

# Using an ANN based computational model to simulate and evaluate Chinese students' individualized cognitive abilities which are important to their English acquisition

J. Yang, M. S. C. Thomas, and X. Liu

**Abstract**—Experience study shows the different experiment plans and scripts in the second language learning researches will diversify the conclusions. More than that, the lacking of the simulation tool also has prevented the researchers from investigating the individualized patterns that cognitive abilities may form to a single learner. Therefore, it is very necessary to evaluate and simulate the cognitive differences that are important to English(L2) learning under a same standard scale. In this paper, we propose an ANN computational model based method to simulate individuals' learning trajectories related to three cognitive abilities. Our experiment result shows, the ANN computational model not only can accurately reflect the participant students' individualized differences, but also can evaluate the compound effects of three cognitive related factors under a same scale. Based on the reliable results produced by this computational model, English teachers are suggested to use this ANN computational model based method in their second language teaching activities to predict the learners' future possible overall English competences based on a virtual base line.

**Index Terms**—cognitive ability, short-term memory, long-term memory, phonological awareness, artificial neural network, computational model

## 1 INTRODUCTION AND RESEARCH BACKGROUND

Individual factors that influence language learning include many cognitive individual differences[1], such as cognitive abilities refers to understanding, perception[2] and memory[3]. Working memory capacity is one of the important cognitive related predictors of success in second language learning(L2)[4], many related studies have testified its importance in English(L2) learning [5-8]. Another important cognitive related factor that matters to English learning is phonological awareness[9-13]. Wagner's research[14] concludes that phonological awareness and short-term verbal memory can be taken as two primary phonological processing skills of second language acquisition. Although experience studies show positive correlations between different English language competences and phonological awareness[9-13], and short-term memory[5-8], the investigating of the compound effects regard to the cognitive factors that contribute to the second language learning is also a long standing research topic[15-17].

However, the results about the degree that different cognitive factors' compound effectiveness to English competences are diverse because of the different experiment

plans that are adopted in the related experiments, e.g., scripts being chosen in the experiments [18] make the research across scripts is very important[19]. Besides the script problem being mentioned above, the missing of a standard evaluation mechanism and instrument is another open problem in the related researching area. More than that, the lacking of the simulation tool also has prevented the researchers from investigating the individualized patterns that cognitive abilities may form to a single learner. Besides the evaluation and simulation requirements, longitudinal influences to language learning brought by the cognitive factors is another important facet to reveal their truly effectiveness to second language learning[6, 9, 11-13, 16, 17, 20, 21]. Therefore, to develop a method that can allow the related experiments being designed and evaluated under a same scale is very necessary.

In general, the new method applied in the English(L2) learning researches need to satisfy the following requirements: the ability to simulate longitudinal trajectories, the ability to simulate the compound effects of the cognitive factors under a same evaluation scale, and the ability to accurately reflect learners' individualized differences.

Based on this research background, computational model would be an optimal choice because it not only reveals the common features shared by the individuals within a same class, but also can simulate the minor differences between individuals. One of the main purposes of using computational models in cognitive related researches is to

• J. Yang is with the Department of Computer Science, Sichuan Normal University, China, E-mail: [jkxy\\_yj@sicnu.edu.cn](mailto:jkxy_yj@sicnu.edu.cn).

• M. S. C. Thomas is with the Developmental Neurocognition Lab, Department of Psychological Sciences, Birkbeck, University of London, UK, E-mail: [m.thomas@bbk.ac.uk](mailto:m.thomas@bbk.ac.uk).

• X. Liu is with the Department of Computer Science, Sichuan Normal University, China, E-mail: [jkxy\\_liux@sicnu.edu.cn](mailto:jkxy_liux@sicnu.edu.cn).

use them as a tool to understand human's learning processes[22], and the computational models of development, particularly those employing artificial neural networks (ANNs), have provided hypotheses about the mechanistic bases of language development[23] and language deficits[22]. Different from classical generative approaches[24, 25] which characterize language in terms of a domain-specific form of knowledge representation called grammar[26], connectionist computational models try to transfer the human's learning behaviors into mapping problems which take grammars as the characterizations of some aspects of the behavior [27, 28]. Building up the correlations between input and output based on the experience (training data) in a black-box way is the basic working mechanism of a connectionist computational model.

In this paper, we propose an ANN computational model based method to simulate individuals' learning trajectories related to their cognitive abilities according to limited discrete data. Our experiment result shows, the ANN computational model can accurately reflect the participant students' individualized differences, evaluate the compound effects of three cognitive related factors under a same scale, and can reflect the significant features among different groups which are consistent with the conclusions of the experience studies. Based on the reliable results produced by this computational model, English teachers are suggested to use this ANN computational model based method in their second language teaching activities to predict the learners' future possible overall English competences based on a virtual base line.

