International Journal for Housing Science and Its Applications
Publish August 3, 2025. Volume 46, Issue 3 Pages 1055-1063

https://doi.org/10.70517/ijhsa46381

Research on the Design and Adaptive Testing System of English Listening Materials in Colleges and Universities Based on Topology Optimization Algorithm

Weifeng Zhang[1,*]
[1] Huanghe Science & Technology University, Zhengzhou, Henan, 450063, China
Corresponding authors: (e-mail: wf-zhang0403@hhstu.edu.cn).

Abstract This paper organically integrates the acoustic model topology, BIC and PSO, and improves the ASR system by optimizing the HMM acoustic model topology. The improved ASR system is used as an auxiliary teaching tool to optimize English listening teaching in colleges and universities by adopting TTS technology. Design the ca-MST model based on Rasch model to realize the accurate assessment of students' English listening ability. Four English major classes of a key university in city A with a total of 120 students were selected for experiments to build an ASR system with optimal acoustic topology. The unidimensionality calibration of the Rasch model is incorporated to ensure the quality of the ca-MST question bank. Putting the proposed system into use, the results of the teaching experiment showed that the mean value of the listening score of the experimental class increased from 18.085 to 22.189 after the experiment, and the t-value of the change in the score was 2.874, whose significance (two-tailed) was 0.003, which was lower than the significance level of 0.05. The control group's achievement improvement is small, and the listening ability does not show significant improvement, which in turn proves that using the system in this paper to assist teaching helps students' listening ability, and provides more choices for the optimization of English listening teaching in colleges and universities.

Index Terms English listening teaching, topology optimization algorithm, ASR system, Rasch model, ca-MST model

I. Introduction

In the listening classroom of college English, in addition to the content outside the textbook, sometimes teachers need to select extracurricular materials for students on their own in order to broaden students' knowledge and cultivate learning interests [1], [2]. Listening itself is a process that requires comprehensive ability, which requires students to collect information, store it, memorize it and respond to it [3], [4]. In response to the current English test's single method of examining students' listening ability, teachers can design additional listening materials to train students' listening ability in a more comprehensive way [5]. How to make the selected materials can help students build up their confidence and interest, as well as conform to their level and help them improve, is a problem that teachers need to consider carefully [6], [7].

With the arrival of information technology in education, integrating computers organically into English language teaching and learning, and even considering how to improve language learning with human-computer interaction based on the potential of computers has become the focus of teaching and learning [8]-[10]. To this end, the design and implementation of a computer-based algorithmic English listening material design and testing system can combine the various factors affecting students' English listening training with each other to form a new networked listening training environment [11]. In this environment, each learner is centered on himself, interacts with the network environment in his own unique way, and achieves good listening training effects through various specific training modes and stage-by-stage graded classification tests [12], [13].

In this paper, we firstly sort out the application of TTS technology in listening teaching activities. The system recognition rate and complexity are chosen through the trade-off coefficient compromise in BIC, and the improved PSO algorithm is used to optimize the acoustic model topology to realize the construction of the hearing ASR system. The basic principle of Rasch model is further elaborated, and the ca-MST test format is proposed by combining the automatic online grouping technology of CAT and the multi-stage presentation test group design of MST. The single-parameter Rasch model is used to estimate the provisional ability of the testees and build the English listening ca-MST system. The control group and experimental group are divided to conduct teaching experiments, running the optimization algorithm of this paper to construct the ASR system as an auxiliary teaching tool. Based on the Rasch model, the unidimensionality test of the question bank was conducted to assess the quality of the question bank.

Compare the students' listening scores before and after the teaching experiment to investigate the feasibility of the proposed system for assisted teaching.

