ingegneria sismica
International Journal of Earthquake Engineering
Anno XLIII - Num. 1 - 2026

Temporal Sequence Characteristics of the Evolution of Chinese Lexical Meanings in Traditional Cultural Canonical Texts

Jianping Xu[1,*]

[1] Pearl River College, Tianjin University of Finance and Economics, Tianjin, 300000, China

SUMMARY: *The evolution of Chinese lexical meanings in traditional cultural texts is closely related to the cultural and epochal background of the language. In this paper, a corpus of Chinese word meanings in traditional cultural texts is constructed by combining Wang Li's Dictionary of Ancient Chinese and Dazidian of Chinese. A two-way LSTM language model and ELMo vector generation algorithm are used to represent word vectors, and a dynamic vocabulary representation learning model is constructed to realize the task of polysemous word recognition. And the cosine similarity is used to calculate the similarity between the sets of related words before and after a certain time node, and then to determine whether the lexical meaning of a word has changed. When the obtained similarity value is larger, it indicates that the change of word meaning is small or does not occur, and vice versa, the change of word meaning is large. The results of the study show that the average accuracy of the Lexical Representation Learning Model (ELMo) is around 75%. Taking the character "Shai" as an example, it can be seen that the character "Shai" is closer to its original meaning from the pre-Qin to Han dynasties, and evolved the meaning of cultural display from the Wei, Jin, Southern and Northern Dynasties to the Tang Dynasty. In the Song, Yuan, Ming and Qing Dynasties, the character "Shai" began to become vernacular. Changes in time and social development are important factors contributing to the evolution of Chinese word meanings in traditional cultural texts.*

KEYWORDS: *LSTM; ELMo; lexical representation learning; cosine similarity; word sense evolution*

1 Introduction

The Chinese language is not a static symbol, but a living history, especially in the traditional Chinese cultural texts, the Chinese language tends to vary in lexical meanings in different dynasties and scenarios, showing clear time-sequence characteristics, and its evolutionary trajectory is closely related to social change and cultural evolution [1-4].

Traditional cultural canonical books are representative books formed in Chinese history and inherited until now. These texts not only reflect the political, economic, cultural, philosophical and religious achievements of ancient Chinese society, but also the crystallization of the wisdom of the Chinese nation [5, 6]. In today's society, the charm and value of traditional cultural texts have become more and more prominent, and they have become valuable assets for us to learn wisdom and inherit culture [7, 8]. The evolution of Chinese lexical meanings in traditional cultural texts refers to the phenomenon that the meaning of Chinese language changes with the influence of time, society, culture and internal factors of language in the process of historical

*shuyan20230809@163.com
https://doi.org/10.65102/is2026248
Published: 30 April, 2026

development [9-11]. This evolution includes the expansion and contraction of lexical meanings, transfer, etc.

Lexical expansion refers to the fact that a word originally refers to a narrower range of objects or concepts and is later extended to a wider range of objects or concepts [12, 13]. Take "river" as an example. In the Pre-Qin period, it specifically referred to the Yellow River, as in the line "Wait for me by the river" in the Book of Songs. Here, "river" is a specific term and does not refer to other rivers in a general sense. With the expansion of geographical cognition and the increase in language usage frequency, since the Han Dynasty, "river" has gradually become a general term for all large waterways. Lexical narrowing, on the other hand, is manifested in the fact that the original meaning covers a wide range of meanings and is later restricted to specific subcategories [14]. Typical example is the word “bird”, which in the oracle bone script refers to flying animals in general, including birds and insects, and gradually refers to domesticated birds, especially chickens, ducks and geese, in the Middle Chinese stage. Such changes are often closely related to the strengthening of social hierarchy and the refinement of semantic division of labor, reflecting the high adaptability of language to the real structure [15-17]. Lexical transfer refers to a change in the basic referential object of Chinese, which no longer points to the original entity but is projected to a new domain through the mechanism of association [18, 19]. Common forms include functional substitution, spatial metaphor and role replacement [20]. For example, the ancient meaning of "walking" is running, Mencius has "walking with armoring and tracers", which means fleeing in defeat. In modern Chinese, it is reduced to walking, and its semantic intensity is significantly weakened, reflecting the social generalization of travel mode and the de-radicalization of vocabulary [21, 22].

In this paper, we adopt a cross-time lexical representation learning method to realize automatic recognition of word meaning evolution. A large-scale corpus is constructed using polysemous words from traditional cultural texts as the corpus, which ensures the richness of the corpus and minimizes the chance of the experiment. After preprocessing the corpus, word vectors are trained by Word2vec to extract related words, and the word sense similarity is calculated. Finally, the word similarity between related word sets is compared to determine the change of word meaning. Through the case study, the law of Chinese word meaning evolution is summarized.