## 2 METHOD

### 2.1 Verb and past tense dataset

Past tense's learning is a commonly used base model for computational models to provide hypotheses of language related cognitive processes[29]. For a long time, verb and past tense has been taken to be a paradigmatic linguistic subsystem exhibiting fundamental properties of language[26]. In this paper, Plunkett & marchman' 19 binary phonological featured coding mechanism[30] is chosen to represent the verbs and their past tenses. Verbs being used in the experiments are 3-phoneme ones with which each phoneme is encoded in 19 binary bits. The corresponding meaning of those 19 binary phonological features can be described as follows: sonorant, consonantal, continuant, voiced, labial, anterior, +coronal, back, strident, nasal, lateral, -coronal, high, central, low, rounded, tense, diphthong. However, Plunkett & marchman's original phonological coding mechanism cannot cover all the phonemes appeared in the 3-phoneme verbs, 6 extra phonemes are added in our experiment, besides that, a 'Null' coding is added to cope with a specific situation. The new added phonemes are: /ə:/, /iə/, /tr/, /ts/, /dr/ and /dz/, and the complete coding mechanism of the phonemes is listed in Appendix A.

Actually, there are only 3 basic forms for those 3-phoneme verbs, they are CVC, CCV and VCC, where 'C' refers to

consonant pronunciation and 'V' refers to vowel pronunciation. The suffix of the past tense are encoded in 5 bits for different kinds of the past tenses: regular, identical irregular, vowel change irregular and arbitrary irregular. Data sets used in this paper are composed by 50 real verbs and their past tenses, the correct mappings from verbs to their past tenses are considered as training data sets, which is used to simulate a virtual learning process of a student. The real mappings of the individuals are used as testing data sets to evaluate their performance differences based on a same benchmark(virtual) learning trajectory which is simulated by a trained artificial neural network.

### 2.2 Designing and implementing the three cognitive ability related experiment tasks

As stated before, the computational model based method being proposed has to equip with the abilities to satisfy the latent demands required by the second language acquisition research. Therefore, a testing that can be executed both by the computational model and the participant students is designed in this part to check whether the proposed method is practical. The testing includes 3 cognitive related tasks: phonological awareness related task (task 1), short-term memory ability related task (task 2) and long term memory ability related task (task 3).

In order to evaluate the cognitive related differences under a same evaluation scale, the verbs and past tenses involved in all 3 tasks follow a same encoding plan, and 50 3-phoneme-verbs and their past tenses compose to a data set for each task. The details of the three tasks are described as follow:

#### (1)Phonological awareness related task (Task 1)

Phonological awareness related task is designed to reflect the basic phonological awareness ability about English(L2) of the students whose native language(Chinese)'s phonological feature space is different from English. This task collects the students' basic phonological awareness performance about English. In this task, teacher only taught students the new phonemes' pronunciation instead of teaching them the past tense's construction rules. For example, before presenting the past tenses which have 'ed' suffix to students, the pronunciation '/id/' will be taught to students first. If the past tense of the given verb is in an arbitrary or vowel-changed way, new phonemes appeared in the past tense will be primarily taught by the teacher. The coding details about the verbs and past tenses being used in this task can be found in Appendix B. 4 steps are involved to implement this task:

- a. English teacher helps the participant students to correctly pronounce the given 8 verbs(last time is 10 verbs);
- b. English teacher teaches the participant students the new pronunciations which will appear in the past tenses of the verbs aforementioned in step a;
- c. Requiring participant students to pronounce the past tenses of the verbs;
- d. Recording each participant students' pronunciations about each verb's past tense.

### (2) Short-term memory ability related task (Task 2)

This task is mainly used to evaluate participant students' short-memory ability related to English pronunciation, 50 three-phoneme verbs will be chosen to compose the data set, and the coding details about the verbs and past tense being used in this task can be found in Appendix C. 3 steps are involved to implement this task:

- English teacher helps participant students to correctly pronounce the given 8 verbs(last time is 10 verbs) and their past tenses;
- Presenting participant student with the 8 verbs (last time is 10 verbs) at a time and requiring them to pronounce their past tenses;
- Recording each participant students' pronunciations of the each verb's past tense.

### (3) Long-term memory ability related task (Task 3)

Long-term memory ability related task is used to evaluate participant students' long-memory ability regard to English pronunciation, and it will be implemented based on short-term memory ability related task, and would share a same data set with the short-term memory ability related task. But the data set used in this task will be one week delay compared with the data sets used in short-term memory ability related task. 2 steps are involved to implement this task:

- Presenting the verbs which were used in last week's short-term memory ability related task to the participant students and requiring them to pronounce those verbs' past tenses;
- Recording each participant students' pronunciations about each verb's past tense.

Those three tasks are repeatedly executive for 7 weeks, and the specific implementing of the 3 tasks scattered on the time line is illustrated in Figure 1. In order to be consistent with the restraints requested for testing short-memory ability, each time only 8 verbs and past tenses are involved(last time involves 10 verbs).