II. Design of English Listening Materials and Adaptive Testing System Based on Topology Optimization Algorithm

In the context of accelerated globalization, the cultivation of English listening ability has become a key part of language education. Traditional listening teaching faces a double challenge: first, static listening materials can hardly meet the needs of individualized learning; second, test systems based on fixed question banks cannot dynamically track the evolution of learners' abilities. Existing TTS technologies cannot fully meet the needs of practical applications, while mainstream adaptive testing systems (e.g., CAT) suffer from the problem of imbalance in question bank exposure, and multi-stage adaptive testing (MST) is limited by the timeliness defect of offline grouping. The aim of this paper is to break through the traditional technical boundaries and propose a dual-engine driven solution for both the technical layer and the assessment layer.

II. A.Application of TTS Technology in English Listening Teaching Activities

II. A. 1) TTS Technology Support for Fine Processed Listening Instructional Activities

TTS technology is the main technical support for the implementation of teaching activities, focusing on the development and utilization of relevant resources in the teaching process, as well as resource support in the process of activity feedback. In the development of teaching resources, based on the needs of the teaching objectives, teachers usually use the technology to adjust the speed of speech and the conversion of the relevant timbre, with the aim of generating different forms of contextualized listening resources for male and female students, so as to meet the needs of students' processing of English listening information. During the activity and the feedback period at the end of the activity, learners need to be provided with the listening resources they need to learn at any time.

II. A. 2) TTS technology support for contextualized communicative listening teaching activities

In the process of English listening teaching, TTS technology can provide relevant queries and feedback for the development of resources needed before teaching activities. At the stage of listening resource development, the technology relies on the MP3 generating function and timbre conversion to actively provide students with communicative resource materials. In the process of activity development, students input and query the pronunciation of listening words according to the technology, aiming to deeply understand the meaning of listening. On the basis of mastering the key words and phrases, the recording feedback function provided by the TTS technology assists the students to think rationally and correct the existing problems in time.

II. B.ASR system design based on topology optimization algorithm

After exploring the specific application of TTS technology in English listening teaching activities, attention needs to be paid to how to further improve the performance and adaptability of the listening system from the technical level. To this end, this paper introduces a topology optimization algorithm, focuses on the structural design and optimization of the acoustic model, and designs an improved ASR system as an English listening assistive teaching tool in this paper.

II. B. 1) Acoustic model topology

HMMs are doubly stochastic processes. The underlying model is an unobservable Markov chain, which can be described by a number of model states and transfer matrices between the states; in addition, each state is associated with a probability density function that generates a sequence of observations, which are usually mixed Gaussian functions when the HMM is used as an acoustic model.

The number of modeled states, the transfer relationships between states, and the number of Gaussian kernels in the mixed Gaussian function of each state together form the topology of the acoustic model. Adopting the state transfer relationship described by the left-right model, the topology referred to in this paper is the combination of the number of model states and the number of Gaussian kernels in each state.

II. B. 2) Bayesian information criterion

BIC is one of the most widely used of the many model selection criteria, defined as:

$$BIC(T) = \ln p\left(O \mid \theta_{ML}, T\right) - \lambda \frac{k}{2} \lg(R) \tag{1}$$

where T is the model topology; $O=\left(O^1,O^2,\cdots,O^R\right)$ stands for a sequence of R observation vectors; k is the number of free parameters in the model; and λ is the trade-off coefficient, which is usually taken in the interval $(0,1]$; and given the O the maximum likelihood parameter estimates for model T are:

$$\theta_{ML}=\arg\max_{\theta}\sum_{r=1}^{R}\ln p\left(O^r\mid\theta,T\right) \tag{2}$$

Eq. (1) is a weighted sum of the likelihood logarithm $\ln p\left(O\mid\theta_{ML},T\right)$ of the training data, which represents how well the model T fits the training data O, and the model complexity penalty expression $-(k/2)\lg(R)$, which is proportional in absolute value to the model complexity.An intuitive interpretation of BIC is that : choose the model with strong and parsimonious portrayal ability, and λ determines the relative importance of the two.

