ISSN 2616-5775 Vol. 6, Issue 10: 48-53, DOI: 10.25236/AJCIS.2023.061008

Recognition and Evaluation Algorithm for English Pronunciation Syllables Based on Neural Prediction Model

Qingjiao Lu

Department of Basic Courses, Hebei Agricultural University, Huanghua, Hebei, 061100, China
lqj0102@163.com

Abstract: *As the most widely used language in the world, English has always had the largest number of learners. Therefore, this study has a practical foundation for the recognition of English stressed syllables. As is well known, listening and speaking are crucial aspects of language learning, as they are directly related to communication. Therefore, this article aimed to design a mature syllable recognition algorithm and assist it based on neural prediction models. In the end, this article used the algorithm system for a month of auxiliary training for a certain English major class, and conducted a comparative test on phrase recognition rate and pronunciation accuracy before and after. The results showed that the phrase recognition rate increased from 89.34% to 96.05%, and the pronunciation accuracy rate increased from 73.65% to 92.84%, comprehensively improving students' English learning ability.*

Keywords: *Neural Prediction Model, Convolutional Neural Network, Recognition of Stressed Syllables, Spatiotemporal Neural Network*

1. Introduction

English listening and speaking training is crucial for English learning, with syllable recognition being even more crucial. Mol C believed that syllable order is very important in song recognition [1]. Flo A believed that tracking the syllables of words can help track the patterns of language learning [2]. Campos A D believed that reading and listening are key directions for children when learning language. Therefore, this article aimed to establish a mature algorithm for recognizing English stressed syllables to improve English listening and speaking learning [3]. Brennan M A pointed out that speech recognition systems need to distinguish between children and adults [4]. In addition to distinguishing the user's voice, this article believed that neural prediction models are also needed to assist the algorithm. Anand C proposed using neural prediction models for stock price simulation [5]. Chae S proposed using neural prediction models to predict and simulate the distribution of air particulate matter [6]. It is not difficult to see that neural prediction models have certain predictive simulation capabilities.

In fact, neural prediction models have been applied in syllable recognition. Ramirez-Mendoza A M E believed that a fuzzy adaptive neuron can be applied to syllable speech recognition [7]. Khanzadi M hoped to use recurrent neural networks for speech awareness assessment and syllable recognition [8]. Pahuja H's work was to establish a syllable recognition system using visual speech recognition using convolutional neural networks [9]. Wang D believed that the study of word syllables is crucial in many alphabetic language texts [10]. Therefore, this article hoped to draw on previous research experience to design algorithms for recognizing and evaluating stressed syllables.

This article first explained the principles of syllable differentiation in English and presented various common pronunciation errors. It also distinguished syllables from various perspectives such as open syllables and closed syllables. Then, based on the principle of distinguishing stressed syllables mentioned earlier, this article constructed a syllable recognition algorithm and introduced a neural prediction model to iteratively train the neural algorithm to adapt to the text library of syllable recognition algorithms. In this way, the neural prediction model can adapt to the work of assisting syllable recognition. Finally, this paper verified the feasibility of the proposed approach through comparative experiments.

ISSN 2616-5775 Vol. 6, Issue 10: 48-53, DOI: 10.25236/AJCIS.2023.061008

2. Rules and Principles of English Syllable Stress

2.1 Principles of English Syllable Differentiation

In order to scientifically and reasonably distinguish stressed syllables in English words, this article believes that it is necessary to divide different syllables in different words. Zhou P proposed pronunciation as a necessary condition for language communication, and it is necessary to classify and analyze the phenomenon of stress and spelling errors in word pronunciation [11].

Table 1: Classification of pronunciations that are prone to errors

Pronunciation type	Fallible point
Consonant	Native language habits lead to choosing consonants that are similar to the native language.
Monophthong	Easy to procrastinate when producing unit tones.
Long vowel	It is difficult to maintain a full mouth shape for a long time when pronouncing long vowels.
Diphthong	Due to the influence of different regional accents, it is difficult to form a diphthong accent.

