Application of Small French Corpus in Computer Aided French Teaching

Qiongfang Qin[1] and Bingbing Pan[1]

1School of Foreign Languages, Guilin Tourism University, Guilin, Guangxi 541006, China, agathe_qin@163.com

Corresponding Author: Bingbing Pan, panbingbing0515@163.com

Abstract. With the popularization of computers and the development of modern educational technology, the link between corpus and French teaching is that corpus was first applied to computer-aided French teaching in French teaching, and the research results in this field are more and more closely. Data Drived Learning (DDL) based on corpus can improve students' language proficiency and cultivate their comprehensive ability. This paper combines corpus and French computer aided teaching in the teaching design, with two parallel class of students as the research object, with the help of corpus resource and technology, apply the teaching mode to mouth often teaching experiment teaching practice, the measurement data before and after the experiment and questionnaire survey data, this paper compares and analyzes to verify the feasibility and effectiveness of this method. After a period of teaching practice, the experimental class students are more recognized for this new teaching mode, the vocabulary level is improved accordingly, the learning effect has obvious advantages compared with the control class. This paper introduces the theoretical basis of the corpus-based computer-aided French teaching -- constructivist learning theory and data-driven learning theory. At the same time, it introduces the design of constructivist learning environment and analyzes the design model of constructivist learning environment. Starting with the six elements of the constructivist learning environment design model, this paper designs a concrete application of the model in the field of corpus-based computer-aided French teaching. The experimental results show that the teaching design based on corpus can help students deepen their understanding of the breadth and depth of vocabulary knowledge and improve their comprehensive language application ability.

Keywords: Corpus; Constructivism; Data-driven learning; Learning environment; French teaching
DOI: https://doi.org/10.14733/cadaps.2022.S1.24-34

1 INTRODUCTION

Corpus can meet the needs of current French teaching, bring new methods to computer-aided French teaching and improve the efficiency of computer-aided French teaching. Corpus is a batch of spoken or written discourse materials collected under certain principles and stored in the computer in the form of electronic version for quantitative investigation and qualitative analysis of language, Lo and Anderson's [1] research has made a great breakthrough in this field. With the popularization and development of computers, corpus plays a certain role in the field of education, such as textbook compilation, outline design, language testing, learner corpus analysis and so on. Corpus can help textbook editors to obtain abundant resources through corpus retrieval and make scientific and reasonable arrangement. Corpus can help teachers make use of abundant corpus resources to prepare lessons and optimize teaching methods. A large amount of real corpus can guarantee the quantity and quality of language input for French learners, and help them form the way of inquiry-based learning and autonomous learning, which conforms to the concept of case-based linguistic research. Therefore, the combination of corpus and French teaching is an inevitable trend.

The computer aided French teaching has its own advantages. Rondon-Melo and Andrade point out that corpus is rich in real corpus resources, which can bring a large number of real corpus to learners and Rondon-Melo [2] enable students to understand vocabulary in real context, so as to master vocabulary. At the same time, the unique context of corpus can provide rich examples for learners to further study a certain vocabulary. Learners can learn independently or with the help of teachers to comprehend the different paradigms of vocabulary, summarize and apply the vocabulary. Ali [3] Corpus-assisted instructional design can really improve the breadth and depth of vocabulary learning. However, at present, the application of corpus in China is still limited to the research of a few experts and scholars, and the corpus has not really entered the French teaching classroom. Corpus technology has not yet been popularized, and most front-line teachers have not yet come into contact with corpus, let alone use corpus to assist classroom teaching. Therefore, this paper tries to combine the method of corpus with the practice of high school French computer-assisted French teaching, design specific teaching plans and put them into practice, and explore the feasibility and effectiveness of the corpus-assisted French computer-assisted French teaching in middle school through the specific operation in teaching.

This paper will use literature analysis method, experimental method, questionnaire survey method, corpus method to carry out research. Of corpus analysis with computer aided French teaching theory, on the basis of the computer aided teaching French design based on corpus were discussed, and taught in the class in the investigation, understanding analyzes the present situation of computer aided teaching French and summarizes the difficulties of the computer aided teaching French, introduces the constructivism learning theory, data driven learning theory, and discusses the model of constructivist learning environments in the application of the data driven learning. This paper aims to find the right entrance and construct a specific corpus-based computer-aided French teaching design for senior high school French by using corpus method, and put it into practice. Finally, an eight-week teaching experiment was conducted to verify the effectiveness of this method and whether it has more advantages than traditional teaching methods.

