Corpus-based English Pronunciation Learning System in English Teaching

Tao Gang

Wenzhou Medical University, Wenzhou, Zhejiang, China

Abstract: Oral English teaching has become an indispensable part of English education in China. The contradiction between traditional teaching methods and the needs of oral English teaching has become increasingly prominent. The development of corpus technology will help solve such contradictions. The purpose of this paper is to study the design of spoken English pronunciation learning system based on corpus. This paper introduces the main technologies designed by this system, including corpus technology, system development technology Visual Studio 2008 and C# language, and introduces the environment used in the system implementation process. This paper proposes an intelligent oral English pronunciation learning teaching solution for both teachers and students. It combines speech recognition, image retrieval, oral evaluation and other functions with oral English teaching, and uses the system to promote the solution of traditional teaching methods and modern oral English conflict between teaching needs. According to the application of corpus in English teaching, the system is designed, and the participants are organized to conduct oral language evaluation experiments. The reasonable rate of model scoring can reach 98%.

1. Introduction

In my country's basic education development strategy, English education has been placed in a prominent position. According to the requirements of the new curriculum standard, English has become a compulsory course. English is a very practical course, and many teachers continue to use traditional teaching methods for teaching, that is, teachers explain pronunciation methods and pronunciation demonstrations, and students follow the method of reading teaching [1-2]. The drawbacks brought by traditional teaching methods, the limitation of learners' own ability, the uneven distribution of educational resources and the lack of learning environment, the accuracy of students' English pronunciation and the fluency of oral expression are severely limited [3-4].

With the development of technology, a variety of spoken English products have emerged in large numbers [5]. Kummin S demonstrates the ways in which combining multimodal materials, students can improve not only their English language skills but also their analytical and creative thinking skills. More importantly, his primary focus is on techniques and tools to improve students' English as a Second Language (ESL) academic performance and learning satisfaction in a classroom setting. More specifically, the study utilizes a variety of multimodal texts as teaching methods, including auditory and visual modalities, including video recordings, YouTube videos, songs, and Adobe

speech [6]. Salins A examines whether incidental orthography promotes oral vocabulary learning in children with hearing loss and whether the benefits are greater than those of hearing children. Word learning was assessed using behavioral and eye-tracking data from picture naming and picture-word matching tasks. Results showed an orthographic boost to oral vocabulary learning in children with hearing loss, and this benefit was maintained throughout the week [7]. Therefore, it is important to study the corpus-based English pronunciation learning system [8].

This paper develops an English accent learning system based on corpus technology. The system has functions such as speech recognition, pronunciation assessment, speech transmission, and oral dialogue, which significantly improves the efficiency of students' oral language learning and allows students to have more time. The advantages of space convenience and rich interactive functions not only help improve learning outcomes, but also enhance learning interest and enthusiasm.

2. Research on the Design of Corpus-based Oral English Pronunciation Learning System

2.1 Corpus

Corpus refers to an operating system consisting of a large number of real corpora and corpus retrieval software stored in a computer or server, that is, the combination of corpus and retrieval software, which can retrieve, calculate, analyze and other functions. According to the structure and development of the corpus, it can be roughly divided into two levels, the non-electronic level: the entire corpus is in a non-electronic form, basically text and cards, and the computer development is very immature. Not to mention that on the application, there are no indexing and control tools, all of which rely on simulation requirements. Electronic corpus stage: At this time, computer technology has been applied, and computers have become an integral part of corpus storage, retrieval and indexing [9-10].

In traditional language learning, most cognitive approaches use inference, where teachers must write out rules and then apply them. Data-guided learning emphasizes learning cognitive thinking, and the development of corpus production software is the use of data-driven learning methods [11-12].

2.2 Principles of System Design

2.2.1 Scientific principles

The system design must be scientific, and the design should be based on the teaching content, following the cognitive laws of students, useing multimedia technology to simplify complex knowledge points and give students interest, attract students' attention, stimulate students' curiosity, and give students initiative opportunities to acquire knowledge in teaching [13-14].

2.2.2 Student-centered principle

The design and use of the oral English pronunciation learning system should be student-centered, focusing on students' autonomous learning, enhancing students' ability to apply language, allowing students to gradually master the English language and flexibly use correct teaching methods under the guidance of teachers [15].

2.2.3 Situational principle

Language is a communication tool, and the use of language always occurs in a specific situation, so the creation of the situation must be as real and natural as possible. It is necessary to use

multimedia technology to provide a large number of multimedia materials such as audio-visual, video, image, animation, etc., to create language situations related to the learning content for students, so that students can immerse themselves in the situation, language environment and make abstract content concrete [16-17].

