SIN-CHN
SCIENTIFIC PRESS PTE.LTD.

Molecular & Cellular Biomechanics 2025, 22(2), 998.
https://doi.org/10.62617/mcb998

Article

Breath control strategies for English coherent pronunciation based on the biomechanics of respiratory muscles

Wenna Xu[1], **Tingyu Luan**[2,*], **Ying Zhang**[1], **Yu Sun**[1]

[1] International Department, Changchun Humanities and Sciences College, Changchun 130000, China
[2] College of Applied Technology, Dalian Ocean University, Wafangdian, Dalian 116300, China
*** Corresponding author:** Tingyu Luan, 15942420327@163.com

CITATION

Xu W, Luan T, Zhang Y, Sun Y. Breath control strategies for English coherent pronunciation based on the biomechanics of respiratory muscles. Molecular & Cellular Biomechanics. 2025; 22(2): 998. https://doi.org/10.62617/mcb998

ARTICLE INFO

Received: 4 December 2024
Accepted: 24 December 2024
Available online: 11 February 2025

Abstract: This paper aims to solve the problem of insufficient breath and lack of coordination between breathing and pronunciation in English learners' coherent pronunciation. By analyzing the biomechanical characteristics of respiratory muscles, scientific breath control strategies are designed to optimize airflow management and improve speech fluency and expression efficiency. English learners of different language levels were recruited and divided into elementary, intermediate and advanced groups. Speech tasks including slow pronunciation, fast pronunciation and long sentence pronunciation were designed. Electromyography was used to record the activity characteristics of respiratory muscles, airflow meter was used to measure airflow output data, and high-sensitivity microphone combined with Praat software was used to extract speech features including speaking rate and pause frequency. This paper studies tasks such as slow pronunciation, fast pronunciation and long sentence pronunciation, clarifies the degree of involvement of key muscles such as the diaphragm, and designs targeted phased breath control strategies, including abdominal breathing training, long sentence reading and coherent reading simulation, and the introduction of speech tasks. The results show that the breath control strategy proposed in this paper significantly improves learners' speech fluency, syllable and phonetic pronunciation accuracy under all task conditions. At a speaking speed of 180 words/minute, the correct pronunciation rate of phonetic symbols and syllables in this strategy intervention was 80%, and the correct pronunciation rate of syllables was 83%, and the breath maintenance rate and pause frequency were improved. The optimization strategy based on respiratory biomechanics can effectively improve language learners' breath management ability and coherent pronunciation level, and provide an innovative theoretical basis and practical path for language learning and voice training.

Keywords: English pronunciation; coherent pronunciation; breath control; respiratory muscles; biomechanical analysis

1. Introduction

In language learning, the fluency and coherence of English pronunciation [1,2] have always been the core problems faced by non-native learners. At present, mainstream English pronunciation training focuses on the control training of speech organs, but pays insufficient attention to the respiratory system. During the pronunciation process, learners often suffer from pronunciation interruptions, unnatural intonation, and even fatigue due to insufficient breath [3,4], excessive or improper ventilation, which directly affects speech fluency and communication effectiveness. This problem is particularly prominent in long-term vocalization tasks. From a biomechanical perspective, the respiratory muscles [5] are the driving force behind speech production. The respiratory system [6] provides stable airflow pressure for vocal cord vibration through the coordinated action of the diaphragm, intercostal muscles, and abdominal muscles, thus supporting the production and regulation of speech. Therefore, studying the biomechanical characteristics of respiratory muscles and designing scientific breath control strategies in combination with linguistic characteristics is the key path to solving the problem of coherent English pronunciation.

Interdisciplinary research has brought new possibilities to the theory and practice of language learning. Biomechanics, as a discipline that studies the laws of biological motion, has been widely used in sports medicine, rehabilitation training and sports science. Introducing biomechanical theory into the field of language teaching, especially the study of English coherent pronunciation, has important innovative significance. The biomechanical properties of the respiratory muscles directly determine the output intensity and stability of the airflow during pronunciation, and these properties show significant differences in different pronunciation situations. Through scientific experimental design and data analysis, the activity pattern of respiratory muscles and their relationship with speech characteristics are revealed, providing a theoretical basis for formulating efficient breath control strategies. Through multidisciplinary integration, the intrinsic relationship between the biomechanical characteristics of respiratory muscles and coherent English pronunciation is systematically studied. This can not only break through the bottleneck of traditional language teaching, but also promote the expansion of speech research from a purely linguistic perspective to a wider scientific field, providing a new solution for improving the speech ability of English learners.

2. Related work

Breathing exercises:

In English pronunciation teaching, the study of coherent pronunciation has always been an important direction in linguistics and phonetics. Language fluency [7,8] not only depends on the movement of speech organs and the mastery of pronunciation rules, but is also closely related to the control of the respiratory system. Biomechanical research shows that during pronunciation, the coordinated action of respiratory muscles such as the diaphragm, intercostal muscles and abdominal muscles determines the stability of airflow output and the continuity of speech. For non-native learners of English, coherent pronunciation requires learners to have good breath control ability, that is, to be able to reasonably distribute breath during speaking and avoid breathing or running out of breath in the middle. Traditional English pronunciation training methods mostly focus on phonetic symbols, syllables, intonation and speaking speed, while ignoring the crucial factor of breath control. In view of the role of respiratory muscles in the pronunciation process, some scholars have studied the relationship between breathing and pronunciation by combining exercise physiology and phonetics.

Abdominal breathing [9,10], as a method widely used in singing and speech training, has been proven to effectively improve the persistence and stability of pronunciation. Some studies have also used electromyography [11,12] to monitor the activity of respiratory muscles and explore breathing patterns in different pronunciation situations. The stability of breathing not only affects the clarity of pronunciation, but also has a significant impact on the emotional richness of language expression. In the expression of long sentences with coherent pronunciation in English, the reasonable distribution and use of breath is particularly important. Current research focuses on the regulation and training of breath, but lacks a systematic discussion of breath control strategies in English pronunciation.

Speech fluency:

In terms of the design and optimization of breath control strategies, many studies based on biomechanics and kinematics principles provide theoretical support for English pronunciation training. Abdominal breathing [13,14] is widely used in singing, public speaking, yoga and other fields. As an effective breath control method, it can help learners increase their lung capacity, improve the endurance of their respiratory muscles, and achieve stability and durability of airflow output. Abdominal breathing [15] training requires learners to control airflow output through the expansion and contraction of the abdomen. The core of this method is to enhance the coordination ability of the diaphragm and abdominal muscles, rather than relying solely on the upper muscles of chest breathing. The advantage of this breathing pattern is that it can minimize the waste of breath and ensure the long-term stability of airflow during pronunciation. In the field of phonetics, scholars have found through experimental research that the stability of breathing has a significant impact on speech fluency. During the pronunciation process [16], breathing not only provides airflow support for the vibration of the vocal cords, but also regulates characteristics such as pitch, speech speed, and speech rhythm. In particular, when speaking at a fast speed and pronouncing long sentences, the reasonable distribution of breath is particularly important. The “breathing-sounding coordination model” in phonetics emphasizes the close relationship between breathing and pronunciation, pointing out that only when the distribution of breath matches the production of speech can the clarity and coherence of pronunciation be ensured. Non-native learners have two typical breath control problems: one is insufficient breath [17] resulting in interrupted pronunciation, and the other is excessive breath consumption [18] resulting in unnatural speech rhythm. Therefore, breath control strategies not only need to focus on the depth and stability of breathing, but also need to consider the reasonable distribution and regulation of breath in different pronunciation situations.

