

## *Retraction*

# **Retracted: Construction of an Assessment System for Business English Linguistics Based on RNN Multidimensional Models**

### **Mathematical Problems in Engineering**

Received 18 July 2023; Accepted 18 July 2023; Published 19 July 2023

Copyright © 2023 Mathematical Problems in Engineering. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

- (1) Discrepancies in scope
- (2) Discrepancies in the description of the research reported
- (3) Discrepancies between the availability of data and the research described
- (4) Inappropriate citations
- (5) Incoherent, meaningless and/or irrelevant content included in the article
- (6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

### **References**

- [1] Y. Shi and H. Shi, "Construction of an Assessment System for Business English Linguistics Based on RNN Multidimensional Models," *Mathematical Problems in Engineering*, vol. 2022, Article ID 8446281, 8 pages, 2022.

## Research Article

# Construction of an Assessment System for Business English Linguistics Based on RNN Multidimensional Models

Yu Shi<sup>1</sup> and Hu Shi<sup>2</sup> 

<sup>1</sup>School of Foreign Languages, Shaoyang University, Shaoyang 422000, Hunan, China

<sup>2</sup>Software Institute, Dalian Jiaotong University, Dalian 116028, Liaoning, China

Correspondence should be addressed to Hu Shi; [shihu@djtu.edu.cn](mailto:shihu@djtu.edu.cn)

Received 23 February 2022; Accepted 4 March 2022; Published 27 March 2022

Academic Editor: Man Fai Leung

Copyright © 2022 Yu Shi and Hu Shi. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The current teaching evaluation system of the course “Introduction to Business English Linguistics” has many problems and should be reformed according to the requirements of the relevant professional syllabus. This paper proposes a recurrent neural network (RNN) modelling technique, which has achieved good results in modelling business English language models. The RNN modelling method is applied to Chinese language modelling according to the characteristics of Chinese language, and the advantages of the two models are combined to propose a fusion model. The experimental results demonstrate that this paper focuses on the process, diversity, and humanity of teaching evaluation and discusses the construction of the evaluation system model of the course “Introduction to Business English Linguistics” in terms of the content and criteria of teaching evaluation and the methods of teaching evaluation, so as to provide a new model for the teaching reform of the course “Introduction to Business English Linguistics”.

## 1. Introduction

The Syllabus for Business English for Business English Majors in Higher Education stipulates that linguistics, as an important course in the professional knowledge category of business English, is a compulsory course for senior undergraduate students of business English. In the current implementation of specific teaching, it has various names, such as General Linguistics, Introduction to Business English Linguistics, and Introduction to Business English Linguistics. The aim of its teaching is to make students aware of the rich achievements of human language research, to raise their awareness of the importance of language in social, humanistic, economic, technological and personal cultivation aspects, to cultivate linguistic awareness, and to develop rational thinking [1].

This syllabus sets out the requirements for students' knowledge of linguistics and their ability to analyse and apply language theory [2]. The main teaching task of Introduction to Linguistics is therefore to enable students to understand and master the basic knowledge and theories of

linguistics, to describe, explain, analyse, and solve practical linguistic problems using the knowledge and theories they have learned, and to improve their linguistic cultivation and language quality [3]. However, to date, the domestic Introduction to Business English Linguistics course does not have an examination syllabus and teaching evaluation system to match the syllabus, and the content and form of the course's teaching evaluation is very subjective and arbitrary, emphasising only the memorisation of knowledge, but not the assessment of ability [4].

The teaching of “Introduction to Business English Linguistics” focuses on the transmission of knowledge but not on the cultivation of ability, and the teaching assessment system is seriously inadequate [5]. At present, the teaching evaluation method of this course mainly adopts a single summative evaluation method, i.e., using examinations, tests, and quizzes as one-off tests to judge students' learning results and teaching quality by the level of achievement, and examinations are equivalent to evaluation [6]. The content of such summative examinations is textbook and formatted, teachers are taught what to test, students are tested on what

to learn, the difficulty of the knowledge assessed is very limited, only in the understanding of terminology and language theory and the reproduction of memory, the examination of the thinking component is small, the linguistic knowledge of the students' ability to use and expand the ability assessment is basically ignored [7]. This has a serious impact on the content and teaching style of the course and consequently on the teaching objectives of the course; i.e., it is not conducive to the development of students' ability to use their linguistic knowledge and theories to analyse linguistic phenomena and solve language learning problems [8]. To change this situation, it is necessary to reconstruct the teaching assessment system of the "Introduction to Business English Linguistics" course, i.e., to change the current single summative assessment method and increase the formative assessment, so that the two are organically combined and a new assessment system is constructed [9].