Table 2: Classification of stressed syllables

Pronunciation type	Characteristic
Simple syllable	A vowel letter containing only one sound
Compound syllable	A vowel letter with more than one pronunciation
Open syllable	Ending with a vowel sound
Closed syllable	Ending with a consonant
Absolute syllable	Distinguish between short sounds that are pronounced or not pronounced according to the convention of vowel letters
Relative syllable	Ending with consonants such as ch, ck, sh, etc

Tables 1 and 2 list common mispronunciations in English words, as well as some basic pronunciation classifications used to distinguish stress. Firstly, the main mistake prone point of consonants in pronunciations is that learners can easily incorporate the pronunciation habits of their mother tongue into English pronunciation, thus choosing similar mother tongue consonants to replace them, resulting in non-standard pronunciation. When pronouncing vowels, it is easy to unconsciously elongate the pronunciation. When pronouncing long vowels, it may be difficult to maintain a full mouth shape for a long time. When pronouncing diphthongs, it may be difficult to fix the special English style diphthong pronunciation due to different regional accents. Quoc T X pointed out that long-term native language habits can form obstacles to English learning [12].

In the classification of stressed syllables, this article also lists some perspectives for classification. To distinguish whether a word is a simple syllable or a compound syllable, how many vowels it has is observed. If there is only one, it is a simple syllable, and vice versa, it is a compound syllable. From another perspective, if the word ends directly with a vowel sound, it is an open syllable; if it ends with a consonant sound (plus a neutral or silent vowel), it is a closed syllable. Minkova D's research was about the pronunciation of open syllables in Middle English, and he believed that studying open syllables could solve the historical problem of Middle English [13]. Moreover, there is a distinction between absolute and relative syllables, which can be applied to any of the preceding classification conditions. The absolute distinction is based on observing the way vowels are pronounced, which depends on whether the syllable is open or closed to observe whether the vowels belong to the original or variant pronunciation. Relativity refers to observing whether it ends with consonants such as "ch", "ck", or "sh".

2.2 Stress Syllable Mind Map

Shang W J pointed out in English learning that establishing mind maps can improve learning efficiency, and the direction of mind maps is English syllables [14].

ISSN 2616-5775 Vol. 6, Issue 10: 48-53, DOI: 10.25236/AJCIS.2023.061008

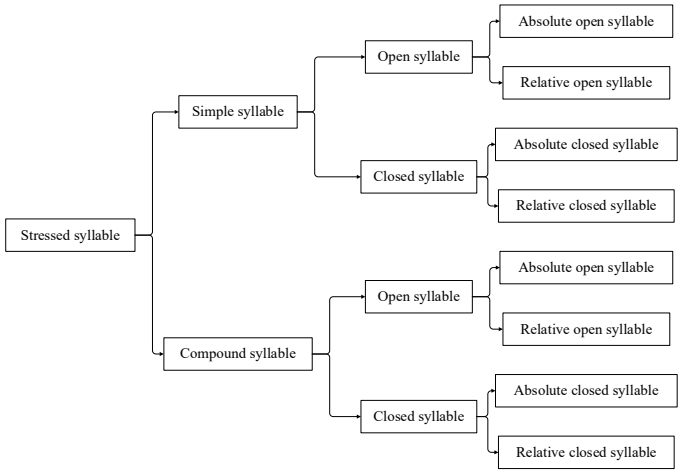

Figure 1: Classification rules for stressed syllables

By establishing a mind map of English syllables in Figure 1, it is clear where the rules for stressed syllables lie. In this map, stressed syllables can be divided into simple syllables and compound syllables, and each category can be divided into open syllables and closed syllables. All types of open and closed syllables have absolute and relative differences.

In order to have a more intuitive understanding of various types of syllables or words, this article provides some examples to assist in understanding. In simple syllables, the word with an absolute opening syllable is "I"; the words with relatively open syllables include "type" and "love"; the words with absolutely closed syllables include "nut" and "cast"; the words with relatively closed syllables include "medial" and "mental". Of course, there are also some special types of words such as "shirt", "cure", etc., which belong to special syllables. In compound syllables, the word with an absolutely open syllable is "die"; the words with relatively open syllables include "house" and "break the"; the words with absolutely closed syllables include "fruit" and "dead"; the words with relatively closed syllables include "cousin" and "other". Similarly, there are also some special syllable words, such as "four" and "bear".

3. Overview of Syllable Recognition Algorithms

In order to facilitate and thoroughly study English or other related research work, and solve the problem of stressed syllables, this article believes that a complete English syllable recognition algorithm is established based on the principle of distinguishing between stressed and prone to errors mentioned earlier. Zhang R P established an algorithm for recognizing Tibetan syllables in his research on Tibetan, and trained the recognition algorithm based on a feature recognition library [15].