2 Method

In ancient Chinese, vocabulary constantly generates new semantic levels with the changes of social systems and ideological concepts, and the meanings of words in early documents are often different from those of later generations, or even generate semantic breaks. In this paper, we propose a time-series-based lexical evolution analysis method, which can trace the trajectory of lexical evolution, clarify the semantic gravity and pragmatic functions of words in different historical stages, and thus restore their original semantic appearance in traditional cultural texts.

2.1 Traditional Culture Canon Corpus Construction

2.1.1 Construction of the basic lexical knowledge base

The target words of this study are commonly used polysemous words in traditional cultural texts. Taking word frequency and academic research needs into consideration, 300 ancient Chinese words were screened.

Dictionaries have an extremely significant impact on the quality of annotation. The selection of dictionaries must be characterized by professionalism, high recognition, and clear descriptions of word meanings. This paper combines the "Wang Li Ancient Chinese Dictionary" and the "Great Chinese Dictionary" to construct a basic word meaning knowledge base, taking into account "comprehensiveness", "contemporanity" and "coverage", which can effectively address the characteristics of long time span and high complexity of word meaning description in ancient Chinese, and meet the needs of a word meaning annotation corpus.

2.1.2 Lexical annotation

After completing the construction of the basic lexical knowledge base, this study annotates the target word's sense items in the corpus based on the lexical knowledge and performs operations such as adding, deleting, merging, and so on, on the sense items in the lexical knowledge base according to the annotation results.

(1) Corpus Sampling and Preprocessing

Starting from the construction needs of the Ancient Chinese Word Sense Annotation Corpus, this study believes that the corpus selection should conform to the following principles: complete sentences and moderate sentence length to provide clearer contextual information. The corpus is balanced, covering different eras and types of documents, and reflecting the use and distribution of lexical meanings as much as possible. No special symbols and markers outside the text content.

According to the above principles, the scope of corpus sampling in this paper is set to the Corpus Online Ancient Chinese Corpus (SLC Corpus) and CCL Ancient Chinese Corpus, both of which are widely used by researchers and processed in simplified Chinese characters, which have the characteristics of large volume, complete collection, and coverage of different dynasties, etc. The corpus sampled from the above corpus is a comprehensive one, covering different periods and types of documents. Sentences containing the target words are extracted from the above corpora, and 300 corpus items are randomly selected for each target word, and their dynastic distribution is ensured to be balanced. Subsequently, the special markers in the corpus are removed.

(2) Lexical Sense Annotation Practice

According to the basic lexical knowledge base, the corpus annotation is carried out by the graduate students majoring in Chinese language philology and classical literature, which follows the following steps:

1) Marking the meaning items. According to the meaning of the target word in the context, select the meaning item number from the meaning item table. For the cases where the corresponding sense cannot be found, the following markings are made: if the target word belongs to a proper noun, it will be marked according to the proper name number mentioned above. If the target word is not included in the knowledge base, it will be marked as “Other”. If the context makes it difficult to determine the meaning of the item, it is marked as “to be determined”. If there are incomplete sentences or ambiguity of the target word in the context, it will be marked as “corpus inappropriate”.

2) Collect annotation feedback, count the frequency information of the items, adjust the list of items in the knowledge base with the dictionary description, and make suggestions for adding, deleting, merging and other operations on the items in the lexical knowledge base.

3) Confirm the revision of the lexical knowledge base by the teachers specialized in Chinese language literacy and Chinese information processing after reviewing the suggestions of addition, deletion, merging, etc. given in the above operations.

4) Revise the corpus annotation results according to the revised lexical knowledge base to ensure the consistency of the revised lexical knowledge base and the meaning items in the corpus annotation. Meanwhile, the example sentences in the dictionary are also added to the corpus as a supplement.

5) Proofreading of the knowledge base and the corpus. Firstly, the senior graduate students of Chinese language and philology proofread the entries such as “to be determined” and “other” in the corpus, and give reasonable suggestions for labeling. Then the students and teachers of the project team proofread the lexical knowledge base and corpus annotation results again.

2.1.3 Overall size of the corpus

The first phase of the Ancient Chinese Word Sense Annotation Corpus contains 300 polysemous words with a total of 3,000 sense items, with an average number of 30 sense items per word. Among them, there are five senses that do not appear in the corpus annotation, and these senses are considered to belong to the original meaning by the "Wang Ligu Chinese Dictionary" or the "Chinese Dictionary", but no example usage is listed, such as the original meaning of "dou", "utense hollow". Considering that these sense items belonging to the original sense have greater significance in the construction of the derivation chain, these low-frequency sense items are retained for examination. At present, the word sense annotation corpus contains 39,634 annotated data, totaling 1,182,000 words. In addition to proper names, the total number of sense items in the annotated corpus is 2995, and each corpus is annotated only for the unique target word.

