

# Research on Second Language Acquisition of Chinese as a Foreign Language under Artificial Intelligence Technology

Yuanyuan Liu<sup>a</sup>

School of Chinese Language and Literature, Anshan Normal University, Anshan 114007, China

<sup>a</sup>nitufenfangnixing@163.com

**Keywords:** Artificial intelligence technology, Chinese as a foreign language, second language acquisition education

**Abstract:** The rapid development of artificial intelligence has brought new development opportunities and challenges to the language teaching of Chinese as a foreign language. As a new type of education, language education APP based on artificial intelligence technology effectively improves the timeliness and interest of language teaching, and supplements and improves the second language education system to a certain extent. Second language acquisition theory integrates multiple disciplines such as linguistics, language education, sociology, and psychology, focusing on studying the mechanisms by which people acquire foreign language abilities. This article combs the relevant academic terminology and research trends in foreign and domestic from the perspective of virtual reality and artificial intelligence, intends to clarify the concept and application breadth of current virtual reality, and provides a basis for its research and application in Chinese as a second language acquisition.

## 1. Introduction

As a new language teaching method, the mobile-assisted language learning app emphasizes interesting elements. With its flexible learning time, refined course content and interesting novel learning models, it has attracted more and more attention. This kind of mobile terminal language learning courses rely on artificial intelligence technology to provide students with independent learning forms. However, many textbooks have the problem of difficulty for beginners to get started. Students with a very low level of second language cannot find the course task orientation that may be unclear, without specific guidance and help, and lose interest in learning prematurely. Through an in-depth observation experience of an artificial intelligence second language education course, the author summarizes the data from practice and uses the "attention hypothesis" theory to provide a basis for adjusting the original teaching model [1].

## 2. Foreign language education and bilingual teaching

From the perspective of language dissemination and application, Indian scholar Kachru proposed three concentric circle theories for the second language in the world, namely inner circle, outer circle (or extension circle), and extended circle (see Figure 1).

The inner circle is a native language user, the outer circle is a second language user, and the extended circle is a foreign language user. The three concentric circles are small from front to back, and partially overlap. Kachru believes that these three circles "show the second language in a cross-cultural, cross-language environment in the form of communication, acquisition patterns, and functional areas." With a small population in the inner circle, the second language is acquired as a mother tongue in a natural way, and its functions and uses are comprehensive. The outer circle has a larger population than the inner circle. The countries to which these populations belong are mainly composed of former British colonies. The second language is mainly obtained through school education and is used in specific functional areas. During the colonial period, this second language skill was still a tool for entering the upper European and American societies and learning Western

cultural traditions. This circle should also include immigrants from non-second language countries living in second language countries, where the second language is their second language. The population in the extension circle is the largest. The second language as a foreign language can be obtained through long-term school education. It plays a role as an auxiliary language in the fields of economy and trade, science and technology, education, academics, and diplomacy. The theory of concentric circles reflects the current situation of second language use, and also reflects the complexity of population changes in the use of second languages around the world. The coincidence between circles and circles fully reflects this point. This theoretical model also shows that the contemporary second language is no longer the "private property" of the second language countries such as Britain and the United States. The second language has appeared or in fact exists a variety of functions with actual social functions and interpersonal communication. , A fully established and reasonable variant in the linguistic sense, such as Indian second language, Singapore second language, Hong Kong second language, South African second language, etc. Contemporary second language can be a carrier of Western culture or a carrier of Oriental culture. As Yan Zhiqiang pointed out: "In a country like India, the cultural connotation of the second language is not only very oriental, but also very Indian. It can reflect the feelings and will of the entire nation more than any other local language.

