

Research Article

Automatic Correction System Design for English Pronunciation Errors Assisted by High-Sensitivity Acoustic Wave Sensors

Yue Sheng¹ and Kai Yang² 

¹Department of Business and Tourism, Henan Transportation Vocational and Technical College, Zhengzhou 450000, China

²School of Humanities and Law, Hebei University of Engineering, Handan, Hebei 056038, China

Correspondence should be addressed to Kai Yang; yangkai@hebeu.edu.cn

Received 23 September 2021; Revised 11 October 2021; Accepted 11 October 2021; Published 26 October 2021

Academic Editor: Guolong Shi

Copyright © 2021 Yue Sheng and Kai Yang. 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.

For correction system of English pronunciation errors, the level of correction performance and the reliability, practicability, and adaptability of information feedback are the main basis for evaluating its excellent comprehensive performance. In view of the disadvantages of traditional English pronunciation correction systems, such as failure to timely feedback and correct learners' pronunciation errors, slow improvement of learners' English proficiency, and even misleading learners, it is imperative to design a scientific and efficient automatic correction system for English pronunciation errors. High-sensitivity acoustic wave sensors can identify English pronunciation error signal and convert the dimension of collected pronunciation signal according to channel configuration information; acoustic wave sensors can then assist the automatic correction system of English pronunciation errors to filter out interference components in output signal, analyze real-time spectrum, and evaluate the sensitivity of the acoustic wave sensor. Therefore, on the basis of summarizing and analyzing previous research works, this paper expounds the current research status and significance of the design of automatic correction system for English pronunciation errors, elaborates the development background, current status and future challenges of high-sensitivity acoustic wave sensor technology, introduces the methods and principles of time-domain signal amplitude measurement and pronunciation signal preprocessing, carries out the optimization design of pronunciation recognition sensors, performs the improvement design of pronunciation recognition processors, proposes the hardware design of automatic correction system for English pronunciation errors based on the assistance of high-sensitivity acoustic wave sensors, analyzes the acquisition program design for English pronunciation errors, implements the parameter extraction of English pronunciation error signal, discusses the software design of automatic correction system for English pronunciation errors based on the assistance of high-sensitivity sound wave sensor, and finally, conducts system test and its result analysis. The study results show that the automatic correction system of English pronunciation errors assisted by the high-sensitivity acoustic wave sensors can realize the automatic correction of the amplitude linearity, sensitivity, repeatability error, and return error of English pronunciation errors, which has the robust functions of automatic real-time data collection, processing, saving, query, and retesting. The system can also minimize external interference and improve the accuracy of acoustic wave sensors' sensitivity calibration, and it provides functions such as reading and saving English pronunciation error signals and visual operation, which effectively improves the ease of use and completeness of the correction system. The study results in this paper provide a reference for the further researches on the automatic correction system design for English pronunciation errors assisted by high-sensitivity acoustic wave sensors.

1. Introduction

In the process of learning English, there is a phenomenon that some learners' spoken language is poor, and as a critical and difficult part of English learning, spoken language has

received increasing attention. Therefore, it is imperative to design a scientific and efficient automatic correction system for English pronunciation errors. The traditional English pronunciation correction system cannot provide timely feedback and correction for learners' pronunciation errors

and has disadvantages such as misleading learners and slow improvement of learners' English proficiency [1]. For the automatic correction system for English pronunciation errors, the level of correction performance and the reliability and practicability of information feedback are the main basis for evaluating its comprehensive performance. The quality of the correction algorithm determines the correction performance, and a reasonable error detection method guarantees [2]. After decomposing and optimizing each subtarget in the multitarget, the high-sensitivity acoustic wave sensor will trade off and coordinate them to make each subtarget. This is because the input information and output information required in the automatic correction of English pronunciation errors are related to the open failure system. The automatic correction system for English pronunciation errors can be divided into two parts: system training and pronunciation correction [3]. The training process of the system is similar to the training in the automatic pronunciation recognition system. The known standard pronunciation information features are extracted and recorded as the standard for pronunciation correction. Pronunciation correction is to correct the pronunciation accuracy of the pronunciation to be tested. The basic process is to extract the features of the pronunciation to be tested, compare its standard pronunciation features, and calculate the score based on the similarity [4].

The high-sensitivity acoustic wave sensor can follow the artificial neural network model, use target tracking to design an automatic correction system, and form an abstract logic layer by combining the characteristics of English pronunciation errors. The similarity between the single target tracking algorithm and the traditional neural network is that they both use a hierarchical structure to construct the logical layer, but the difference is that the three-layer construction mode is the most suitable for automatic correction system [5]. Relying on the optimized design of the pronunciation recognition sensor and the improved design of the pronunciation recognition processor, the software design of the system is completed based on the design of the English pronunciation acquisition program and the extraction of English pronunciation error signal parameters. In this process, although the amount of data is large and the calculations are more complicated, the calculation process of each sentence is the same [6]. It is necessary to use analog-digital signal conversion to improve the data sampling efficiency, and the sampling efficiency is not less than a certain value and the single-target tracking algorithm is used to continuously perform repeated iterative calculations [7]. In pronunciation recognition, a multifrequency oscillator is designed to automatically calibrate the pronunciation accuracy, while the calibration of the circuit conversion is the key to realize the conversion of the English printing information mode. By collecting and controlling the original pronunciation information of the circuit, the accuracy of system's automatic correction data can be improved [8].

Based on the summary and analysis of previous research results, this paper expounds the current research status and significance of the design of automatic correction system for English pronunciation errors, elaborates the develop-

ment background, current status, and future challenges of high-sensitivity acoustic wave sensor technology, introduces the methods and principles of time-domain signal amplitude measurement and pronunciation signal preprocessing, carries out the optimization design of pronunciation recognition sensors, performs the improvement design of pronunciation recognition processors, proposes the hardware design of automatic correction system for English pronunciation errors based on the assistance of high-sensitivity acoustic wave sensors, analyzes the acquisition program design for English pronunciation errors, implements the parameter extraction of English pronunciation error signal, discusses the software design of automatic correction system for English pronunciation errors based on the assistance of high-sensitivity sound wave sensor, and finally, conducts system test and its result analysis. The detailed chapters are arranged as follows: Section 2 introduces the methods and principles of time-domain signal amplitude measurement and pronunciation signal preprocessing; Section 3 proposes the hardware design of automatic correction system for English pronunciation errors based on the assistance of high-sensitivity acoustic wave sensors; Section 4 discusses the software design of automatic correction system for English pronunciation errors based on the assistance of high-sensitivity sound wave sensor; Section 5 conducts system test and its result analysis; Section 6 is the conclusion.