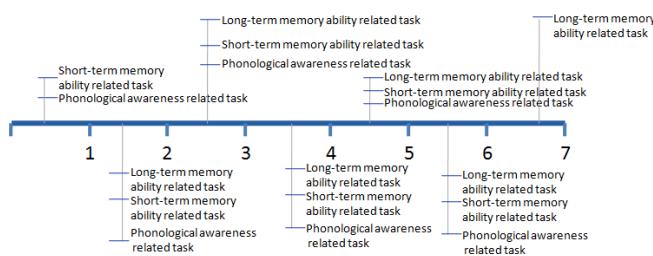


Fig. 1. Three tasks' executive timeline

### 2.3 A simulating model based on three artificial neural networks

As illustrated in Figure 2, 3 artificial neural networks (ANNs) will be constructed to simulate a virtual learning process of an individual of the 3 cognitive ability related tasks, and at the same time to evaluate different learners' learning performances under a same scale. The training data set for each ANN is composed by 50 three-phoneme-

verbs and their past tenses.

ANN used to simulate the process of task 1 is a 62\*40\*62 three layered one which has 62 neurons on the visible layer, 40 neurons on the hidden layer and 62 neurons on the output layer. ANNs used to simulate the process of task 2 and 3 are in the form of 57\*30\*62, where number 57, 30, 62 correspond to 57 neurons on the visible layer, 30 neurons on the hidden layer, and 62 neurons on the output layer.

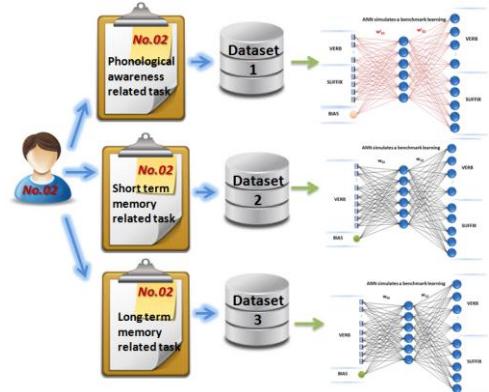


Fig. 2. A sample student numbered 2 is evaluated by 3 trained artificial neural networks

We use 100 epochs to simulate each week, and the training data sets at the beginning of each  $100*i+1$ 'th (where  $i \in [0,5]$ ) epoch are the pre-defined 8(last time is 10) verbs and their past-tenses. After successfully trained the ANNs, participants' actual performance on those tasks will be input into those ANNs as testing data sets to evaluate the simulation abilities of the ANN based computational model.

### 2.4 Data generating and collecting

15 students are carefully selected out from over 200 grade one students in Zongbei experimental middle school, Chengdu, Sichuan. The participant students include 9 girls and 6 boys (age:12-13, mean=12.33, sd=0.471), all of them have not learned the rules of generating past tenses from verbs before, but are ready to learn, all of them performs normal on their first language. 5 of them are evaluated by their English teacher as learning disorder students (whose overall English competences are extremely poor, and are labeled as individual 02, 04, 05, 11 and 12), 5 are evaluated as normal (labeled as individual 01, 08, 09, 10 and 15) and the rest 5 are evaluated as excellent students (labeled as individual 03, 06, 07, 13 and 14). As stated before, there is an extra virtual student who has a convergent learning trajectory. This virtual student is considered as a base line to provide a benchmark scale for all participants to compare and simulate their individualized cognitive abilities. The lines labeled as 'Base\_line' which converges best (with the smallest root mean squared error) in the following figures represent the performances of this virtual student.

In Figure 3-5, the participant student  $i$ ' performances of

each task at each learning epoch  $m$  is computed through the following formula, which actually is a relative learning performance compared with a virtual learner who is allowed to make mistakes during his/her learning process. In Formula (1),  $S_i$  refers to the student  $i$ 's pronunciation on each testing verbs and their past tense, and  $SV$  refers to the training data set(verbs and the corresponding past tenses used in the experiment),  $|S_i|$  refers to the number of the patterns that ANN is processing at that epoch.  $Net_3^m(\cdot)$  is the sigmoid output of the final layer of an ANN at epoch  $m$ ,  $w_j^m$  refers to the weight matrix of the ANN on layer  $j$  at epoch  $m$  based on the training data set  $SV$ .  $|S_i| = |SV|$  is different at different learning phase,  $|S_i| = 8 * (\lfloor m/100 \rfloor + 1)$ , if  $m \in [1,500]$ , otherwise,  $|S_i| = 50$ .

$$RMSE_i^m = \frac{1}{|S_i|} \sqrt{\sum \|Net_3^m(S_i) - Net_3^m(SV)\|^2} \quad (1)$$

$$Net_j^m(Data) = \begin{cases} \text{sigm}(Data, w_1^m), & j = 1 \\ \text{sigm}(Net_{j-1}^m, w_j^m), & j > 1 \end{cases} \quad (2)$$

According to formula (1) and (2), the output of the ANNs regard to each individual is actually the standard deviation of the benchmark line.

### 3 RESULTS OF THE EXPERIMENT AND THE RELATED DATA ANALYSIS

#### 3.1 Simulating participants' overall longitudinal performances on 3 tasks

Figure 3 to 5 are the participant students' current and expected future performances regard to phonological awareness related task, Short-term memory ability related task and Long-term memory ability related task, respectively. The first 500 epochs simulate the experimental process of the students while the last 400 epochs simulate their future possible performances. We can tell from those three charts that almost all testing students' basic phonological awareness ability will be convergent, while their short-term memory related ability and their long-term memory ability diverge in the future. We can observe participant students' differences in those overall performance charts based on the benchmark line, but we still don't know which kind of ability will make what influences to learners to which degree, and what those individualized simulations' can do. These questions will be discussed in the following section.