Interdisciplinary research:

In terms of the practical application of breath control strategies, many studies have tried to combine breathing training with voice training to improve the fluency and naturalness of language expression. Taking singing and speaking as examples, breath control [19,20] is not only related to the loudness and persistence of the voice, but also directly affects the speaker’s emotional transmission and the audience’s understanding. In English pronunciation, good breath control plays an important role in the clarity, coherence and emotional expression of pronunciation. With the development of multidisciplinary cross-disciplinary research, the combination of biomechanics, linguistics and computer technology has provided new ideas for intelligent training of breath control.

Intelligent speech training systems based on artificial intelligence and big data can monitor learners' breathing conditions in real time [21] and provide targeted adjustment suggestions based on their speech output. These intelligent systems obtain learners' breath data through sensors and accurately evaluate the efficiency of breath use and the fluency of speech by comparing and analyzing the learners' speech output, helping learners to achieve self-adjustment and optimization during the pronunciation process. Compared with traditional manual guidance, intelligent training platforms can more efficiently track learners' progress, adjust training strategies in a timely manner, and improve learners' breath control ability and pronunciation fluency [22]. Although existing research has achieved some results in breath control training, it still faces some challenges, including how to accurately control breath in changing speech situations and how to design personalized training programs based on individual differences. These issues still need further exploration and optimization.

3. Methods

3.1. Selection and grouping of experimental subjects

Breath control in English coherent pronunciation is the core factor in improving speech fluency and clarity of expression. By properly managing breathing rhythm and airflow distribution, it can reduce problems such as stagnation of speech speed and unnatural punctuation caused by insufficient or unstable breath, and enhance the coherence of pronunciation and the naturalness of intonation. Breath control is particularly important for the pronunciation of long sentences, rapid communication or complex expressions. It not only helps learners master the skills of voice connection, but also improves their confidence in language expression and communication effectiveness. It is especially suitable for the practical needs of language teaching and oral training.

In order to ensure the scientificity and representativeness of the research results, this study plans to recruit 180 English learners whose native language is not English. These participants can be used as research samples. All participants must have some experience in learning English, but have not yet reached the level of a native speaker. Participants need to receive at least two years of formal English learning to ensure that they have basic phonetic knowledge and pronunciation skills.

The basic information of the participants is shown in **Table 1**.

The participants collected were evenly distributed in terms of gender, age, native language, etc. Males and females accounted for half of the total number, and the age range was 18–40 years old, with 45, 48, 42, and 45 people aged 18–25, 26–30, 31–35, and 36–40, respectively. Native languages include Chinese, Korean, Japanese, and Spanish, and people with different native language backgrounds are evenly distributed.

Table 1. Basic information of participants.

Basic information	Category	Number of people	Percentage (%)
Gender	Male	90	50.0%
	Female	90	50.0%
Age (years)	18–25	45	25.0%
	26–30	48	26.7%
	31–35	42	23.3%
	36–40	45	25.0%
Native language	Chinese	45	25.0%
	Korean	45	25.0%
	Japanese	45	25.0%
	Spanish	45	25.0%

During the registration stage, participants are required to provide their most recent English proficiency test scores, and grouping is based on the oral score of IELTS (International English Language Testing System). Participants can be divided into elementary group (60 people), intermediate group (60 people), and advanced group (60 people) according to their English speaking ability.

The oral part of IELTS [23,24] consists of three parts: Part 1 is a self-introduction and discussion on familiar topics; Part 2 is an independent statement after a short preparation; Part 3 is an in-depth discussion with the examiner. The grouping logic is shown in **Table 2**.

Table 2. Grouping logic.

Category	IELTS Speaking section score	Ability	Disadvantages
Beginner Group	4.0–5.5	Basic communication skills	There are obvious pronunciation problems
Intermediate Group	6.0–7.0	Can have a relatively fluent conversation	There are still deficiencies in complex expressions and speech coherence
Advanced Group	7.5–8.5	Shown in fluent and clear language output	There are slight deficiencies in the naturalness of speech

The score range of the IELTS speaking section is 0–9 points, with 0.5 as the unit. This paper selects 4.0–5.5 as the beginner group, 6.0–7.0 as the intermediate group, and 7.5–8.5 as the advanced group.

In order to comprehensively analyze the performance of participants' breath control in English coherent pronunciation, the paper followed the principle of gradual progress, set specific tasks for different language usage situations, and combined multi-dimensional speech feature extraction methods to ensure the scientificity and comprehensiveness of the experiment.

The task scenarios include slow pronunciation, fast pronunciation, and long sentence pronunciation, which are designed to simulate the complete process from basic language practice to complex speech output, and help researchers systematically analyze the mechanical characteristics and regulatory capabilities of respiratory muscles in different pronunciation scenarios.

In the slow pronunciation task, participants read a set of words and short sentences. The words include commonly used monosyllabic and disyllabic words, and the short sentences contain simple grammatical structures. This stage aims to eliminate the interference of complex language output on breathing patterns and observe the basic breathing rhythm of participants under slow speech speed and low pronunciation pressure, including the coordination of inhalation and exhalation.

In the rapid pronunciation task, a list of sentences in common life conversation situations is provided, and participants are required to read aloud at a natural speed and try to read them together. The rapid pronunciation task simulates a real communication situation and examines whether the participants can coordinate breathing and pronunciation under increased breath pressure.

In the long sentence pronunciation task, a paragraph-level text is provided, with a text length of 50–80 words. Participants are required to read the entire content in a clear and coherent tone. The paragraph content focuses on long sentences and complex grammatical structures. The long sentence pronunciation task tests the participants' extreme performance in breath distribution and control, especially their ability to cope with complex language output supported by a single breath for a long time.

3.2. Data collection and technical support

In order to study the role of the biomechanical characteristics of the respiratory muscles in English coherent pronunciation, this paper uses tools such as electromyography [25], airflow meter and high-sensitivity microphone to comprehensively collect data related to breathing and speech. These data can be used to analyze respiratory muscle activity, airflow output characteristics and speech characteristics, revealing the scientific basis of breath control strategies.

The display of respiratory muscles [26,27] is shown in **Figure 1**.