2.2 Prediction-based learning algorithm for lexical representations

2.2.1 Neural network language model

The neural network language model mainly consists of a three-layer feed-forward neural network, and its main architecture is shown in Figure 1. The model mainly consists of three parts: the input layer, the hidden feature layer, and the output layer. The overall training process of this model is to predict what the next word will be based on the previous word in the given corpus. In the training process the input layer of the model first gives each word a vector in continuous space (word vector), through the shared weights of this layer, the input word numbers are transformed into vectors in the trainable space, and then through the subsequent hidden layer neural network to learn this distributed representation, and then finally through the softmax layer to predict the probability of each word, and then take the output word with the highest probability as the predicted result.

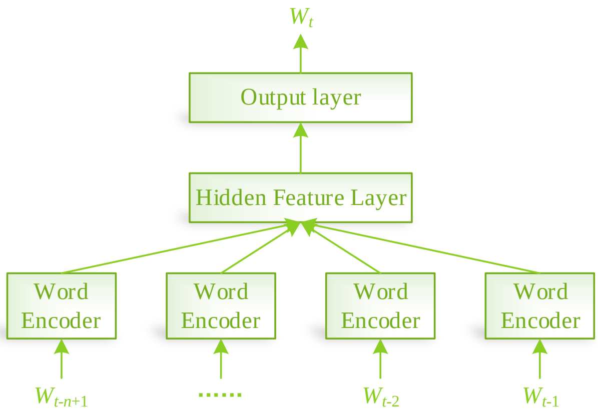

Figure 1: Neural network language model

2.2.2 Word2vec

The Word2vec tool contains two models: one is called the Continuous Bag of Words (CBOW) model and the other is the Skip-gram model. In both models, all words are learned to be represented in the form of corresponding word vectors through lexical representation learning, which can be seen as a mapping method to map discrete words into continuous space. During the training process, the model first initializes a vector for each word, which can be regarded as a point in the vector space, and when there is a certain error between the predicted next word and the actual next word in the training set, this error value is used to adjust the vector value of the corresponding word in the network. Through the training of a large amount of training data, words with the same or similar context will be “closer and closer” in the vector space, thus realizing the word semantic representation. Since this paper explores the time-series characteristics of the evolution of Chinese word meanings in traditional cultural texts, which involves infrequent words, the Skip-gram model is chosen for modeling.

The structure of the Skip-gram model is shown in Figure 2. Given the word w_t as prior knowledge to predict other words in the context, the training process of the Skip-gram model is to maximize Equation (1):

$$w_t = \arg\max\left(p\left(i \mid w_{t-n+l},\ldots,w_{t-2},w_{t-l}\right)\right), i \in V \tag{1}$$

$$w_{t+j} = \arg\max\left(\frac{1}{|V|}\sum_{t=l}^{|V|}\sum_{-c\ll j\ll c, j\neq 0}\log p\left(w_{t+j} \mid w_t\right)\right) \tag{2}$$

In Equation (2) $|V|$ refers to the total number of words in the word list, and c represents the number of words in the context, in Equation (2), the larger c is, the longer the computation time will be, and the result will be better. $P(*)$ is the softmax regression function, which is represented as shown in equation (3):

$$p\left(w_o \mid w_i\right) = \frac{\exp(vv_{w_o} v_{w_i})}{\sum_{w=1}^{|V|}\exp\left(v_{w_o}^{T} v_{w_{w_i}}\right)} \tag{3}$$

In Eq. (3) v_w and v_w represent the vector representations of the input and output layers, respectively. Both models are forward-feedback neural network models, which simplify the model and reduce the time required for training by taking away the nonlinear hidden layers compared to the neural language model.

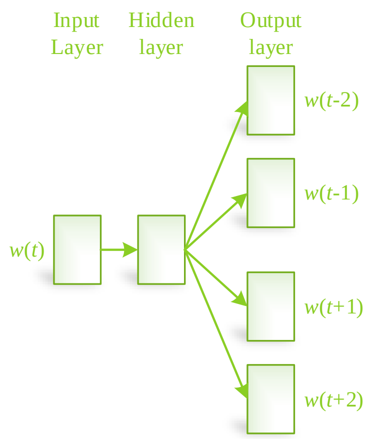

Figure 2: Skip-gram model structure

2.2.3 Long- and short-term memory networks

Since the traditional recurrent neural network will have serious forgetting phenomenon when the sequence is very long, even affecting the recent input data, in order to alleviate this problem gate-based recurrent neural network improvement has appeared, of which the widely used ones are mainly the long-short time memory network (LSTM).

The main structure of LSTM is shown in Fig. 3, in which σ is the sigmoid function, tanh is the hyperbolic tangent function, and C is the memory cell, and its core improvement is the introduction of the input gate, output gate, and forgetting gate to dynamically control long-term and short-term information, and to realize the mining of long-term input and short-term input for the final data in the longer input sequences through the control of the gate. By controlling the gates, it is realized that the long-term inputs and short-term inputs can be mined in the long input sequence, which has a greater influence on the final results, so that the final results can not only reflect the influence of short-term inputs, but also ensure the influence of long-term inputs on the model results to a certain extent.