#### 3.2 Testing the restoring ability of the ANN computational model to the individual cognitive differences

In order to testify whether the ANN computational model can effectively restore the individual cognitive differences, in this section we conduct an ANOVA analysis both on participants' relative RMSE performances (deviation of the benchmark line) and their absolute RMSE (deviation of the correct verbs and their corresponding past tenses) performances, details are presented in Table 1 and

2.

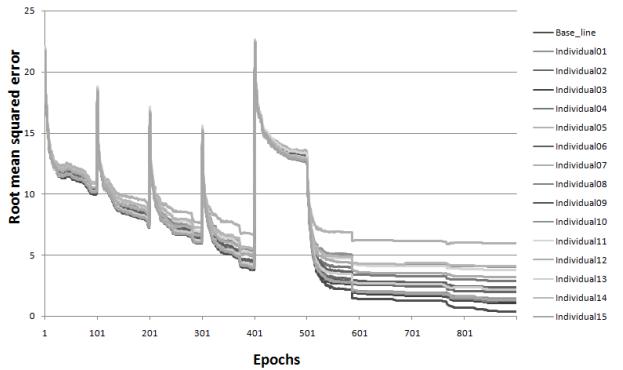


Fig. 3. Individualized performance on Task 1

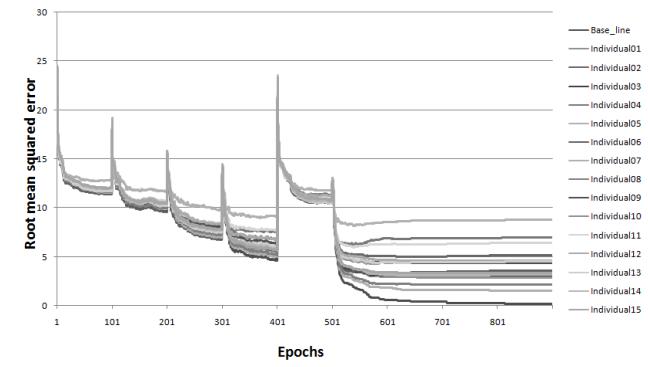


Fig. 4. Individualized performance on Task 2

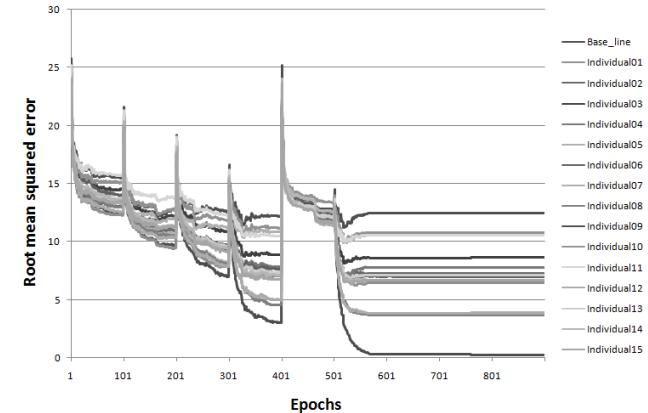


Fig. 5. Individualized performance on Task 3

ANOVA comparisons conducted in ANN computational model(Table 1) and traditional experimental method(Table 2) show that the significant results are highly similar, except the appearing time points in Table 2 are in the current while them in Table 1 are in the future. The delay of the appearing of the significant differences in ANN can be explained by formula (1), the benchmark line is a virtual learner who is allowed to make some mistakes during his/her learning process but finally will be convergent to an error-free state. This simulated process is more compliant with a truly learning process. Therefore, the individuals' deviation range of the benchmark line is narrower than the deviation range of the correct answer. But, with the performance of the virtual learner becomes error-free, the differences are appeared

among different groups because the  $Net_3^m(SV)$  approaches to 0.

TABLE 1  
ANOVA comparison of the participants' relative RMSEs of each experimental phase

Task	Disorder	Normal		Excellent		Sig.	$\eta^2$
		M	SD	M	SD		
1st week	T1	21.8	.03	21.84	.02	.21.85	.01
	T2	24.38	.06	24.35	.01	.24.42	.01
2nd week	T1	18.48	.06	18.53	.02	.18.52	.01
	T2	18.92	.03	18.89	.01	.18.82	.01
3rd week	T3	24.22	.33	24.53	.57	.24.42	.19
	T1	16.77	.06	16.64	.00	.16.67	.01
4th week	T2	15.49	.05	15.62	.03	.15.54	.02
	T3	20.22	.45	20.59	.45	.18.23	.36
5th week	T1	15.22	.05	15.09	.01	.15.17	.01
	T2	14.19	.04	14.27	.01	.14.07	.02
6th week	T3	17.8	.67	18.22	.37	.18.33	.07
	T1	22.42	.04	22.49	.00	.22.48	.01
7th week	T2	23.43	.01	23.35	.01	.23.43	.00
	T3	15.44	.21	15.78	.29	.15.84	.16
8th week	T1	12.52	.01	12.41	.03	.12.49	.03
	T2	12.84	.02	12.88	.01	.12.64	.02
9th week	T3	23.66	.06	24.14	.36	.23.88	.05
	T1	3.98	1.89	2.44	1.05	.2.75	.96
10th week	T2	6.27	2.39	3.47	.88	.3.19	.88
	T3	13.17	.23	13.73	.32	.13.59	.16
Overall	T1	3.87	1.96	2.32	1.07	.2.66	1.12
	T2	6.29	2.66	3.46	.99	.3.13	1.06
average	T3	8.01	7.75	8.02	12.57	.6.66	2.98
	T1	13.19	.18	12.64	.17	.12.77	.16
Overall	T2	14.24	.22	13.31	.11	.13.16	.09
	T3	15.39	1.13	15.58	2.27	.15.19	.39

