Over the past 30 years, neural computational models have achieved great success in the field of language learning, bringing about remarkable improvements in the study of individual differences across languages. However, most of the existing research on neural computational models is about the cognitive development of mother tongue, and there is no correlation model of phonological ability, short-term memory ability, and long-term memory ability with individual differences in the research. Therefore, the intervention suggestions in English learning disabilities cannot be targeted. Through neural network simulation and t-test, it can be seen that the factors influencing the overall English performance of learners are not phonological awareness, but rather short-term memory and long-term memory. It can be seen from the simulated individual ANN that the role of working memory is particularly obvious. This conclusion is consistent with neuroscience theories on bilingual control in the brain. Individual ANN can provide a basis for predicting learners’ learning ability and risk analysis of learning disabilities by simulating the learning path of learners, and at the same time providing more accurate clues and directions for subject teachers to carry out personalized teaching.

1. Introduction

Computer model refers to a powerful concept mode that makes use of the capacity for a big and high-speed information processing, sets a certain environment in the computer, implements certain rules or rules in the objective system, and runs at high speed, in order to observe and predict the condition of the objective system. Computational models have become an increasingly common tool for investigating mechanisms of cognitive developmental change [1]. Over the past 30 years, neural computational models have achieved great success in the field of language cognition, bringing about a world-renowned change [2]. For example, used artificial neural networks to explore the network learning ability of four types of verbs, compared the performance of a single-layer perceptron with the multilayer perceptron used by Rumelhart and McClelland, and outlined the performance of several systems. U-shaped learning performance, or the deep systematic learning mode, is a theory of computer learning process based on comprehension and explores the many ways in which multilayer networks capture the characteristics of children’s U-shaped learning; Michael S. C. Thomas and Annette Karmiloff-Smith use computational models to study language developmental disorders in children with Williams syndrome [3]. The developing computational models, especially the artificial neural network (ANN), each of which contains parameters that adjust the performance of the system, and these parameters represent the explanation mechanism for performance changes, so they can be used for cognition and related hypotheses that provide defect mechanisms in language development.

Neural computational models are the core principles that attempt to abstract and mathematically elucidate the basis of information processing in biological nervous systems or their functional components. The application of neural computing models in education is mainly reflected in the field of educational neuroscience. In 1999, the Organization for Economic Cooperation and Development launched the "Learning Science and Brain Science Research" project, which aims to establish a close cooperative relationship between education researchers, educational...
2. Relevant Theoretical Basis

2.1. BP Neural Network Computation Model. Computational models are simulations of a set of processes observed in nature in order to understand those processes and predict the outcomes of natural processes based on a specific set of input parameters [10]. The neural computing model is a computing model with artificial neural network as the core algorithm, the longest-used algorithm in artificial intelligence and machine learning. Artificial neural network (ANN or neural network for short) is a data processing model that was inspired by biological neural networks. As far as neural networks research is concerned, there has not been a unified definition. As an internationally renowned research expert in the field of neural networks, Hecht Nielsen gives a definition of neural network: an artificial neural network is a dynamic system, which is built manually and has a directed graph as its topology, which processes information by responding to continuous or discontinuous input states.

2.1.1. Artificial Neural Network Terminology. The common terms of artificial neural network mainly include neurons, weights, thresholds, activation functions, and learning rules.

1. Neurons: Neurons are the basic units in the neural network, also known as nodes or units. Neurons are connected to each other to form a neural network.

2. Weight: It is also known as connection weight; in biological neurons, it is the connection strength of the cell body connecting the synapses between different neurons.

3. Threshold: It is also known as threshold intensity; it refers to the minimum stimulus intensity required to release a behavioral response. It is represented by \( b \), which is also called bias in mathematical models. The signal sent by the cell body will be sent to other neurons through the axon. Otherwise, it will be in a suppressed state. The input signal calculation formula is

\[
z = \sum_{i=1}^{n} w_i x_i + b.
\]

4. Activation function: It is also known as the transformation function, or the transfer function, which is used to make the artificial neuron generate the output of the neuron based on the input received. It is represented by \( y = (z) \); \( y \) is the output of the neuron.