2 RELATED WORK

With the help of corpus technology and quantitative mathematical methods such as statistics and probability, an unprecedented large-scale quantitative research on French and Chinese characters has been carried out. The results not only basically solve the problem of Chinese characters entering into computers, but also promote the improvement of French teaching and research level. Through the establishment of teaching French oriented modern French corpus and interlanguage corpus, and multi-angle analysis and research, make the French teaching syllabus to formulate

and from early estimates of qualitative analysis and by experience analysis, to the more accurate and objective quantitative analysis [4], so a more scientific and targeted. The development of French corpus is unbalanced. At present, Zhao [5] introduced there are more written corpus and fewer spoken corpus. The above-mentioned modern French corpus and French interlanguage corpus for foreign French teaching have played an important role in teaching and scientific research. In contrast, Gao [6] described the construction of the spoken French interlanguage corpus is still in the tentative stage. All kinds of corpus processing techniques are very important to the design and application of French computer aided teaching system and courseware. For example, French word processing and sentence processing technology can be applied to the assessment of the writing level of learners, that is, by judging the number of words, words and sentences used in students' compositions, as well as the grade and recurrence rate of these words, words and sentences, the writing level of students can be objectively and automatically assessed.

The use of speech synthesis and recognition in the French phonetics teaching multimedia and phonetic experiment is adopted to improve the phonetic teaching in the classroom, students can not only hear the sound, you can also see the voice map, you can also use the voice and speech analysis software graphics one-to-one correspondence, Gilakjani and Rahimy [7] point out that the interpretation of students don't need a lot of language, can read the sound, not only can well solve the theory explanation and the contradiction between language communication, more major is accords with the cognitive law of adult learning speech. According to the function of human-computer interaction, we divide the interaction mode into two types: one is reactive interaction. This type of interaction is the learner's ability to respond to stimuli presented by the system. This is a passive response. The courseware maker controls the choice of teaching content, the difficulty of exercises and the order of presentation. Such interaction is used for general data presentation and systematic introduction of professional knowledge [8]. The second type is active interaction. In this kind of interaction, learners can receive knowledge information and actively reflect the process of thinking, which is the interaction provided according to the needs of learners. This is a two-way interaction, the system can be adjusted according to the learner's reaction and the interaction, the different reaction to different Suggestions, through a variety of interactive communication with learners and tracking of thinking way, determine the condition of the learners' learning, and provide timely advice [9], learners can depend on this to adjust their learning strategies. This type is the highest level of human-computer interaction, and is also a form of higher difficulty and more complex process. In recent ten years, the theory and practice of French computer aided instruction have made great progress. In theory, it is mainly reflected in the development of French information processing and French computer-assisted instruction, the combination of multimedia and network technology and French computer-assisted instruction, and Liu [10] introduced the development of French computer-assisted instruction under the background of modern educational technology. In practice, it is embodied in computer aided teaching and supporting environment, multimedia classroom teaching design method, distance teaching design method, multimedia teaching materials and resources construction, teaching experiment and testing and teaching management, etc. It is of great significance to sort out and show the achievements in this field for the better development of French computer-aided teaching in the information age. French teaching has experienced a transformation from traditional media to modern media in terms of media teaching, which can be roughly divided into three stages: blackboard + chalk stage; Audio + video stage; Multimedia computer and network teaching stage [11]. Average word length refers to the arithmetic mean of the word length of all the symbols that appear in the corpus to be counted. The average sentence length is the arithmetic mean of the number of word symbols in all sentences in the corpus to be counted. These two statistical parameters can also reflect the characteristics of the corpus to a certain extent. For example, in general, the average sentence length of corpus from beginners in French is smaller than that of corpus from higher-level French learners, and also smaller than that of corpus from authors who are native speakers of French [12]. There is no clear and widely accepted definition of collocation, "the co-occurrence of two or more words within a short distance of a text. Using corpus indexing software to judge the collocation word of a certain word and the collocation force between the