TABLE 2  
ANOVA comparison of the participants' absolute RMSEs of each experimental phase

Task	Disorder	Normal		Excellent		Sig.	$\eta^2$
		M	SD	M	SD		
1st week	T1	2.92	2.69	2.05	.84	.2.25	.61
	T2	4.82	15.59	4.23	8.06	.4.24	.645
2nd week	T1	2.61	7.59	1.92	2.42	.2.79	.639
	T2	4.49	12	2.59	7.15	.1.58	.82
3rd week	T3	6.64	24.52	4.01	1.86	.3.73	16.3
	T1	2.67	8.08	.67	.84	.14	.1
4th week	T2	4.81	5.63	3.45	1.36	.2.97	.86
	T3	5.86	5.29	6.02	22.64	.4.01	10.06
5th week	T1	1.83	.43	.54	1.15	.2.33	.516
	T2	8.67	14.29	1.27	.6	.2.73	.656
6th week	T3	7.51	8.27	8.16	25.69	.6.01	8.23
	T1	3.38	9.85	1.82	10.1	.35	.62
7th week	T2	4.9	4.64	2.88	2.73	.1.95	.36
	T3	6.46	14.92	7.84	28.35	.4.9	16.95
Overall	T1	4.33	5.39	1.66	1.15	.2.06	.2.63
	T2	6.31	10.82	3.24	4.67	.3.12	.3.76
average	T3	8.29	13.84	9.45	12.53	.8.26	.3.56
	T1	10.36	9.72	8.89	24.53	.7.78	14.69
Overall	T2	2.96	2.45	1.44	1.62	.1.65	.71
	T3	5.69	2.71	2.94	.85	.2.76	.91

\*Sig.<.05,  $\eta^2=.01$  (small effect);  $\eta^2=.06$  (medium effect);  $\eta^2=.14$  (large effect)

Table 1 and Table 2 indicate that the differences between learning disorder, normal and excellent learners which appear at 4th and 5th week during the experimental process can be simulated by the ANN's outputs at the future epochs. If the contents in Table 1 and Table 2 are not persuasive enough, the following Bonferroni Post-hoc analysis on the significant results (Table 3 and Table 4) can provide more convincing evidences. Both in Table 3 and Table 4, disorder learners' mean RMSE is significantly higher than normal learners and excellent learners. Table 1-4 indicates that if we allow a machine learning process being taken as a benchmark line to reflect the differences of the individuals, the results produced by ANN is as reliable as the results concluded from the traditional methods.

If we plot the individual 02, 04, 05, 11, 12 (labeled as learning disorder)'s and 03, 06, 07, 13,14(labeled as excellent)'s ANN trajectories about task 2 into two charts (figure 7a and 7b), respectively, we can get the following

conclusion:

*If the learning trajectory of an individual's short-term memory phonological related ability is above the average line, that individual may be a learner with poor English(L2) overall competence.*

This conclusion is consistent with the results concluded from most of the English(L2) acquisition researches[5-8], and based on this reliable result simulated by the ANN computational model, we can further infer some more detailed implications from Figure 3, 4, and 5:

*If the learning trajectory of an individual's short-term memory phonological related ability diverges instead of converges to the virtual student's simulating learning trajectory in the future, that individual may be a learner with poor English(L2) overall competence.*

This implication suggests that teachers can use this ANN computational model based method to predict the learners' future possible overall English competences with a partial data set collected at the beginning stage of the teaching period, and can predict students' differences based on the benchmark line.

