Development Research on Deep Learning System of College Oral English under the Cross-Cultural Communication Background

Ai Zhao, Wenchao Liu

School of Foreign Languages, Zhangjiakou University, Zhangjiakou 075000, China

Abstract

With the gradual promotion of globalization and the international standing of China, the cross-cultural communication is growing vigorously, and the social requirements on English oral talents are also increasing. However, the overall teaching quality of college oral English is seriously affected by defects in the teaching environment and conditions etc. of many colleges in China, the English oral capability of modern college students cannot meet the actual demands of the society, and the students’ learning of oral English is also under a difficult circumstance. However, with the rapid development of network information technology, and according to the deep research of network learning field in deep study and the accumulation of many language information, the spoken pronunciation and evaluation technology have been developed rapidly. Therefore, the comprehensive research aims on the development of deep learning system of college oral English based on cross-cultural communication background in the Thesis; the deep learning thought, practicing methods and belief network are described in detail, and the algorithms of appraisal is also introduced; therefore, the system model for college oral English learning is built up based on the development indicator of deep learning system, which aims on proposing scientific and reasonable opinions for students’ promotion in oral English, thus the cross-cultural communication capability of college students can be increased so as to meet the talents needs of the society.

Keywords: Cross-cultural, Oral English, Deep Study, System Development.

1. RESEARCH BACKGROUND

1.1 Literature review

During the process of learning and practicing oral English, the college students in China usually learn by using traditional recording software in their mobile phones or the special mobile software for oral English practicing based on the standard pronunciation; however, the software cannot identify the pronunciation of students rapidly, students can only read follow the software and cannot get the effective pronunciation opinions and feedback directly (Liu, 2013). In addition, due to the limits in technology, many domestic CALL systems more focus on English words, sentence pattern and grammar, but the function of which is not complete, students can only get the total score from these systems. Due to the limit of the comprehensive ability of students in English, they cannot understand their insufficient in oral English only from the total score, thus they cannot make corrections promptly and effectively (Zhao, 2013). What’s more, as for the evaluation aspect of oral English, the college oral English test only focuses of many colleges only focuses on the subjective consciousness, and uniform evaluation indicators are not established; the comprehensive and systematic evaluation on students’ oral English cannot be achieved, which affects the promotion of college students’ cross-cultural communication capability severely; and such evaluation method usually costs more human resource and materials. For all of these situations mentioned above, the new comprehensive system software beneficial for students’ oral English practicing should be developed as soon as possible.

1.2 Research objectives

The college students are main research objectives in the Thesis, the deep learning theory is applied to the system development for college oral English sufficiently so as to help students identify the oral pronunciation; and the deep learning thought, practicing methods, belief network and system algorithm are described and introduced (Yang, 2017). The deep learning system for college oral English is set up based on the cross-cultural communication background, and the indicators and process of system development are also introduced specifically; therefore, the evaluation which can meet the actual conditions can be provided for students.
effectively and scientifically; positive guidance can be provided to the wrong pronunciation, and the enthusiasm of students in independent learning can be driven; thus their oral expression capability can be improved, and the solid foundation can be laid to the cross-cultural communication capability; in the meanwhile, the college teaching quality in oral English can be promoted, which is of important realistic significance to the English teaching and college students’ future development.

2. OVERVIEW ON DEEP LEARNING THEORY OF COLLEGE ORAL ENGLISH

2.1 Deep learning thought and practice

2.1.1 Deep learning thought

The main thoughts of deep learning include the following: (1) the independent learning applies to the preculture in any networks; (2) Any practice layer is monitored with independent learning and supervision each time, and the final result is taken as the information input of a higher level; (3) All levers are readjusted with monitor learning (Tan, 2017). The main adjusting methods shall include: (1) The single-layer nerve cell is set up layer by layer, and only one single network layer can be practiced every time; (2) when all practices are completed, the dormant algorithm is used for optimum adjustment of parameters. The weights among other practice layer excluding the top layer are transferred into two-way levels, thus other layers can also transfer the neural network of single layer into graph model after the top layer is excluded; however, the weights from the bottom to the top can be applied to cognition, and the weight from the top to the bottom can be applied for generation (Yan, 2017). Then the dormant algorithm can be applied to the uniform adjustment of the entire weight. The mutual coordination between cognition and generation can be achieved, and the top layer expression generated can be used to recover the node of the bottom layer to the maximum accuracy.

