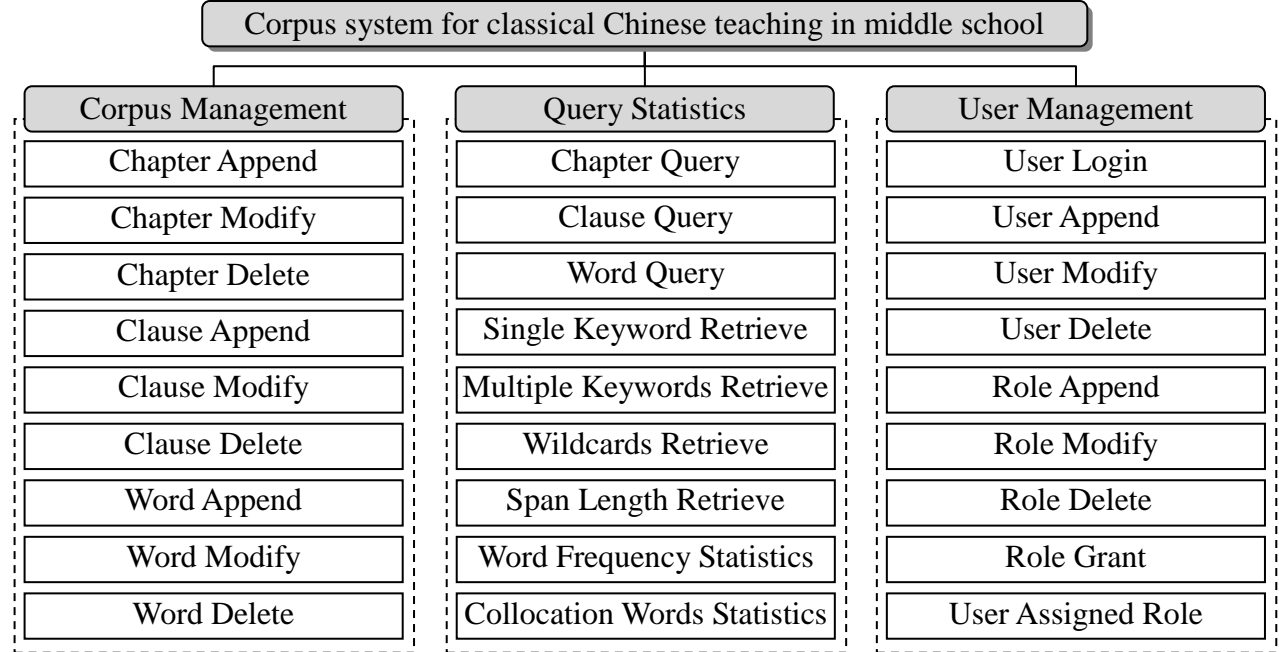


Fig. 1. Function on corpus system for classical Chinese teaching in middle school

Acknowledgement

This work is supported by teaching reform project of Bohai university (BDJG-15-YB-C-015): Linking Practice on Ancient Chinese Teaching and Classical Chinese Teaching in Middle School based on Rotation Classroom.

References

- [1] Velislava Stoykova, "Teaching Corpus Linguistics," *Procedia - Social and Behavioral Sciences*, vol. 143, no. 8, pp. 437-441, 2014.
- [2] Blanka Frydrychova Klimova, "Using Corpus Linguistics in the Development of Writing," *Procedia - Social and Behavioral Sciences*, vol. 141, no. 1, pp. 124-128, 2014.
- [3] X. J. Zhu, "On how to carry out the teaching of classical Chinese in middle school," *Language Teaching In Middle School*, vol. 27, no. 10, pp. 59-62, 2005.
- [4] X. Han, "Knowledge Discover Based on Law FrameNet Ontology Corpus System," Master's degree of Shanxi University, 2010.
- [5] W. Zhao, "The Construction of Small-sized Learning Resources Corpora based on Web Concordance," Master's degree of Xi'an Electronic and Science University, 2013.

- [6] J. G. Duan, "The present situation of classical Chinese teaching in middle schools and the teaching strategies," *Teaching Research for Primary and Middle Schools*, vol. 15, no. 9, pp. 11-12, 2014.
- [7] Y. Qian, X. Luo, "The present situation and suggestion of classical Chinese teaching in middle school," *Journal of Inner Mongolia Normal University (Educational Science)*, vol. 23, no. 12, pp. 99-101, 2010.
- [8] X. L. Xiong, "The present situation and outlet of classical Chinese teaching in middle school," *Chinese Language World*, vol. 21, no. 12, pp. 39-40, 2013.