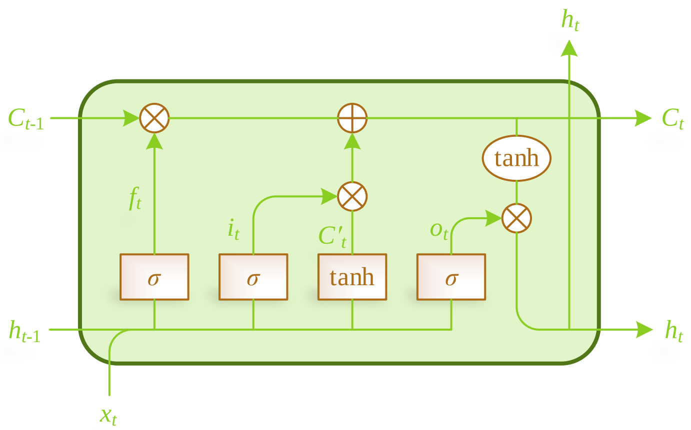

Figure 3: Long time memory network framework

The long and short term memory network learning process is abstracted as the following equation:

$$\begin{aligned}
i_t &= \sigma\left(W_i \cdot [h_{t-1}, x_t] + b_i\right) \\
f_t &= \sigma\left(W_f \cdot [h_{t-1}, x_t] + b_f\right) \\
o_t &= \sigma\left(W_o \cdot [h_{t-1}, x_t] + b_o\right) \\
c_t' &= \tanh\left(W_c \cdot [h_{t-1}, x_t] + b_c\right) \\
c_t &= f_t \odot c_{t-1} + i_t O c_t' \\
h_t &= o_t \odot \tanh\left(c_t\right)
\end{aligned} \tag{4}$$

LSTM has the ability to acquire and maintain long term contextual information. In LSTM, Cell can be seen as a mechanism specialized for error transfer across time, where the error of historical data during the training process is transferred on Cell and passed to the next moment of training, and merged with the error of the next moment to generate the error of the next moment. It is due to the mechanism of Cell transfer across time that the historical error (i.e., historical state) can be retained for a long time using LSTM, and this mechanism fits well with the tasks related to natural language processing, in which we usually have to consider the relationship between certain words in a sentence or even multiple sentences, and the traditional natural language representation models usually have a window to limit us to consider the relationship between Traditional natural language representation models usually have a window to limit us to consider the relationship between words in a large range, while using LSTM can jump out of this limitation, and training LSTM models on the corpus can naturally react to the relationship between words that are separated by a long period of time.

2.2.4 ELMo

ELMo is a dynamic lexical representation learning model, the core concept of dynamic lexical representation learning model is that the lexical representation of the same word is not unique, it changes with the given conditions, and it is an LSTM-based language model.

ELMo consists of two main parts, which are the bidirectional LTSM language model, and the construction of ELMo vectors. The framework of the bi-directional LSTM language model is shown in Fig. 4.ELMo uses multi-layer LSTM networks in the forward and backward directions.When we use L-layer LSTM networks, we obtain $2L$ features generated by the LSTM network at each input state.Together with the learning of representations of the words at the input of the model, we can, through the bi-directional LSTM language model obtain $2L+1$ representations for the current word, which is formulated as follows:

$$R_k = \left\{ X_k^{LM}, \vec{h}_{k,j}^{LM}, \overleftarrow{h}_{k,j}^{LM} \mid j = 1, \ldots, L \right\} \tag{5}$$

$$R_k = \left\{ h_{k,j}^{LM} \mid j = 0, \ldots, L \right\} \tag{6}$$

In the simplified formulation, $h_{k,0}^{LM}$ represents the representation of the input layer for words, and $h_{k,j}^{LM} = \left[\vec{h}_{k,j}^{LM}, \overleftarrow{h}_{k,j}^{LM} \right]$ denotes the output of the two-way LSTM language model at each layer.

While constructing the ELMo vector we need to integrate the representations obtained earlier into a single vector, for this purpose we can just use the output of the last layer of the bi-directional LSTM language model, i.e.:

$$ELMo_k = E\left(R_k; \theta_e\right), E\left(R_k\right) = h_{k,L}^{LM} \tag{7}$$

Task-specific ELMo vectors can also be computed by using Eq. (8):

$$ELMo_k^{task} = E\left(R_k; \theta^{task}\right) = \gamma^{task} \sum\nolimits_{j=0}^{L} s_j^{task} h_{k,j}^{LM} \tag{8}$$

where s^{task} is the weight after softmax regularization, that is, the weight vector whose weights sum to l. y^{task} can be regarded as the parameter that scales the vector, and the introduction of γ^{task} in the actual optimization process can guarantee the optimization to a certain extent. In this way, through a pre-trained bidirectional LSTM language model combined with the ELMo vector generation algorithm, we can quickly generate dynamic word vectors according to a specific task, and then complete the subsequent natural language processing tasks.