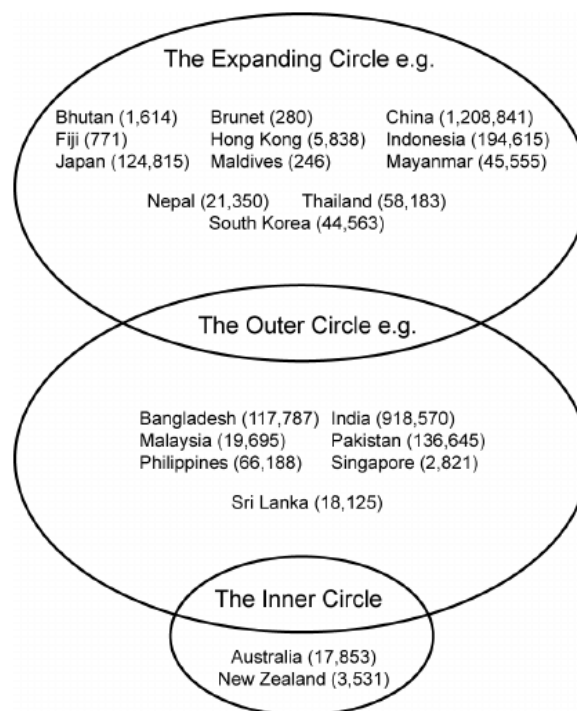


Figure 1. The theory of three concentric circles in the second language of the world

There is an essential difference between bilingual teaching and foreign language education. There are many definitions of bilingual teaching in academic circles. In short, we can say that bilingual teaching refers to the use of two languages for the courses set in the syllabus (except for language classes and foreign language classes). Bilingual teaching is often divided into several different types, mainly including: (1) Immersive bilingual teaching. Refers to the school's teaching in a language different from the students' mother tongue. The Canadian French immersion course often mentioned in the literature is this teaching model. The results of the Canadian French Immersive Bilingual Teaching Course show that through seven or eight years of school education, the French level of those who are native speakers of second language is almost the same as that of students who are native speakers of French. Compared with students who learn French as a second foreign language, their French level far exceeds the latter. (2) Maintenance bilingual teaching. Students are taught in their mother tongue at the beginning of school, and then gradually transition to some subjects to learn in another language prescribed by the school. (3) Transitional bilingual teaching. At the beginning of school, students partially or completely use their mother tongue, and later switch to

teaching in another language prescribed by the school. In the United States, this model is mainly for immigrant students to integrate into the mainstream second-language society as soon as possible, so that students who are not native speakers of the second language and other students have equal opportunities for education and work. In addition, there are of course political reasons for maintaining national unity.

### **3. Experimental research**

#### **3.1 Theoretical assumptions**

Based on the theoretical principles, this experiment uses the three experimental textbooks of "Looking at Pictures"; "Retelling and Short Answer" and "Situational Dialogue" to compare the teaching effect of the original curriculum model to verify the validity of the experimental textbook using the "attention" theory. Two experimental hypotheses are proposed: (1) Experimental teaching improves students' ability to understand course content and autonomous learning ability. (2) Level 1 students can well understand and master the course content. The teaching effect of experimental textbooks was used to verify the guiding role of the "attention hypothesis" theory in second language acquisition [2].

#### **3.2 Research methods**

Experimental scheme: The method of combining quantitative research and qualitative research is used to examine the change of experimental teaching materials compared to the original teaching materials and different learning experiences. The questionnaire and test score data analysis are used to measure the students' understanding ability and autonomous learning before and after experimental teaching. Changes in ability and second language level. The pre-test is conducted two weeks before the launch of the experimental textbook. Due to the uncoordinated class time of the mobile phone students, the post-test is scheduled to be conducted 3 months after the experimental teaching in order to ensure that the research conclusions are true and reliable. Research tools: The questionnaire designed by the author mainly includes students' understanding and opinions on artificial intelligence language teaching; the conformity of the course content with the students' second language proficiency; the "task" element of the textbook design and the subjective attention of the students. There are 7 questions in the category ("views, improvement, conformity" and "intentional attention"). Some of the questions in the questionnaire are a comprehensive summary of the questions asked by students in the teaching and guidance work. They are highly reliable and widely representative. The three functions of "Attention", "Frequency" and "Communication" for understanding ability and autonomous learning ability are measured by 6 questions in the "Intentional Attention" part of the questionnaire. This experiment only focuses on the relationship between attention degree and understanding ability and autonomous learning ability. Intentionally pay attention to this part of data for analysis and research. The questionnaires were distributed 300 points in the program WeChat group at the time point in the plan. Because of the artificial intelligence mobile phone teaching mode, the conditions for submitting questionnaire reward courses were set in order to better recycle the questionnaires. All 300 questionnaires were collected.