collocation word and the word provides operability. However, it must be noted that this definition is not strict enough. This is a passive response. The courseware maker controls the choice of teaching content, the difficulty of exercises and the order of presentation. This kind of interaction is used for general data presentation, systematic introduction of professional knowledge, etc. The second is active interaction. In this kind of interaction, learners can receive knowledge information and actively reflect the process of thinking, which is the interaction provided according to the needs of learners. This is a two-way interaction, the system can be adjusted according to the learner's reaction and the interaction, the different reaction to different Suggestions, through a variety of interactive communication with learners follow and ways of thinking, determine the condition of the learners' learning, and provide timely advice, learners can depend on this to adjust their learning strategies. This type is the highest level of human-computer interaction, and is also a form of higher difficulty and more complex process. The application of computer-aided teaching in classroom teaching, in the original has been formed in the teaching system has added a new element - multimedia courseware, the formation of a new teaching model we call the multimedia teaching model. Multimedia technology has triggered a series of changes in language classroom teaching, such as teaching materials, teachers, students and tests.

3 THEORETICAL BASIS OF FRENCH TEACHING AIDED BY COMPUTER WITH SMALL FRENCH CORPUS

3.1 Design of Learning Environment for Computer-Aided French Teaching based on Constructivism

Understanding any problem requires learners to have certain experience in the problem and be able to construct the corresponding mental model. However, for beginners in general, what they lack most is experience, which is very important for them to solve the problem. Therefore, it is very important for constructivist learning environments to provide a series of relevant examples that learners can refer to. Relevant examples in constructivist learning environments can support learners' learning in two ways.

Relevant examples in constructivist learning environments can enhance learners' cognitive flexibility. The cognitive flexibility theory, a branch of constructivism theory, holds that traditional teaching often simplifies the complex real background of problems, which tends to make students have a one-sided understanding of problems. Therefore, the theory advocates providing multiple representations and interpretations of relevant contents to express the complexity inherent in the knowledge domain itself, the connection between a certain point of view and the concept and the connection within the concept, and using multiple and related examples to convey multiple viewpoints on many issues. Therefore, in order to emphasize students' cognitive flexibility, relevant examples should provide a variety of viewpoints and perspectives on the problem to be solved. However, in the actual computer-aided French teaching process, the operation of these steps in the human brain may appear jump, that is, do not need to go through some steps, and there may be backtracking, that is, repeat some steps. Moreover, further empirical research is needed to determine whether semantic representation is separated from specific linguistic pairs in the translator's mind. Therefore, his theoretical model of the computer-aided French teaching process can only be an idealized theoretical hypothesis. This is shown in Figure 1.

As shown in figure 1, the background and meaning of the learners to understand the problem, to construct their own model and puts forward problems solving assumption, detailed background need-to-know about the problem, and need to learn the necessary preparatory knowledge, thus the rich learning resources is a part of constructivism learning environment is necessary, in order to understand the problem, what kind of information students need. Constructivist learning environment should provide learners with a variety of information resources (including text, graphics, sound and video, etc.) that are available for learners to choose, rich and readily available, and related to problem solving.