2. Methods and Principles

2.1. Amplitude Measurement of Time Domain Signal. From the perspective of the characteristics of the automatic correction system for English pronunciation errors; the assistance of the high-sensitivity acoustic wave sensor is actually a system function to obtain the required frequency response characteristics, and the same is true for digital filtering. For a linear time-invariant causal simulation system, the relationship between its input and output is

$$A(x) = a^n + \sum_{x=1}^n \frac{b(x)}{a}, \quad (1)$$

where x is the input of the system; $A(x)$ is the output response of the system; a is the continuous time component; $b(x)$ is the transfer function of the system; n is the number of convolution operators.

For the input English phoneme x of the system, given the observation vector x_i of each frame of the i th segment of pronunciation related to it, calculate its frame-based posterior probability as

$$B(x_i) = c(x) - \int \frac{d(x)}{A(x)} dx, \quad (2)$$

where $B(x_i)$ is the probability distribution of the observation vector x_i for a given phoneme x ; $c(x)$ is the prior probability of phoneme x ; $d(x_i)$ is the summation function of all text independent phonemes.

The design of the high-sensitivity acoustic wave sensor-assisted automatic correction system for English pronunciation errors has passed the first-level calibration to measure the sensitivity of the standard acoustic wave sensor, so the final calculation formula for the sensitivity $C(x)$ of the sensor under test is

$$C(x) = \frac{e(x) - f(x)}{g(x) - h(x)}, \quad (3)$$

where $e(x)$ is the sensitivity of the standard acoustic wave sensor; $f(x)$ is the sensitivity of the acoustic wave sensor to be measured; $g(x)$ is the amplitude of the acoustic wave sensor to be measured; $h(x)$ is the amplitude of the reference acoustic wave sensor.

The development of the correction system first recognizes the English pronunciation error signal and then performs dimensional conversion on the collected English pronunciation error signal according to the channel configuration information. Then, the high-sensitivity acoustic wave sensor is embedded in the correction system. The measurement process first filters the pronunciation signal to filter out the interference components in the output signal of the acoustic wave sensor; the system performs real-time spectrum analysis on the filtered English pronunciation error signal and evaluates the sensitivity of acoustic wave sensor [9]. Therefore, the system can minimize external interference and improve the accuracy of sensor sensitivity calibration when there is interference in the on-site environment. In addition, the software provides auxiliary functions such as reading and saving the pronunciation error signal and the operation of the visualization area to improve the ease of use and completeness of the system. The system uses a control signal source and an oscilloscope to complete the task of sending and collecting English pronunciation errors signals. Due to the limitation of the number of measurements, a loop control structure is added to measure the sensitivity of the sensor under test to achieve a certain number of cycles, and the oscilloscope collects signals are added to the program to ensure the integrity of signal reception and finally realize the task of channel triggering and channel reception.

2.2. Pronunciation Signal Preprocessing. After the high-sensitivity acoustic wave sensor calculates the ratio of signal input to output, the system function can be obtained by pulling and transforming the comparison value. The acoustic wave sensor is designed by the impulse response method, and the general form of the function $D(x)$ for pronunciation error correction is

$$D(x) = \frac{d[(j_{ix} + k_{ix} - l_{ix})e^{m_i}]}{dx}, \quad (4)$$

where j_{ix} is the acoustic wave sensor coefficient of the x th state at time i ; k_{ix} is the cumulative output probability of the x th state at time i ; l_{ix} is the previous state number of the x th state at time i ; m_i is the optimal state sequence at time i status.

The logarithm of the posterior probability of the phoneme x_i in the i th segment of pronunciation for each meal of the English pronunciation error signal is taken, and then, the logarithmic posterior probability $E(x_i)$ score of the phoneme x_i under the i th segment of pronunciation can be obtained:

$$E(x_i) = o(x_i) + \frac{1}{r} \int_i^n [p(x_i) - q(x_i)] dx, \quad (5)$$

where $o(x_i)$ is the duration of the i th time period corresponding to phone x_i ; $p(x_i)$ is the normalized function of the i th time period of phone x_i ; $q(x_i)$ is the likelihood of the segment of the i th time period of phone x_i ; r is the final output probability.

The sound wave sensor-assisted automatic correction system regards English pronunciation errors as a common pronunciation classification problem and uses a classification model to solve this problem. This model is based on a four-layer feed-forward network, which includes a pronunciation vector mapping table; the formula for inputting the input layer vector into the feed-forward network for forward calculation is as follows:

$$F(x) = \frac{s(x)}{t(x)} = \frac{v[u(x) + 1]}{w(x) + y(x)}, \quad (6)$$

where $f(x)$ is the network weight; $s(x)$ is the bias value; $t(x)$ is the activation function; $u(x)$ is the output value of the corresponding layer; $v(x)$ is the learning rate; $w(x)$ is the dimension of the pronunciation error; $y(x)$ is the size of the vector table of the pronunciation error.

English pronunciation error preprocessing includes sampling of English pronunciation errors, antialiasing band-pass filtering to remove individual pronunciation differences and noise effects caused by equipment and environment. English pronunciation error is an unstable random process, so it needs to use high-sensitivity acoustic wave sensor for short-term processing and involves primitive selection and endpoint detection of pronunciation recognition [10]. Endpoint detection refers to determining the start and end of pronunciation from English pronunciation errors, which is an important part of preprocessing. The process of pronunciation recognition is a process of digitally processing English pronunciation errors. Before processing English pronunciation errors, they must be digitally processed, and this process is analog-to-digital conversion. The analog-to-digital conversion process has to go through two processes, sampling and quantization, to obtain discrete digital signals in time and amplitude, and preemphasis is usually performed before transformation and after antialiasing filtering. After the system obtains learner's follow-up pronunciation, it extracts its characteristics and calculates the similarity between it and the standard pronunciation in the test question bank and finally maps the similarity to a grade score that is easier for the learner to understand and accept.