II. B. 3) Particle cluster optimization

PSO is a set of intelligent optimization algorithms for finding the optimal solution of a problem by performing global, parallel and stochastic searches in a predefined solution space, and is a powerful tool for solving complex multi-peaked optimization problems.

The PSO algorithm is a simulation of the collective foraging behavior of a flock of birds or a school of fish, where the individual birds or fish are referred to as particles and the aggregate is referred to as a swarm. The position of each particle in the search space is encoded with a problem-dependent coding scheme that represents a potential solution to the problem. In addition, each particle has its own flight speed.

The algorithm is initialized with a randomly chosen position and velocity for each particle in the search space. In each subsequent iteration, the flight speed of each particle is adjusted by combining the experience of the particle itself and the experience of the best-performing particles in the swarm, while refreshing the position of each particle in the search space based on the current position and the new flight speed. Usually, the algorithm terminates after completing a certain number of iterations and gives the searched optimal solution.

II. B. 4) Improvement of ASR system creation

Their optimized topologies are obtained by optimizing the BIC function of selected trade-off coefficients of each acoustic model, which leads to the establishment of an efficient and parsimonious ASR system. The specific steps of the method are:

(1) Select the trade-off coefficients λ of BIC according to the relative importance of recognition rate and complexity.

(2) Obtain the topology optimization corpus of each model to be optimized.

(3) With BIC as the objective function, run the PSO-based acoustic model topology selection algorithm on each model topology optimization corpus separately to obtain the optimized topology of each model.

(4) Construct the ASR system using the optimized topology.

II. C.Rasch model based ca-MST test model construction

II. C. 1) Basic assumptions and requirements of the Rasch model

Rasch measurement applications are subject to unidimensionality and local independence testing requirements and contain stronger assumptions than other measurement models. Unidimensionality refers to the assumption that a given test measures only one latent trait, whereas local independence implies that performance on a test question does not covary with performance on other questions after conditioning on the latent trait. Controlling for ability values and question difficulty, this means that the probability of a test taker answering a question correctly is independent of the results of answering other questions, i.e., conditional independence of the question. The local independence condition is related to the number of dimensions of the psychological profile, which is empirical evidence of unidimensionality.

However, if the Rasch model includes the ability parameter (b) for each dimension, the same local independence of questions can be obtained for multidimensional data. In short, local independence of topics is a necessary but not sufficient condition for their unidimensionality, but if topics are unidimensional, then they necessarily reflect local independence. Violation of these assumptions reduces the plausibility of the interpretation of the results, since calibration of the data to the psychometric trait estimates produced by the Rasch model presupposes that the assumption that the model fits the data well holds.

II. C. 2) Rasch parametric modeling

Rasch's one-parameter model constructs the respondent's response process with only one location parameter, which is a representation of the gap between the respondent's ability parameter (b) and the question difficulty parameter (d). Where the difficulty parameter (d) refers to the position of the question on the mental continuum of features constructed for that test. The upper end of the continuum feature indicates a higher degree of target ability value for the tested group than for the test takers located at the lower end. Thus, by placing the examinee and the topic on a common, continuous scale, the response of the examinee group on a given topic is inferred accordingly.

Specifically, for a topic n with a difficulty of d_n, the probability P_{in} that a subject i with an ability of b_i will answer the question correctly (assuming that a correct answer is rated at 1 point and an incorrect answer is rated at 0 points) can be computed by using Equation (3).

$$\ln\left(\frac{P_{in}}{1-P_{in}}\right)=b_i-d_n \tag{3}$$

Equation (3) is a mathematical expression of the dichotomous Rasch model, which describes the relationship between the probability P of a subject answering a question correctly and the difference (b_i-d_n) between the subject's ability value b_n and the question difficulty d_i. The greater the difference between the respondent's ability and the difficulty of the question, the greater the value of P; and when $b_i-d_n=0$, $P=0.5$.

In addition to the Rasch one-parameter model, there are also partial rating models, rating scale models, multi-faceted Rasch models, and other extensions derived in the Rasch model. In this study, it was decided to use the one-parameter Rasch model while controlling for the prior knowledge of the respondents, test scenarios, and other relevant variables.