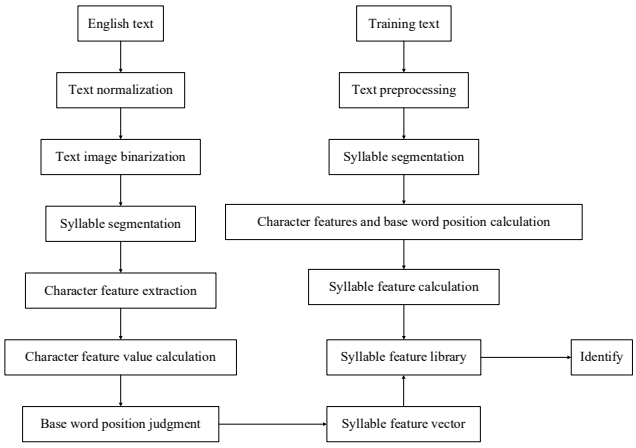

Figure 2: Process of syllable recognition algorithm

Figure 2 shows the process of syllable recognition algorithm. Firstly, it is necessary to standardize the English text, and then binarize the text image to perform syllable segmentation. Afterwards, the

ISSN 2616-5775 Vol. 6, Issue 10: 48-53, DOI: 10.25236/AJCIS.2023.061008

algorithm also needs to extract character features from the image and calculate the features. Then, it is necessary to determine the position of the base character based on the stressed syllables, determine the syllable feature vector, and search for the corresponding syllable features in the syllable feature library, in order to finally complete the recognition. From another perspective, the algorithm can also absorb a large amount of data from a text library, conduct text training, and then perform text preprocessing. Through syllable segmentation, the calculation of character features and base word positions can be completed. Finally, through syllable feature calculation, eligible syllables can also be found in the syllable feature library.

In order to perform image binarization, it is necessary to binarize a text image with a size of a×b. If the pixel grayscale matrix of the image is set to K, the pixel elements in the matrix are set to x, and P is set to x_{ij}, a specific threshold whose pixels must be between 0 and 255, that is, between black and white, then:

$$K = \begin{bmatrix} x_{11} & \cdots & x_{1b} \\ \cdots & \cdots & \cdots \\ x_{a1} & \cdots & x_{ab} \end{bmatrix} \tag{1}$$

$$x_{ij} = \begin{cases} 0, x_{ij} < P \\ 255, x_{ij} \geq P \end{cases} \tag{2}$$

Subsequently, the algorithm can extract the parameters of English character features after preprocessing the image. The column projection vector of matrix K is set as S, and the projection vector of a certain element x_{5b} is set as s_b, then:

$$s_j = \sum_{i=1}^{5} x_{ij}, j = 1,2,3,\ldots,b \tag{3}$$

$$S_j = f(s_j) = s_j^{\,2}, j = 1,2,3,\ldots,b \tag{4}$$

Then, the final feature parameter λ is:

$$\lambda = \sum_{i=1}^{5} \frac{S_i}{5} \tag{5}$$

After obtaining the feature parameters, the corresponding syllables can be found in the syllable feature library.

4. Neural Prediction Model

The neural prediction model is based on neural networks and can perform predictive analysis of time series data. This article believes that using the neural strategy model can effectively assist syllable recognition algorithms in identifying specific syllables and identifying dynamic features of a certain language signal. In order to enhance the matching degree between the neural prediction model and the syllable recognition algorithm, it is necessary to repeatedly train the model. Fan W proposed iterative training of neural network algorithms in his research on using convolutional neural networks to recognize Chinese character verification codes [16].

Table 3: Iterative training results

Iterations	Precision
25	81.37%
50	90.08%
75	97.63%
100	99.95%

According to Table 3, it can be seen that after multiple iterations of training for neural network algorithms, the testing accuracy has reached 99.95%, and its accuracy can be basically guaranteed.

After the iterative training of neural network algorithms is implemented, this article needs to establish a mature neural prediction model to assist syllable recognition algorithms. Fan H established

ISSN 2616-5775 Vol. 6, Issue 10: 48-53, DOI: 10.25236/AJCIS.2023.061008

neural prediction models in his research on spatiotemporal neural networks [17].