5. Learning rules: It is a learning mechanism in the neural network. Different neural networks have different learning rules, that is, to use the error between the output of the neural network and the actual output of the corresponding target to learn. The general learning method is mainly: adjust the connection weight to reduce the error and finally make the error not greater than the target error requirement. The commonly used learning rules are as follows:

(1) Hebb’s learning rule: the vector formula and element formula are

\[
\Delta w_j = \eta f(w_j^TX)X,
\]

\[
\Delta w_{ij} = \eta f(w_{ij}^TX)X_i.
\]

(2) Perception learning rules: vector and element expressions are

\[
\Delta w_j = \eta [d_j - \text{sgn}(w_j^TX)]X,
\]

\[
\Delta w_{ij} = \eta [d_{ij} - \text{sgn}(w_{ij}^TX)]X_i.
\]

(3) The delta (Delta) learning rule: the vector formula and the element formula are

\[
\Delta w_j = \eta [d_j - \text{sgn}(w_j^TX)]X,
\]

\[
\Delta w_{ij} = \eta [d_{ij} - \text{sgn}(w_{ij}^TX)]X_i.
\]
**2.1.2. BP Neural Network Algorithm.** In this study, the artificial neural network that will be used in this paper is a three-layer BP neural network (as shown in Figure 1). BP neural network is the error backpropagation algorithm (Error Backpropagation, referred to as BP), which was proposed by scientists such as Rumelhart and McClelland in 1986. Compared with the single-layer perceptron, the main point is that it can solve the problem of linear inseparability.

BPANN includes input layer, hidden layer (referred to as hidden layer), and output layer. The main learning idea is the forward propagation of the input signal and the backpropagation of the error (as shown in Figure 2). In the forward propagation direction, the input signal starts to pass through the hidden layer from the input layer and is processed between layers in the hidden layer and finally reaches the output layer of the network. The state of each layer of neurons in the BP neural network can only play a role in the state of the neurons in the next layer and has no effect on the state of the neurons in the other layers [11]. When the network output of the jth neuron is not equal to the actual output, that is, the expected output, there will be an output error, and the network will go to the backpropagation stage of the error. In this BP neural network, the average squared error is used to calculate the error E, and the calculation formula of the error E is

$$E = \frac{(d - 0)^2}{N} = \frac{\sum_{j=1}^{m} (d_j - O_j)^2}{N}. \tag{6}$$

In formula (6), N is the number of samples and m is the number of neurons in the output layer; the weight adjustment is generally based on the delta learning rule, and the weight adjustment formula is

$$\Delta w_j = \eta(d_j - O_j) f'(net_j)X. \tag{7}$$

**2.1.4. Parameter Selection.** When a task is determined, the number of nodes in the input layer and output layer will be determined accordingly. Therefore, the adjustment of parameters here mainly includes the number of network layers, the number of hidden layer nodes, the selection of activation function, the determination of initial weights, and the setting of learning rate. There are three empirical formulas obtained: Equation (11), the nodes of the hidden layer can be regarded as the adjustment parameters of the structure and function of the brain [13]:

$$\sum_{i=0}^{n} C_M > k, \tag{11}$$

$$M = \sqrt{m + n + a},$$

$$M = \log_2 n.$$
In the above three empirical formulas, $M$ is the number of neurons in the hidden layer, $k$ is the number of selected samples, $m$ is the number of neurons in the input layer, and $n$ is the number of neurons in the output layer, and $a$ is $[0, 10]$. Selection of activation function: both the hidden layer and the output layer are selected Log-S-type functions, and the activation function can be regarded as the adjustment parameter of the neural function [14]. Determination of initial weight: generally, the value of random function is used as initial weight. Common random functions that can be used as initial weight include zeros(), ones(), eye(), rand(), and randn() (see Table 1 shown); here, randn() is selected as the initial function of the weight between the input layer and the hidden layer, and zeros() is selected as the initial function of the weight between the hidden layer and the output layer. The initial weight can be regarded as the initial learning ability. Status tuning parameters are set.

The setting of the learning rate is generally a decimal between $(0.01, 1)$; the smaller the learning rate, the slower the convergence speed of the neural network; that is, the running time will also become longer, and the learning rate can be regarded as the learning speed. Adjust parameters.

