Application Analysis of Artificial Intelligence in Oral English Assessment

QiaoLing Yong*
Yantai Vocational College, Shandong, 264670

*Corresponding author e-mail: Yongqiaoling@163.com

Abstract. During oral English learning, assessment allows students to understand and correct their problems in oral English to achieve the effect of “Promoting learning through assessment”. The traditional manual assessment has many limitations such as low efficiency and high cost, while automatic assessment of oral English can be implemented based on artificial intelligence (AI) technology, which is a crucial reform direction of oral English teaching and assessment in the future. In this paper, the significance of oral English assessment is first introduced. The existing problems of manual assessment are analyzed. Subsequently, the principle of automatic assessment of oral English based on AI technology is expounded. The development status is explained through two products (Chisheng and Zhiling). Finally, the existing technology and problems in use are analyzed, and targeted suggestions are proposed.

Keywords: Artificial Intelligence, Oral English, Assessment

1. Introduction

English is the most widely used language in the world. As an international common language, many countries have set up special English learning courses as part of basic education. In the process of English learning [1-2], oral practice is an important part. Learners can grasp the key points of English use and improve their expression ability by communicating with others. Currently, the oral English level of Chinese people is uneven, and most people can only use a little common communication language [3]. In classrooms, the focus of the teacher's teaching is mainly on the interpretation of books and the explanation of grammar knowledge [4]. Meanwhile, many schools have not even set a special oral English course. As a result, the students can only write and speak “Dumb English”. On the other hand, many people in China speak Chinese English (Chinglish), and some even speak English in a dialect [5-6]. This leads to the problem that foreigners cannot understand us when we talk to them in English.

In the process of oral English learning, assessment is the key step. Through assessment, students
can understand their oral English level and their shortcomings, and then correct them to achieve the effect of “promoting learning by assessment”. In the traditional oral English assessment, teachers judge and grade students' oral English level according to their personal experience. Without a very standard assessment criteria, it is difficult for students to understand the specific problems of their pronunciation and expression. The level of Chinese English teachers is also uneven, and their pronunciation may have problems. They can neither assess the oral English level of students accurately, not provide guidance in detail. In addition, in large-scale oral English tests, due to the differences of personal subjective feelings and experience, the assessment criteria are not the same among teachers, which will affect the reliability of the test. In addition, the number of people taking the test is large. Hence, there are inevitable scoring errors by teachers due to fatigue and other factors not related to the actual oral English level. Given the above problems and the rapid development of artificial intelligence (AI) technology in recent years, people began to design and develop intelligent oral language assessment tools to help teachers assess the oral English level of students via computers.

2. Principle of intelligent assessment of oral language
Currently, the oral English test mainly includes self-introduction, reading articles, situational questions, listening to essays and answering questions, etc. It can be seen that the main problems to realize intelligent assessment are speech recognition and content understanding. To implement the automatic assessment, it should be performed from the oral expression and content of the evaluator. In the aspect of expression assessment, the first step is to cut off the speech segment of the oral language tester, then analyze the cut-off content, obtain the characteristics, and make a comprehensive assessment from the aspects of sound quality, voice color, tone, etc., and finally summarize them to get a reasonable score. For links such as self-introduction that require content check, the principle of the assessment is more complex. The system needs to learn and train a scoring model from a large number of self-introduction templates in advance, to give a reasonable score;

Hence, the structure of the original oral English assessment system is improved in this paper, and a feature layer is added in the input layer. Through the weight matrix $F$ and $G$, it is connected with the hidden layer and the output layer, respectively. The structure of the oral English assessment system is shown in Figure 1, where the input of the AI oral English assessment system becomes the following:

$$x(t) = w(t) + s(t-1) + f(t).$$

Figure1. RHLM structure combined with feature layer
In the feature layer, input vectors $f(t)$ is the context sensitive vector of words, which contains more multimedia oral English assessment, which is a supplement to the input vector $w(t)$, making the calculation of word probability more accurate:

$$s(t) = f(Uw(t) + Ws(t - 1) + Ff(t)),$$
$$y(t) = g(Vz(t) + Gf(t)).$$  

(2)

Where $f(z)$ is a supplement, which makes the calculation of word probability more accurate:

$$f(z) = (1 + e^{-z})^{-1},$$  

(3)

Where $f(z)$ activate the function for SoftMax.

$$g(z_m) = \frac{e^{z_m}}{\sum_i e^{z_i}}.$$  

(4)

The assessment principle of Situational Questions and listening essays needs to formulate a standard answer in advance. During the assessment, voice recognition technology is used to convert the examinee's audio file into text, and then the difference between the examinee and the standard answer is compared. Finally, the score of the question is given. Certainly, beside the answer, teachers should set some additional scoring standards. For example, some students do not understand the question and how to deal with temporary answers.

3. Typical oral English assessment products

Suzhou Chisheng Information Technology Co., Ltd. (referred to as “Chisheng”) is an educational technology company specializing in intelligent voice technology research and industrialization. The company mainly developed an intelligent English learning product based on big data and deep learning, which can test students' oral English online. This product mainly evaluates the oral ability of the tester from three aspects: pronunciation, pragmatics, and expression: pronunciation ability assessment: it mainly evaluates the pronunciation ability of the tester in the context of reading words and sentences, paragraphs and paragraphs, and following reading, aiming at the pronunciation, intonation and other multiple dimensions, and gives the overall assessment of pronunciation through the assessment, and can carry out personalized diagnosis and multi-dimensional assessment Degree feedback pronunciation problems; pragmatic ability assessment: through the design of specific tasks under a certain scene, let the tester use the language knowledge to complete specific language tasks, to evaluate the pragmatic ability of learners; expression ability assessment uses natural language processing and other technologies to achieve complex oral questions based on open answers (e.g., picture reading, oral composition, reason for evaluating the language expression ability of learners.