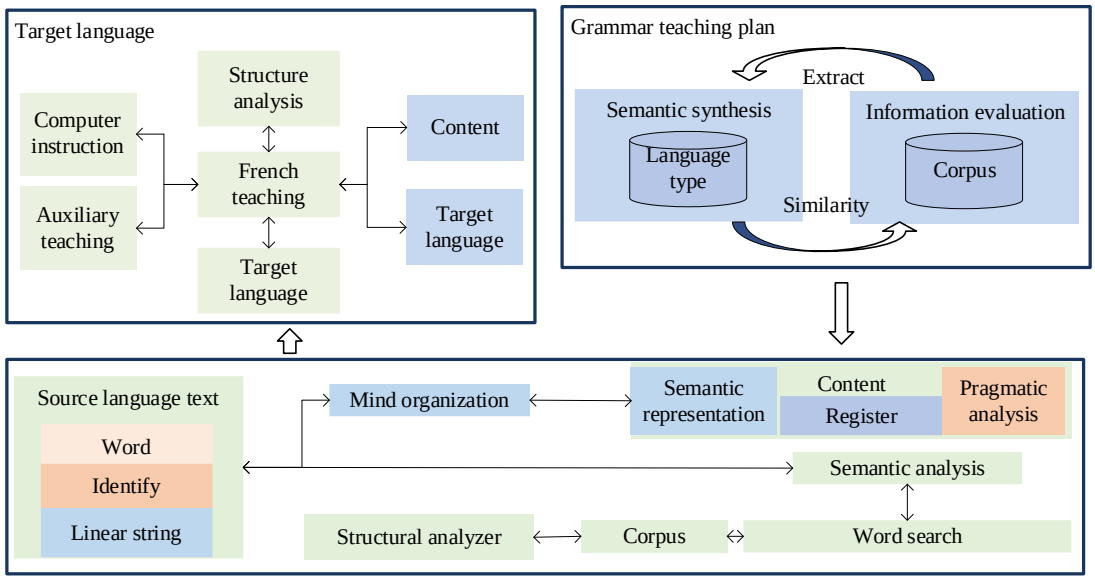

Figure 1: Theoretical model of small corpus in computer-aided French teaching.

3.2 Data-Driven Learning French from a Small French Corpus

The content of instructional design mainly includes five aspects :(1) the determination of teaching objectives and teaching contents. For example, in the pronunciation process of pronunciation teaching, computer images can be used to display the pronunciation schematic diagram of syllables, the direction of airflow and the graph of sound waves. (2) Analysis of learner characteristics. Basic knowledge level, motivation, learning attitude, will, cognitive structure, etc. (3) the design of knowledge structure. Research teaching content, teaching process, teaching methods, etc. (4) Diagnostic evaluation design. The courseware raises the question, the student answers, the system gives the feedback information. (5) Selection of media information. Use different media to present different teaching content, in order to get the best teaching effect "less. The design and production process of teaching courseware is generally as shown in Figure 2.

The content of computer aided French teaching design mainly includes the following four aspects :(1) Screen interface design. The elements of the screen interface are: menus, Windows, buttons, ICONS, dialog boxes, hotkeys and so on. Attention should be paid to the design: the layout of the screen information should be symmetrical, concise and balanced, the display effect of the screen information should be bright, the display density should be appropriate, and the collocation of the color should be paid attention to. (2) Cover introduction design. The cover guide is like the preface of the book, which is a brief description of the functions and Windows of the courseware. (3) Design of navigation strategy. Most of the courseware uses hypertext technology, the relationship is complex, learners are easy to get lost. The specific ways of navigation are: navigation charts, ICONS, buttons, keyword labels, etc. The position of the navigation system in the courseware, the way of navigation and so on should be considered in the design. (4) Design of interactive mode. Frequent human-computer interaction: menu dialogue, question-and-answer dialogue, function keys, menu interaction, etc. The interaction mode should be determined according to the specific teaching content.

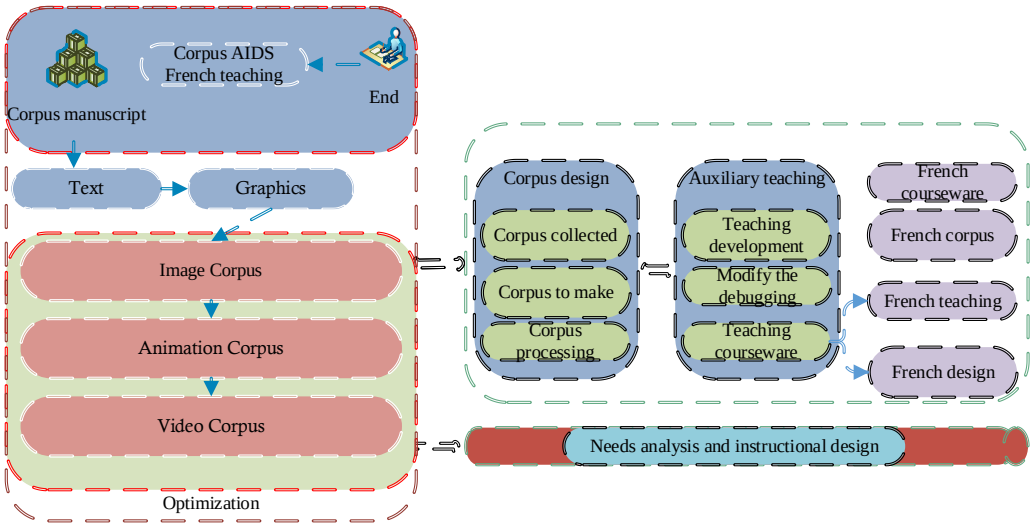

Figure 2: The application process of small French corpus in computer-aided French teaching.