TABLE 3  
Bonferroni comparison of the absolute RMSEs

Comparisons	Mean difference	Std. error	95% CI		
			Lower bound	Upper bound	
4th week	D vs. N	.7.41*	.1.69	.2.7	12.11
	D vs. E	.5.94*	.1.69	.1.23	10.64
	N vs. E	-.1.46	.1.69	-.6.16	3.24
5th week	D vs. N	.2.01	.1.01	-.79	4.83
	D vs. E	.2.95*	.1.01	.13	5.76
	N vs. E	.93	.1.01	-.1.88	3.74
Overall	D vs. N	.2.72*	.77	.57	4.87
	D vs. E	.2.90*	.77	.75	5.05
	N vs. E	-.18	.77	-.1.97	2.32

TABLE 4  
Bonferroni comparison of the relative RMSEs

Comparisons	Mean difference	Std. error	95% CI		
			Lower bound	Upper bound	
7th week	D vs. N	.2.81*	.74	.74	4.88
	D vs. E	.3.08*	.74	1.01	5.15
	N vs. E	.27	.74	-.1.79	2.34
8th week	D vs. N	.2.83*	.79	.63	5.04
	D vs. E	.3.16*	.79	.96	5.37
	N vs. E	.33	.79	-.1.87	2.53
9th week	D vs. N	.2.89*	.79	.70	5.09
	D vs. E	.3.17*	.79	.97	5.36
	N vs. E	.27	.79	-.1.93	2.46

### 3.3 ANN computational model's simulation ability to the participants' individualized trajectories

In this section, we will investigate the ANN computational model's simulation ability to the participants' individualized trajectories of 3 cognitive tasks, the relative ANN RMSE results of each individual is illustrated in Figure 8. Under a same encoding and script mechanism, although most participants' long-term memory performances are obviously poorer than their phonological awareness and short-term memory performances, there still exists some exceptions. The computational model

clearly reflects the longitudinal trajectories of each participant student on 3 cognitive ability related tasks.

Although, no evidence suggests that the long-term memory capacity is matter to English(L2) learning, it still cannot deny that long-term memory capacity could have some influence to some student's second language acquisition. By using the ANN simulation method proposed in this paper, the exploring of the diversified possible patterns of individuals' second language acquisition is feasible. As illustrated in Figure 8, although individual 02 and 12 are both evaluated as learning disorder, two months after we finished the experiment, student 12's overall English competence is much better than 02. The individual's simulation trajectories may provide more hypotheses and explanations to diversified English(L2) learning patterns.

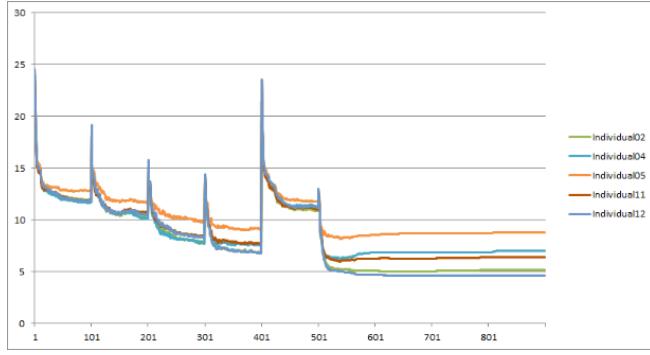


Fig. 7a. Disorder participants' performances on task 2

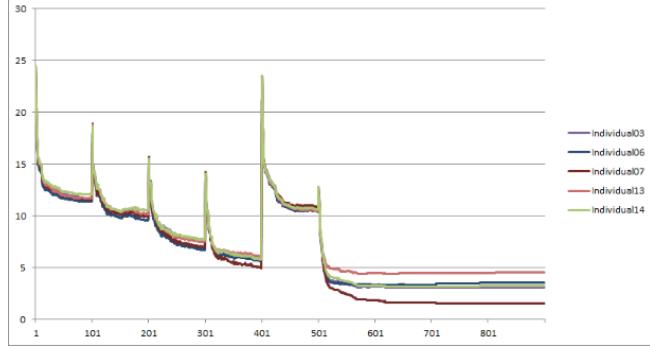


Fig. 7b. Excellent participants' performances on task 2

## 4 CONCLUSIONS

In this paper, we proposed an ANN computational model based method to evaluate and simulate the cognitive abilities of the learners which are important to their English(L2) acquisition. The ANN computational model not only equip with the ability to simulate longitudinal trajectories, the ability to simulate the compound effects of the cognitive factors under a same evaluation scale, but also equip with the ability to accurately reflect learners' individualized differences.

The experiment results show that the ANN has a good restoring ability to the significant features among differ-

ent clustered (learning disorder, normal and excellent) individuals. Based on the reliable conclusions rebuilt by the ANN computational model, another practical using of the ANN computational model is being suggested in the related English(L2) teaching activities, that is to predict the students' future possible overall English competences and students' differences based on the benchmark line. More than filtered out the cognitive features which are