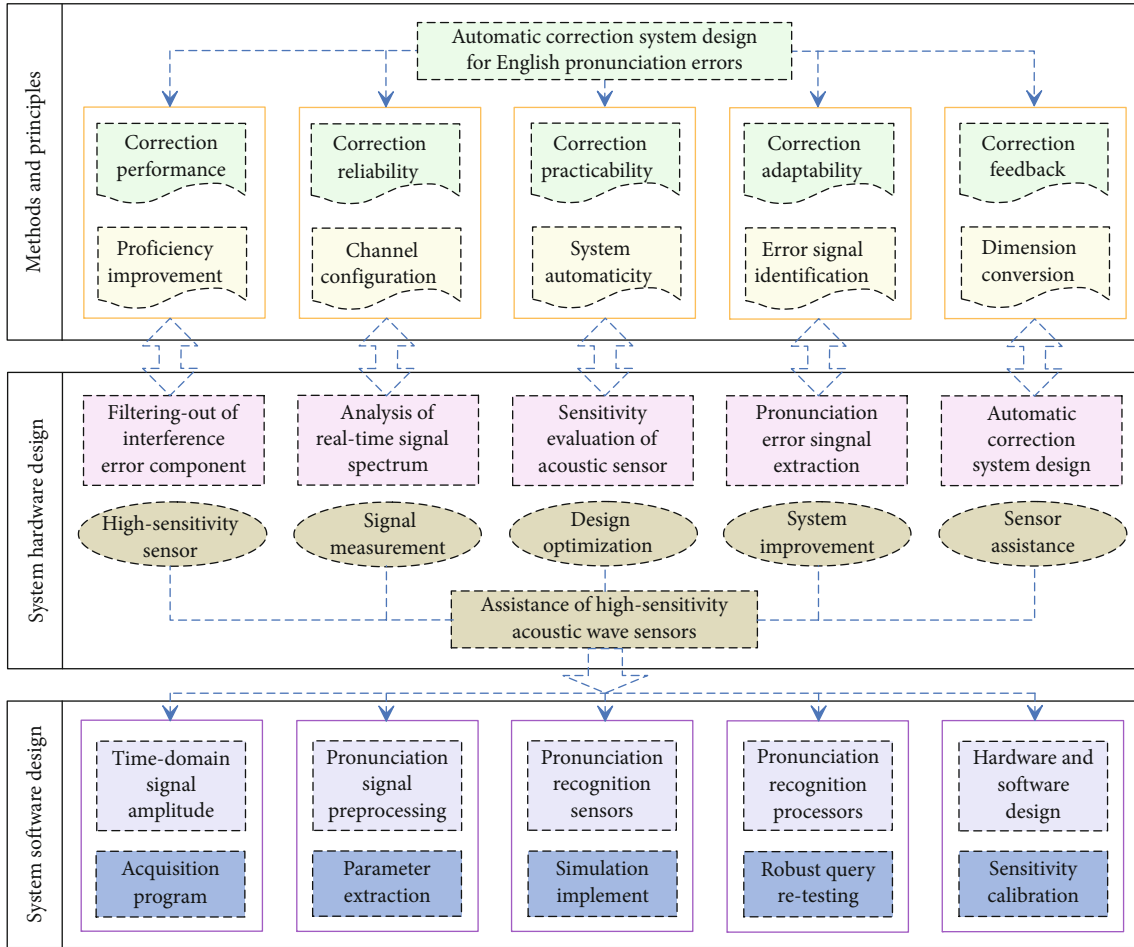


FIGURE 1: Automatic correction system design framework for English pronunciation errors assisted by high-sensitivity acoustic wave sensors.

3. Hardware Design of Automatic Correction System Based on High-Sensitivity Acoustic Wave Sensors

3.1. Optimization Design of Pronunciation Recognition Sensors. In order to ensure the accuracy, reliability, unity, and self-adaptability of English pronunciation errors correction and to adapt to the development trend of automatic correction, the system hardware design must carry out effective measurement supervision on the accuracy and reliability of the sound wave sensor's measurement value transmission to standardize and perfect the calibration of the sensor. The main components include raster data conditioning module, sensor output conditioning module to be calibrated, and acquisition device and computer system. This module can, respectively, realize the correction of the amplitude linearity, sensitivity, repeatability error, and return error of English pronunciation errors, and has the functions of automatic real-time data collection, data processing, storage, query, and remeasurement. The grating ruler is converted into the corresponding electrical signal through its conditioning circuit, and the corresponding processing is carried out by the formant acquisition card and the English pronunciation

signal is input into the system. The sensor to be calibrated outputs the corresponding voltage or current through its conditioning circuit, through data acquisition device, use the interface to achieve serial communication, set the data acquisition device in the reset state, establish the trigger condition, and initialize the control settings; the model can enter the working state, open the serial port, input the output signal into the computer system, and finally, respond with the collected data analysis and processing [11]. Figure 1 shows the automatic correction system design framework for English pronunciation errors assisted by high-sensitivity acoustic wave sensors.

The automatic correction of English pronunciation errors is more complicated than pronunciation recognition, but it can also be divided into front-end and back-end designs. The front-end processes the audio stream and undergoes preprocessing to separate the sound segments that may be vocalized and convert them into a series of values that can represent vowels in the signal. The back end is a dedicated search engine, which gets the output generated by the front end and uses the following three databases to search an acoustic model, a dictionary, and a language model to find the words that best match the sound

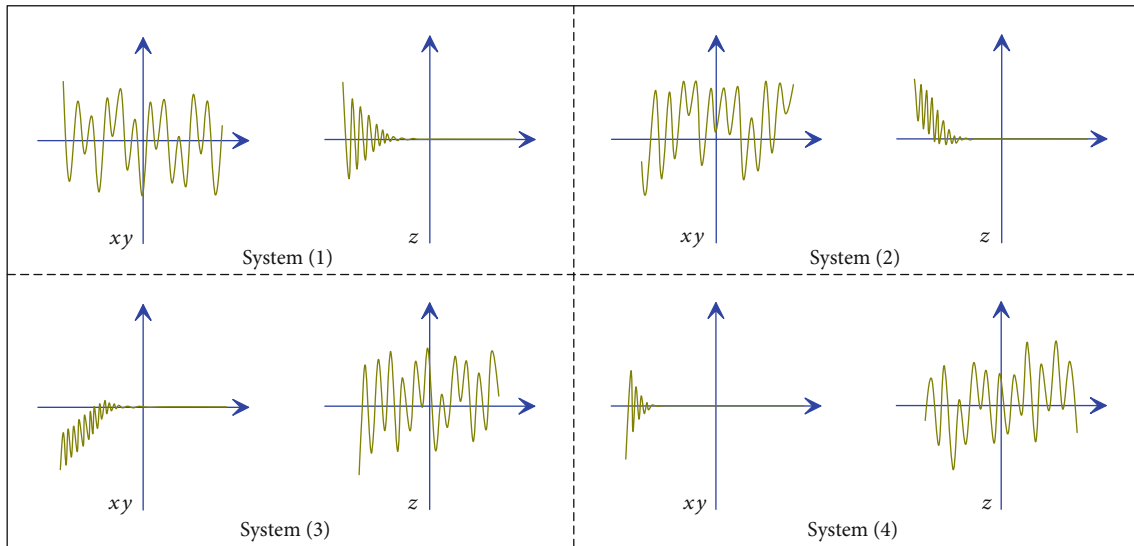


FIGURE 2: Optimization and improvement design of pronunciation recognition sensors and processors.

segment. The acoustic model represents the pronunciation characteristics of a language, which can be obtained by training the pronunciation patterns of a specific user and the characteristics of the pronunciation environment. For large vocabulary pronunciation recognition, subword units such as phonemes are used to build acoustic models. The language model represents how the words of a language are combined into phrases or sentences, that is, the language relationship between them. The purpose of language modeling is to effectively combine the knowledge of grammar and semantics in the process of pronunciation recognition, improve the recognition rate, and reduce the scope of search. It can be divided into a rule-based language model and a statistical-based language model. The rule-based language model is to summarize grammatical rules and even semantic rules and then use these rules to exclude the results of grammatical rules or semantic rules in acoustic recognition.