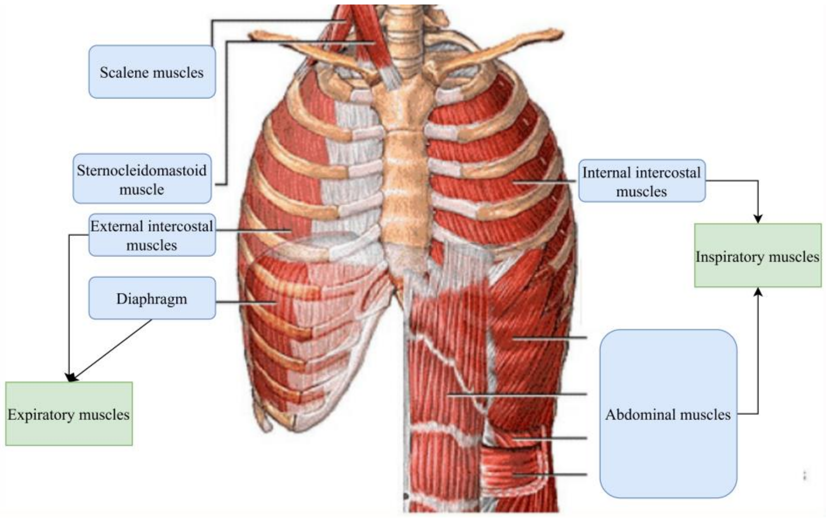

Figure 1. Respiratory muscles.

The respiratory muscles are the core muscle group responsible for driving respiratory activities, which are divided into inspiratory muscles and expiratory muscles. Inspiratory muscles include the diaphragm [28,29], external intercostal muscles, sternocleidomastoid muscles, and scalene muscles. Their main function is to expand the volume of the chest cavity through diaphragm contraction and rib expansion, creating negative pressure to drive air into the lungs. The expiratory muscles are mainly composed of the intercostal muscles and abdominal muscles. They push air out of the lungs by compressing the chest cavity and reducing the volume of the abdominal cavity. During quiet breathing, the inhalation process is active muscle contraction, and exhalation mainly relies on passive elastic retraction; while during deep breathing or vocalization, the expiratory muscles can actively participate to enhance the strength and control of exhalation.

The process of inhalation and exhalation is shown in **Figure 2**.

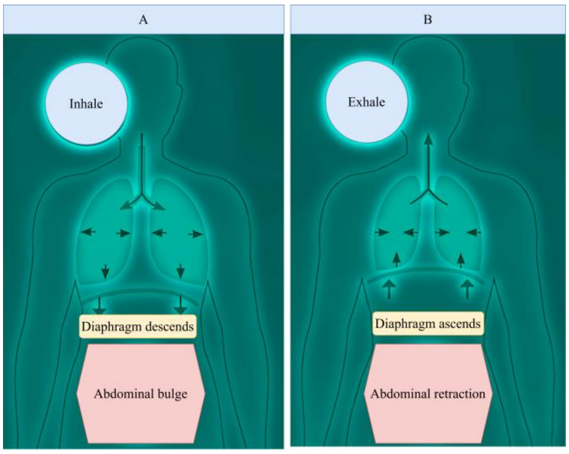

Figure 2. Inhalation and exhalation process: **(A)** Inhalation process; **(B)** Exhalation process.

Inhalation draws air in through the lungs, the abdomen expands, and the diaphragm descends. In exhalation, the air is expelled from the body, the abdomen shrinks, and the diaphragm rises. The main function of inhalation is to expand the chest cavity volume through the downward movement of the diaphragm and the contraction of the external intercostal muscles, reduce the pressure in the chest cavity, and allow the outside air to enter the lungs to complete the absorption of oxygen. The function of exhalation is to reduce the chest cavity volume through the relaxation of the diaphragm, the contraction of the internal intercostal muscles and the abdominal muscles, increase the pressure in the chest cavity, and expel the waste gas in the lungs to complete the emission of carbon dioxide. This process ensures the normal gas exchange and maintains the body’s metabolic needs.

Electromyography is used to record the contraction strength and frequency changes of the respiratory muscles to analyze the patterns of muscle activity in different pronunciation situations. The portable electromyography recorder was set to a sampling frequency of 1000 Hz to ensure sufficient time resolution. The electrodes were placed in the upper middle of the abdomen to record the contraction of the diaphragm during inspiration and phonation.

The electrodes were placed along the mid-axillary line between the 5th and 7th intercostal spaces on the lateral wall of the thorax to record the dynamic contraction characteristics of the intercostal muscles during inspiration and exhalation. The original signal is filtered through a 50 Hz notch filter to remove power frequency noise. A high-pass filter with a cutoff frequency of 20 Hz is used to filter low-frequency baseline drift, and envelope detection is used to extract the amplitude characteristics of muscle activity.

The equation for the root mean square value of the electromyographic signal is:

$$RMS = \sqrt{\frac{1}{N}\sum_{i=1}^{N} x_i^2} \tag{1}$$

In Equation (1), x_i represents the electromyographic signal value of the i-th sampling point. N represents the number of sampling points, and RMS reflects the overall strength of muscle contraction.

The equation for muscle activation frequency is:

$$f_{mean} = \frac{\sum_f P(f)\cdot f}{\sum_f P(f)} \tag{2}$$

$P(f)$ represents the value of the signal power spectrum density at frequency f. f_{mean} represents the contraction frequency characteristics used to characterize the respiratory muscle group.

A bidirectional digital airflow meter was used to measure the stability and intensity changes of airflow output to evaluate the breath distribution ability of participants in different pronunciation situations. The range is ±200 L/min and the time resolution is 10 ms. The airflow meter is calibrated using a standard airflow generator to ensure measurement accuracy. The maximum airflow rate during inspiration, the maximum airflow rate during exhalation, and the duration of exhalation are recorded.

The equation for expiratory volume is:

$$V_e = \int_0^T Q(t)dt \tag{3}$$

$Q(t)$ represents the airflow rate, and T is the duration of exhalation.

The voice features are collected through a highly sensitive microphone to evaluate the impact of breath control on voice output, including analysis of pitch, volume, speech speed, voice rhythm and other data.

4. Biomechanical characteristics analysis

This paper analyzes the activity patterns of the respiratory muscles under different tasks to reveal the biomechanical characteristics and regulatory mechanisms of the respiratory muscles when completing various speech tasks, which is of great significance for understanding the generation and regulation of breath during vocalization. By studying slow pronunciation, fast pronunciation, and long sentence pronunciation tasks, this paper can clarify the involvement, contraction pattern, and coordination changes of key muscles such as the diaphragm, and provide a scientific basis for identifying the specific demands of different language tasks on the respiratory muscles. It can promote the development of scientific breath control methods, help English learners enhance the stability and continuity of speech output, and improve the efficiency and quality of language learning. It also provides new perspectives and data support for speech pathology research. The beginner, intermediate, and advanced groups were tested for slow pronunciation, fast pronunciation, and long sentence pronunciation, 10 times each, and the load on different muscles in the task was assessed by recording the amplitude of electrical activity. The results of the amplitude of electrical activity during slow pronunciation are shown in **Figure 3**.