### 3. Design of English Teaching Model Based on 3-Layer Neural Network Technology

#### 3.1. Association Model Assumptions
Based on the ANN individual simulation, this paper establishes a hypothetical model of the association between learners’ phonological awareness and memory ability (as shown in Figure 3) and analyzes the interaction of factors that may affect second language learning [15].

In Figure 3, the speech awareness data set is the training set of ANN. Each ANN training set is different and consists of 150 patterns, which are mainly derived from the real data of the participants who participated in the experiment to complete task 1 and the artificial data constructed according to their error characteristics. Data: the input is the correct past tense pronunciation of real data and artificial data and is encoded as 62 bits, and the output is the student’s pronunciation of real data and the past tense pronunciation corresponding to the artificial data and is encoded as 62 bits. The short-term memory data set is the test set of ANN, which consists of 50 patterns. It comes from the real data of the participants who participated in the experiment to complete task 2. The input is the correct and white past tense pronunciation of the real data and is encoded as 62 bits, and the output is the real data [16]. The student’s pronunciation and encoding are 62 bits.

In the hypothesis of the association model proposed in this paper, both the training set and the test set take the pronunciation of the past tense as samples and simulate them with ANN. The transformation of English verbs is from their original form to the past tense. The rule is +ed after the original form of the verb (like → liked, play → played), but there are also many irregular verbs, for example, arbitrary (catch → caught, go → went), consistent (let → let, put → put), and vowel change (come → came, get → got).

#### 3.2. The Basic ANN Model in the Association Model
Taking the BP neural network as the core algorithm of the calculation model, we will use a 3-layer BP neural network in the simulation of each subject’s voice awareness, working
memory, and long-term memory, including an input layer, a hidden layer, and an output layer [17]. For each subject, an ANN is established corresponding to each task in the experiment (voice awareness, working memory, and long-term memory). The 62 nodes correspond to the coding of the past tense of triphones; here we use 19-bit (5-bit suffix) coding scheme; each phonetic symbol corresponds to 19 features in the coding scheme, and the suffix corresponds to 5 bits. Therefore, the coding of each past tense is 62 bits. The difference is that the input layer is the 6-bit encoding of the correct past tense pronunciation of the verb, and the output layer is the 62-bit encoding of the students’ past tense pronunciation collected during the experiment. The encoding scheme is based on the encoding scheme of Plunkett and Marchman (detailed encoding content). See Appendix 2 of this article [18]. The variation of the error with time during the experiment is shown in Figure 4:

### 3.3. Experimental Task Design

#### 3.3.1. Subjects

The subjects selected for this project are 15 university students [19], who have not systematically studied the past tense of verbs (age 18–23, mean = 18.33, standard
3.3. Vocabulary Selection. Choose 50 words for each task. Since task 2 and task 3 are all about memory ability, the long-term memory of task 3 is completed on the basis of the working memory of task 2, so task 2 and task 3 are to choose the same vocabulary; the proportion of words is regular: irregular (arbitrary: consistent: vowel change) = 4:1 (1:2:7); the selected words are all junior high school entrance examination vocabulary of People’s Education Press, and both are 3-phoneme verbs.

3.3.3. Experimental Procedure. Simulation data consists of data generated after learners complete three related tasks. The three related tasks involved here are as follows: task 1, phonological awareness test, task 2, short-term memory test, and task 3, long-term memory test. Before starting task 1, explain the classification rules of the past tense of verbs (Table 2, the construction rules of the past tense of verbs) and the pronunciation of the past tense suffix, and explain the pronunciation rules of the past tense suffix to the students. For the classification rules of the past tense, it will be divided into two major categories: regular and irregular, and irregular verbs are divided into three subcategories: arbitrary, consistent, and vowel change. The rules taught are slightly different. When regular verbs construct the past tense, the suffix has three pronunciations /t/, /d/, /id/. For the specific suffix pronunciation rules, see Table 2: Suffix Pronunciation Rules.