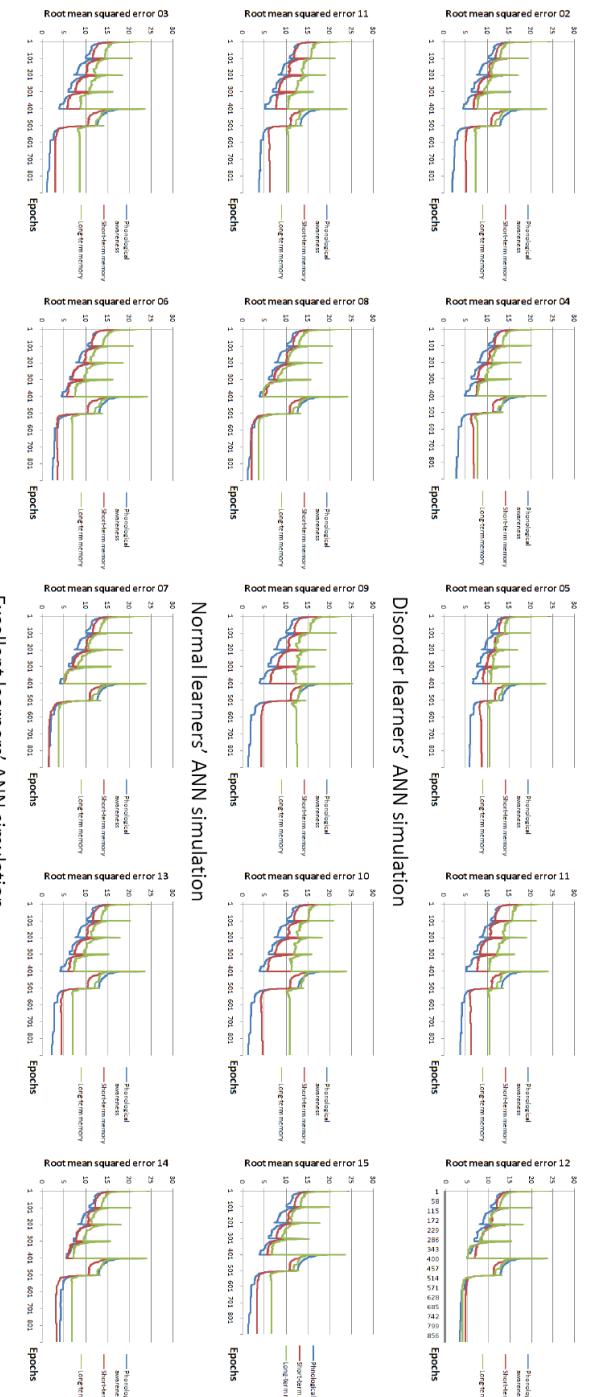


Figure 8. ANN simulations of the participant students' learning trajectories regard to 3 cognitive tasks

sensitive to English acquisition, ANN computational model also can accurately reflect the cognitive factors' compound effects to different individual's English acquisition under a same encoding mechanism, which makes

the providing of new hypotheses about personalized English acquisition possible.

Of cause, the proposed computational model based method also has its limitation in simulating other kinds of cognitive factors because not all cognitive abilities can be modded by verbs and their past tenses. If we want to evaluate and investigate individual differences covered more kinds of the cognitive factors related to second language acquisition, new methods need to be explored.

## ACKNOWLEDGMENT

This research is supported by the National Natural Science Foundation of China(61402309) and UK ESRC grant RES-062-23-2721.