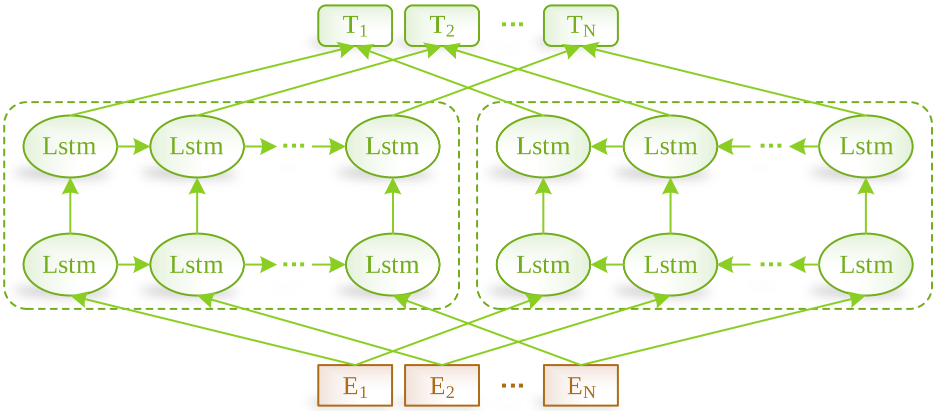

Figure 4: Two-way LSTM language model framework

2.3 Calculation of lexical evolution

In this study, we adopt a corpus-based lexical similarity computation method and analyze the lexical similarity using Word2vec tool. In the corpus, semantically similar words always have similar contexts, and by extracting the feature information of the context, words can be represented as vectors, which can be efficiently trained on thousands of datasets, and the similarity between words can be calculated by word vectors. The core of word vector representation is to use contextual information to represent words, and words with the same or similar contextual information are highly likely to have the same or similar word representation. Word2vec can analyze the possible values of words based on the context of the words, or it can analyze the words to derive possible contexts.

Once the feature vectors of the context are constructed, the semantic similarity between words is converted into the calculation of the similarity between word vectors. The cosin metric is widely used in the calculation of similarity, and its formula is as follows:

$$\cos(w_1, w_2) = \cos(\vec{v_1}, \vec{v_2}) = \frac{\vec{v_1} \times \vec{v_2}}{|\vec{v_1}| \times |\vec{v_2}|} = \frac{\sum_{c \in T(w_1) \cap T(w_2)} \left(weight(w_1, C) * weight(w_2, C) \right)}{\sqrt{\sum_{c \in T(w_1)} weight^2(w_1, C)} + \sqrt{\sum_{c \in T(w_2)} weight^2(w_2, C)}} \tag{9}$$

where $\vec{v_1}, \vec{v_2}$ is the feature vector of the target word w_1, w_2, $T(w)$ is the word appearing in w context, C is the context word, and $weight$ is the weight.

On the basis of calculating the similarity of word meanings, the comparison can determine whether the word meanings have changed before and after a certain time point. If the intersection of the most relevant words in the two time periods is larger, the lexical similarity of the words in these two time periods is higher, and the lexical change is smaller. Conversely, the lower the lexical similarity and the greater the lexical change.

3 Results and discussion

3.1 Model validation

3.1.1 Dispersion experiments

For the LSTM and ELMo models proposed in the previous section, this chapter sets up word sense disambiguation experiments to verify whether the models can recognize the correct word sense from multiple word sense lists.

Three groups of controlled experiments are set up for LSTM and ELMo models respectively, and the models are trained in a supervised form, and the training and test corpora are used to validate the models using the corpus constructed in the previous section. Word form and lexical features (Type 1), word form and semantic class features (Type 2), and word form, lexical and semantic class features (Type 3) were used to verify the accuracy of word sense recognition, respectively.The experimental data of the LSTM model is shown in Fig. 5. Twenty polysemous words were extracted from the previously constructed corpus for experimental validation, and the average recognition accuracies of the LSTM model on the three types of corpus features were 64.07%, 64.29%, and 65.55%, respectively.

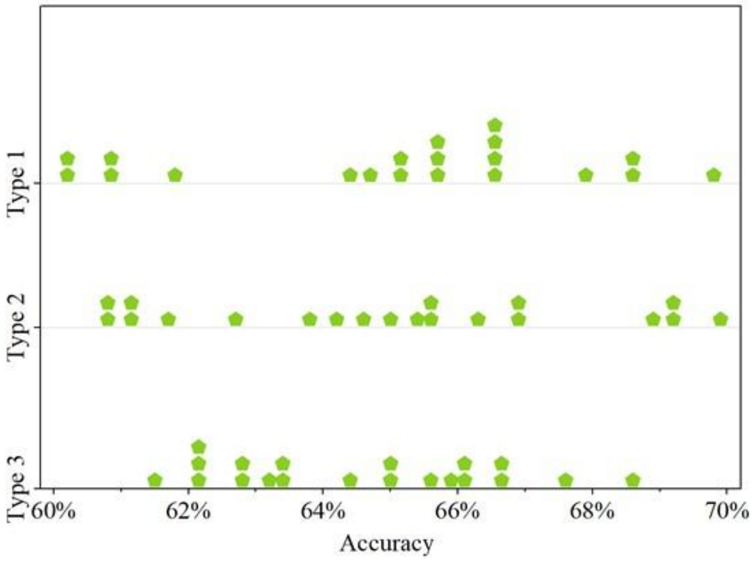

Figure 5: Experimental data of LSTM model

The experimental data of ELMo model is shown in Fig. 6.The average accuracy of ELMo model is 74.52%, 74.63%, and 75.96%, respectively. It shows that ELMo model has stronger disambiguation ability than LSTM model.