The "Introduction to Business English" course is designed to give students not only basic knowledge and skills but also the ability to explore on their own, through a variety of learning activities, such as cooperative group learning and classroom discussions [10]. In this way, students can learn and master the scientific method of exploring knowledge while gaining a deeper understanding of it. Therefore, in the process of teaching evaluation, special attention should be paid to the dynamic learning process of students, guiding them to pay more attention to, recognise, grasp, and improve the microprocesses of their own learning, and through teachers and students jointly monitoring, reflecting and regulating the whole process of teaching, prompting changes in teachers' teaching styles and students' learning methods [11].

The plurality of teaching evaluation is manifested in several aspects such as the content, evaluation subjects, and evaluation criteria. In terms of content, the "Introduction to Linguistics" course not only evaluates students' level of mastery of basic linguistic knowledge but also the development and improvement of individual students' interests, attitudes, and strategies in the learning process; in terms of the subject of evaluation, it changes from a single teacher's evaluation to a combination of teacher evaluation, student self-evaluation, and student mutual evaluation. In terms of evaluation criteria, it has both evaluation criteria for students' basic knowledge and skills and evaluation criteria for students' practical application ability, thinking ability, critical ability and innovation ability [12]. The two are combined and complement each other to evaluate students' learning status.

## 2. Related Work

It is an essential part of the teaching and learning process and is designed to check and facilitate teaching and learning. The famous American educator Bloom divides teaching evaluation into three main categories: diagnostic evaluation, formative evaluation, and summative evaluation. These three types of assessment correspond to the distinction that

is made before, during, and at the end of a sequence of educational activities [13]. Formative assessment was introduced in 1967 by M. Scriven, an American expert in evaluation, and later by Bloom, an American educationalist, into the field of teaching. Formative assessment emphasises the evaluation of students' knowledge acquisition and competence development in the educational process, including the evaluation of learning outcomes, the evaluation of various input conditions, and the evaluation of educational programmes and instructional methods. Its aim is to enable both teachers and students to receive timely feedback to improve the teaching and learning process and enhance the quality of teaching and learning. Howard Garden-er, a developmental psychologist at Harvard University, put forward the theory of multiple intelligences in 1983, and he advocated an educational evaluation that is conducted through multiple channels and in multiple forms [14]. Wiegrefe and Pinter [15] chose test questions similar to the wrong ones for practice. This is a mock exposure design that attempts to control the exposure of the overall question pool and remove highly exposed questions, but there is a gap with the real exposure design, which is still not practical enough. Reference [16] on the other hand starts from the evaluation aspect of the exercise and applies a knowledge map to manage the exercise questions.

While the theoretical results of adaptive testing are fruitful, the practical applications are also very extensive [17]. For example, the Graduate Record Examination (GRE) is used to find more suitable students and to provide students with the opportunity to choose a more suitable school and major; the US Army Vocational Aptitude Battery (ASVAB) allows people to enter a more suitable branch of the military and to perform at the highest level of operational efficiency; the National Assessment of Educational Progress (NAEP) optimises the entire US education system and allows students to learn more efficiently [18]. Domestic educational Internet companies are also beginning to consider such methods to high overall efficiency, reduce costs, and high the possibility of supplying the market with more cost-effective options for more children [19]. In addition, CAT has gained the attention of experts, scholars, and frontline engineers in various fields, and Internet education companies are devoting more attention to this area, which has good practical value and broad development prospects [20].

## 3. RNN Model in This Paper

**3.1. N-Gram Language Model.** The n-gram language model with statistical rules, introduced in 1980, is a widely used language model that uses the Markov assumption that the probability of occurrence of each predictor variable is related only to the context of length  $n - 1$ . If the historical information of word  $w_i$  is expressed as  $h_i = w_1 w_2 \dots w_{i-1}$ , then the probability of occurrence of word  $w_i$  and sentence  $s$  according to the conditional probability formula and Markov's hypothesis are as follows:

$$\begin{aligned}
p(w_i|w_1, w_2, \dots, w_{i-1}) &= p(w_i|h_i) \\
&= p(w_i|w_{i-n+1}^{i-1}), \\
p(s) &= p(w_1, w_2, \dots, w_N) \quad (1) \\
&= \prod_{i=1}^N p(w_i|h_i).
\end{aligned}$$

For n-gram language models, a large corpus is usually trained, and the more frequent words in the corpus tend to be trained better, while low frequency words are not trained as well.