II. C. 3) Rasch-based ca-MST measurement model construction

The two existing forms of computational adaptive testing (CAT and MST) have obvious limitations of their own. Therefore, this paper combines the automatic online grouping technology of CAT with the design of multi-stage presentation of test question sets of MST, and proposes the ca-MST test format to make up for the shortcomings of the existing adaptive tests. ca-MST refers to the stage-by-stage adaptive presentation of test questions of MST, and only makes adaptive adjustments between stages. However, the adaptive presentation of test questions in each stage of ca-MST is in the context of pre-assembly of the corresponding discourse, and the test questions in each stage of ca-MST are dynamically selected online from the question bank according to the ability level of the test takers.

The process of constructing the Rasch-based ca-MST measurement model is shown in Figure 1. The stages of the ca-MST measurement model are designed as 1-3-3-3, i.e., up to four stages including the initial stage. Among them, the initial stage contains one module, and each of the remaining stages contains modules of three difficulty levels, namely, hard-medium-easy (H, M, and E), and each module contains one discourse. Since there is only one module in the initial stage, in order to reduce the exposure of discourse and topics in this module, ca-MST is designed to have multiple parallel panels, i.e., a parallel test design is used so that the high-information topics in the question bank can be evenly distributed in different panels, thus reducing the exposure of topics.

In the question bank, five questions are attached under each discourse. According to the design of the ca-MST measurement model, three of the five questions are adaptively presented by the ca-MST program based on the built-in question selection strategy algorithm to the participant when he/she takes the reading test for a particular discourse within the module. In more detail, first, the test panel of ca-MST is assembled according to the stage design, and the English reading discourse to be tested is assembled at the stage level in advance. Since the length of the modules in the initial stage determines whether the estimation error reduction effect is significant or not. Therefore, the length of topic sets in the initial stage of ca-MST was designed to be a complete set of five topics to enhance the measurement accuracy. Then, the ca-MST system will calculate the provisional ability estimate of the subject based on his/her response to the questions in the initial phase based on the Rasch model, route it to one of the modules in the next phase that is closest to his/her ability value, and assemble a new set of questions as the second phase to match the calculated provisional ability estimate. After the subject completes the second stage, ca-MST will continue to update the ability estimates based on the Rasch model based on the subject's response to answering all the questions that were presented adaptively and assemble a new stage (third stage) to match the updated ability estimates.