## REFERENCES

- [1] Gardner, R.C., Social Psychology and Second Language Learning: The Role of Attitudes and Motivation. 1985: Edward Arnold.
- [2] Yang, J., et al, Dynamic learning style prediction method based on a pattern recognition technique. *IEEE Transactions on Learning Technologies*, 2014. 7(2): p. 165-177.
- [3] Chrysafiadi, K. and M. Virvou, Student modeling approaches: A literature review for the last decade. *Expert Systems with Applications*, 2013. 40(11): p. 4715-4729.
- [4] Sawyer, M., Ranta, L., Aptitude, individual differences and instructional design, in *Cognition and Second Language Instruction*(pp.319-353), P. Robinson, Editor. 2001, Cambridge University Press.
- [5] Papagno, C. and G. Vallar, Verbal short-term memory and vocabulary learning in polyglots. *Quarterly Journal of Experimental Psychology A Human Experimental Psychology*, 1995. 48(1): p. 98.
- [6] Service, E. and V. Kohonen, Is the Relation between Phonological Memory and Foreign Language Learning Accounted for by Vocabulary Acquisition? *Applied Psycholinguistics*, 1995. 16(2): p. 155-172.
- [7] Speciale, G. and N.C. Ellis, Phonological sequence learning and short-term store capacity determine second language vocabulary acquisition. *Applied Psycholinguistics*, 2004. 25(2): p. 293-321.
- [8] Kormos, J. and A. Safar, Phonological Short-Term Memory, Working Memory and Foreign Language Performance in Intensive Language Learning. *Bilingualism Language & Cognition*, 2008. 11(2): p. 261-271.
- [9] Comeau, L., et al, A longitudinal study of phonological processing skills in children learning to read in a second language. *Journal of Educational Psychology*, 1999. 91(1): p. 29-43.
- [10] Cunningham, A.E., Explicit versus implicit instruction in phonemic awareness. *Journal of Experimental Child Psychology*, 1990. 50(3): p. 429-444.
- [11] Ho, S.H. and P. Bryant, Phonological skills are important in learning to read Chinese. *Developmental Psychology*, 1997. 33(6): p. 946-51.
- [12] Huang, H.S. and J.R. Hanley, A Longitudinal Study of Phonological Awareness, Visual Skills, and Chinese Reading Acquisition among First-graders in Taiwan. *International Journal of Behavioral Development*, 1997. 20(2): p. 249-268.
- [13] Gottardo, A., et al, The impact of language on the relationships between phonological awareness and word reading in different orthographies: A test of the psycholinguistic grain size theory in bilinguals. *Applied Psycholinguistics*, 2016. 37(5): p. 1083-1115.
- [14] Wagner, R.K. and J.K. Torgesen, The nature of phonological processing and its causal role in the acquisition of reading skills. *Psychological Bulletin*, 1987. 101(2): p. 192-212.
- [15] Gottardo, A., et al, Factors related to English reading performance in children with Chinese as a first language: More evidence of cross-language transfer of phonological processing. *Journal of Educational Psychology*, 2001. 93(3): p. 530-542.
- [16] Mann, V.A. and I.Y. Liberman, Phonological awareness and verbal short-term memory. *Journal of Learning Disabilities*, 1984. 17(10): p. 592.
- [17] Chow, W.Y., C. McBridechang, and S. Burgess, Phonological Processing Skills and Early Reading Abilities in Hong Kong Chinese Kindergarteners Learning to Read English as a Second Language. *Journal of Educational Psychology*, 2005. 97(1): p. págs. 81-87.
- [18] Wagner, R.K., et al, Changing relations between phonological processing abilities and word-level reading as children develop from beginning to skilled readers: a 5-year longitudinal study. *Developmental Psychology*, 1997. 33(3): p. 468-79.
- [19] Wimmer, H, From the perspective of a more regular orthography. *Issues in Education*, 1995. 1: p. 101-104.
- [20] Solari, E.J., et al, LONGITUDINAL PREDICTION OF 1ST AND 2ND GRADE ENGLISH ORAL READING FLUENCY IN ENGLISH LANGUAGE LEARNERS: WHICH EARLY READING AND LANGUAGE SKILLS ARE BETTER PREDICTORS? *Psychology in the Schools*, 2014. 51(2): p. 126-142.
- [21] Service, E., Phonology, working memory, and foreign-language learning. *Quarterly Journal of Experimental Psychology A Human Experimental Psychology*, 1992. 45(1): p. 21-50.
- [22] Mareschal, D. and M.S. Thomas, Computational modeling in developmental psychology. *IEEE Transactions on Evolutionary Computation*, 2007. 11(2): p. 137-150.
- [23] Christiansen, M.H. and N. Chater, Connectionist psycholinguistics: Capturing the empirical data. *Trends in Cognitive Sciences*, 2001. 5(2): p. 82-88.
- [24] Chomsky, N, Knowledge of language: Its nature, origin, and use. 1986: Greenwood Publishing Group.
- [25] Chomsky, N, Aspects of the Theory of Syntax. Vol. 11. 2014: MIT press.
- [26] Joanisse, M.F. and M.S. Seidenberg, Impairments in verb morphology after brain injury: A connectionist model. *Proceedings of the National Academy of Sciences*, 1999. 96(13): p. 7592-7597.
- [27] Rumelhart, D.E. and J.L. McClelland, Parallel distributed processing: Explorations in the microstructure of cognition: Foundations (Parallel distributed processing). 1986, MIT Press, August.
- [28] Seidenberg, M.S. and J.L. McClelland, A distributed, developmental model of word recognition and naming. *Psychological review*, 1989. 96(4): p. 523.
- [29] Yang, J. and M. Thomas. C. Simulating intervention to support compensatory strategies in an artificial neural network model of atypical language development. in Airenti G. Proceedings of the 4th Euro Asian Pacific joint conference on cognitive science (EAPCog Sci 2015). 2015. Torino: CEUR-WS.
- [30] Plunkett, K. and V. Marchman, From rote learning to system building: Acquiring verb morphology in children and connectionist nets. *Cognition*, 1993. 48(1): p. 21-69.

**J. Yang** received her PhD in Artificial Intelligence from Southwest University, China, in 2007. She is a researcher in the Learning Technology Lab of the Department of Computer Science at Sichuan Normal University, China. She used to be a visiting scholar of the Developmental Neurocognition Lab, Department of Psychological Sciences, Birkbeck, University of London. Her research interests include learning style in intelligent learning systems, simulations in e-learning, and learning data analysis.

**M. S. C. Thomas** Received his BSc(Hons.) degree in Psychology at the University of Exeter, MSc degree in Cognitive Science at the University of Birmingham and completed D.Phil. in Experimental Psychology at the University of Oxford, in 1995. Between 1995 and 1998, he lectured at the University of Winchester, until joining Prof. Annette Karmiloff-Smith's research lab at the Institute of Child Health as a research fellow. He joined Birkbeck College in 2002, and is currently a Professor of Cognitive Neuroscience. He established the Developmental Neurocognition Laboratory in 2003, within Birkbeck's world-leading Centre for Brain and Cognitive Development. In 2006, the lab was the co-recipient of the Queen's Anniversary Prize for Higher Education. In 2010, he became Director of the Centre for Educational Neuroscience, a tri-institutional research centre which aims to further translational research between neuroscience and education, and establish new transdisciplinary accounts in the learning sciences. He is a Chartered Psychologist, Fellow of the British

Psychological Society, and Fellow of the American Psychological Society.

**X. Liu** was a graduate student at Sichuan Normal University, and received her master's degree of educational technology in 2016.