In addition, the higher the order  $n$ , the more binding the model is, but as  $n$  increases, the size of the model grows exponentially, increasing the computational complexity of training and placing greater demands on storage space. The appropriate value of  $n$  is therefore a compromise between accuracy and complexity of the language model [21]. In a practical recognition system,  $n=3$  is generally chosen to construct a Tri-gram language model; i.e., the probability of each word occurring in a sentence from the training data is only related to its first two words, which can be expressed as

$$p(w_i|h_i) \approx p(w_i|w_{i-2}, w_{i-1}). \quad (2)$$

In this paper, we improve the language model in the Chinese speech recognition system and use the RNN language model to re-score the initial recognition candidates and perform postrecognition processing to complete the recognition process of the whole system, while the other modules of the baseline system remain unchanged.

**3.2. RNN Language Models.** Reference [14] states that a recurrent neural network, also known as an Elman network, has the structure shown in Figure 1 and consists of three network layers, the input layer, the implicit layer, and the output layer, with the storage layer acting as part of the input and preserving the state of the implicit layer at the previous moment [22]. After the text corpus is trained by this RNN structure, the probability of the current word  $w_i$  occurring is expressed as

$$\begin{aligned}
p(w_i|h_i) &= p_{rmn}(w_i|w_{i-1}, h_{i-1}) \\
&= p_{rmn}(w_i|h_i). \quad (3)
\end{aligned}$$

Reference [15] also states that the input word sample of the network at time  $t$  is assumed to be  $w(t)$ , i.e., the vector of current words; dimensionality is determined by the number of word samples in the corpus  $|V|$ ; the state of the implicit layer  $h(t)$  is determined by both the input current word vector  $w(t)$  and the state of the implicit layer at the previous moment, i.e., the history information  $h(t-1)$ , through the connection from the implicit layer to the input layer, and the state of the implicit layer at time  $t-1$  as part of the input at time  $t$ ; the output layer  $y(t)$  represents information about the probability distribution of the subsequent words under the current history; and the number of nodes in the output layer is the same as the number of nodes in the input layer also  $|V|$ .

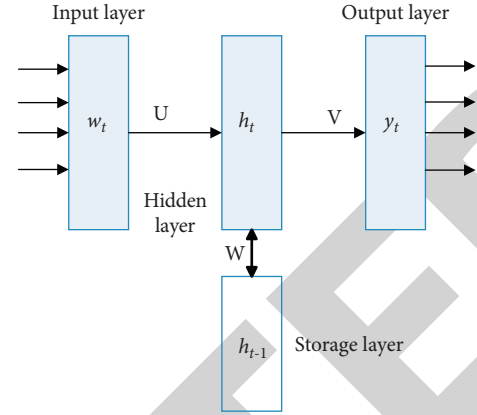


FIGURE 1: Recurrent neural network framework.

The computational relationship between the layers is represented by the following equation:

Inputs to the implicit layer :  $x(t) = w(t) + h(t-1)$ ,

$$\text{Implicit layer state : } h_j(t) = f\left(\sum_i x_i(t)u_{ji}\right). \quad (4)$$

Compared to business English speech, Chinese speech recognition is more complex. There are more than 6,000 commonly used characters in Mandarin Chinese, with about 60 phonemes, 407 untuned syllables, and 1,332 toned syllables, each of which consists of a vowel, a rhyme, and a tone, and each character represents a syllable. At the same time, there are also a large number of homophones and polysyllabic characters in Chinese, which must be constrained by high-level nonacoustic knowledge such as contextual background in order to complete recognition [23].

In terms of language, business English utterances focus on structure, while Chinese utterances focus on semantics, where the same word has different meanings in different contexts and the long-distance dependencies between words are relatively tight. When RNN is used to train the language model, more high-level semantic information is taken into account, which can better reflect the binding relationship between Chinese words. Therefore, RNN modelling techniques will be more suitable for training Chinese language models.

In addition, there are obvious spaces between words in the English corpus, so the corpus can be trained directly with a little processing. Unlike the Chinese corpus, there is no clear boundary between words in the Chinese sentences, and the training corpus needs to be divided into subword units according to the word separation model. After a series of processing to obtain a pure text corpus, the model can be trained.