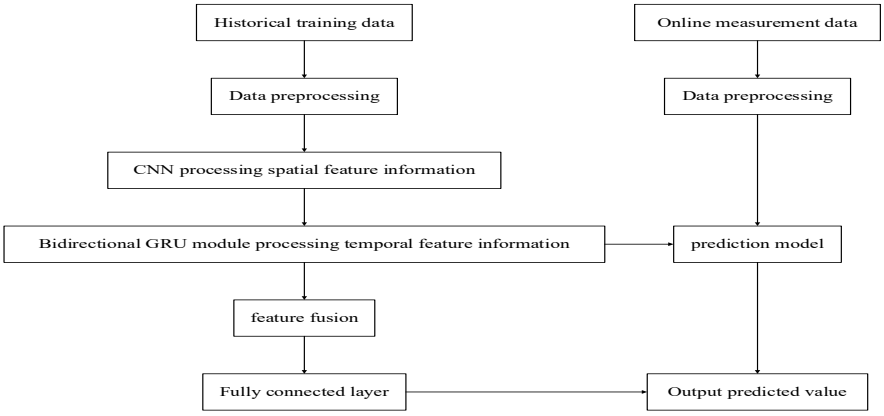

Figure 3: Spatiotemporal neural network prediction model

Figure 3 shows the workflow of the spatiotemporal neural network prediction model. In this model, it is first necessary to preprocess the historical training data and online measurement data separately, and then use CNN (Convolutional Neural Network) to process spatial feature information on the preprocessed historical training data. Sarvamangala D R proposed that convolutional neural networks are very suitable for processing image information [18]. Then, the bidirectional GRU (Gated Recurrent Unit) module is utilized to process temporal feature information. Ni Q used the GRU module to handle the prediction of bearing rolling [19]. At the same time, the processed information is also combined with the pre processed online measurement data, and the prediction is completed through a prediction model. It is also used for feature fusion and outputs predicted values through fully connected layers.

5. Comparative Experiments on Syllable Recognition Algorithms

Table 4: Comparative experimental results

Assessment criteria	Before training	After training
Pronunciation accuracy	73.65%	92.84%
Phrase recognition rate	89.34%	96.05%

In order to test the performance of the English stressed syllable recognition algorithm based on neural prediction models, this article selected students from a certain class of English majors in a certain university as the experimental subjects. The class was trained in English listening and speaking based on the algorithm for a period of one month. The daily duration was consistent with the previous routine learning, and the extracurricular assignments and daily tests were also consistent with the routine learning hours. Atmowardoyo H pointed out that English learning for the millennial generation needs to focus on listening, speaking, reading, and writing, with listening and speaking training being the most important [20]. The experimental group in this article conducted a phrase recognition and reading assessment on the class before and after training. The assessment subjects were the pronunciation accuracy of these phrase texts (especially for the stressed syllable part of the phrase) and the recognition rate of the phrase. The phrase was randomly selected from an algorithmic text library (the difficulty of the selected part basically met the teaching level of the major).

Table 4 shows the comparison of the results of the class before and after the experiment. From it, it can be seen that the pronunciation accuracy has increased from 73.65% to 92.84%, indicating that this algorithm has indeed greatly improved students' pronunciation problems. Moreover, the phrase recognition rate has also increased from 89.34% to 96.05%. As English majors, the experimental subjects' ability to recognize phrases is beyond doubt. The original 89.34% is already high enough, but through algorithm training, the recognition rate can be further improved. This indicates that novel and powerful syllable recognition algorithms can also stimulate students' learning enthusiasm and improve users' learning ability from other perspectives [21].

6. Conclusions

The idea of constructing syllable recognition algorithms based on neural prediction models in this article has been studied by many predecessors, so it has sufficient feasibility. However, the innovative

ISSN 2616-5775 Vol. 6, Issue 10: 48-53, DOI: 10.25236/AJCIS.2023.061008

approach of this article is based on the special principle of distinguishing stressed syllables, allowing syllable recognition algorithms to perform different recognition methods. However, there are also some shortcomings in the final experiment of this article, which is that it did not completely balance all variables in the two experiments. For example, the phrases used in the assessment were randomly selected, and the difficulty may fluctuated.

However, overall, the design concept of this article is still mature and usable, and research on the recognition of English stressed syllables would inevitably become more mature in the future, leading to a shift in more educational ideas from "reading and writing" to "listening and speaking", strengthening future students' English communication skills.