The dormant algorithm is mainly divided into two parts, the first part is the “sober” status. During the entire process of cognition, the weights generated in all layers are adjusted with step down according to the node status of each layer of abstraction occurring in the external features and upward weights. The second part is the dormant status (Wei, 2012). During the generation process, the status presented in the bottom layer is generated according to the expression of the top layer and the downward weight, and in the meanwhile, the upward weight among layers is also modified.

2.1.2 Training process of deep learning

Firstly, the non-monitoring learning from the top to the bottom should be used. The data is prepared without standards, and the parameters of each layer are practiced without monitor layer; compared to the traditional neutral network, the most obvious difference is that such procedure is equal to the characteristic learning process. In particular, the first layer is practiced with the data free of specific rating data, and each parameter of the first layer are learned firstly in the training. Due the limit in model storage volume and sparseness, the data structure can be learned through the model, thus the characteristics with more presentation capability can be obtained (Xu, 2012). After the layer of x-1 is learned, the input of x-1 is deemed as the input of layer x-1, then the layer x is practiced, and parameters of different layers can be obtained. Secondly, the supervision from top to bottom. According to parameters of each layer can be obtained from the first step, then the effective supervision in learning can be used to complete the parameters of the entire layer model. It is of obvious difference to the entire process of random original value in the traditional neutral network; as the initial parameter of DL is obtained by learning the input information structure in the first step, which is not the initial random information; therefore, the expected effect can be achieved finally according to the more closer relation between such initial value and the overall optimization. Therefore, the entire effect of the DL is sourced from the process of characteristic learning in the first step to a great extent.

2.2 Belief network of deep learning

The belief network in the deep learning is called DBN for short, the core of which is the hierarchical ascending learning algorithm not requiring the supervision; such method is used to predict the model weight value generated in the practice; then the network is subject to adjustment in small scope with the algorithm of inverse transmission, thus the network model of optimum performance can be obtained (Li, 2012). A lot of researches indicate that the actual weight value of the initial multi-layer cognition of DBN with relevant configuration is much better than the initial result achieved without random. Under the premises of Boltzmann influence, in case
of increase and decrease in the quantity of hidden layers, the non-supervision layer ascending method can be used to achieve the deep Boltzmann, which means that first RBM is practiced sufficiently, and the weight and amount of move can be fixed; then it is deemed as the input vector of the RBM in the next layer based on its hidden nerve cell. By parity of reasoning, in case one part in the visible layer can be approached, the use is also the belief, which can transmit the error value inversely; and the RBM in the visible layer in the farthest area can be achieved, finally the deep belief network can be obtained, as shown in Figure 1. In essence, the DBN application layer ascending can get the relative optimum initial parameter in RBM practice, thus the network performance can be improved (Yang, 2012). A lot of experiments also verify that the DBM can be used to solve problems in BP neutral network effectively.

![Figure 1. DBN Model](image)

2.3 Algorithm of deep learning system

As for the RBM under learning or the learning completed, the most direct and effective evaluation indicator should be the logarithmic similarity of RBM against the practice data; however, because the algorithm complexity and difficulty are increased due to the normalized molecule, main advantages of RBM can only be evaluated with the similarity method (Zhao and Feng, 2014). The common similarity method is to rebuild the structure error, which means that the practice sample is as the initial state, and the numerical difference between the single Gibbs transformation and the initial data can be obtained through RBM arrangement, and the specific operation steps are as follows:

\[
\text{error} = 0 \\
\text{FOR} \forall v \in S \text{ DO} \\
\quad j = \text{sample}_j \text{ given } k(k, \text{RBM} (W, a, b)) \\
\quad k' = \text{sample}_k \text{ given } j(k, \text{RBM} (W, a, b)) \\
\quad \text{error} = \text{error} + \|k - k'\|
\]

The rebuilt structure error can be used to evaluate the similarity of RBM to the practice sample, but the stability of which is still low; as a whole, such operation method is simple and of low cost, thus it is of great application value in the practice.
3. DEEP LEARNING SYSTEM CONSTRUCTION OF COLLEGE ORAL ENGLISH UNDER CROSS-CULTURAL COMMUNICATION BACKGROUND