The Chinese training corpus is shown in Figure 2, where the text corpus is cleaned to remove noisy information such as letters and punctuation marks from the coarse corpus and to remove redundant information; there are a large number of numbers in the corpus, and the regularisation of the numbers is completed to convert the Arabic numbers in the corpus into Chinese characters; thus, only Chinese

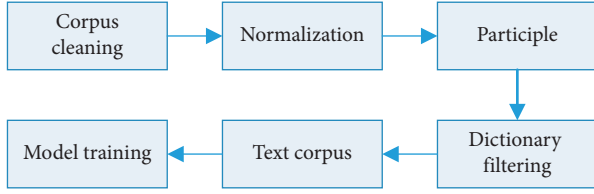


FIGURE 2: Processing flow of the Chinese corpus.

characters exist in the corpus. The lexical filtering removes English boundary characters and nonlexical sentences from the corpus. This results in a corpus that can be used for training. The model is then trained.

**3.3. RNN Model and n-Gram Model Fusion Modelling.** The higher the frequency of words in the corpus, the better the n-gram modelling technique can be trained, while the opposite is true for lower frequency words. When using RNN to train the corpus, some words in the corpus can be trained well despite their low frequency, which can effectively complement the n-gram model. In order to fully exploit the advantages of both modelling techniques and obtain better recognition results, this paper investigates a fusion modelling method based on the RNN model and the n-gram model.

As shown in Figure 3, for the same speech, the n-best list can be obtained from the word map (lattice) generated by the decoder, and the trained RNN language model is then used to re-score the n-best list. The n-gram model score information of the n-best list is then interpolated and fused with the re-score information of the RNN model to calculate a new language model score for each candidate unit.

Among the fusion algorithms for models, linear interpolation fusion is currently the more common approach, predicting the probability of the current word  $w_i$  based on the context  $h_i$ :

$$p_x(w_i|h_i) = \sum_j^L \lambda_j p_j(w_i|h_i), \quad (5)$$

$$p(w_i|h_i) = \lambda_1 p_{rm}(w_i|h_i) + \lambda_2 p_{ng}(w_i|h_i).$$

where  $L$  is the number of interpolation models and the interpolation weights  $\lambda_j$  of each model are nonnegative and sum to 1; i.e.,  $\sum_{j=1}^L \lambda_j = 1$ . The log-likelihood score is re-scaled for each n-best list sentence after model fusion:

$$\lg L(s) = \sum_{i=1}^n asc_i + \text{lms} \sum_{i=1}^n \lg p_x(w_i|h_i) + n.wp, \quad (6)$$

where  $n$  is the number of words in the sentence;  $wp$  is the penalty score of the word;  $asc_i$  is the word  $w_i$  acoustic model score;  $\text{lms}$  is the model size; and  $p_x(w_i|h_i)$  represents the fusion score of the n-gram and RNN model for each word. The overall score of each list is calculated by combining the linguistic and acoustic model scores and the penalty score, and the highest score is selected as the final recognition result of this n-best list.

## 4. Teaching Evaluation Methods and Experimental Analysis

**4.1. Model Evaluation Criteria.** The evaluation of the performance of the language model is based on information-theoretic knowledge. The performance of a language model is measured by calculating the magnitude of its perplexity on the test text. The perplexity is the inverse of the geometric mean of the probabilities of occurrence of each word in a given text set when predicted by the language model. Assuming that there are  $M$  words in the test text, the perplexity is

$$\text{perplexity} = \frac{1}{\sqrt[M]{\sum_{i=1}^M \log_2 P(w_i|w_{i-n+1}, \dots, w_{i-1})}} \quad (7)$$

In general, the smaller the value, the more binding the language model is to the language and the better the performance of the trained model. In addition to measuring language models in terms of perplexity, the most intuitive idea is to apply the model to a system and measure it by testing the system's Word Error Rate (WER). In general, a well-trained model will result in a high recognition rate. In this paper, the two evaluation criteria are combined to test the language model.

**4.2. Experimental Design.** Experiment 1 was used to verify the effectiveness of RNN language models in Chinese speech recognition. In this experiment, RNN and n-gram models were trained on the same dataset, and different RNN language models were trained by varying the number of nodes in the implicit layer of the RNN to investigate the changes in the perplexity and recognition rate of the RNN model with different parameters and to compare the performance with that of the n-gram model. Experiment 2 was used to verify the effectiveness of the proposed model fusion algorithm. The n-gram model trained in Experiment 1 was fused with the RNN model by linear interpolation, and the change in recognition rate was tested. In order to accelerate the training of RNN models, the training of RNN models was performed on a GPU (NVIDIA GTX 650) server with CUDA Toolkit 5.5, which is two to three times faster than the training on a CPU and shortens the training time of RNN models [24].