4 APPLICATION OF SMALL FRENCH CORPUS IN COMPUTER AIDED FRENCH TEACHING

In order to develop such a theoretical model with measurable indexes, it is best to start from the teaching theory of computer-aided French teaching which is close to the practical practice of computer-aided French teaching or translator training. "Not computer aided teaching process of detailed and accurate description of the" French, just for the purpose of "teaching", the computer aided teaching process model including the French "understand" and "refactoring" two stages: the first translator of the meaning of computer aided teaching French units make assumptions, and then take advantage of existing knowledge of the language and paralanguage test this hypothesis is correct, if not appropriate, it returns to significance hypothesis, until the translator sees fit. Then comes the reconstruction stage, the translator reconstructs the expression of the computer-aided French teaching unit in the target language, and then tests it according to its fidelity and language acceptability. If the result is not satisfactory, the target language expression of the computer-aided French teaching unit will be rewritten, and then tests it until it is satisfactory. Compared with Bell's theoretical model of CAI process, this model is relatively simple and close to the actual operation steps of CAI process.

The subjects of this experiment are the students majoring in French in 2006, who have entered the third grade at that time. Before participating in this experiment, all of them had studied French for more than 8 years (3 years in junior high school, 3 years in senior high school, and 2 years in university), and Chinese was their mother tongue. They all passed the Chinese language test of the National College Entrance Examination, and took part in the National Professional French Test Band 4 in April 2008. A total of 52 students from two nature classes were selected to participate in this experimental study. One nature class (Class 5) was taken as the experimental group and the other nature class (Class 6) was taken as the control group. The basic information is shown in Table 1 and Table 2.

Number	Age	Chinese Scores in College Entrance Examination	French score in college entrance examination	French Major

Male 9	Female 29	No. 38	Max 22	Min 19	Avg 21.2	Max 135	Min 105	Avg 118	Max 136	Min 113	Avg 123	Max 85	Min 62	Avg 71.5

Table 1: Basic information of subjects in the experimental group.

Number			Age			Chinese Scores in College Entrance Examination			French score in college entrance examination			French Major		
Male 9	Female 29	No. 38	Max 22	Min 18	Avg 20.2	Max 132	Min 104	Avg 115	Max 133	Min 112	Avg 121	Max 86	Min 61	Avg 72.5

Table 2: Basic information of subjects in the control group.

Establishment of a mini-computer-aided French teaching corpus: In the actual teaching of computer-aided French teaching, the teacher can download various Internet resources and software tools from the Internet, and then show the method of creating a computer-aided French teaching corpus to the students in class. The teaching of students' computer aided French teaching should not only improve their basic ability of computer aided French teaching, but also provide professional training of computer aided French teaching for related majors such as medicine, electricity, law and mechanical industry. The computer-aided French teaching (CAFT) corpus is used to fill the CAFT words and sentence patterns of related industries. On the Internet, there are a lot of resources, thus providing a reliable data source for the creation of a variety of corpora. In actual learning, we can also use some small computer aided French teaching software such as Wordsmith and HTTrack to collect corpus. The study of computer-aided French teaching using corpus can greatly improve students' autonomous learning ability.

There are a total of 9 modal words that can be used in the imperative sentence. The highest frequency of "ba" is 32 times, while the lowest frequency of "ya" is only once. From the perspective of corpus style, it mainly focuses on general self-statement and dialogue. It is difficult for us to find out the correlation between these modal words and imperative mood from the statistical frequency, and the support of a larger corresponding corpus is needed. This is shown in Figure 3.