The main function of the English pronunciation formant graphic display module is to dynamically display the inconsistencies between the standardized English and learners' pronunciation audio with the help of the image form. The learners can clearly see the changes in this image. Know where your pronunciation is wrong and correct your pronunciation mistakes. The main working methods of this module are as follows. The first is to use preprocessing methods to initially label the audio; the second is to parse the audio using a fast conversion calculation method; the third is to find the formant of the audio and then convert it to visualize the image and accurately find out the sound that needs to be corrected [12]. The English pronunciation scoring module mainly uses pronunciation scoring technology to judge learners' pronunciation audio score and generate parameters; scientific scoring and automatic parameter generation are the core parts of this module. This module can compare the pronunciation audio of the learners with the standard pronunciation and preprocess the pronunciation audio of the learners by implementing endpoint detection and other methods. The purpose of this is that the

preprocessed audio can be converted into visual data information with more obvious characteristics, and then, the model is matched to obtain the frame average matching distance between speaker's pronunciation audio and the standard pronunciation.

3.2. Improvement Design of Pronunciation Recognition Processors. The pronunciation recognition processor assisted by the high-sensitivity acoustic wave sensor is improved and designed based on the original optimization. According to the high-sensitivity acoustic wave sensor, the analysis and recognition of English pronunciation is completed, which emphasizes the real-time nature of English pronunciation processing. The selection of the acoustic wave sensor is also very important to the improved design of the pronunciation recognition processor. It is not only related to the processing speed of English pronunciation but also involves the difficulty and progress of processor's improvement. Therefore, in addition to the control function and high-speed operation of the ordinary processor, the pronunciation recognition processor also improves the processor mechanism, instruction system, and flow for the requirements of real-time English pronunciation processing and improves the accuracy of correction. Under normal circumstances, the floating-point acoustic wave sensor has a high correction accuracy rate in the automatic correction of English pronunciation errors (Figure 2). The improvement of the pronunciation recognition processor should consider the processing speed and recognition accuracy of the chip. For an automatic English pronunciation system, the processing speed of pronunciation recognition is the most important, and the processor must be required to complete the corresponding processing tasks within a limited time. Otherwise, it will be difficult to guarantee the real-time processing of English pronunciation. When the pronunciation recognition processor is designed, the recognition speed of the processor chip is determined according to the high-sensitivity acoustic wave sensor and the processing time requirements.

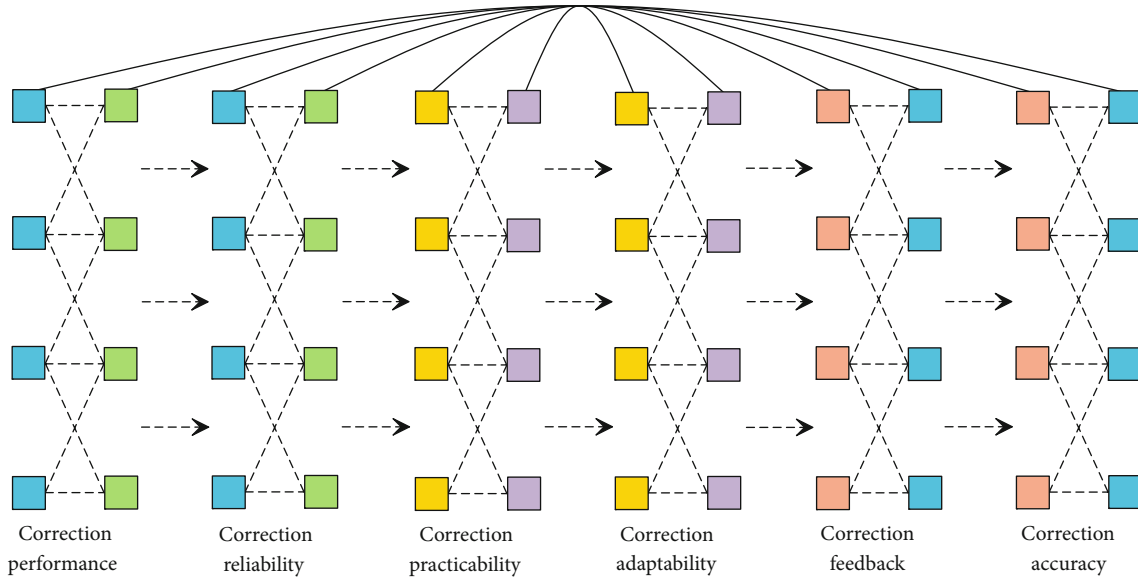


FIGURE 3: Hardware design of automatic correction system based on high-sensitivity acoustic wave sensors.

The models judge the file type according to the suffix name, use the corresponding function to open the file, and read the basic information of the file; if the file does not belong to the supported type, an error message will be given. The pronunciation database is the input content of the pronunciation automatic evaluation system, which mainly includes the standard phonetic database, dictionary database, and test phonetic database. The training process has a large amount of data and complex calculations, but the training process for each sentence is basically the same. The core of this process is an information processing process assisted by a high-sensitivity acoustic wave sensor, which obtains knowledge about English pronunciation errors by processing the available information of the monitored system. They read each sample value of the audio file cyclically, calculate the ordinate value according to the number of quantization bits, calculate the abscissa value according to the unit time length, combine the two to get a coordinate value, and then, draw points and lines (Figure 3). The training of the pronunciation model requires the use of a large amount of known audio information, and iterative and repeated training is completed, including audio reading, preprocessing, parameter extraction, text segmentation, and parameter estimation [13]. The standard pronunciation database stores the pronunciation data files of sentences and some tag information, and the dictionary database stores the pronunciation specification files of words, both of which are used to train the pronunciation model. The pronunciation library to be tested is used to test the performance of the pronunciation evaluation system.

The biggest flaw of the traditional automatic correction system of English pronunciation errors is that it tries to use limited rules to regulate the complex real language, which is determined by the characteristics of the English language itself. In addition, the problem of information conflict between the acoustic wave sensors also makes the rule-making after reaching a certain number, and then,

adding rules is not conducive to grammar checking. The conflict of rules of English pronunciation errors is very common. Some usages are established by convention and may not conform to the usual rules of English [14]. The conventional method uses writing more rules to distinguish the difference, but this may cause more conflicts, making the expansion of the rules infinite. Due to the potentially complex interactions between the effects of related features, there is no way to directly calculate the model parameters that maximize the possibility of training. Therefore, the high-sensitivity acoustic wave sensor uses iterative optimization technology to select model parameters; this technology uses random values to initialize the parameters of the model and then iteratively optimizes these parameters to make them closer to the optimal solution. These iterative optimization techniques ensure that every time they participate in optimization, they will be closer to the optimal value, but they may not be able to provide a method to determine whether the optimal value has been reached.