2.3 Visual Studio 2008

Visual Studio has always been a popular and versatile application development environment under the Windows platform. Visual Studio 2008 adds some features over older versions of VS.NET. For example: Visual Designer. The reason why the .NET Framework can quickly build great, object-oriented applications is that the .NET Framework has a variety of building blocks to solve common system development tasks, such as prefabricated software. The rational construction of closely related applications in the business process of the .NET Framework model facilitates the integration of the system into other complex system development environments [18].

2.4 C# Language

As an object-oriented programming language evolved from C and C++, C# is not only safe and stable, because it has the power of C and C++, and does away with complex features such as macros and multiple C++ inheritance. In addition, C# also combines the advantages of simple operation of VB and high performance of C++. With its powerful functions, fluent syntax, innovative language capabilities and convenient support for component-oriented programming, it has become the preferred language for .NET development.

3. Investigation and Research on the Design of Corpus-based Oral English Pronunciation Learning System

3.1 System Environment

Development tools: use the language C# and JSP. The IDE tools used are Visual Studio. Net2008+DirectX SDK and Eclipse. The operating system is mainly Windows, and the Streaming server needs to be installed for the client to access the video. The software required on the client computer includes .net framework 1.1 and IE browser.

3.2 Corpus Description

The total size of the corpus is 100,000 words, of which the spoken sub-base and the written sub-base each account for half. In addition, it is equipped with a variety of research tools designed by the research group. Since the corpus of this study is the spoken language database, the written language database will not be introduced. The entire corpus of the spoken language database comes from the National English Major Oral Test (TEM4-Oral) between 2020 and 2022 and the National English Major Oral Test (TEM8-Oral) between 2020 and 2022, with about 1 million words. The spoken language library includes two folders, Zhuan 4 and Zhuan 8. In order to facilitate searching and matching, each folder contains two subfolders corresponding to the structure, namely AUDIO and TEXTS, which store the audio files and the texts transcribed according to the audio files respectively. . Each sub-file further includes the audio files and transcribed texts of each year corresponding to the structure. All the audio files are monophonic, the rate is set to 44, 100Hz, and the bit rate is 16bit.

3.3 Scoring Model

3.3.1 Integrity

Completeness detects the ratio of the number of words actually pronounced by the user that match the text to the number of words in the text. For words that are rarely read by the user or whose pronunciation accuracy score is very low, they are considered to be missing words relative to the text. Depending on the status corresponding to each phoneme output by the alignment model, the completeness score is the sum of correctly pronounced phonemes (or words) and incorrectly pronounced phonemes (or words) divided by the total number of phonemes (or words). The calculation formula is shown in formula (1).

$$completeness = \frac{cor + sub}{cor + sub + del} \tag{1}$$

3.3.2 Fluency

The calculation of fluency is determined based on the duration of each phoneme of a word, and a model is established for the pronunciation duration of each phoneme through the statistics of a large amount of speech data, and the mean value of the duration probability of each phoneme contained in the word is calculated, as shown in formula (2).

$$fluency(W) = \frac{1}{N} \sum_{p \in W} P(t_p) \tag{2}$$

where t_p denotes the duration of phoneme p, and N denotes the number of phonemes contained in word W.

For the fluency of a sentence, in addition to considering the fluency of each word in the sentence, it is also necessary to consider the pauses in the middle of the words. If there are pauses of more than a certain duration between words, the pronunciation is not considered fluent enough, and the overall fluency score needs to be increased.

4. Analysis and Research on the Design of Corpus-based Oral English Pronunciation Learning System

4.1 System Function Modules

The functional module diagram of the corpus-based spoken English pronunciation learning system is shown in Figure 1:

Login: The application can implement functions related to user system login. The functions provided include: new users can log in to the system normally after the user enters the user name and password normally, and provides the function of resetting the password if the user forgets the password.

Speech Recognition: It can realize the function of speech to text. It includes modules such as speech acquisition, speech data preprocessing, speech data feature extraction, codebook generation, model training and model recognition.

Voice evaluation: The core function of this module is that the system will evaluate and score the user's follow-up recording by reading an English word or an English sentence. The system makes a self-made recorder, and uses the volume algorithm to ensure that the voice input by the user can be

better recognized and the evaluation results are fed back to the user.

Voice broadcast: The core function of this module is that the user inputs a piece of text, and the system can convert text to speech and broadcast in natural, accurate and fluent voice.