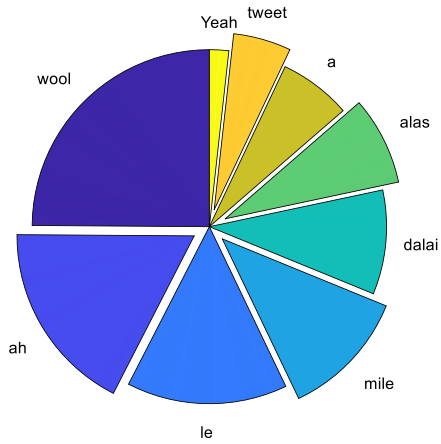

Figure 3: Inflammations used in exclamatory sentences in the corpus.

There are a total of 5 modal words that can be used in a sentence, and the frequency of "ah" and "terteric" is shown in Figure 4. It can be seen that modal words used in a sentence appear more frequently in natural discourse, which is related to the short sentences of natural discourse and the flexible conversion and switching. This is shown in Figure 4.

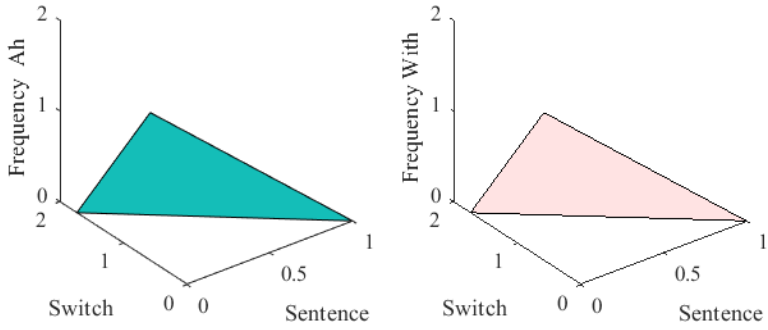

Figure 4: Frequency statistics of modal words used in sentences.

Program debugging process, the two corpora parameters increased gradually, we found that the figure curve of tortuous changes gradually reduced, until we will choose the number of iterations and the initial center of mass is set to 500 and 1000 respectively, clustering to try many times found that the curve of the phenomenon of high and low to reduce gradually, the curve start in one direction, as shown in Figure 5 and figure 6, suggesting that each input the same K value, the classification results tend to be stable, gradually close to the criterion function, and the results of classification index line analysis found that the clustering result is good. In the debugging stage, the number of index rows in the corpus is all 1000, but when the amount of index row data is larger, in order to improve the clustering effect, the above two parameters should be set to be larger.

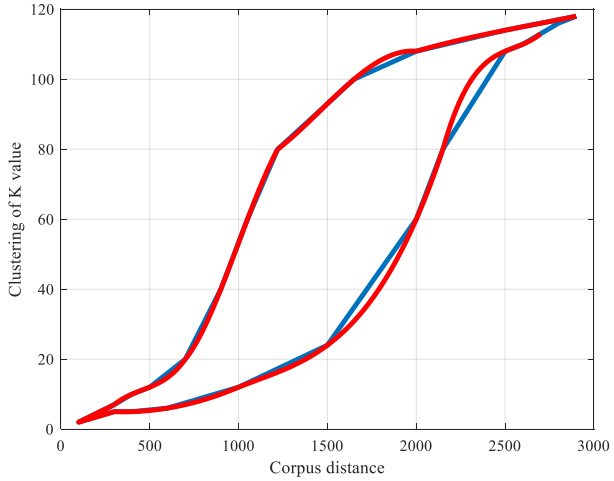

Figure 5: Curve of the sum of squares of distances within a cluster and K value.

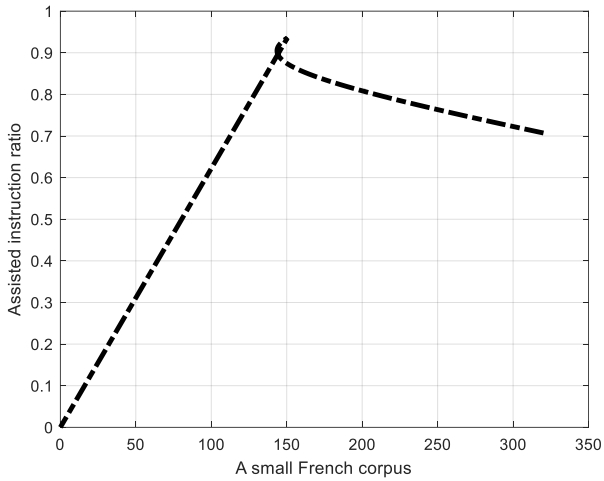

Figure 6: Ratio curve of BETWEEN SS/TOTSS.