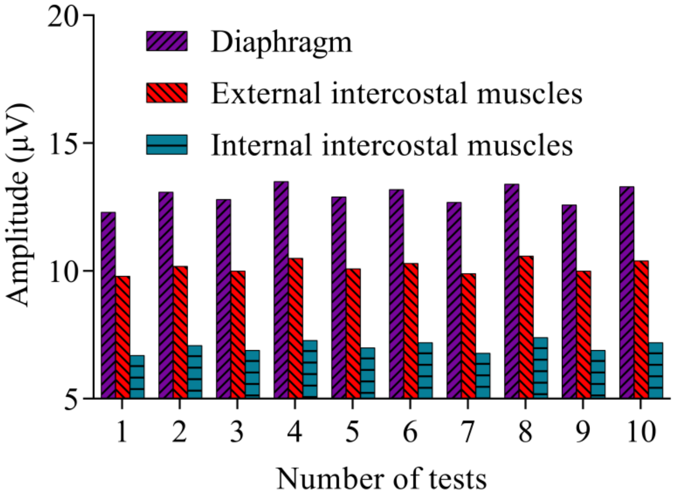

Figure 3. Results of electrical activity amplitude in slow pronunciation.

In the slow pronunciation task, the average electrical activity amplitude of the diaphragm was 13.0 μV, indicating that its activity intensity was high and stable, and it was the main source of power during breathing. The average electrical activity amplitude of the external intercostal muscles was 10.2 μV, indicating that they play an important role in expanding the volume of the chest cavity and assisting the diaphragm to complete inspiration, but their activity level is slightly lower than that of the diaphragm. The average electrical activity amplitude of the internal intercostal muscles was 7.1 μV, reflecting that their contribution to slow pronunciation is mainly concentrated in the exhalation stage. The low activity amplitude indicates that this task does not require high exhalation control.

The electrical activities of the various muscle groups showed a certain degree of coordination: the correlation between the activity levels of the diaphragm and the external intercostal muscles was strong, indicating that they had a significant synergistic effect during the inspiration phase. The correlation between the internal intercostal muscles and other muscle groups was low, which was consistent with their secondary role. From the data distribution, the range of changes in the amplitude of muscle group activity during the task execution is small, reflecting that slow pronunciation has a lighter load on the respiratory muscles and is suitable for training breath control for elementary language learners. The consistency of multiple test results is high, indicating that the reliability of the experimental design and measurement tools is good.

The results of the amplitude of electrical activity in fast pronunciation are shown in **Figure 4**.

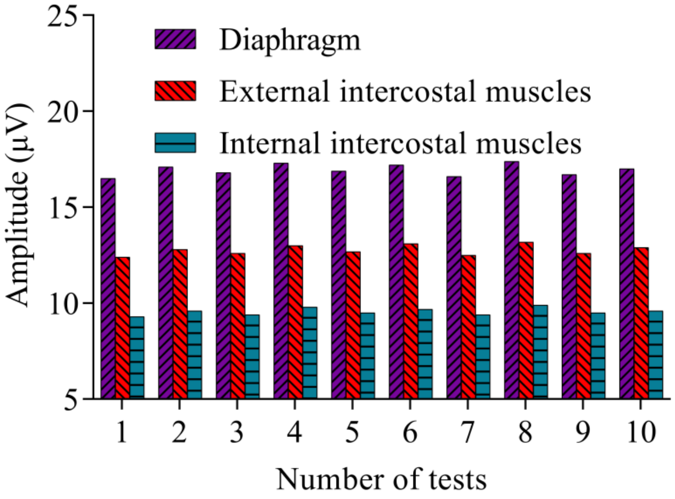

Figure 4. Results of electrical activity amplitude in rapid pronunciation.

In the rapid pronunciation task, the average electrical activity amplitude of the diaphragm was 17.0μV, showing that it exhibited stronger activity intensity in the high-speed speech task, indicating that rapid pronunciation requires higher inspiratory power and rapid airflow supply support. The mean amplitude of the external intercostal muscles was 12.8 μV, with a significant increase in their activity, highlighting their synergistic role in chest expansion; the mean amplitude of the internal intercostal muscles was 9.6 μV, showing a significantly enhanced need for expiratory control to achieve fast speech output compared to slow speech. The correlation between the activities of the diaphragm and the external intercostal muscles is still strong, reflecting the highly coordinated characteristics of the rapid inhalation process, indicating that more precise coordination is required between the respiratory muscles to cope with the high breath requirements of rapid pronunciation. From the test data of multiple times, the activity fluctuation range of each muscle group is slightly higher than that of slow pronunciation, indicating that there are certain challenges in maintaining breath stability in the rapid pronunciation task.

The results of the amplitude of electrical activity in long sentence pronunciation are shown in **Figure 5**.

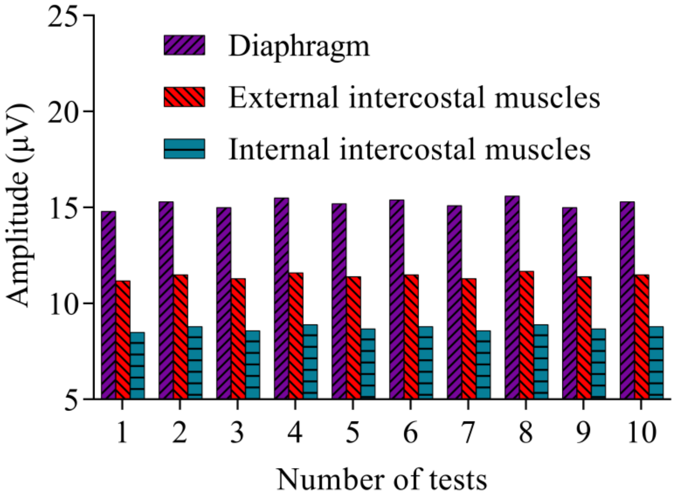

Figure 5. Results of electrical activity amplitude in long sentence pronunciation.

In the long sentence pronunciation task, the average electrical activity amplitude of the diaphragm was 15.2 μV, indicating that the diaphragm has a stable and high activity level in maintaining inspiratory control for a longer period of time. The average amplitude of the external intercostal muscles was 11.4 μV, and the activity level was slightly lower than that of the diaphragm, but it still showed an important auxiliary role, mainly reflected in expanding the volume of the chest cavity to meet the inhalation needs of long sentences. The average amplitude of the internal intercostal muscles was 8.7 μV, and its activity was slightly weakened compared with fast pronunciation, but still higher than slow pronunciation, indicating that long sentence pronunciation has higher requirements for the continuity and intensity of airflow during the exhalation phase. The test results of long sentence pronunciation show that the range of muscle group activity amplitude changes is relatively stable, highlighting the continuity and coordination requirements of long sentence tasks for breathing control.

The duration of inhalation and exhalation under different tasks is shown in **Table 3**.

Table 3. Duration of inhalation and exhalation.