#### **3.3 Results**

This chapter organizes the data of the pre- and post-test questionnaires and the results of the second language proficiency test to analyze the changes in the level of students' "intentional attention" and knowledge of the textbooks after three months of original textbooks and three months of experimental textbooks [3].

##### **3.3.1 Student learning data analysis**

The average number of learning days per month is 22.7 days after the test (9.2 days higher than the previous test); the average score (out of 100 points) is a comprehensive evaluation of the course taken by 300 students by the automatic scoring system, and the post-test is 67.35 points (49.28 points higher than the previous test). As shown in Table 1.

Table.1. Student learning data

Time	measurement standard			
	Average learning days		average score	
	Pretest	Posttest	Pretest	Posttest
The first month	18.29	25.09	16.79	55.55
The second month	13.11	22.50	17.35	69.41
The third month	9.15	20.61	20.07	77.10

### 3.3.2 Analysis of students' intentional attention

The comparison methods of the pre- and post-test data in Table 2 and Table 3 are analyzed to verify the first experimental hypothesis. The questionnaire function of the two questions "I will pay attention to gradually improve the correct rate when speaking pictures" and "Attention to actively create a second language learning environment (such as watching an original movie)" is to consider the students' awareness of "intentional attention". 490 points, 420 points; average score 1.63 points, 1.40 points, the highest score. The post-test total score was 830 points and 1010 points; the average score was 2.76 and 3.36 points. The post-test score increased by 102.2% compared with the pre-test score. This result indicates that the "attention" emphasized in the experimental teaching can help understanding, and let students' subjective consciousness be actively related to learning tasks. The interface of the experiment textbook is bright, lively and friendly, and the content design deliberately guides students to pay attention to improve the students' ability to understand the content of the course. The characteristics of the question type are two of the four characteristics of "the best language input in Krashen" "monitoring theory" This is "understandable and interesting and related".

Table.2. Student learning behavior data

Student learning performance	Total score	The average score
When doing "look at pictures", you will pay attention to gradually improve the accuracy	490	1.63
Pay attention to actively create a second language learning environment (such as watching an original movie)	420	1.40
When repeating the short answer, I will practice repeatedly and gradually approach the original sound	390	1.30
Actively retrieve possible errors after finishing homework	370	1.23
Actively pay attention to learning a second language with foreigners	260	0.86
Try to use the second language you learn in communication	270	0.90

Table.3. Second Language Proficiency Test Results

Choose to fill in the blanks	Look at the picture and speak	Recap	Situational dialogue
The average score	The average score	The average score	The average score
25.15	19.79	12.23	5.75

### 3.3.3 Analysis of second language test scores

The second language test score is an objective comprehensive evaluation index, which measures the changes in the teaching effect of the original textbook and the experimental textbook and verifies the second experimental hypothesis. Table 3 and Table 4 show that the average score of each subject in the second language test is 15.73 points in the pre-test and 77.61 points in the post-test, which is significantly higher than the pre-test. This shows that experimental teaching enables level1 students to understand and master the course content well. It is worth mentioning that in the questionnaire, the communicative attributes of "learning a second language with a foreigner" and "deliberate use of the learned" questions currently have the largest and most significant post-test scores. However, in the second language test scores, the "situational dialogue" courses that are homogenous (communicative attributes) with the questionnaire question have the smallest horizontal progress compared with the adjacent courses, and the average score is also the lowest. The gap between subjective and objective

conclusions. Explained by the principles of psychology, the high score of the post-test questionnaire reflects the students' preference for question forms, psychological expectations and confidence in completing communication (output). The test scores more objectively and truly reflect the difficulty index of "output" in second language acquisition. But generally speaking, after receiving three months of experimental teaching, students' intentional attention to the course tasks and their ability to understand have been improved and the range is very large, and the teaching effect is obvious [4].