**4.3. Experimental Data.** The training data were obtained from the annotated data of the Chinese telephone speech transcription task provided by KODA XUNFE, with a total of 16 M, containing 550 thousand sentences and 4342 thousand words. The model perplexity test set is 9332 sentences containing 23 thousand words, and the speech test set is a 100-best list of 3433 sentences decoded from telephone speech, with a size of 87 kB. The number of nodes in the implicit layer for the RNN training model is set with six sets of parameters [25].

**4.4. Experiment 1.** The n-gram language model was trained using the Kneser-Ney backward smoothing algorithm with good smoothing performance, and the model order was 3,

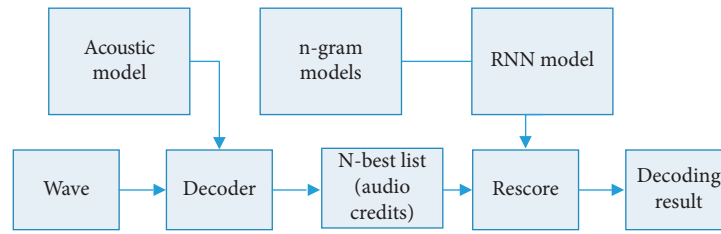


FIGURE 3: English model.

i.e., 3-gram, constructed from the SRILM toolbox. The number of n-grams in the model was  $30274 + 6568660 + 13830707 = 20429641$ , and the model size was 463599 kB. The RNN language model was then trained with the same training data, and the changes in the PPL values and the WER of the recognition system were tested. The experimental results are shown in Table 1.

As can be seen from Table 1, the RNN and 3-gram language models were trained with the same training data, and the confusion level of the RNN language model was reduced by about 7%; the error rate of speech recognition was reduced by about 5%, which proved the effectiveness of the RNN language model in Chinese speech recognition. In general, the lower the perplexity of the generated model, the higher the recognition rate of the system.

In addition, Table 1 also shows that (1) with the increase of the number of nodes in the hidden layer, the perplexity of the RNN language model and the system error rate gradually decrease, indicating that the learning ability of the network increases with the increase of the number of nodes; (2) when the number of nodes in the hidden layer increases to a certain degree, the perplexity of the generated model increases and the system recognition rate decreases, indicating that the more the number of nodes in the hidden layer. The more complex the network structure is, the smaller the error of the training samples can be reduced to a sufficient size through learning; however, excessive pursuit of learning on the training samples will produce overtraining. With a limited number of training samples, if the average training relative error of the learning sample set continues to decrease after a certain stage of learning, while the average test relative error (generalisation error) of the test sample set increases, the generalisation ability of the network will be reduced, affecting the performance of the trained model. Therefore, for different sizes of training corpus, the parameters need to be adjusted when using RNN training in order to achieve better experimental results.

The guiding principle of promoting the all-round development of all students through assessment, as shown in Figure 4, treats students as individuals with individual characteristics and different interests and needs, recognises their individual differences in their development levels, and examines all aspects of their knowledge, intelligence, and emotional factors; making horizontal comparisons fully considers students' vertical development; while assessing students teaches them to self-assess. The course is designed to help students develop a learning style that is truly effective and in line with their individual characteristics, fully

reflecting the characteristics of the student as the main subject and making them the master of learning. As a humanities subject, "Introduction to Business English Linguistics" should strive to create a relaxed and friendly atmosphere, reduce students' tension, evaluate students' achievements with a developmental and humanistic perspective, focus on students' subjectivity, and make students feel their own progress and development, so that their interest and creativity are stimulated and their self-confidence is enhanced. Self-confidence is enhanced, fully reflecting the humanistic spirit of respecting students' individuality.

**4.5. Experiment 2.** On the basis of the recognition results of Experiment 1, a linear interpolation method was used to interpolate and re-score the language model score of each word in the 100-best list by both sets of models, and then the probability of each sentence was calculated according to equation (11) combined with the acoustic model score, from which the highest score was selected as the recognition result of that speech, as shown in Table 2. The interpolation coefficient in the experiment is 0.6, i.e.,  $\lambda_1 = 0.6$  achieved a better recognition result.

As can be seen from Table 2, after linear interpolation, the recognition rate of the model decreases by about 8% compared to the 3-gram model and by about 3% compared to the RNN model, and the recognition rate after model fusion improves more significantly compared to that of the n-gram model because the RNN model is better trained for low-frequency words and can effectively solve the data sparsity problem. It can also be seen that the recognition rate of the fused models is higher than that of either model alone, indicating that the two models are complementary and demonstrating the effectiveness of the model fusion approach. In a practical recognition system, an n-gram language model with a large corpus can be trained and used as a general model and then interpolated and fused with the RNN model trained with a small corpus in the post-processing module of the speech recognition system, using this method to postprocess the decoding results to improve the recognition rate of the system. In this experiment, due to the limited training corpus, the overall recognition rate of the system is not very high, but it can still show the superiority of RNN in modelling Chinese language models and the effectiveness of the model fusion construction method.