Task	**Inspiratory duration (s)**	**Exhalation duration (s)**	**Ratio**
Slow pronunciation	1.2	3.6	0.33
Fast pronunciation	0.8	1.5	0.53
Long sentence pronunciation	1.5	5.4	0.28

Judging from the duration of inhalation and exhalation and their ratio in different tasks, the duration of exhalation is significantly longer than that of inhalation, which is consistent with the characteristics of normal speech production. Inhalation is a rapid process, while exhalation needs to be gradually released during the process of speech production to maintain the stability of the sound flow. In the slow pronunciation task, the average duration of inhalation was 1.2 s, and the average duration of exhalation was 3.6 s. The inhalation/exhalation ratio was 0.33, indicating that the inhalation efficiency was high during slow pronunciation, and the exhalation phase was stable and lasted longer, which provided learners with more time to complete the clear pronunciation of each syllable. In the rapid pronunciation task, the duration of inhalation was shortened to 0.8 s, while the exhalation time was also reduced to 1.5 s, and the inhalation/exhalation ratio increased to 0.53, indicating that rapid pronunciation requires faster inhalation rhythm and higher exhalation intensity to match the increased speech rate.

The inhalation duration of the long sentence pronunciation task increased to 1.5 seconds, the exhalation duration reached 5.4 s, and the inhalation/exhalation ratio dropped to 0.28, indicating that the pronunciation of long sentences significantly increases the demand for sustained exhalation control. Learners need to store enough airflow in a single inhalation to complete longer speech output. Different tasks have significant differences in the dynamic requirements for inhalation and exhalation. The coordination and efficiency of the respiratory muscles are crucial to the speech performance under different tasks.

5. Design and effect evaluation of breath control strategies

5.1. Experimental environment

In order to ensure the reliability and repeatability of the experimental results, the experimental scene was strictly designed and recorded in detail.

Laboratory environment:

The experiment was conducted in a specially designed laboratory located in the inner area of the scientific research building, away from external noise sources. The laboratory has been acoustically optimized and equipped with high-efficiency sound insulation materials. The walls are made of multi-layer sound-absorbing panels, the ceiling is equipped with noise-reducing suspension devices, and the floor is covered with shock-absorbing pads to minimize external interference. The background noise level in the laboratory was measured by a professional sound level meter and controlled below 30 decibels, which meets the requirements of psychology and neuroscience experiments for a quiet environment. In addition, the light in the laboratory is soft and stable, and non-stroboscopic LED lights are used to avoid interference to participants due to light changes.

Experimental equipment and layout:

The experimental equipment includes high-precision EEG acquisition equipment (model Brain Products actiCHamp) for recording EEG signals; high-definition monitors (resolution 1920 × 1080) for presenting experimental tasks; and high-fidelity headphones (Beyerdynamic DT 770 Pro) for playing auditory stimuli. The experimental table is height-adjustable to ensure that participants can maintain a comfortable sitting posture during long experiments, and the experimental chair is equipped with lumbar support to reduce fatigue. The connecting cables between the devices are shielded to prevent electromagnetic interference.

Experimental environment of participants:

Before the experiment, participants were guided to the laboratory and adapted to the environment in a quiet preparation room for 10 min to stabilize their physiological and psychological state. During the experiment, participants wore high-fidelity headphones to ensure the consistency of the sound quality of auditory stimulation while avoiding interference from external noise. The volume of the headphones was pre-calibrated and set between 60–70 decibels, which can clearly present the stimulation without causing discomfort to the auditory system.

In order to eliminate visual interference, all areas in the laboratory except the monitor were covered with matte black cloth, and the participants’ line of sight was limited to the screen area. The temperature in the laboratory was kept constant at 22 °C and the humidity was kept at around 50% to provide a comfortable experimental environment. During the experiment, participants were asked to turn off electronic devices such as mobile phones to avoid electromagnetic signal interference.

Task execution environment:

During the task execution, the position of the participants’ hands and heads was fixed to reduce the influence of unnecessary motion artifacts on data acquisition. The head fixation device adopts a flexible design, which can limit large-scale movements without causing pressure on the participants. The instructions for the experimental tasks are presented on the screen in the form of text and voice to ensure that the participants accurately understand the task requirements.

During the experiment, a “experiment in progress” indicator light was set outside the laboratory to prevent other people from entering the laboratory by mistake. Only one experimental administrator was allowed to enter the laboratory, and his position was behind the participants to avoid direct visual contact to reduce interference. The experimental administrator observed the experimental progress in real time through the remote monitoring system to ensure that the participants’ performance in the experiment met expectations and avoid excessive manual intervention.

Data acquisition and monitoring:

Experimental data acquisition adopts real-time monitoring and redundant storage mechanism. The EEG signal acquisition device is connected to an independent experimental computer equipped with a high-performance processor and a large-capacity hard disk to ensure real-time processing and storage of data. The performance of the participants was recorded by high-definition cameras throughout the experiment. The camera position was carefully arranged so as not to cause visual pressure on the participants, but to capture their facial expressions and movement changes for later behavioral analysis.

The execution interface of the experimental task was controlled by specially developed software, which recorded the participant’s response time, accuracy and error type for each response. All experimental data are backed up to external storage devices in real time and uploaded to a secure cloud server to ensure data integrity and security.

Environmental calibration and testing:

Before the formal experiment began, the experimental environment was calibrated and tested several times. The background noise level, equipment operating status, and consistency of experimental task presentation were strictly checked. The stability of the laboratory environment was verified in 7 consecutive days of testing to ensure that the experimental conditions would not fluctuate due to external factors.

5.2. Breath control strategy design

This paper designs breath control strategies for different tasks to effectively optimize the coordination and adaptability of the respiratory muscles and improve the fluency and accuracy of pronunciation. By analyzing the activity characteristics of the respiratory muscles in slow, fast and long sentence pronunciation tasks, learners are guided to accurately adjust the breath intensity and distribution to ensure smooth airflow output. This enables learners to achieve natural and coherent pronunciation in complex contexts and enhance their practical application ability in spoken English.

The breath control strategies are shown in **Table 4**.

Table 4. Breath control strategies.

Stage	Strategy	Times/day	Duration (min)
Beginner	Abdominal breathing training	2–3	10
Intermediate	Long sentence reading and coherent reading simulation	2	20
Advanced	Introduction of speech tasks	2	15

The primary stage training is mainly for beginners. Through abdominal breathing, the control ability of the diaphragm and abdominal muscles is strengthened to lay the foundation for pronunciation stability. The training content includes static breathing exercises and basic pronunciation combined exercises. Static breathing exercises require students to sit or lie on their backs with their hands on their abdomens. They should feel the movement of the diaphragm by slowly inhaling to expand the abdomen and slowly exhaling to retract the abdomen. It is recommended that each exercise last for 10 min, 2–3 times a day. The basic pronunciation combined exercises combine simple vowel and consonant pronunciations, observe the coordination of inhalation and pronunciation, and guide students to extend the duration of single sounds by moderately controlling the airflow.

At the intermediate stage, the training difficulty is appropriately increased, focusing on the stability of the airflow and the precise control of the exhalation stage. The training methods include long sentence reading and coherent reading simulation. The long sentence reading training requires students to select texts containing multi-syllable words and complex sentence patterns for reading, and gradually increase the sentence length to practice the breath distribution strategy. Students are required to complete the continuous pronunciation of no less than 15 syllables after a single inhalation, while maintaining a stable tone and volume. The coherent reading simulation uses high-frequency sentences in daily conversation situations to guide students to naturally connect vocabulary and optimize the coordination of breath and coherent pronunciation.