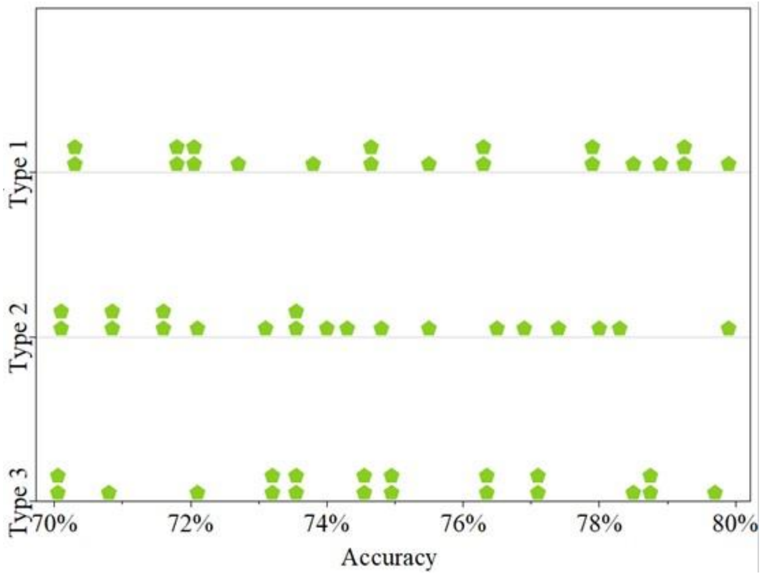

Figure 6: Experimental data of ELMo model

3.1.2 Validation of Dynamic Word Vector Representation Validity

Assuming that the word vectors at different times computed by the ELMo model have been mapped into the same vector space, given a word, the similarity of the word vectors at two different moments can be measured by the cosine similarity. If the result complies with the assumption that "the closer the semantics of the words, the higher the cosine similarity of the vectors", then the word sense representation is accurate. The test words and their contextual sentences are selected, and the contextual word vectors are obtained by ELMo to see whether the semantic distance between words and the cosine similarity are positively proportional to each other.

"Complement" is used in the corpus with the following samples and corresponding semantic items:

$\cos_1$:

Context: It's not too late to mend the dungeon after a sheep is lost (excerpted from "Strategies of the Warring States · Chu Ce").

Meaning: denoting mending.

$\cos_2$:

Context: Bravely mending the peacock fur (excerpted from <<Meng of the Red Chamber>>).

Meaning: A metaphor for Yang Sichang's mending of the mountains and rivers in the Ming Dynasty, with "mending" becoming a political symbol.

$\cos_3$: Zhuang Zi was dressed in a large cloth and mended it, and he passed the king of Wei with his feet tied (excerpted from <<Zhuang Zi · Shanmu Chapter>>).

Meaning: denoting mending.

The cosine similarity values of the lexical items in the different usage samples are shown in Table 1. It can be observed that $\cos_1$ and $\cos_3$ have the largest cosine similarity values because they share the same lexical item, which means "sewing". $\cos_2$, on the other hand, means "political symbol", so its similarity with the other two samples is lower. It is worth pointing out that, due to the structure of ELMo network, the absolute size of the similarity value based on the word vectors generated by ELMo network cannot directly reflect the degree of similarity, that is to say, the cosine similarity value of 0.8 (the upper limit is 1) does not mean that it has a high degree of similarity, but the relative size of the cosine similarity can be used to compare the degree of similarity between the words. Therefore, 0.89 > 0.85 > 0.71 indicates that "complement" in $\cos_1$ and $\cos_3$ has the closest meaning. This also indirectly verifies the validity and accuracy of the lexical meaning expression of dynamic word vectors generated based on ELMo.

Table 1: The cosin of "bu" in different contexts

	cosin
$\cos_1$	0.89
$\cos_2$	0.71
$\cos_3$	0.85

3.2 Case Studies

The study uses Word2vec to train word vectors, which needs to load the discrete processed corpus data into the model.Because the corpus set is large, Text8Corpus is used to construct Sentences in the model.Then the nodes in the model are set. Since the corpus set size of each research subject is different, in order to balance the experimental effect, the dimension size of the feature vector in the model is set to 100. The maximum distance between the current word and the predicted word in the same sentence window is set to 8, and the words that occur less frequently than 8 in the same corpus will be filtered out to improve the efficiency and quality of the experiment.