The content of the course "Introduction to Business English Linguistics," as shown in Figure 5, is determined by its teaching objective, which is "to enable students to

TABLE 1: Performance comparison of RNN language model and 3-gram language model.

Number of hidden layer nodes	n-gram (KN3)		RNN		Decline rate of ppl (%)	Decrease rate of WER (%)
	PPL	WER (%)	PPL	WER (%)		
100	166.3	42.09	158.3	40.2	4.81	4.49
200	166.3	42.09	157.6	40.1	5.23	4.72
300	166.3	42.09	156.6	40.08	5.83	4.77
400	166.3	42.09	155.4	39.96	6.55	5.06
55	166.3	42.09	154.5	39.87	7.1	5.27
600	166.3	42.09	155.2	39.39	6.67	5.13

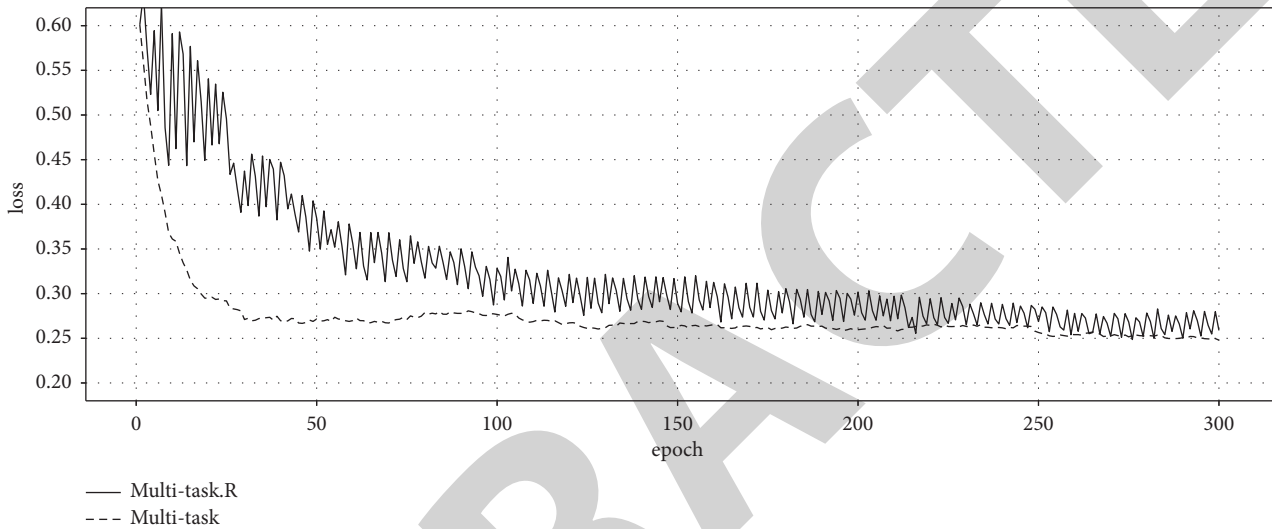


FIGURE 4: Student education learning outcomes.

TABLE 2: Comparison of recognition performance after fusion of the two models.

Number of hidden layer nodes	RNN + n-gram	n-gram	RNN	Decline rate compared with n-gram (%)	Decline rate compared with RNN (%)
	WER%	WER %	WER %		
100	39.23	42.09	40.2	6.97	2.41
200	39.14	42.09	40.1	7	2.39
300	38.6	42.09	40.08	8.29	3.69
400	38.62	42.09	39.69	8.24	3.28
500	38.62	42.09	39.87	8.38	3.28
600	38.59	42.09	39.39	8.31	3.35

understand and master the basic knowledge and theories of linguistics, to describe, explain, analyse, and solve practical linguistic problems using the knowledge and theories they have learned, and to improve their language skills and quality.” The aim of the course is to “enable students to understand and master the basic knowledge and theories of linguistics, to use their knowledge and theories to describe, interpret, analyse and solve practical language problems, and to improve their language skills and quality.” The students

are assessed in two ways: their mastery of basic linguistic theory is assessed in terms of the syllabus and the examination syllabus, and their ability to grasp the knowledge structure as a whole is assessed in a quantitative way. The ability to apply theoretical knowledge in linguistics should be assessed in a comprehensive way (combining in-class and out-of-class) to evaluate students’ ability to apply theoretical knowledge, in a variety of forms, using a qualitative assessment approach.