Considering that students cannot remember so many words each time, every 8 words are divided into groups (the last group is 10 words) during the test. The word grouping situation is shown in Figure 5, and the test is carried out according to the grouping of words. The whole process of the experiment was filmed with a camera. Only one or two sets of words for each task were performed, so the experimental front was drawn over a long period of 4 weeks. The specific implementation steps of each task are shown in Figure 5:

During the experiment, after the teacher has taught the students the pronunciation of a group of words, he will ask each student to read each word in the group; that is, student 1 reads all the words in a group, followed by student 2 reading the same word, then student 3, and finally student 15. Since the pronunciation of the three tasks obtained should be implemented to each subject; that is, each subject will have the past tense pronunciation of the three tasks, after the experiment, the pronunciation of the three tasks should be corresponding to each student. Of course, it is unavoidable to encounter some problems during the experiment. The main problems are shown in Table 3: Main problems encountered during the experiment and the solutions adopted.

4. Experimental Results and Their Analysis

4.1. Data Analysis Based on Consistency of Task Completion. The so-called high consistency between tasks means that when the learner completes one task well (or poorly), the other task is also completed well (or poorly); conversely, low consistency between tasks occurs when a learner does one task well and another poorly. For example, in this study, if the three tasks of phonological awareness, working memory, and long-term memory are all completed well, then the learner’s task completion consistency is high; then the learner’s task completion degree is low. The error (Summation of Square Error, SSE) of the three tasks of the tested object is shown in Figure 6. In these figures, the abscissa epoch is the training times, the ordinate is the SSE, the “- -” line in the legend is the SSE of the baseline, the “-” line is the SSE of phonological awareness, the “- - -” line is the SSE of short-term memory, and the “- o” line is the SSE for long-term memory.

It can be seen from Figure 6 that the SSE error values of learners 6, 7, and 8 during epoch periods [100, 200] and [200, 300] did not exceed ε1 and ε2, respectively. To sum up, learners 6, 7, and 8 completed tasks well, and no one belongs to the group of learning disabilities. For students with high completion of learning tasks, they have typical characteristics, either poor in three aspects or strong in three aspects.

4.2. t-Test Analysis. In this section, t-test will be used to examine the effect of learners’ phonological awareness, short-term memory, and long-term memory on learning effect. In this section, the data collected in the experiment is subdivided, and the subdivision method is evolved according to the literature.

Table 4 shows the mean scores (M) and standard deviations (SD) of the two groups of subjects on the three tasks (phonological awareness, working memory, and long-term memory) on words, consonants, vowels, and suffixes, respectively. M is accurate. To the percentile, SD is accurate to the thousandth.
Table 5 is a supplement to Table 4. TTEST-SD is the standard deviation (SD) of the three tasks (phonological awareness, working memory, and long-term memory) on words, consonants, vowels, and suffixes, respectively, on the value of T-TEST, accurate to the percentile; “∼” is approximately equal. In Table 4, it can be seen from the independent data of the t-test that both short-term memory and long-term memory have statistically significant effects on learners’ learning disabilities. In Table 5, it can be seen from the independent data of the t-test that the effects of working memory and long-term memory on learners’ learning disabilities are statistically significant, and in long-term memory, the suffix part is the key effect.

4.3. Proposed Intervention Methods

(1) Because both working memory and bilingual switching involve the DLPFC (Dorsolateral Prefrontal Cortex) area of the brain, for learners with poor working memory, the learning effect can be improved by strengthening the attention training of bilingual switching. For example, the mother tongue interference method proposed in the literature can be used. Carry out DLPFC function training, given a picture; let students find the correct option from the multiple corresponding words on the other side, and the provided options include native language distractors (such as pinyin similar pronunciation).

(2) For long-term memory, this paper proposes two methods: (a) the method of questioning from time to time: during class, teachers can ask students with poor long-term memory to ask questions about what they have learned before; (b) repeating the grammar of the lesson: the teacher can let the students repeat the text from a few days ago, reach the habit of letting students read after class, and ultimately enhance students' long-term memory.

(3) For phonetic awareness, the phonetic training method proposed can be used. For example, a word can be added with phonetic symbols, decreased with phonetic symbols, or its sound modified after phonetic symbols, or synthesized words, although
Table 3: The main problems encountered in the experimental process and the solutions adopted.