References

[1] Mol C, Bolhuis J J, Moorman S. Vocal learning in songbirds: the role of syllable order in song recognition. Philosophical Transactions of the Royal Society B, 2021, 376(1836): 1-7

[2] Flo A, Benjamin L, Palu M, et al. Sleeping neonates track transitional probabilities in speech but only retain the first syllable of words. Scientific reports, 2022, 12(1): 1-13

[3] Campos A D, Oliveira H M, Soares A P. Syllable effects in beginning and intermediate European-Portuguese readers: Evidence from a sandwich masked go/no-go lexical decision task. Journal of Child Language, 2021, 48(4): 699-716

[4] Brennan M A, McCreery R W, Massey J. Influence of audibility and distortion on recognition of reverberant speech for children and adults with hearing aid amplification. Journal of the American Academy of Audiology, 2022, 33(3): 170-180

[5] Anand C. Comparison of stock price prediction models using pre-trained neural networks. Journal of Ubiquitous Computing and Communication Technologies (UCCT), 2021, 3(2): 122-134

[6] Chae S, Shin J, Kwon S, et al. PM10 and PM2. 5 real-time prediction models using an interpolated convolutional neural network. Scientific Reports, 2021, 11(1): 1-9

[7] Ramirez-Mendoza A M E, Yu W, Li X. A new spike membership function for the recognition and processing of spatiotemporal spike patterns: syllable-based speech recognition application. Mathematics, 2023, 11(11): 1-28

[8] Khanzadi M, Veisi H, Alinaghizade R, et al. Persian phoneme and syllable recognition using recurrent neural networks for phonological awareness assessment. Journal of AI and Data Mining, 2022, 10(1): 117-126

[9] Pahuja H, Ranjan P, Ujlayan A, et al. Convolution Neural Network Based Visual Speech Recognition System for Syllable Identification. Recent Advances in Computer Science and Communications (Formerly: Recent Patents on Computer Science), 2022, 15(1): 139-150

[10] Wang D, Zeng M, Zhao H, et al. Effects of syllable boundaries in Tibetan reading. Scientific Reports, 2023, 13(1): 1-10

[11] Zhou P. Analysis and Solution of English Majors' Oral Pronunciation Errors and Vocabulary. Modernization of Education, 2018, 5(29):170-171

[12] Quoc T X, Thanh V Q, Dang T D M, et al. Teachers' perspectives and practices in teaching English pronunciation at Menglish center. International Journal of TESOL & Education, 2021, 1(2): 158-175

[13] Minkova D, Lefkowitz M. Middle English Open Syllable Lengthening (MEOSL) or Middle English Compensatory Lengthening (MECL)? English Language & Linguistics, 2021, 25(1): 155-180

[14] Shang W J, Wang G H. Application of Mind Mapping in English Learning Strategies. Journal of Changchun Normal liniversily, 2018, 37(1):183-186

[15] Zhang R P, Jiang Z C. Tibetan syllable recognition algorithm based on the features. Electronic Design Engineering, 2018, 26(20):137-142

[16] Fan W, Han J G, Gou F, Li S. Chinese character CAPTCHA recognition based on convolution neural network. Computer Engineering and Applications, 2018, 54(3):160-165

[17] Fan H, Zhang X M, Mei S W, Zhang Z L. A spatiotemporal neural network-based ultra-short term wind speed prediction model for wind farms. Automation of Electric Power Systems, 2021, 45(1):28-35

[18] Sarvamangala D R, Kulkarni R V. Convolutional neural networks in medical image understanding: a survey. Evolutionary intelligence, 2022, 15(1): 1-22

[19] Ni Q, Ji J C, Feng K. Data-driven prognostic scheme for bearings based on a novel health indicator and gated recurrent unit network. IEEE Transactions on Industrial Informatics, 2022, 19(2): 1301-1311

[20] Atmowardoyo H, Weda S, Sakkir G. Learning Strategies in English Skills used by Good Language Learners in Millennial Era: A Positive Case Study in Universitas Negeri Makassar. ELT Worldwide, 2021, 8(1): 28-40

[21] Sujanta Kazemanzadeh. Distributed System Integrating Virtual Reality Technology in English Teaching. Distributed Processing System (2022), Vol. 3, Issue 1: 62-70.