The use of the product can reduce the teaching burden of the oral English teacher. Students can test anytime and anywhere according to their own schedule, and get feedback immediately, which is very
convenient. The product can meet the needs of students' oral English practice in different scenarios, such as pre-class, after class, stage assessment, simulation test, oral competition, etc. by building an environment for students' oral English learning and use. For teachers, it is difficult to find the materials needed for oral practice in different scenes in a short time, and the product can well meet the needs of teachers, covering a variety of types of oral English tests, including listening questions, oral questions, and oral composition, picture reading and speaking, story retelling and other open types of questions. Currently, the oral service customers of Chisheng technology have covered 132 countries and regions. The cooperative customers include online education, the training industry, the publishing industry, Internet intelligent equipment, intelligent software, and government agencies.

Intelligent listening oral assessment can support the assessment of words, sentences and other modes. Word mode is only used to evaluate the pronunciation of a word. The assessment results are required to be more detailed. The output indicators include phoneme accuracy, word accuracy, fluency, and stress position. Sentence mode is used to evaluate a sentence. The assessment results are more focused on the overall effect. The output indicators include word accuracy, sentence integrity, and fluency information. Compared with other oral assessment products, the product can cover a wider age range and support all people over 3 years old. In particular, the product also provides differentiated assessment ability support for the particularity of children's oral assessment, such as follow-up reading for children's English picture book words and sentences, and for follow-up reading.

Intelligent listening oral assessment can quickly understand students' oral English ability, provide multi-dimensional assessment results, and provide data support for teachers' curriculum arrangements. After the practice, students can support reading and comparison of data in the background to give feedback on the students' overall learning and mastery of English words and sentences in the classroom, which is very convenient for training institutions and teachers.

4. Case analysis
Although AI technology has many applications in oral English assessment, there are also deficiencies in technology and use currently. Compared with human beings, machines are challenging to achieve the same understanding as human beings. For example, when checking the students' situational dialogue, if the students' answers are inconsistent with the answers, how to judge the machine quickly and accurately is challenging. Currently, although the development and application of natural language processing technology are very extensive, its ability to understand language is limited, and the gap with human expression level is apparent. Through communication, we can easily distinguish between machine and human. On the other hand, the speech analysis technology is not perfect, it is difficult to analyze the emotion and tone of the tester accurately. Therefore, if the machine wants to replace the teacher totally to evaluate the students' oral English, it is still at the primary level at this stage and needs to make a breakthrough in technology.

To observe the influence of the word vector on the performance of the model, the dimension of the word vector is 50, and the window length is 30. The size of the hidden layer is 30, 100, 200, and 300. The confusion of different circulation Multimedia English writing intelligent marking system model on the development set is shown in Figure 2 below.
Figure 2. Error degree of various methods in oral English assessment

Figure 2 shows that the introduction of word vectors in both methods can significantly reduce the error degree of oral English assessment, and the performance of AI algorithm is slightly better than that of the other two models. When the sentences are long, the decrease of this method is more obvious for oral English assessment, while with the increase of the number of Multimedia English writing, the increase of algorithm performance is reduced by adding feature vectors. This shows that in the sentences In a long time, the influence of feature vector on oral English assessment system is greater, and the performance of the algorithm can be improved better.

Education is not just about “learning knowledge”, but more importantly, the cultivation of students' moral character and the improvement of their overall quality, such as self-confidence, communication, questioning, sharing, etc. which are tough to be accomplished by AI. But teachers can use AI as a kind of educational assistant tool and use it to understand students' mastery of knowledge more precisely. For example, teachers can use the above oral assessment software to understand the weak points in students' oral English learning and make more targeted learning plans for students accordingly.

Conclusions

In this paper, the development background and principle of AI in oral English assessment are introduced. The development status is described through two products (Chisheng and Zhiling), and the existing technology and application problems are analyzed. With the continuous development of AI technology, the existing technical problems in oral English assessment will be resolved one by one, and there will be increasing numbers of users. In the future, aided by oral assessment technology, oral English teaching will be more targeted, and the efficiency of learning and testing of students will be significantly improved.

References

[1] Seweryn Kokot, Zbigniew Zembaty, & Piotr Bobra. (2013). An analysis of the effectiveness of application of rotation rate sensors in non destructive damage assessment. Key Engineering Materials, 569-570, 783-790.

[2] Robert Bogue. (2014). The role of artificial intelligence in robotics. Industrial Robot, 41(2), 119-123(5).

[3] Michele De Gennaro, Elena Paffumi, Harald Scholz, & Giorgio Martini. (2014). Gis-driven
analysis of e-mobility in urban areas: an assessment of the impact on the electric energy grid. Applied Energy, 124(1), 94 - 116.

[4] Gómez-Rodríguez, Carlos, Alonso-Alonso, Iago, & Vilares, David. (2017). How important is syntactic parsing accuracy? an empirical assessment on rule-based sentiment analysis. Artificial Intelligence Review(4), 1-17.

[5] Qizheng Ye, Dahai Yun, Fuli Yang, & Dan Tan. (2013). Application of the gray-level standard deviation in the analysis of the uniformity of dbd caused by the rotary electrode. IEEE Transactions on Plasma Science, 41(3), 540-544.

[6] Flora, June A., Saphir, Melissa, Lappé, Matt, Roser-Renouf, Connie, Maibach, Edward W., & Leiserowitz, Anthony A. Evaluation of a national high school entertainment education program: the alliance for climate education. Climatic Change, 127(3-4), 419-434.