<table>
<thead>
<tr>
<th>Problems encountered during the experiment</th>
<th>Solution taken</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1) The teacher habitually corrects the students, which interferes with the normal pronunciation of the students.</td>
<td>The teacher corrects the student before the valid data.</td>
</tr>
<tr>
<td>(2) When collecting data, some students cannot remember the pronunciation of words; other students will remind them in a low voice below, or other students will read what they read, resulting in the same mistakes in the pronunciation of the past tense.</td>
<td>Ask students not to remind them to read what they think they should read, and not to care about the pronunciation of other students.</td>
</tr>
<tr>
<td>(3) The sound of the camera recording is too low or unclear.</td>
<td>Phone recording as backup sound.</td>
</tr>
<tr>
<td>(4) The bell after class interferes with the recording.</td>
<td>Temporarily interrupt the school bell when it rings, and let students read it after the bell.</td>
</tr>
</tbody>
</table>

Figure 6: SSE for learners 6, 7, and 8.

Table 4: t-test analysis of words, consonants, vowels, and suffixes on each task for two types of students.

<table>
<thead>
<tr>
<th>Task</th>
<th>Word score</th>
<th>Consonant score</th>
<th>Vowel score</th>
<th>Suffix score</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>M</td>
<td>SD</td>
<td>M</td>
<td>SD</td>
</tr>
<tr>
<td>Phonetic awareness</td>
<td>Learning disability</td>
<td>3.59</td>
<td>0.734</td>
<td>1.81</td>
</tr>
<tr>
<td></td>
<td>Study normally</td>
<td>3.79</td>
<td>0.495</td>
<td>1.86</td>
</tr>
<tr>
<td>Working memory</td>
<td>Learning disability</td>
<td>*3.16</td>
<td>1.131</td>
<td>*1.74</td>
</tr>
<tr>
<td></td>
<td>Study normally</td>
<td>*3.66</td>
<td>0.626</td>
<td>*1.94</td>
</tr>
<tr>
<td>Long-term memory</td>
<td>Learning disability</td>
<td>*2.80</td>
<td>1.242</td>
<td>1.64</td>
</tr>
<tr>
<td></td>
<td>Study normally</td>
<td>*3.18</td>
<td>1.258</td>
<td>1.70</td>
</tr>
</tbody>
</table>

Note: * p < 0.05, *0.05 < p < 0.1.

Table 5: t-test analysis of SD on each task of words, consonants, vowels, and suffixes of two types of students.

<table>
<thead>
<tr>
<th>Task</th>
<th>Subject</th>
<th>Word score TTEST-SD</th>
<th>Consonant score TTEST-SD</th>
<th>Vowel score TTEST-SD</th>
<th>Suffix score TTEST-SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Phonetic awareness</td>
<td>Learning disability</td>
<td>0.125</td>
<td>0.219</td>
<td>0.583</td>
<td>0.192</td>
</tr>
<tr>
<td></td>
<td>Study normally</td>
<td>0.000</td>
<td>0.000</td>
<td>0.003</td>
<td>0.000</td>
</tr>
<tr>
<td>Working memory</td>
<td>Learning disability</td>
<td>0.000</td>
<td>0.000</td>
<td>0.003</td>
<td>0.000</td>
</tr>
<tr>
<td></td>
<td>Study normally</td>
<td>0.910</td>
<td>0.844</td>
<td>0.206</td>
<td>0.045</td>
</tr>
</tbody>
</table>

Note: * p < 0.05, *0.05 < p < 0.1.
adding or subtracting prefixes and stems can accomplish the same result. Of course, as mentioned above, the factors that cause learners’ learning disabilities may also be compounded, so educators need to design compound teaching strategies that better meet the characteristics of learners’ disabilities.

5. Conclusion

Neural computational models play an important role not only in the field of language acquisition of native speakers, but also in English education and teaching (such as the transfer of English teaching, subject teaching, semantic analysis), and have made breakthroughs in the field of education. However, there is no correlation model involving individual differences in speech ability, short-term memory ability, and long-term memory ability in the existing research. Therefore, intervention recommendations in second language learning disabilities cannot be targeted. The main research purpose of this paper is to establish and verify the association hypothesis model by using the individual ANN simulation and, based on the corresponding data analysis, to provide corresponding intervention suggestions for students with English learning disabilities as their follow-up English learning guidance.

Data Availability

The labeled data set used to support the findings of this study is available from the corresponding author upon request.

Conflicts of Interest

The author declares that there are no conflicts of interest.

Acknowledgments

This work was supported by the Jiangsu Vocational College of Medicine.

References