Oral dialogue: The core function of this module is to realize the function of human-computer dialogue.
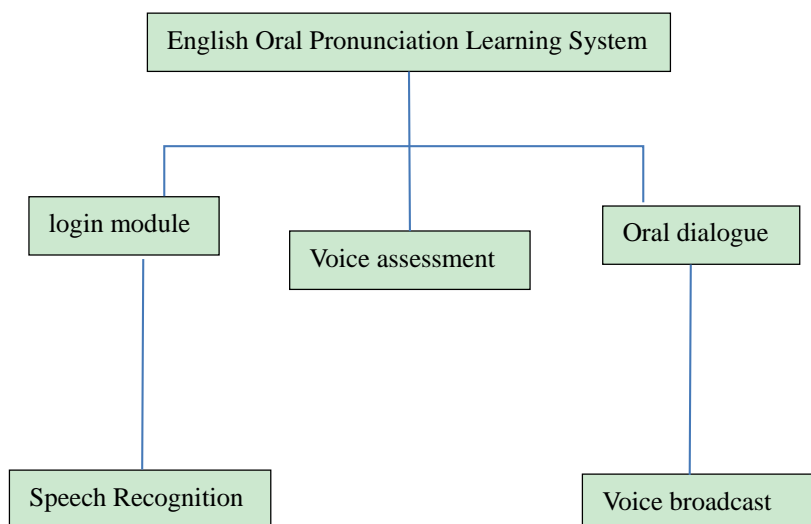

Figure 1: Functional module diagram of spoken English pronunciation learning system

4.2 The Effect of the Spoken Language Evaluation Model

After the evaluation model gives a score for the audio, it is necessary to manually check whether the score is reasonable. The acceptance personnel are all English language experts with English proficiency of CET6 and above. They will evaluate whether they belong to this interval or adjacent interval according to the model score and audio evaluation. If the evaluation of audio is in the range of 41-100, it is in the reasonable range. If the evaluation of audio is in the range of 0-40, it is in the unreasonable range. The scores are divided into 3 intervals, namely 0-40 (poor pronunciation), 41-70 (average pronunciation), and 71-100 (good pronunciation). The pronunciation quality is judged according to completeness and accuracy.

The final rendering of the spoken language evaluation model is shown in Table 1. In the standard American pronunciation evaluation set, the reasonable rate of the model scoring is above 95%, and it can even reach 98% reasonable in the online sentence pronunciation evaluation set nonnative-sen evaluation set. The performance on the word nonnative-word evaluation set is relatively weak, and there is still room for optimization, as shown in Figure 2.

Table 1: The effect of the spoken language evaluation model

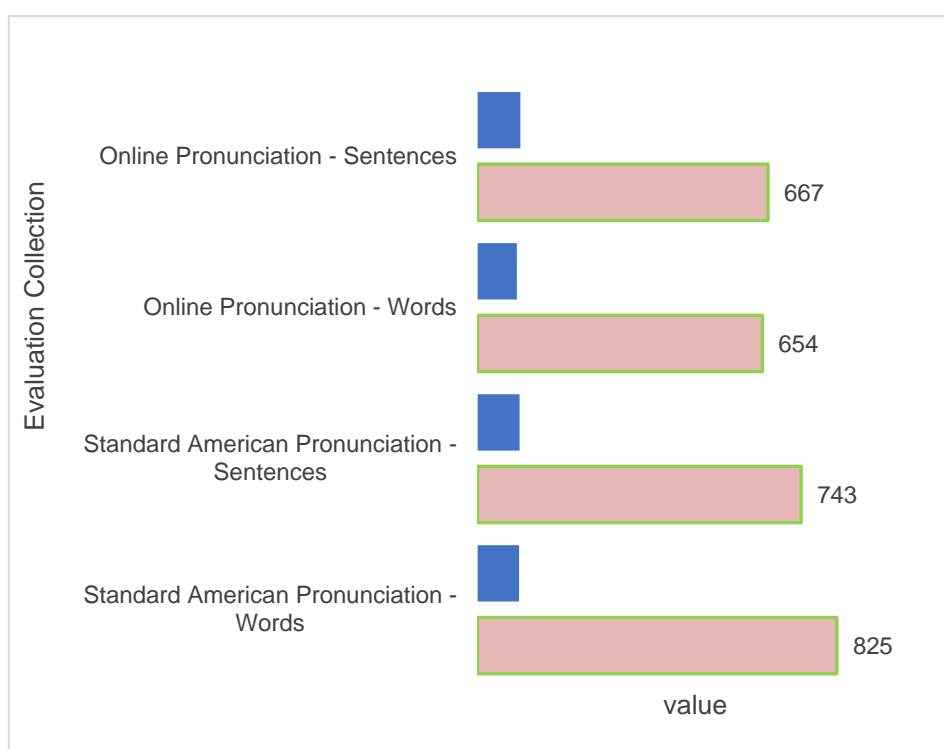| Evaluation Collection	Introduction to the evaluation set	evaluation measure	Reasonable rate (%)
native-word	Standard American Pronunciation - Words	825	95
native-sen	Standard American Pronunciation - Sentences	743	96
nonnative-word	Online Pronunciation - Words	654	90
nonnative-sen	Online Pronunciation - Sentences	667	98