Table.4. Second Language Proficiency Test Results

Choose to fill in the blanks	Look at the picture and speak	Recap	Situational dialogue
The average score	The average score	The average score	The average score
86.56	81.75	79.90	62.22

#### 4. Research on second language learning software for foreign language majors based on artificial intelligence

This article proposes that after training the DNN-HMM model with all fine standard data and unsupervised data, only the fine standard data is used to modify the DNN-HMM model parameters during the last model parameter update, which can improve the recognition performance. All configurations of this experiment Consistent with the semi-supervised experiment based on perplexity, the acoustic model training uses 3 h fine standard data and unsupervised data training based on perplexity. The language model is generated from the text corresponding to the 3 h fine standard data. 2 different training methods the recognition results are as follows: When all the data are trained together in the DNN model and iterated multiple times, the WER is 56.66%, and when the last iteration uses only fine-scale data to train the DNN-HMM model, the WER is 55.79% [5].

##### 4.1 WFST decoding

This paper proposes a general decoding method based on WFST. WFST is a finite state acceptor (FSA) in which each conversion has input symbols, output symbols, and weights. The input symbol sequence can be obtained through the WFST decoding path and the output symbol sequence can be sent out. The decoding method in this paper represents the CTC label, dictionary and language model as separate WFST. The highly optimized FST library can effectively merge WFST into a single search graph. This article is illustrated in the English context, but has the same applicability to other languages. The construction of a single WFST is described as follows:

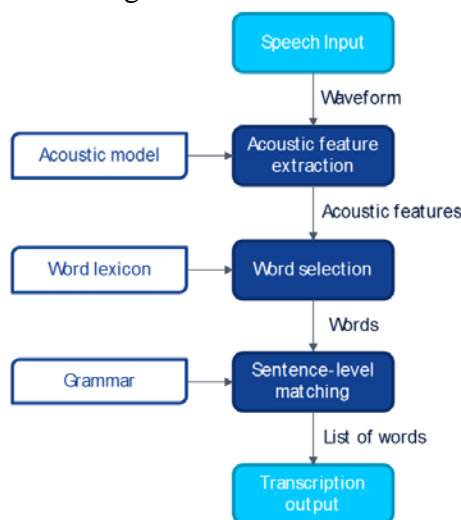


Figure 2. WFST for grammar (language model)

Grammatical WFST encodes the sequence of words allowed in the language and register. Figure 2 shows a WFST example of grammar, which contains two sentences "how are you" and "how is it". The WFST symbol is a word, and the value on the line represents the probability of the language

model, that is, the probability of linking the next word given the previous word. Node 0 is the initial node, and double circle node 4 is the ending node. Using the WFST representation of grammar, CTC decoding can in principle use any language model converted to WFST.

The dictionary's WFST encodes the mapping of dictionary unit sequences to words. According to the labels modeled by RNN, this paper considers two label situations: If the label is a phoneme, the dictionary is the standard word used in the usual hybrid method; when the label is a character, the dictionary contains only the spelling of the word. The main difference between these two cases is that a dictionary containing only spellings of words can be expanded to contain any non-normative terms (OOV). Since phonemes may have multiple meanings, the expansion of phonemes is not easy to achieve. Represent the dictionary's WFST as L. Figures 3 and 4 show examples of constructing these two types of L. In Figure 3, the "<eps>" symbol indicates no input or no output. In Figure 4, the "<space>" symbol indicates that words begin and end with a space character [6].