In this section, we take "Shai" as an example, load the corpus of "Shai" into the model by time period, and calculate the similarity of words related to "Shai".

The preprocessed corpus of "Shai" is loaded into the Skip-gram model of Word2vec, and the lexical similarity of the words related to "Shai" in this time period is obtained after running the model. The operation results are shown in Table 2. The meanings of "Pu" from the Pre-Qin period to the Han Dynasty, "Fuxuan" from the Warring States period to the Han Dynasty, "Shaishu" in the Wei and Jin Dynasties, "Zhibei" in the Tang Dynasty, "Liang" in the Song Dynasty, and "Shaiwang" in the Ming and Qing Dynasties are highly similar to "sai".

Table 2: Related words and similarity

	Relevant word		Similarity
1	曝	Pu	0.99907
2	负暄	Fuxuan	0.99925
3	晒书	Shaishu	0.99943
4	炙背	Zhibei	0.99975
5	晾	Liang	0.99983
6	晒网	Shaiwang	0.99902
7	熯	Han	0.99979
8	献曝	Xianpu	0.99991
9	晒秋	Shaiqiu	0.99998
10	晒腹	Shaifu	0.99994

In the experiments, cosin is used to calculate the similarity of related sets of words in neighboring time periods, and larger values obtained indicate smaller changes in word meanings in neighboring time periods. The smaller the value, the greater the change in word meaning. By comparing the similarity of the related word sets and plotting the word meaning change curve, we can see at a glance whether the word meaning has changed or not. By calculation, the lexical change curve of “Shai” is shown in Figure 7. From the curve of lexical change, it can be seen that from the pre-Qin to Han Dynasties (Stage 1), “Shai” refers to drying or warming oneself in the sun, which is closer to its original meaning. By the time of the Wei, Jin, North and South Dynasties to the Tang Dynasty (Stage 2), it mostly means cultural display. During the Song, Yuan, Ming and Qing Dynasties (Stage 3), the word “Shai” began to be dialectized and was often used to refer to drying clothes. From the pre-Qin to Han Dynasties and the Song, Yuan, Ming and Qing Dynasties, the similarity value is relatively flat, which shows that its lexical meaning has not changed much. By the Wei, Jin, North and South Dynasties to the Tang Dynasty, the similarity value of the related words dropped to the lowest point (30%), which indicates that the lexical meaning of “Shai” has changed compared to the previous one.

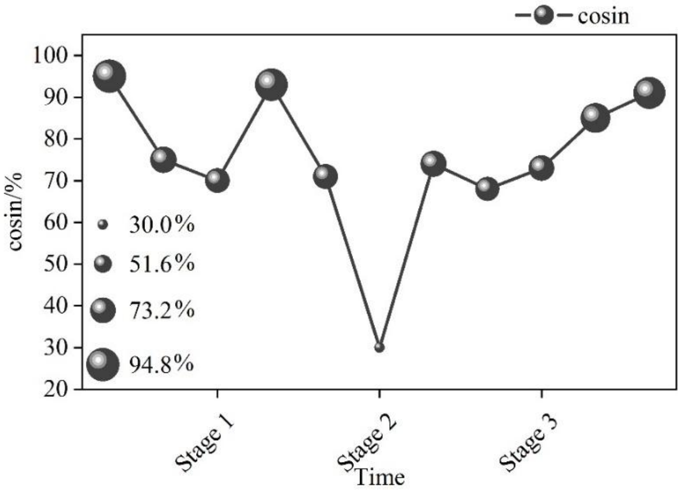

Figure 7: The meaning of "Shai" changes the curve

3.3 Summary of evolutionary features

The evolution of Chinese lexical meanings is jointly determined by external factors such as social development and internal factors of the language system's own development. From the experimental point of view, the evolution of Chinese lexical meaning over time is closely related to the change of time, mainly in the aspect of social development. The development of society breaks the relative balance between vocabulary and society, and people need more words to express the new things and new phenomena that keep appearing in the society, and they urgently need more words to express their inner thoughts, so there is a contradiction between the existing vocabulary meanings and people's needs for expression, and this contradiction pushes forward the development of word meanings, which is a solid foundation for the emergence of the phenomenon of Chinese word meanings. This conflict drives the development of word meaning and is the solid foundation for the phenomenon of Chinese word meaning.

4 Conclusion

For the lexical representation learning method across time, this paper adopts two-way LSTM language model combined with ELMo vector generation algorithm for word sense recognition.The average accuracy of ELMo model is around 75%, and it has stronger classification ability than LSTM model. The Word2vec tool is used to analyze the similarity of word meanings, and the word “Shai” is used as an example for the case study. The evolution of Chinese word meanings in traditional cultural texts is affected by time changes and social development, and new things and phenomena have prompted more words to express new ideas, and when there is a conflict between the word meanings of vocabulary words and the expression of new ideas, the evolution of the word meanings is pushed forward. When the lexical meanings of words conflict with the expression of new ideas, the evolution of lexical meaning is promoted. In addition, the shortcomings of the study are mainly in the Word2vec parameter settings, because of the need to take into account all the research subjects and the amount of corpus in each time period, some of the parameters may not be appropriate for the study of some words, which leads to a certain degree of error in the experiment.