The advanced stage training is aimed at students who already have basic breath control ability, focusing on improving their breath distribution ability and multi-tasking ability in complex speech tasks. The training content includes fast pronunciation tasks, situational dialogue simulations, and intonation change training. The rapid pronunciation task requires students to complete the pronunciation of the specified text at a high speed, focusing on observing the efficient coordination ability of the respiratory muscles. The situational dialogue simulation combines actual application scenarios, such as role-playing tasks, to enhance the flexibility and accuracy of exhalation distribution by constantly changing the voice content and emotional expression. Intonation training uses texts with rich emotions, and requires students to express different meanings and emotions by adjusting the depth of breathing and the intensity of airflow.

5.3. Fluency analysis

In order to verify the design effect of the breath control strategy, the strategy in this paper is compared with the traditional breath control strategy. The traditional breath control strategy is a simple chest breathing training, which relies on the expansion of the chest cavity for breath supply and lacks effective training of the abdominal muscles and diaphragm.

The pause frequency analysis was performed using Praat [30,31] to measure speech fluency, which measures language fluency by counting the number of pauses per minute in a speech segment. The proposed strategy and the traditional strategy were implemented for all participants. The fluency of the 10-min English speech is shown in **Figure 6**.

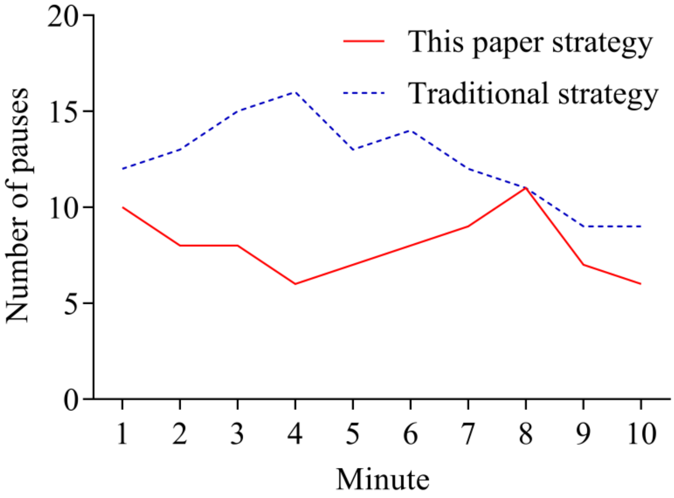

Figure 6. Fluency analysis.

The proposed strategy is significantly better than the traditional strategy in terms of the number of pauses per minute, showing higher speech fluency. In the 10-min speech task, the proposed strategy has an overall lower number of pauses per minute, with an average of 8.0 times, while the average number of pauses of the traditional strategy is 12.4 times. In the initial stage of the speech task (1st to 3rd min), the number of pauses of the proposed strategy was 10, 8, and 8, which was lower than the 12, 13, and 15 of the traditional strategy. This shows that the proposed strategy can quickly help learners adapt to the speech task and optimize the initial breath distribution and output stability. In the middle stage of the task (4th to 7th min), the proposed strategy further reduced the number of pauses to between 6 and 9, indicating that the proposed strategy has initially shown its strengthening effect on the endurance of the respiratory muscles, and learners can complete speech output more smoothly.

Especially in the 4th and 5th min, the number of pauses of the proposed strategy dropped to 6 and 7 times, while that of the traditional strategy was 16 and 13 times respectively, indicating that in tasks with higher complexity, the proposed strategy significantly improved learners' breath control ability. In the later stages of the task (8th to 10th min), the difference in the number of pauses between the proposed strategy and the traditional strategy gradually narrowed, but the proposed strategy still showed a certain advantage. In particular, in the 9th and 10th minutes, the proposed strategy had 7 and 6 pauses, while the traditional strategy was stable at around 9, which still has room for improvement. This phenomenon may be related to the accumulated fatigue of learners' breath control strategies in the later stages of the task.

The strategy in this paper successfully delayed the appearance of fatigue effects by optimizing airflow stability and muscle coordination. By introducing targeted respiratory muscle training, such as abdominal breathing and breath distribution optimization, breathing efficiency and speech coherence were effectively improved. The strategy in this paper combines the design of specific speech tasks to strengthen learners' breath control ability in complex contexts and reduce unnecessary pauses caused by lack of proficiency. In contrast, traditional strategies are not targeted enough for respiratory muscle training and lack a dynamic adjustment mechanism, which leads to learners' poor performance in high-intensity or long-term tasks, and high and unstable pauses. The proposed strategy has significant advantages in improving speech fluency and shows stronger fluency in English speech tasks.

5.4. Breath stability analysis

This paper measures the breath holding time of learners in pronunciation and sets the number of words in the sentence to 20, 25, 30, 35, and 35. The breath maintenance ratio can be observed. The breath maintenance ratio is the ratio of the time that can support continuous pronunciation after a single inhalation to the total time. The breath stability during pronunciation is shown in **Figure 7**.

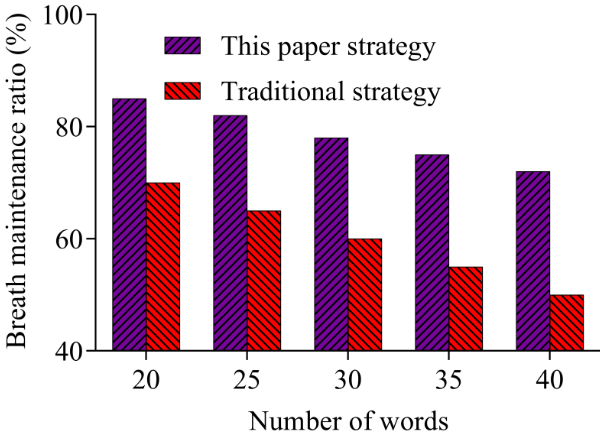

Figure 7. Breath stability.

The proposed strategy is significantly better than the traditional strategy in all word count conditions, showing higher breath efficiency and control ability. When the number of words is 20, the breath maintenance ratio of the proposed strategy is 85%, while that of the traditional strategy is only 70%, indicating that the proposed strategy can more effectively use breath to support phonation in short sentence pronunciation. As the number of words increases, the breath maintenance ratio gradually decreases, but the decrease in this strategy is relatively small. When the number of words is 40, this strategy maintains 72%, while the traditional strategy is only 50%.

The reason for this difference is that this strategy optimizes the endurance and coordination of the respiratory muscles by introducing abdominal breathing training and strengthening airflow stability, thereby reducing airflow leakage and improving the ability to continuously output breath. The traditional strategy relies on chest breathing and lacks sufficient exercise of the diaphragm and abdominal muscles, resulting in poor airflow stability and easy interruption. The proposed strategy can better support learners to complete the expression of complete semantic units in medium and long sentence tasks and reduce the frequency of pauses caused by insufficient breath, which plays an important role in improving speech fluency and expression coherence.