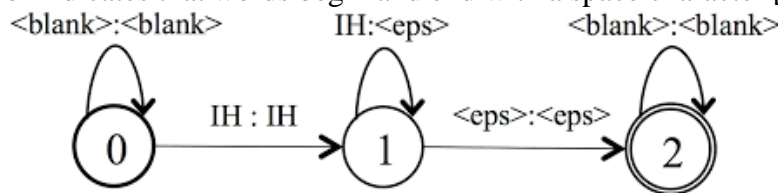


Figure 3. WFST of "is IH Z" in phoneme vocabulary

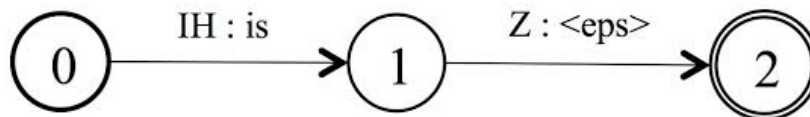


Figure 4. WFST with the word spelling "is"

## 4.2 Experimental research

In order to compare the impact of language models on recognition performance, Table 1 lists the performance when using different precision data to train language models. As can be seen from Table 1, language models have a greater impact on performance. [7] The more data the training language model has, the better the performance the better. However, even if the language model is only trained with 1 h text, the performance of the method of this article is significantly improved, as shown in the last two rows of Table 5.

Table.5. WER comparison of training data of acoustic model and language model in different situations

Acoustic model training data	Language model training data	WER/%
3h precision standard data baseline system	3h fine text	61.17
3h + selected data + selected unsupervised data based on perplexity	3h fine text	56.66
3h + selected data + selected unsupervised data based on perplexity	1h fine text	57.6
1h + carefully selected data + unsupervised data based on perplexity	1h fine text	58.61

Experiments show that whether it is to select data based on the confidence of the decoding result or to select data based on the degree of confusion, speech recognition performance will be improved when resources are scarce. The method of confusion focuses on the selection of data similar to the training set scene, and the filtering is not reliable or Unsupervised annotations that are not related to the collection can better reflect the information of the training set or the conversation scene. If the selected unsupervised data is used to train the acoustic model together with the original fine standard data, the performance of the acoustic model is improved. From the recognition results, we can see that Compared with the DNN system based on confidence, the WER based on the perplexity method decreased by 0.66% [8].

## 5. Conclusion

This article uses the Eesen framework to build an end-to-end ASR system. Eesen uses RNN as the acoustic model and CTC as the training target function. Only the RNN model needs to be trained directly to reduce the complexity of the ASR system calculation. Decoding through WFST can be efficient and effectively combine dictionaries and language models. Due to the open source nature of Eesen, it can be used as a shared benchmark platform for end-to-end ASR research.

## References

- [1] Booth, J., Di Eugenio, B., Cruz, I. F., & Wolfson, O. Robust natural language processing for urban trip planning. *Applied Artificial Intelligence*, Vol.29 (2015) No. 7, p.859-903.
- [2] Osa, T., Pajarinen, J., Neumann, G., Bagnell, J. A., Abbeel, P., & Peters, J. An algorithmic perspective on imitation learning. *Foundations & Trends in Robotics*, Vol.7 (2018) No. 2, p. 1-179.
- [3] Chen, D., & Mak, B. Multi-task learning of deep neural networks for low-resource speech recognition. *IEEE/ACM Transactions on Audio Speech & Language Processing*, Vol. 23 (2015) No. 7, p.1-1.
- [4] Mautner, J., Makmal, A., Manzano, D., Tiersch, M., & Briegel, H. J. Projective simulation for classical learning agents: a comprehensive investigation. *New Generation Computing*, Vol. 33 (2015) No. 1, p.69-114.
- [5] Inkpen, D., Liu, J., Farzindar, A., Kazemi, F., & Ghazi, D. Location detection and disambiguation from twitter messages. *Journal of Intelligent Information Systems*, Vol. 49 (2017) No. 2, p. 237-253.
- [6] Miller, Z. F., & Godfroid, A. Emotions in incidental language learning an individual differences approach. *Studies in Second Language Acquisition*, Vol. 42 (2020) No. 1, p.115-141.
- [7] Dekel, Ofer, Gentile, Claudio, & Sridharan, Karthik. Selective sampling and active learning from single and multiple teachers. *Journal of Machine Learning Research*, Vol. 13 (2016) No. 1, p. 2655-2697.
- [8] Turney, P. D., & Mohammad, S. M. Experiments with three approaches to recognizing lexical entailment. *Natural language engineering*, Vol. 21 (2015) No. 5, p. 437-476.