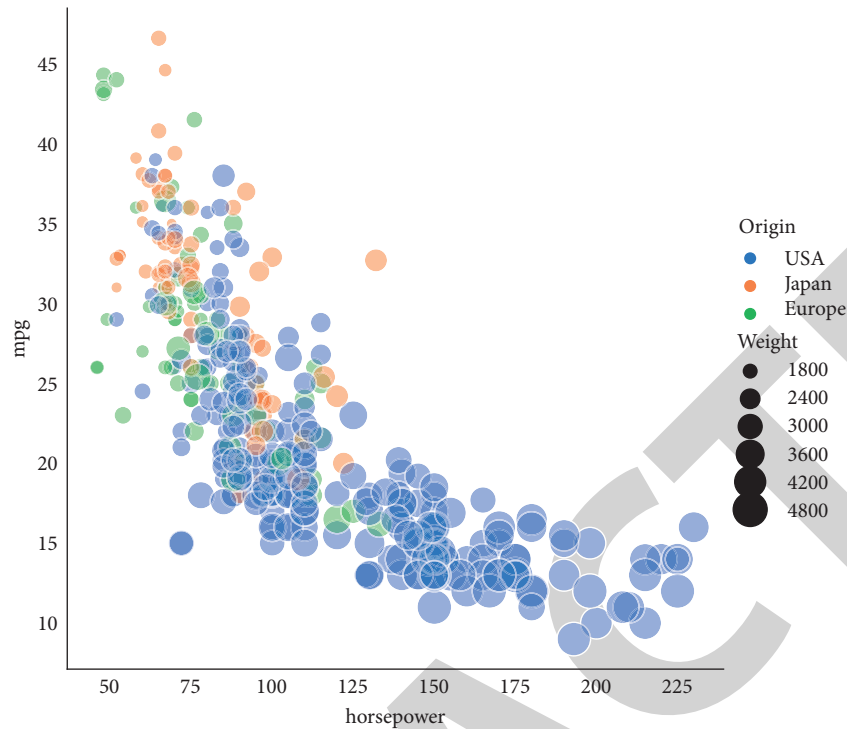


FIGURE 5: Effectiveness of teaching evaluation.

## 5. Conclusions

This paper applies RNN language model modelling to English language teaching assessment and verifies the effectiveness of the RNN modelling approach in Chinese language processing by comparing the generated model with the traditional n-gram model, which reduces the perplexity by about 7%. The reconstruction of the teaching evaluation system model of the Introduction to Business English course is a complex systemic project, which cannot be established without the related reform of teaching management, teaching content, teaching methods, teaching materials, and other aspects.

## Data Availability

The raw data supporting the conclusions of this article will be made available by the authors, without undue reservation.

## Conflicts of Interest

The authors declare that they have no conflicts of interest regarding this work.

## References

- [1] N. Basnin, L. Nahar, and M. S. Hossain, "An integrated CNN-LSTM model for Bangla lexical sign language recognition," in *Proceedings of the International Conference on Trends in Computational and Cognitive Engineering*, pp. 695–707, Malaysia, Asia, January 2021.
- [2] Z. Chen, Y. Wu, F. Yin, and C. L. Liu, "Simultaneous script identification and handwriting recognition via multi-task learning of recurrent neural networks," vol. 1, pp. 525–530, in *Proceedings of the 2017 14th IAPR international conference on document analysis and recognition (ICDAR)*, vol. 1, pp. 525–530, IEEE, Kyoto, Japan, November 2017.
- [3] S. Ghosh, M. Chollet, E. Laksana, L. P. Morency, and S. Scherer, "Affect-lm: a neural language model for customizable affective text generation," 2017, <https://aclanthology.org/P17-1059>.
- [4] I. Serban, A. Sordoni, R. Lowe et al., "A hierarchical latent variable encoder-decoder model for generating dialogues," in *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 31, no. 1, Vancouver, Canada, 2017 February.
- [5] F. Klubička, A. Toral, and V. M. Sánchez-Cartagena, "Quantitative fine-grained human evaluation of machine translation systems: a case study on English to Croatian," *Machine Translation*, vol. 32, no. 3, pp. 195–215, 2018.
- [6] M. Richardson, C. Abraham, and R. Bond, "Psychological correlates of university students' academic performance: a systematic review and meta-analysis," *Psychological Bulletin*, vol. 138, no. 2, pp. 353–387, 2012.
- [7] K. Adnan and R. Akbar, "An analytical study of information extraction from unstructured and multidimensional big data," *Journal of Big Data*, vol. 6, no. 1, pp. 1–38, 2019.
- [8] J. T. Downer, R. C. Pianta, X. Fan, B. K. Hamre, A. Mashburn, and L. Justice, "Effects of web-mediated teacher professional development on the language and literacy skills of children enrolled in prekindergarten programs," *NHSA dialog*, vol. 14, no. 4, pp. 189–212, 2011.
- [9] K. Zhu, K. L. Kraemer, and J. Dedrick, "Information technology payoff in e-business environments: an international perspective on value creation of e-business in the financial services industry," *Journal of Management Information Systems*, vol. 21, no. 1, pp. 17–54, 2004.
- [10] X. I. E. Tao, C. Zhang, and Y. Xu, "Collaborative parameter update based on average variance reduction of historical