4. Software Design of Automatic Correction System Based on High-Sensitivity Acoustic Wave Sensor

4.1. Collection Program Design for English Pronunciation. English pronunciation error signal collection is the first process of the automatic correction system of English pronunciation errors. The acquisition program can automatically search for the wrong pronunciation signal, which not only improves the efficiency of correcting English pronunciation errors but also improves the accuracy of correction. The high-sensitivity acoustic wave sensor combines the requirements of real-time processing of English pronunciation, selects the appropriate processor software, accelerates the speed and accuracy of recognition, completes the improved design of pronunciation recognition processing, and realizes

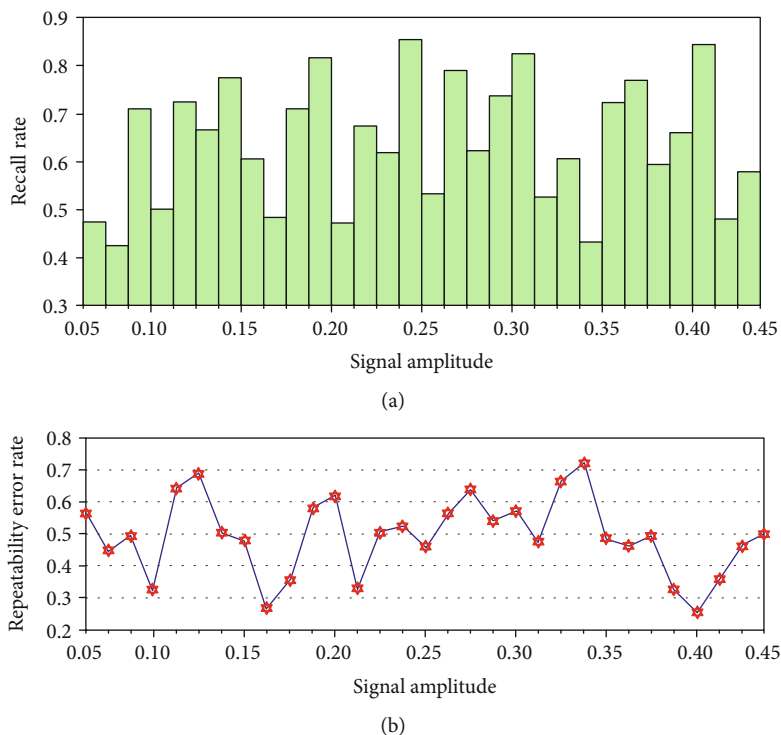


FIGURE 4: Recall rates (a) and repeatability error rates (b) with different signal amplitudes.

the software design of the system. They allocate storage space for the recognized English pronunciation buffer, start calculating error signal parameters, and finally, stop signal collection. If there is no pronunciation device in the program, the system will display a parameter error [15]. On the basis of installing the pronunciation device, initialize the English pronunciation error signal parameters to ensure that the received English pronunciation is in the format supported by the system. The recall rates and repeatability error rates with different signal amplitudes are shown in Figure 4. Relying on the optimized design of the pronunciation recognition sensor and the improved design of the pronunciation recognition processor, the software design of the system is completed based on the design of the English pronunciation acquisition program and the extraction of English pronunciation error signal parameters. Based on the high-sensitivity acoustic wave sensor, the design of an automatic correction system for English pronunciation errors has been realized.

The sensitivity calibration of the acoustic wave sensor is a program that realizes the collection, processing, and output functions of the English pronunciation error signal. The calibration system design is composed of signal collection, signal display, and sensor calibration. The signal collection part of English pronunciation errors is mainly to set the parameters of the collected signal; the signal display part is to display the amplitude, frequency, and amplitude-frequency curve of the collected signal; the sensor calibration part is set by the parameters of the instrument and select the sensor and sensitivity to be calibrated. The output shows three parts. The virtual signal generator generates the excitation signal, and its program is composed of the English pronunciation error signal control part and the display part.

The type, amplitude, and frequency of the generated waveform can be selected through the signal control part; the display part displays information such as the amplitude, frequency, and waveform of the generated waveform. The automatic correction system for English pronunciation errors installs a standard acoustic wave sensor with a known sensitivity and a calibrated acoustic wave sensor on the same interference table successively, so that they are subjected to the same frequency and the same magnitude of interference. Then, it measures the voltage output of the two acoustic wave sensors separately, calculates the interference acceleration of the interference table surface, and calculates the sensitivity value of the calibrated sensor from the output voltage value of the calibrated sensor.

The purpose of automatic correction of English pronunciation errors is to detect and correct the English pronunciation errors of interest and their causes as early as possible. The core of this process is an information processing process assisted by a high-sensitivity acoustic wave sensor, which obtains knowledge about English pronunciation errors by processing the available information of the monitored system. Model-based automatic correction of English pronunciation errors is based on the mathematical model of the high-sensitivity acoustic wave sensor; when the acoustic wave sensor cannot accurately describe the system or the deviation between the model and the actual object during the system operation, the designed correction system is difficult to guarantee expected diagnostic performance requirements. However, in a complex automatic correction system for English pronunciation errors, model errors and noise disturbances are inevitable. Therefore, it is necessary to establish a robust automatic correction algorithm for English

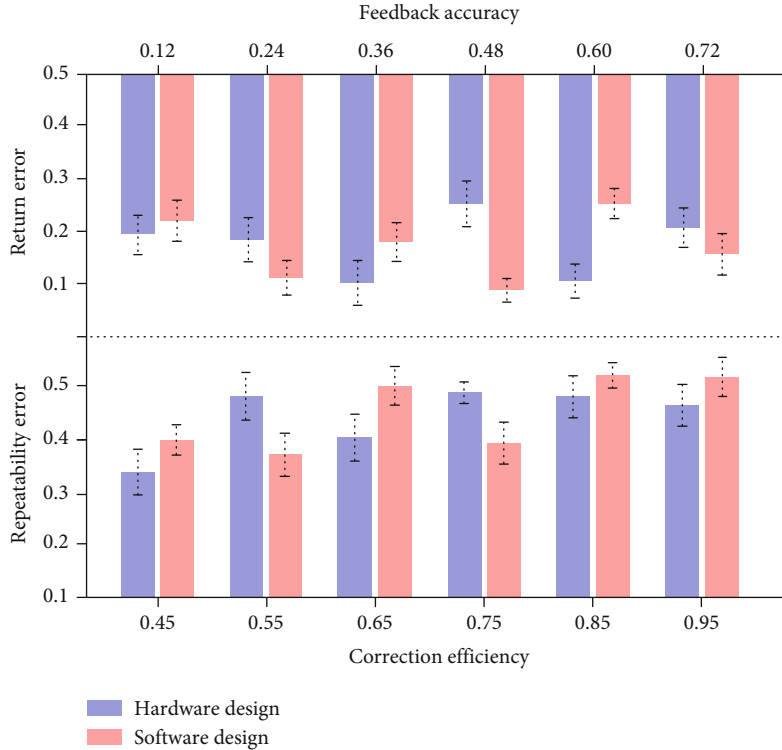


FIGURE 5: Return error and repeatability error in hardware and software design for automatic correction of English pronunciation errors assisted by high-sensitivity acoustic wave sensors.

pronunciation errors. The robustness of the automatic correction system for English pronunciation errors reflects the sensitivity of the high-sensitivity acoustic wave sensor to express English pronunciation errors signals. The parameter changes and interference affect the real process in an uncertain way, so it is necessary to design for English, which is very difficult for an automatic correction system for English pronunciation errors that is highly sensitive to pronunciation errors and at the same time insensitive to uncertainty and nonmodeling interference [16].