Figure 7 is the waveform diagram of ten fundamental frequency curve of "now luowen must be lifted up". This sentence is a recording of Zhao Yongqin, the speaker. From the point of fundamental frequency, at the end of "it", we can see an obvious downward trend, which should be a falling tone.

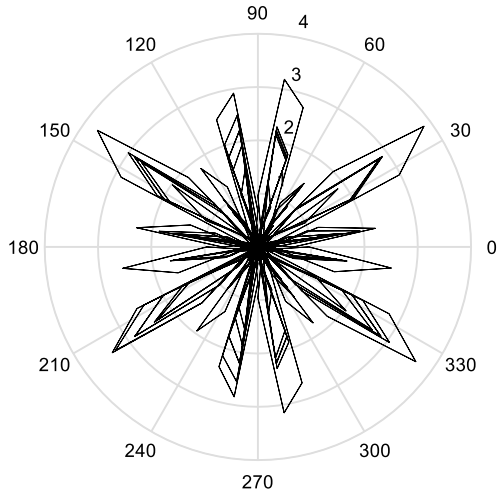

Figure 7: Waveform diagram and fundamental frequency curve.

The selection of K value can be determined according to the research purpose. In order to find the most common type of this node word, the K value corresponding to the first inflection point in the graph can be selected. As shown in Figure 8, K corresponding to the first inflection point is 9.5. In order to find as many types of this node word as possible, the K value corresponding to the second inflection point can be selected. The K value corresponding to the second inflection point in Figure 8 is 12.5. It is even possible to select the value of K for the fourth or fifth inflection, and it is possible to find a very useful form of the verb.

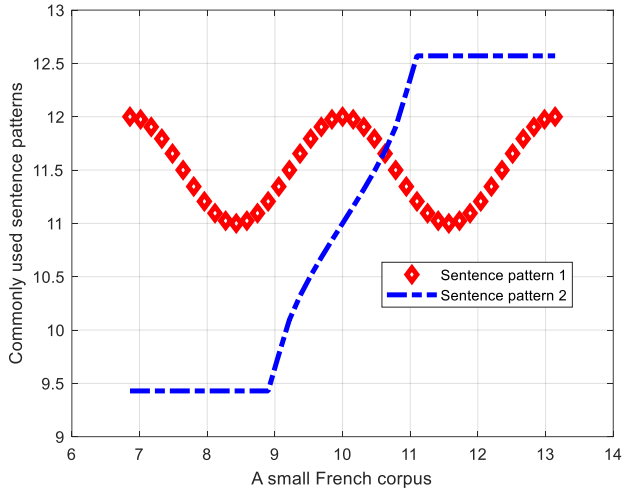

Figure 8: Curve of residual sum of squares and K values.

Through the analysis of the test scores of the experimental class and the control class, it can be found that there is no significant difference between the test scores of the experimental class and the control class before the teaching experiment, and the vocabulary level of the students in the two classes is about the same. After the experiment, the control class before and after the test result has no obvious change, and the experimental classes after test result contrast a pretest result has made significant progress, and the experimental class of measurement result is significantly higher than control class after test scores, to show that the experimental classes after based on corpus of computer-aided French teaching experiment, students' vocabulary levels had significantly improved, and exceed before test level of the control class, has obvious advantages.