Figure 2: Reasonable rate of oral assessment

5. Conclusions

Good English study is not only a means to understand and learn foreign culture and knowledge, but also an essential tool for future work. More and more foreign companies are established in China. However, in everyday English vocabulary, learning to speak is the biggest problem. According to the characteristics of language learning, this paper summarizes the corpus of spoken English pronunciation, and designs and implements a spoken English pronunciation auxiliary learning system that integrates pictures, videos of real people's standard pronunciation, and comparison between real people's pronunciation and standard pronunciation. The system is a software that attaches great importance to learners' English pronunciation, mouth shape, pronunciation foundation and timely detection of learners' pronunciation. The system is stable, practical and effective.

References

[1] Sydorenko T, Smits T, Evanini K, et al. Simulated speaking environments for language learning: insights from three cases [J]. Computer assisted language learning, 2019, 32(1-4):17-48.
[2] Vargas I B. A Language-in-Use Study of EFL Students' Social Discourses in Project-Based Learning[J]. Colombian Applied Linguistics Journal, 2018, 14(1):108-126.
[3] Kandeel R. Using online games in a hybrid context for the purpose of developing the oral comprehension learning in French as a foreign language [J]. Porta Linguarum, 2018, 2018(30):89-102.
[4] Castelo A. Chinese voices in portuguese: call needs for the oral skills learning[J]. Revista Diadorim, 2020, 22(3):46-64.
[5] Wurth A, Tigelaar D, Hulshof H, et al. Key elements of L1-oral language teaching and learning in secondary education. A literature review[J]. L1-Educational Studies in Language and Literature, 2019, 19(Running Issue):1-23.
[6] Kummin S, Surat S, Kutty F M, et al. The Use of Multimodal Texts in Teaching English Language Oral Skills [J]. Universal Journal of Educational Research, 2020, 8(12):7015-7021.
[7] Salins A, Leigh G, Cupples L, et al. Orthographic Facilitation of Oral Vocabulary Acquisition in Children With

Hearing Loss[J]. Journal of Speech Language and Hearing Research, 2021, 64(3):1-13.

[8] BD Diego-Lázaro, Pittman A, Restrepo M A. Is Oral Bilingualism an Advantage for Word Learning in Children With Hearing Loss? [J]. Journal of Speech Language and Hearing Research, 2021, 64(3):1-14.

[9] Canals L. The effects of virtual exchanges on oral skills and motivation [J]. Language, Learning and Technology, 2020, 24(3):103–119.

[10] Taggart G, Moore G. The Use of Closed-Circuit Television for Oral Tests in the Language Laboratory[J]. IALLT Journal of Language Learning Technologies, 2019, 7(1):27-30.

[11] Gray S, Restrepo M A, Yeomans-Maldonado G, et al. The Dimensionality of Oral Language in Kindergarten Spanish–English Dual Language Learners [J]. Journal of Speech Language and Hearing Research, 2018, 61(11):2779-2795.

[12] Kirsch C. Young children capitalising on their entire language repertoire for language learning at school[J]. Language Culture & Curriculum, 2018, 31(1):39-55.

[13] Ayunon C C, Haloc L I. How far have we gone? Integration of intercultural language learning principles in Philippine ESL classrooms [J]. Journal of Education and Learning (EduLearn), 2021, 15(1):144-152.

[14] Rugo M, Ordulj A. Cross-linguistic transfer in the oral L2 production of Croatian L1 speakers learning Italian as a foreign language [J]. Journal of Foreign Language Teaching and Applied Linguistics, 2021, 2(3):139-148.

[15] Emelogu N U, Nwafor C K, Chigbu G U, et al. Perceived effects of examination special centres on teaching and learning of English language and quality of education in Nsukka local government area, Enugu state, Nigeria[J]. International Journal for Educational Integrity, 2021, 17(1):1-18.

[16] Kamal M I, Zubanova S, Isaeva A, et al. Distance learning impact on the English language teaching during COVID-19 [J]. Education and Information Technologies, 2021, 26(6):7307-7319.

[17] Green C. The oral language productive vocabulary profile of children starting school: A resource for teachers: [J]. Australian Journal of Education, 2021, 65(1):41-54.

[18] Malec A, Peterson S S, Elshereif H. Systematic Review of literature on Assessing Childrens Oral Language [J]. Canadian Journal of Education, 2020, 40(3):362-392.