About the Author

Jianping Xu was born in Yuncheng, Shanxi, P.R. China, in 1993. I graduated with both my undergraduate and master's degrees from Jilin Normal University, majoring in Teaching Chinese to Speakers of other Languages. At present, I am working as a teaching secretary in the Tianjin University of Finance and Economics Pearl River College. My main research direction is International Chinese Language Education.

References

[1] Chen, J., Chersoni, E., & Huang, C. R. (2022, May). Lexicon of changes: Towards the evaluation of diachronic semantic shift in Chinese. In Proceedings of the 3rd Workshop on Computational Approaches to Historical Language Change (pp. 113-118).

[2] Tang, X., Qu, W., & Chen, X. (2016). Semantic change computation: A successive approach. World Wide Web, 19(3), 375-415.

[3] Xing, J. Z. (2012). Semantic change in the grammaticalization of classifiers in Mandarin Chinese. Newest trends in the study of grammaticalization and lexicalization in Chinese, 169-214.

[4] Liu, M., Liang, Y., & Wan, Y. (2022, May). Novel lexical semantic change and interactivization. In Workshop on Chinese Lexical Semantics (pp. 125-138). Cham: Springer Nature Switzerland.

[5] Fredriksson, M. (2022). A Reflection on the Cultural Significance of the Protection of Classics. Stockholm Intellectual Property Law Review, (2), 8-13.

[6] Zhang, Y., Lou, J., & Cheng, Z. (2022). A study on the translation of cultural classics based on deep learning methods. Scientific Programming, 2022(1), 1026926.

[7] Chunlei, Y., & XinRu, L. (2022). Thinking on innovative mode of chinese traditional culture communication in the new era--A case study of China in the classics. Amazonia Investiga, 11(54), 189-198.

[8] Mou, H. (2017, October). Research on the Influence of Traditional Culture Classics on

Contemporary Education Based on Principal Factor Analysis. In 2017 International Conference on Robots & Intelligent System (ICRIS) (pp. 278-281). IEEE.

[9] Zhang, B., & Haoxuan, Z. (2025). The study of classic Chinese literature from the perspective of historical poetics. Neohelicon, 1-13.

[10] Wang, Y., & Zheng, D. (2021). The Birth of Classical Chinese and its Development―A Classical Case of 5E Imagination. Humanities Center Booklet, 9, 174-199.

[11] Shen, Y. (2015). Rethinking “Traditions”: Reading the Classics as Ritual. Journal of Chinese Humanities, 1(2), 363-383.

[12] Yu, R. (2020). On expansion and features of word-formation between English and Chinese new words. Theory and Practice in Language Studies, 10(8), 928-932.

[13] Qin, M. X., & Guo, H. (2022). Word formation of Chinese English words: Evidence from the Chinese English dictionary. World Englishes, 41(2), 151-168.

[14] Walker, S. (2014). A narrowing of meaning: loss of narrative unity and the nature of design change. International Journal of Sustainable Design, 2(4), 283-296.

[15] Yan, X., & Minnhagen, P. (2015). Maximum entropy, word-frequency, Chinese characters, and multiple meanings. PloS one, 10(5), e0125592.

[16] Kuo, L. J., Ku, Y. M., Chen, Z., & Shih, C. Y. (2024). Acquisition of Chinese characters: the impact of character properties and the contribution of individual differences. Applied psycholinguistics, 45(6), 1114-1146.

[17] Liu, H., & Lin, S. (2024, July). A Study on Chinese Character Components and Transfer Learning in Literacy Instruction: A Case Study of Elementary School Students in Lower Grades. In 5th International Conference on Language, Art and Cultural Exchange (ICLACE 2024) (pp. 198-209). Atlantis Press.

[18] Li, X. L. (2022). Learning Chinese Characters: Academic vs Dynamic Approach. KEMANUSIAAN: The Asian Journal of Humanities, 29(2).

[19] Anderson, R. C., Ku, Y. M., Li, W., Chen, X., Wu, X., & Shu, H. (2013). Learning to see the patterns in Chinese characters. Scientific Studies of Reading, 17(1), 41-56.

[20] Zhang, T., & Ke, C. (2018). Research on L2 Chinese character acquisition. The Routledge handbook of Chinese second language acquisition, 103-133.

[21] Hsiang, T. P., Graham, S., Wang, Z., & Gong, Y. (2021). How Chinese characters are taught: An analysis of three popular textbooks used in Macao. Technology, Knowledge and Learning, 26(2), 355-381.

[22] Chen, T. (2019). Joint contributions of multilevel linguistic knowledge to character meaning retention in L2 Chinese. Journal of Psycholinguistic Research, 48(1), 129-143.