3.1 Develop indicators for deep learning system of oral English

The entire level of Chinese students in oral English pronunciation has been promoted obviously in recent years, especially the breakthrough has been achieved in single tone, which attributes to large input in education resource and high enthusiasm of students in learning English in China. Therefore, as for most students learning English, they can master the pronunciation of English words well (Zhang and Wang, 2015). However, the English statement rhythm is the key point in cross-cultural communication capability in real life. Many students still face problems of different degrees after master the pronunciation skills of English words, they can speak each standard syllable, but their English expression can hardly be called the real British accent (Wang and Liu, 2015). These students include the students learning English initially, the students with middle and high level in learning English and also postgraduate students of English major. The rhythm of language pronunciation is the key factor in sentence pattern, and different languages are of obvious characteristics in their own rhythm; the expression of the entire sentence may be rusty and awkward-sounding if these characteristics cannot be mastered.

The entire quality evaluation of phonetic pronunciation usually focuses on the accuracy of phonetic pronunciation, syllable length and rhythm etc. During such process, the pronunciation accuracy is the main indicator for factors and words, however, the English sentence pattern and paragraphs not only focuses on the main content expressed by the pronunciation, the rhythm features of which also affects the real connotation of the entire sentence (Zhu, 2016). As for judgment on intonation of sentences and English paragraphs, the key point is that if students can understand the core factor of the entire sentence accurately, and the relative secondary information, length and quality etc. in the sentence should be considered as well. To be specific, as for the entire pronunciation quality evaluation of the English sentence, the best pronunciation quality should be complete content, clear and smooth, free of obvious errors, the suitable speed and pitch degree, clear rhythm and natural intonation.

3.2 Development process of deep learning system of oral English

The subjective evaluation mainly refers to the scientific and effective evaluation of professional linguists on the quality of oral English pronunciation, and the basic process is divided into three steps. The first step is the listening test. The test voice is compared with the standard voice in memory based on the experiences accumulated in past the languages, thus the difference of each level can be found out. Finally, the gaps of different layers are combined so as to make right comprehensive evaluation on the voice (Ge and Wang, 2016). Generally speaking, the evaluation result of linguists on the phonetic pronunciation can reflects the accuracy of oral pronunciation and the real level of students’ oral English. However, due to the large gap between the knowledge structure and personnel experiences of these linguistics, large opinion deviation may occur in the same language tested. In addition, as the evaluation of oral pronunciation is closely related to the spoken language theory and language theory, and in the meanwhile, it also involves the knowledge of physiology and mentality, which means that the same expert may provide different evaluation result in different stages for the same voice (Lei, 2016). Therefore, the truth of pronunciation quality can be ensured in the subjective evaluation result, but the defect of subjective evaluation may be presented.

The objective evaluation mainly means the automatic evaluation made to the entire quality of oral English pronunciation. The computer software is applied to the objective evaluation on the pronunciation quality of oral English of students, which can avoid the defects in the subjective evaluation, reduce the evaluation error and promote the entire evaluation efficiency. The objective evaluation is subject to the same evaluation standards, which can show the advantages of objective evaluation in facing many spoken English testing tasks. The design of objective evaluation system should comply with the specific procedures provided by English experts for spoken English pronunciation tests.

4. CONCLUSION

In case the network software is used to help students learn college oral English, the speech recognition and its evaluation technology are key points. The most important factor is the voice recognition technology, which plays an important role to the promotion of students’ cross-cultural communication capability; the primary cause
is that the voice recognition is the fundamental and important premise to oral English evaluation. Therefore, the deep learning system is developed for the college oral English based on the cross-cultural background; thus the scientific and effective evaluation result and reference basis can be provided for students’ oral English learning, and students can also recognize their weakness in oral English and take effective measures to make promotion promptly; thus the oral English expression capability of students can be promoted, more excellent professionals can be trained for the cross-cultural communication of China.

ACKNOWLEDGMENTS


REFERENCES

Tan Y.Y. (2017). The study of college English oral teaching based on cultural content, Literature Education, (04), 158.
Wei L. (2012). Research and practice of college English intercultural communication ability training system, Open Education Research, 18(01), 118-123.