4.2. Parameter Extraction of English Pronunciation Error Signal. The high-sensitivity acoustic wave sensor can ensure the reasonable decomposition of the pronunciation segment and ensure that the pronunciation evaluation module of the automatic correction system for English pronunciation errors can receive information such as the extraction of evaluation parameters, the parameter association process, and the evaluation mechanism. The standard English pronunciation and the formant change of learner's pronunciation will be fully reflected in the image display module of the pronunciation formant in a graphical manner, clearly showing the difference between the two and identifying the difference between the correct pronunciation in the process and correct the wrong English pronunciation. The acoustic wave sensor compares the two results through the formant, which can clearly express the difference between learner's English pronunciation and the standard pronunciation and correct the wrong pronunciation. The system integrates the English pronunciation scoring module. English pronunciation scor-

ing and parameter generation constitute two core parts of the module both play an irreplaceable role in scientific scoring and adaptive generation of scoring parameters for English speakers (Figure 5). The English phoneme error detection program is a key part of the entire system, which can measure the weights of various parameters in English pronunciation, extract the wrong syllables in English pronunciation, and give feedback to the speaker to correct errors and improve, so that English speakers can improve their pronunciation level.

The high-sensitivity acoustic wave sensor-assisted automatic correction system for English pronunciation errors has an efficient correction algorithm to make system's correction performance close to the level of manual correction. At present, most systems basically reach this level after introducing advanced correction algorithms, if the system only gives a string digital correction which does not help learners improve their learning ability and improve pronunciation quality. Therefore, this method can only be used in the English pronunciation test system to make sense; finally, when the system provides English pronunciation error information feedback function, it is necessary to ensure these [17]. The accuracy of the feedback information, otherwise the learning system will not help users improve their language skills, but will make users become accustomed to repeating mistakes many times, so that system users cannot achieve the purpose of oral practice, completely deviating from the original intention of the correction system design (Figure 6). In addition to learner's own pronunciation errors, the individual pronunciation and pronunciation

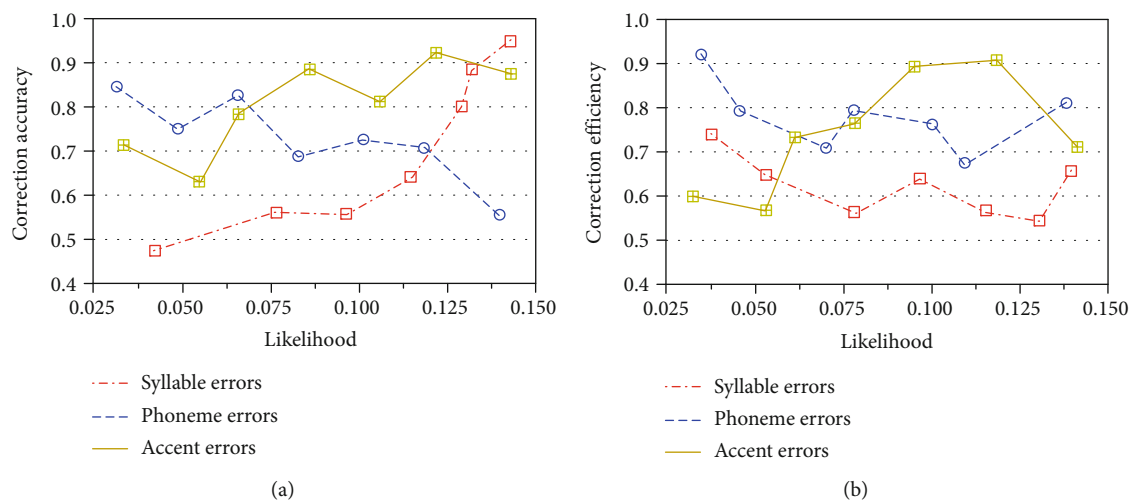


FIGURE 6: Relationship between automatic correction accuracy and pronunciation likelihood in English syllable, phoneme, and accent errors.

errors of the standard pronunciation, the pronunciation difference between the acoustic reference model of the recognizer and the standard pronunciation, and the complexity of the pronunciation text content may all cause the learner to obtain the result through the learning system. If only the correction of the learning system is compared with the correction of the standard pronunciation, the wrong judgment result that may be caused by other factors is not considered; for the learner, the purpose of learning English has not been achieved.

The high-sensitivity acoustic wave sensor-assisted automatic correction system for English pronunciation errors uses some information of the current state to automatically determine the specific method of pronunciation error correction. In the traditional correction system for English pronunciation errors, the information used generally includes some words, syllables, phonemes, or tones in English sentences, as well as some information in the current state stack and buffer [18]. In natural language processing, syllable or phoneme information will be very sparse if it is represented by a sound wave sensor. On the other hand, even very experienced linguists may not be able to summarize these signals and apply them to the analyzer. Most importantly, extracting the signal features of these English pronunciation errors can be very time-consuming. The predecessors tried a corrector that searches to find the best pronunciation tree. The main time of this type of corrector is spent on feature extraction, and the correction speed is also very slow. The high-sensitivity acoustic wave sensor-assisted automatic correction system for English pronunciation errors uses the very popular pronunciation vectors in the field of natural language processing to mention the above-mentioned features such as syllables or phonemes, so that the resulting features are much smaller in dimension than the original features, and neural networks can be easily used to train the data.

5. System Test and Result Analysis

5.1. System Test Experiment. From the perspective of training templates, the corpus is the fundamental source of

pronunciation knowledge required by the pronunciation recognition engine: from the perspective of performance evaluation. The quality of the corpus will directly affect the effectiveness of the evaluation results. In this test experiment, the effective reading time of each person participating in the system corpus recording is 45 minutes, the reading sentence is 1200 sentences, the total number of people is 120, 60 people are male and female, and the sentences cover 2517 daily English words. And a dedicated person will mark the time of each sentence at the word level. The setting of the initial value of the iteration in this algorithm is very important. Improper selection of the initial value will not only increase the number of iterations but also may cause the algorithm to diverge or iteratively converge to a nonglobal optimum. This paper assumes that there is no restriction on the transfer between models, that is, the export state of a model can be transferred to the population state of any model with equal probability. Therefore, this system first performs state segmentation and dynamic clustering on the feature vector of the pronunciation signal. Then, use fuzzy statistics to find out the parameters and then iteratively reevaluate. This is more evidence-based and effective than comparing the set parameters repeatedly based on experience. Not only can the number of iterations be reduced, but more importantly, the divergence of the model is avoided to a certain extent, and the parameters converge to the global optimum.