5 CONCLUSION

The computer aided French teaching design based on corpus is feasible in classroom teaching, and the students welcome this teaching method. As students are more interested in this teaching method, their learning motivation, interest, attitude and learning strategies have also been comprehensively improved. Students can dare to make assumptions and verify in the learning process. In the process of language learning, students can also cultivate their independent thinking ability and critical thinking, thus killing two birds with one stone. The computer aided French teaching based on corpus can effectively improve students' vocabulary level, make up for the deficiency of traditional teaching, and improve the effectiveness of teaching. This computer-aided French teaching method can attract students' attention more. With its unique features of abundant corpus resources, regular presentation patterns and clear keywords, it can lead students to explore the rules of language, obtain the fun and sense of achievement of independent learning, and master vocabulary more accurately and consistently. Students' vocabulary breadth and depth of knowledge can be improved, and can be flexibly used in practice. The research carried out in this study on the characteristics of student translators' thinking in computer-aided French teaching with the intervention of bilingual correspondence corpus is an enrichment and supplement to the theory of decision-making process in computer-aided French teaching. The general trend of the research on the process of computer aided French teaching is the empirical shift, that is, the use of psychological methods and even the high-end equipment of brain science to study the subjects' thinking activities.

Qiongfang Qin, https://doi.org/0000-0003-2433-2635
Bingbing Pan, https://doi.org/0000-0001-9817-169X

ACKNOWLEDGEMENTS

The phased achievement of the key project "Research on Translation Strategies towards the Tourism Culture of North Guangxi Ethnic Minorities" supported by Chinese Ministry of Education during the 13th Five-year plan (No. JKY10314).

REFERENCES

[1] Lo, Y.-Y.; Anderson, A.-L.: Building Vocabulary of English Learners With Reading Disabilities Through Computer-Assisted Morphology Instruction, Intervention in School and Clinic, 52(3), 2016, 1053-1067. https://doi.org/10.1177/1053451216644829

[2] Rondon-Melo, S.; Andrade, C.-R.: Computer-assisted instruction in Speech-Language and Hearing Sciences: impact on motivation for learning about the Orofacial Myofunctional System, Codas, 28, 2016, 45-67. https://doi.org/10.1590/2317-1782/20162015143

[3] Ali, H.: Use of Computer Assisted Instruction to Improve Students' Reading Skill in English Language, Journal on English Language Teaching, 9, 2019, 219-232. https://doi.org/10.1080/09588221.2014.1000933

[4] Zaini, A.; Mazdayasna, G.: The impact of computer-based instruction on the development of EFL learners' writing skills, Journal of Computer Assisted Learning, 31(6), 2015, 490-496. https://doi.org/10.1111/jcal.12100

[5] Zhao, L.: How to foster Learner Autonomy in Computer-Assisted Instruction Environment, World Journal of English Language, 5(4), 2015, 57-68. https://doi.org/10.5430/wjel.v5n4p57

[6] Gao, X.: New Mode of Distance Learning in Primary Schools in the Environment of Multimedia Computer Assisted Instruction, Journal of Physics: Conference Series, 1544(1), 2020, 12031 -12037. https://doi.org/10.1088/1742-6596/1544/1/012031

[7] Gilakjani, A.-P.; Rahimy, R.: Using computer-assisted pronunciation teaching(CAPT) in English pronunciation instruction: A study on the impact and the Teacher's role, Education and information technologies, 25(2), 2020, 1129-1159. https://doi.org/10.1007/s10639-019-10009-1

[8] Adigun, O.-T.: Computer-assisted instruction, project-based learning and achievement of Deaf learners in Biology, Journal of E-Learning and Knowledge Society, 16(1), 2020, 23-32. https://doi.org/10.20368/1971-8829/1135190

[9] Nakamura, J.: On the Use of Chernoff Faces in the Teaching of English (Computer Assisted Instruction), Arele Annual Review of English Language Education in Japan, 5, 2017, 186-200. https://doi.org/10.1017/S0261444815000245

[10] Liu, M.: Hypermedia-Assisted Instruction and Second Language Learning: A Semantic-Network-Based Approach, Computers in the Schools, 10(3-4), 2016, 293-312. https://doi.org/10.1300/J025v10n03_04

[11] Leidy, J.: Using Computer Assisted Instruction in an ESL Language Program, IALLT Journal of Language Learning Technologies, 15(1), 2019, 13-24. https://doi.org/10.17161/iallt.v15i1.9074

[12] Xu, C.; Qi, Y.: Analyzing pauses in computer-assisted EFL writing-A computer keystroke-log perspective, Educational Technology & Society, 20(4), 2017, 24-34. https://doi.org/10.26858/ijole.v4i2.10106