- gradients,” *Journal of Electronics and Information Technology*, vol. 43, no. 4, pp. 956–964, 2021.
- [11] Z. Zhengwan, Z. Chunjong, L. I. Hongbing, and X. I. E. Tao, “Multipath transmission selection algorithm based on immune connectivity model,” *Journal of Computer Applications*, vol. 40, no. 12, p. 3571, 2020.
- [12] Z. Zhang, D. Wu, and C. J. Zhang, “Study of cellular traffic prediction based on multi-channel sparse LSTM,” *Computer Science*, vol. 48, no. 6, pp. 296–300, 2021.
- [13] V. Sorin, Y. Barash, E. Konen, and E. Klang, “Deep learning for natural language processing in radiology-fundamentals and a systematic review,” *Journal of the American College of Radiology*, vol. 17, no. 5, pp. 639–648, 2020.
- [14] Y. Long, R. Xiang, Q. Lu, C. R. Huang, and M. Li, “Improving attention model based on cognition grounded data for sentiment analysis,” *IEEE transactions on affective computing*, vol. 12, no. 4, pp. 900–912, 2019.
- [15] S. Wiegrefe and Y. Pinter, “Attention is not not explanation,” 2019, <https://arxiv.org/abs/1908.04626>.
- [16] W. Zhai and C. Cheng, “A long short-term memory approach to predicting air quality based on social media data,” *Atmospheric Environment*, vol. 237, Article ID 117411, 2020.
- [17] C. Van Hee, E. Lefever, and V. Hoste, “Exploring the fine-grained analysis and automatic detection of irony on Twitter,” *Language Resources and Evaluation*, vol. 52, no. 3, pp. 707–731, 2018.
- [18] R. N. Jones, “Identification of measurement differences between English and Spanish language versions of the minimal state examination,” *Medical Care*, vol. 44, no. Suppl 3, pp. S124–S133, 2006.
- [19] H. Elfaik, “Deep bidirectional lstm network learning-based sentiment analysis for Arabic text,” *Journal of Intelligent Systems*, vol. 30, no. 1, pp. 395–412, 2021.
- [20] A. Sakalle, P. Tomar, H. Bhardwaj, D. Acharya, and A. Bhardwaj, “A LSTM based deep learning network for recognizing emotions using wireless brainwave driven system,” *Expert Systems with Applications*, vol. 173, Article ID 114516, 2021.
- [21] Y. Zhang, J. Zheng, Y. Jiang, G. Huang, and R. Chen, “A text sentiment classification modeling method based on coordinated CNN-LSTM-Attention model,” *Chinese Journal of Electronics*, vol. 28, no. 1, pp. 120–126, 2019.
- [22] F. Ye and J. Yang, “A deep neural network model for speaker identification,” *Applied Sciences*, vol. 11, no. 8, p. 3603, 2021.
- [23] L. A. Erdodi, “Aggregating validity indicators: the salience of domain specificity and the indeterminate range in multivariate models of performance validity assessment,” *Applied Neuropsychology: Adultspan*, vol. 26, no. 2, pp. 155–172, 2019.
- [24] L. Buyse, “The business of pragmatics. The case of discourse markers in the speech of students of business English and English linguistics,” *Languages in Business Education: Introduction*, vol. 161, no. 1, pp. 10–30, 2011.
- [25] A. V. Litvinov, S. A. Burikova, and D. S. Khrumchenko, “Phrasal verbs as learning material in Business English courses for students majoring in Linguistics,” *A nation’s culture resides in the hearts and in the soul of its people*, vol. 1, no. 4, p. 91, 2017.