5.5. Analysis of the impact of pronunciation accuracy

This paper analyzes the correct pronunciation ratio of phonetic symbols and syllables under different speaking speed conditions, reveals the impact of speech speed changes on pronunciation accuracy, and provides targeted practice directions for language learners. The speech accuracy results are shown in **Figure 8**.

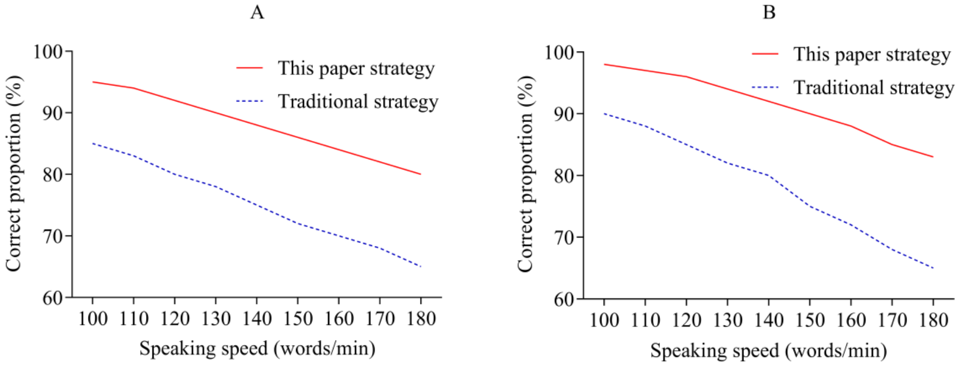

Figure 8. Speech accuracy: **(A)** Proportion of correct pronunciation of phonetic symbols; **(B)** Proportion of correct pronunciation of syllables.

In **Figure 8A**, the proposed strategy maintains a high accuracy of phonetic pronunciation at different speech speeds, and the accuracy is always higher than that of the traditional strategy. As the speaking speed increases, the accuracy of phonetic symbols gradually decreases for both the proposed strategy and the traditional strategy. The accuracy of the proposed strategy decreases at a slower rate, indicating that it can still ensure the accuracy of phonetic pronunciation at a higher speaking speed. The strategy in this paper optimizes breath control and coordination of respiratory muscles, ensuring that good pronunciation quality can be maintained at a higher speaking speed. The pronunciation accuracy of traditional strategies decreases significantly when the speaking speed increases. Traditional breath control training lacks systematicity and cannot effectively support accurate pronunciation at a fast speaking speed. The proposed strategy not only performs well in slow pronunciation, but also adapts well to changes in speech speed, ensuring speech clarity and accuracy, while the traditional strategy is prone to unclear or incorrect pronunciation at higher speech speeds. This difference further shows that optimizing breath control and training of respiratory muscles play an important role in improving speech pronunciation accuracy.

In **Figure 8B**, as the speaking speed increases, the accuracy of the syllables generally shows a downward trend. Under the traditional strategy, as the speaking speed increases, learners may not be able to effectively control their breathing and breath when pronouncing, resulting in a gradual decrease in the accuracy of syllable pronunciation. The strategy in this paper can maintain high syllable accuracy at a higher speaking speed by optimizing breath control and coordination of the respiratory muscles. Although the accuracy decreases with the increase in speaking speed, the decrease is relatively small. This shows that the breath control strategy proposed in this paper can effectively improve the pronunciation stability and accuracy of learners at a faster speaking speed, while the traditional strategy cannot maintain a good pronunciation effect under high speaking speed conditions. Breath control and pronunciation coordination are particularly important at a fast speaking speed, and the strategy proposed in this paper shows obvious advantages in this regard.

6. Discussion

In view of the potential impact of different language backgrounds on the experimental results in this study, the following improvement suggestions are proposed: First, the applicability of cross-language strategies can be explored, such as introducing multilingual pre-training models to enhance the model's ability to understand data in different languages. By jointly training on multilingual corpora, the model can capture the common features between languages, thereby improving cross-language adaptability and generalization. Secondly, intelligent technologies such as transfer learning and federated learning can be introduced. In transfer learning, the training results of high-resource languages are used as the initial model of low-resource languages, and fast adaptation is achieved through fine-tuning; while federated learning can collect distributed data in multiple language environments for joint optimization while protecting data privacy, thereby improving the overall performance of the model.

The study can also design cross-cultural experiments to evaluate the specific impact of language and cultural factors on model performance. For example, by comparing the experimental results in single-language and multi-lingual backgrounds, the performance differences of the model in the process of language transfer can be analyzed. To further optimize the effect, intelligent technologies can be combined with linguistic theories, such as semantic similarity analysis or language-specific feature extraction, to help the model understand the subtle differences between languages more accurately.

In future research, we will build multilingual annotated datasets and develop models based on multimodal information, combining language features with other information such as images and speech to achieve more comprehensive cross-language applicability. These improvements will provide more reliable theoretical and practical support for the application of intelligent technology in a multilingual context.

7. Conclusions

This paper systematically studies the activity patterns of respiratory muscles and breath control strategies under different task conditions, and verifies the effectiveness of breath control strategies based on respiratory biomechanics optimization in improving English coherent pronunciation. The research results show that the strategy in this paper not only significantly improves the accuracy of learners' syllable and phonetic pronunciation under high-speed speaking conditions, but also improves the breath maintenance ratio and pause frequency, providing theoretical support and practical guidance for non-native learners to improve their English oral fluency. This paper systematically combines respiratory biomechanics with speech pronunciation tasks for the first time and provides a phased breath control training method. This method and tool can be widely used in language learning, speech therapy and the design of related intelligent teaching systems, and has important practical significance.

The limitations of this study are mainly reflected in the fact that different language backgrounds may affect the universality of the experimental results. Since the experimental subjects are concentrated in a single language environment, cultural and language differences may lead to a decline in the performance of the model in cross-language applications. In addition, the sample size and diversity are limited, which may affect the generalizability of the results. Future research should expand the sample range to cover multilingual and multicultural backgrounds and explore the specific impact of language background on model adaptability. At the same time, the cross-language transfer ability of the model can be further optimized. It is assumed that the combination of multilingual data training and language-independent feature extraction methods can improve the versatility and robustness of the model, providing more extensive application support for related fields.

Author contributions: Conceptualization, WX and TL; methodology, YZ; software, YS; validation, WX, TL and YZ; formal analysis, YS; investigation, WX; resources, WX; data curation, WX; writing—original draft preparation, WX; writing—review and editing, WX; visualization, TL; supervision, YZ; project administration, YS; funding acquisition, TL. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by "Successful Education" Demonstration Classroom Project of Changchun Humanities and Sciences College [RWJY2024K28].

Ethical approval: Not applicable.

Conflict of interest: The authors declare no conflict of interest.