In order to reduce the impact of the incoherence of the English pronunciation error signal on the boundary of two adjacent frames, generally before processing each frame of pronunciation, the high-sensitivity acoustic wave sensor will multiply a window function. The main purpose of using the window function is to increase the continuity of the signals of two adjacent frames and to better preserve the overall characteristics of the signal. The short-term characteristics of English pronunciation error signals are closely related to the selection of window functions [19]. The selection of window functions can greatly affect the results of short-term analysis. A suitable window function should be able to better conform to the short-term characteristic parameters of the

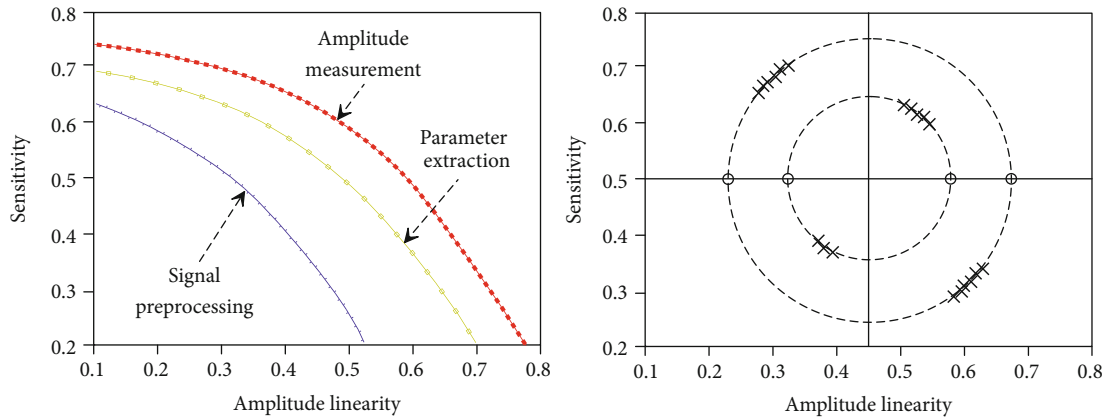


FIGURE 7: Sensitivity analyses of high-sensitivity acoustic wave sensors in the automatic correction system design for English pronunciation errors.

signal. Generally speaking, the shape of a good window should change slowly at the edge of the window to reduce the truncation effect of the English pronunciation error signal frame; from the perspective of the window length, the wider the window, the better the smoothing effect on the signal, but it is also easy to lose the signal details of the information. The acoustic wave sensor makes the signal lose the high frequency part, and it is easy to lose the signal details, but the spectral smoothness is better, which is more suitable for the time domain analysis of the system. The automatic correction system for English pronunciation errors adopts a rectangular window function in the time domain end point detection and adopts a window function in the short-time frequency domain transformation processing. Endpoint detection refers to the use of digital processing technology and related algorithms to find the start and end points of each segment of the pronunciation signal and exclude nonpronunciation segments.

5.2. Result Analysis. Almost all English pronunciation errors are handled based on this assumption. The process of dealing with English pronunciation errors requires the introduction of smooth process methods and theories into the short-term processing of English pronunciation errors. The automatic correction system assisted by the high-sensitivity acoustic wave sensor divides the English pronunciation errors into thousands of short segments, each of which is pronounced as one frame (Figure 7). Processing the analysis frame will process thousands of pairs of continuous pronunciations with fixed characteristics. After processing, a new time-dependent sequence will be generated from the original pronunciation sequence, which is used to describe the characteristics of English pronunciation errors. English pronunciation errors can be roughly divided into two parts: voiced and unvoiced [20]. At the beginning of the pronunciation activation period, it is often a low-level unvoiced sound. When the background noise is large, the unvoiced sound level is almost the same as the noise level. The high-sensitivity acoustic wave sensor adopts the traditional pronunciation endpoint detection method, which can easily cause the missed pronunciation of the pronunciation activa-

tion. The unvoiced segment of the English pronunciation signal plays a very important role in the quality of the pronunciation. For example, it may happen that weak fricatives and nasal sounds are cut off, the silent section of the blasting sound or the interval between words are mistakenly regarded as the end of pronunciation, and the impact noise is mistakenly judged as pronunciation.

The English pronunciation error detection system based on the equivalent space design of the optimal configuration of the acoustic wave sensor can be aimed at the linear time-invariant dynamic system. After analyzing the principle of the observer-based English pronunciation error detection, the problem of the optimal configuration of the sensor is described and based on the sensor. The optimal configuration of English pronunciation errors is in detection multiobjective optimization problem description. After decomposing and optimizing each subtarget in the multitarget, the high-sensitivity acoustic wave sensor will trade off and coordinate them to make each subtarget. This is because the input information and output information required in the automatic correction of English pronunciation errors are related to the open failure system. Therefore, in designing an automatic correction system for English pronunciation errors, it is not necessary to consider a controller, which is consistent with the separation principle in control theory, because automatic correction of English pronunciation errors can be widely regarded as an observation problem. Once the input to the actuator is available, the problem of automatic correction of English pronunciation errors is the same regardless of whether the system is operating in open loop or in closed loop. In addition, the high-sensitivity acoustic wave sensor-assisted automatic calibration is performed during system operation and is an online diagnosis, because the required system input and output information can only be used when the system is running [21].

English pronunciation demonstration means that when learning pronunciation, first, play the standard pronunciation lip animation video or standard pronunciation sound and at the same time cooperate with the pronunciation lip structure diagram and introduction text, so that learners have a correct understanding of the pronunciation and can

fully understand the pronunciation. The preprocessing of the high-sensitivity acoustic wave sensor includes preemphasis of the English pronunciation error signal, frame and window, and endpoint detection [22]. After the test pronunciation and standard pronunciation are preprocessed, feature extraction and pattern matching calculation are performed. The system uses feature parameters and dynamic time warping methods and the main points of pronunciation, the movement characteristics of mouth shape and tongue position, etc. After the above processing, the average matching distance of the frame between the test pronunciation and the standard reference pronunciation can be obtained. The acoustic formant display module of the acoustic wave sensor is mainly responsible for graphically depicting the change of the standard reference pronunciation and the formant change of learner's pronunciation over time, so as to qualitatively reflect the difference in the pronunciation of the two. If the sound wave sensor performs a pronunciation scoring operation, the system will substitute the average frame matching distance of the test pronunciation and the standard reference pronunciation into the scoring function and finally get the pronunciation score. For pronunciation recognition systems, accurate detection of pronunciation endpoints is the key to realizing the functions of the entire system.