References

1. Utami, Tri Kurnia Putri. "Vietnamese English Accent Vs American English Accent: Locating the Phonological Variation." Linguistics and Literature Journal 3.2 (2022): 90-97.
2. Yuwita, Mia Rahmawati, and Nisrina Dewi Ambarwati. "Exploring University Students' Attitudes Towards Their English Accent and Native English Accents." Linguistics and Literature Journal 4.1 (2023): 21-27.
3. Spring, Ryan, and Ryuji Tabuchi. "Assessing the practicality of using an automatic speech recognition tool to teach English pronunciation online." Journal of English Teaching through Movies and Media 22.2 (2021): 93-104.
4. Hoque, Mohammed Shamsul, et al. "Untangling the Tongues: Exploring Pronunciation Development Strategies for Learners of English at High Schools in Oromia Region of Ethiopia." World Journal of English Language 14.5 (2024): 257-257.

5. Kholis, Adhan. “Elsa speak app: automatic speech recognition (ASR) for supplementing English pronunciation skills.” Pedagogy: Journal of English Language Teaching 9.1 (2021): 01-14.
6. Qiao Wei, Liu Su, Wang Ying, Hou Tingting, Song Sijin, Wang Siye, et al. “Effect of external diaphragmatic pacing therapy combined with abdominal muscle electrical stimulation on respiratory function in stroke patients.” Chinese Journal of Rehabilitation Theory and Practice 29.1 (2023): 104-109.
7. Sardegna, Veronica G. “Evidence in favor of a strategy-based model for English pronunciation instruction.” Language Teaching 55.3 (2022): 363-378.
8. Burri, Michael, and Amanda Baker. “‘I Feel… Slightly out of Touch’: A longitudinal study of teachers learning to teach English pronunciation over a six-year period.” Applied linguistics 42.4 (2021): 791-809.
9. Yeldham, Michael. “Abdominal enhancement techniques for L2 English phoneme pronunciation: Comparing the effectiveness of the techniques for Mandarin and Cantonese speakers.” Language Teaching Research 28.3 (2024): 1231-1254.
10. Chen, Zixin. “Analysis of The Role of Emphasizing the Diaphragm on Breath Control in The Teaching of Beginning Children’s Voices.” International Journal of Education and Humanities 12.3 (2024): 183-186.
11. AbuNurah, H. Y., D. W. Russell, and J. D. Lowman. “The validity of surface EMG of extra-diaphragmatic muscles in assessing respiratory responses during mechanical ventilation: a systematic review.” Pulmonology 26.6 (2020): 378-385.
12. Ramli, Muhammad Imran, Nur Azah Hamzaid, and Julia Patrick Engkasan. “Monitoring Breathing Muscle Performance during Singing Noninvasively Using Mechanomyography and Electromyography.” Journal of Voice 34.6 (2020): 862-869.
13. Li Ya, Xue Cui, Wu Yabo, Wu Meijuan, Wu Ye. “Effects of pursed lip-abdominal breathing combined with breathing exercises on exercise tolerance and dyspnea symptoms in patients with chronic obstructive pulmonary disease.” Nursing Practice and Research 17.12 (2020): 63-64.
14. Ahmad, Muhammad Faisal, Z. Abdullah, and A. W. Ali. “Theatre in Education: The Implementation of Improvisation in Dialogue Pronunciation Technique in Teaching.” International Journal of Academic Research in Business and Social Sciences 11.8 (2021): 726-735.
15. Wang, Yixuan. “The effectiveness of innovative technologies to manage vocal training: The knowledge of breathing physiology and conscious control in singing.” Education and Information Technologies 29.6 (2024): 7303-7319.
16. Kissova, Olga. “Contrastive analysis in teaching English pronunciation.” SWS Journal of Social Sciences and Art 2.1 (2020): 39-65.
17. Piotrowska, Magdalena, Andrzej Czyzewski, Tomasz Ciszewski, Grazina Korvel, Adam Kurowski, Bozena Kostek, et al. “Evaluation of aspiration problems in L2 English pronunciation employing machine learning.” The Journal of the Acoustical Society of America 150.1 (2021): 120-132.
18. Yan Meiwu. “A brief discussion on the basic skills of “articulation” and “word articulation” in song singing.” Teaching Method Innovation and Practice·Scientific Research 3.2 (2024): 7-9.
19. Zhu Zhiyuan. “On the importance and training methods of breath in vocal singing.” Creativity and Innovation (Chinese Version) 6.3 (2022): 159-160.
20. Caihua Hu. “On the breath and emotion issues in vocal teaching.” Teaching Method Innovation and Practice 3.10 (2020): 149-151.
21. Reyes, Alvaro, Javiera Castillo, Isabel Cornejo, Travis Cruickshank. “The effects of respiratory muscle training on phonatory measures in individuals with Parkinson’s disease.” Journal of Voice 34.6 (2020): 894-902.
22. He Jia, and Liu Jiannan. “The role of mobile-assisted language learning in oral English teaching.” Journal of Hubei Adult Education Institute 27.2 (2021): 35-40.
23. Read, John. “Test review: The international English language testing system (IELTS).” Language testing 39.4 (2022): 679-694.
24. Peltekov, Peter. “The International English Language Testing System (IELTS): A Critical Review.” Journal of English Language Teaching and Linguistics 6.2 (2021): 395-395.
25. Wang Xiaochen, Zhu Mingxing, Yang Zijian, Wang Xin, Huang Jianping, Chen Shixiong, et al. “Study on the contribution of symmetrically positioned articulatory muscles to speech recognition based on high-density electromyography.” Integrated Technology 9.1 (2020): 55-65.

26. Severin, Richard, Ross Arena, Carl J. Lavie, Samantha Bond, Shane A. Phillips. “Respiratory muscle performance screening for infectious disease management following COVID-19: a highly pressurized situation.” The American journal of medicine 133.9 (2020): 1025-1032.
27. Piatnytska-Pozdnyakova, Irina, Alina Korshunova, and Andrii Galaidin. “Role And Functions of The Torso in The Process of Vocalization.” Scientific Journal of Polonia University 51.2 (2022): 125-134.
28. Spiesshoefer, Jens, Simon Herkenrath, Carolin Henke, Lisa Langenbruch, Marike Schneppe, Winfried Randerath, et al. “Evaluation of respiratory muscle strength and diaphragm ultrasound: normative values, theoretical considerations, and practical recommendations.” Respiration 99.5 (2020): 369-381.
29. Hannan, Liam M., Rebekah De Losa, Nicholas Romeo, Sanjeevan Muruganandan. “Diaphragm dysfunction: a comprehensive review from diagnosis to management.” Internal Medicine Journal 52.12 (2022): 2034-2045.
30. Osatananda, Varisa, and Wipawan Thinchan. “Using Praat for English Pronunciation Self-Practice outside the Classroom: Strengths, Weaknesses, and Its Application.” LEARN Journal: Language Education and Acquisition Research Network 14.2 (2021): 372-396.
31. Larassati, Anisa, Nina Setyaningsih, Valentina Widya Suryaningtyas, Setyo Prasiyanto Cahyono. “Using Praat for EFL English pronunciation class: Defining the errors of question tags intonation.” Language Circle: Journal of Language and Literature 16.2 (2022): 245-254.