6. Conclusions

This paper carries out the optimization design of pronunciation recognition sensors, performs the improvement design of pronunciation recognition processors, proposes the hardware design of automatic correction system for English pronunciation errors based on the assistance of high-sensitivity acoustic wave sensors, analyzes the acquisition program design for English pronunciation errors, implements the parameter extraction of English pronunciation error signal, discusses the software design of automatic correction system for English pronunciation errors based on the assistance of high-sensitivity sound wave sensor, and finally, conducts system test and its result analysis. The accuracy of the feedback information, otherwise the learning system will not help users improve their language skills, but will make users become accustomed to repeating mistakes many times, so that system users cannot achieve the purpose of oral practice, completely deviating from the original intention of the correction system design. The high-sensitivity acoustic wave sensor-assisted automatic correction system for English pronunciation errors uses the very popular pronunciation vectors in the field of natural language processing to mention the above-mentioned features such as syllables or phonemes, so that the resulting features are much smaller in dimension than the original features. The study results show that the automatic correction system of English pronunciation errors assisted by the high-sensitivity acoustic wave sensors can realize the automatic correction of the amplitude linearity, sensitivity, repeatability error, and return error of English pronunciation errors, which has the robust functions of automatic real-time data collection, processing, saving, query, and retesting. The system can also minimize external interference and improve the accuracy

of acoustic wave sensors' sensitivity calibration, and it provides functions such as reading and saving English pronunciation error signals and visual operation, which effectively improves the ease of use and completeness of the correction system. The study results in this paper provide a reference for the further researches on the automatic correction system design for English pronunciation errors assisted by high-sensitivity acoustic wave sensors.

Data Availability

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

Conflicts of Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This work was supported by the Social Science Foundation Project of Hebei Province (grant no. HB19YY023) and Innovation of English Teaching Project of Hebei Province (grant no. 2019YYJG034).

References

- [1] X. Liu, M. Xu, M. Li et al., "Improving English pronunciation via automatic speech recognition technology," *International Journal of Innovation and Learning*, vol. 25, no. 2, pp. 126–140, 2019.
- [2] L. F. S. Sidgi and A. J. Shaari, "The effect of automatic speech recognition EyeSpeak software on Iraqi students' English pronunciation: a pilot study," *Advances in Language and Literary Studies*, vol. 8, no. 2, pp. 48–54, 2017.
- [3] Y. Jin, "Design of students' spoken English pronunciation training system based on computer VB platform," *International Journal of Emerging Technologies in Learning*, vol. 14, no. 06, pp. 41–52, 2019.
- [4] O. Kolesnikova, "Comparative analysis of American English and Mexican Spanish consonants for computer assisted pronunciation training," *Revista Signos*, vol. 50, no. 94, pp. 195–216, 2017.
- [5] Y. H. Jung, S. K. Hong, H. S. Wang et al., "Flexible piezoelectric acoustic sensors and machine learning for speech processing," *Advanced Materials*, vol. 32, no. 35, p. 1904020, 2020.
- [6] C. Agarwal and P. Chakraborty, "A review of tools and techniques for computer aided pronunciation training (CAPT) in English," *Education and Information Technologies*, vol. 24, no. 6, pp. 3731–3743, 2019.
- [7] A. Kholis, "Elsa speak app: automatic speech recognition (ASR) for supplementing English pronunciation skills," *Pedagogy: Journal of English Language Teaching*, vol. 9, no. 1, pp. 01–14, 2021.
- [8] P. H. Tsai, "Beyond self-directed computer-assisted pronunciation learning: a qualitative investigation of a collaborative approach," *Computer Assisted Language Learning*, vol. 32, no. 7, pp. 713–744, 2019.

- [9] S. M. McCrocklin, "Pronunciation learner autonomy: the potential of automatic speech recognition," *System*, vol. 57, pp. 25–42, 2016.
- [10] J. Li, "An empirical study on reading aloud and learning English by the use of the reading assistant SRS," *International Journal of Emerging Technologies in Learning (iJET)*, vol. 15, no. 21, pp. 103–117, 2020.
- [11] S. Fu, H. Gu, and B. Yang, "The affordances of AI-enabled automatic scoring applications on learners' continuous learning intention: an empirical study in China," *British Journal of Educational Technology*, vol. 51, no. 5, pp. 1674–1692, 2020.
- [12] S. Ding, C. Liberatore, S. Sonsaat et al., "Golden speaker builder - an interactive tool for pronunciation training," *Speech Communication*, vol. 115, pp. 51–66, 2019.
- [13] Y. Wang and P. Zhao, "A probe into spoken English recognition in English education based on computer-aided comprehensive analysis," *International Journal of Emerging Technologies in Learning (iJET)*, vol. 15, no. 03, pp. 223–233, 2020.
- [14] C. Tejedor-García, D. Escudero-Mancebo, E. Cámara-Arenas, C. González-Ferreras, and V. Cardeñoso-Payo, "Assessing pronunciation improvement in students of English using a controlled computer-assisted pronunciation tool," *IEEE Transactions on Learning Technologies*, vol. 13, no. 2, pp. 269–282, 2020.
- [15] L. F. S. Sidgi and A. J. Shaari, "The usefulness of automatic speech recognition (ASR) Eyespeak software in improving Iraqi EFL students' pronunciation," *Advances in Language and Literary Studies*, vol. 8, no. 1, pp. 221–226, 2017.
- [16] M. Piotrowska, A. Czyżewski, T. Ciszewski, G. Korvel, A. Kurowski, and B. Kostek, "Evaluation of aspiration problems in L2 English pronunciation employing machine learning," *The Journal of the Acoustical Society of America*, vol. 150, no. 1, pp. 120–132, 2021.
- [17] M. Ramasari, "Students pronunciation error made in speaking for general communication," *Linguistic, English Education and Art Journal*, vol. 1, no. 1, pp. 37–48, 2017.
- [18] X. Qian, H. Meng, and F. Soong, "A two-pass framework of mispronunciation detection and diagnosis for computer-aided pronunciation training," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, vol. 24, no. 6, pp. 1020–1028, 2016.
- [19] P. Yu, Y. Pan, C. Li et al., "User-centred design for Chinese-oriented spoken English learning system," *Computer Assisted Language Learning*, vol. 29, no. 5, pp. 984–1000, 2016.
- [20] J. Phomprasert, "Creating awareness of incorrect English pronunciation in Thai elementary school by using the Detect Me English application," *Journal of Legal Entity Management and Local Innovation*, vol. 6, no. 4, pp. 181–191, 2020.
- [21] H. Bozorgian and E. Shamsi, "Computer-assisted pronunciation training on Iranian EFL learners' use of suprasegmental features: a case study," *Computer-Assisted Language Learning Electronic Journal*, vol. 21, no. 1, pp. 93–113, 2020.
- [22] G. G. Lee, H. Y. Lee, J. Song et al., "Automatic sentence stress feedback for non-native English learners," *Computer Speech and Language*, vol. 41, pp. 29–42, 2017.