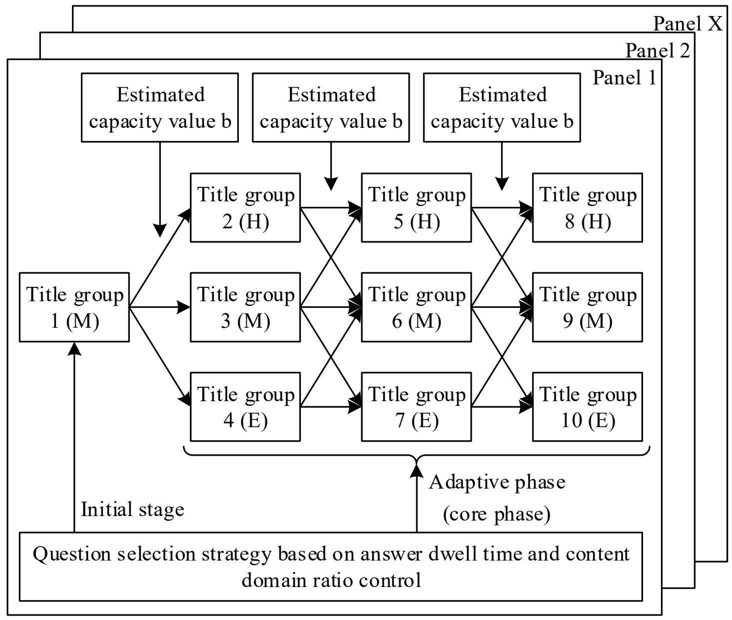

Figure 1: Construction of the Rasch-based ca-MST measurement model

III. Analysis of the application of ASR and ca-MST systems for English listening instruction

The listening material question bank constructed in this paper contains a total of 266 items, each item measures only one cognitive attribute, but the same listening learning attribute can be measured by multiple items. In order to reduce students' testing burden, this paper splits the question bank into five listening materials according to the idea of equivalence design, which is divided into four parallel questions, T1, T2, T3, T4, and one anchor question, T0. Each listening material contains a certain number of questions, which are divided into equal proportions according to the examined content, difficulty, and type of factors to ensure the difficulty and quality of the questions. In the anchor question T0, this paper selects items of equal proportion and equal difficulty, and at the same time contains a sufficient number of items that are the same as other listening materials to ensure the fairness and accuracy of the test.

In this paper, a total of 120 students in four English major classes in a key university in city A were selected for the experiment, of which 60 students were in the experimental group and 60 students were in the control group. The experimental group adopts the system of this paper to assist teaching, and the control group is taught according to the conventional method.

III. A. Construction of ASR system based on topology optimization algorithm

The creation of the baseline system started with a single Gaussian kernel speech recognition system, and a candidate baseline system with an even distribution of other Gaussian kernels was obtained by gradually increasing the number of Gaussian kernels for each state in the alphanumeric model. The alphanumerics in the baseline system were modeled with a 10-state, left-to-right, whole-word HMM, while the silence model and the short-stop model were modeled with a 3-state, left-to-right, single-Gaussian HMM and a 1-state, left-to-right, single-Gaussian HMM, respectively.

Preparing the corresponding speech training corpus for each model to be optimized is a necessary preparation for model topology optimization. Considering that the speech corpus is the continuous pronunciation of the corresponding speech of the model to be optimized, the following segmentation of continuous speech is adopted to obtain the model topology optimization corpus: the best performing one on the test corpus is selected from the multiple baseline systems that have already been created and used to decode the training corpus, obtaining the time-stamped recognition results; after that, the speech segments that correspond to the respective numerical models to be optimized are extracted from the correctly recognized results As the model topology optimization corpus, the specific values of the control parameters of the optimization algorithm are shown in Table 1.

Table 1: Control parameters of the optimization algorithm

Parameter	Value
BIC tradeoff factor	0.6
Population size	100
Evolutionary frequency	500
Maximum possible value of Gaussian kernel in each state	80
Maximum possible value of the number of states	20
Selection ratio	0.7
Crossover rate	0.8
Variation rate	0.062
Selective pressure difference	3.5

The model topology of the optimized speech recognition system was also created from a single Gaussian kernel system, but more Gaussian kernels were added at a time in order to avoid "over-training" the digit model with a generally low number of Gaussian kernels for each state. In the optimized system, the mute and short-stop models have exactly the same structure as in the baseline system, while the topology of the English numerical model is derived from the output of the optimization algorithm.

The overall topologies of the optimized numerical models are: all are topologies with a non-uniform distribution of Gaussian kernels; three models have a state number of 14, and the remaining eight models have a state number of 15. As an example, the optimized topologies of the three models are given as shown in Table 2.

Table 2: Examples of topology optimization results

Model	Condition number	Gaussian kernel number by state
"three"	13	7,12,38,15,15,8,9,9,16,15,22,14,12
"six"	14	22,16,15,15,8,8,9,11,19,37,2,7,51,16
"ten"	15	31,11,6,25,33,16,16,28,30,32,46,7,7,5,16

The graphs showing the optimization process of the 11 numerical models show that the algorithm of this paper is an effective tool for topology optimization of acoustic models. Among them, the optimization process of model "three" is shown in Fig. 2. The model converges at an iteration number of 322, and the minimum value of -BIC is 58672.

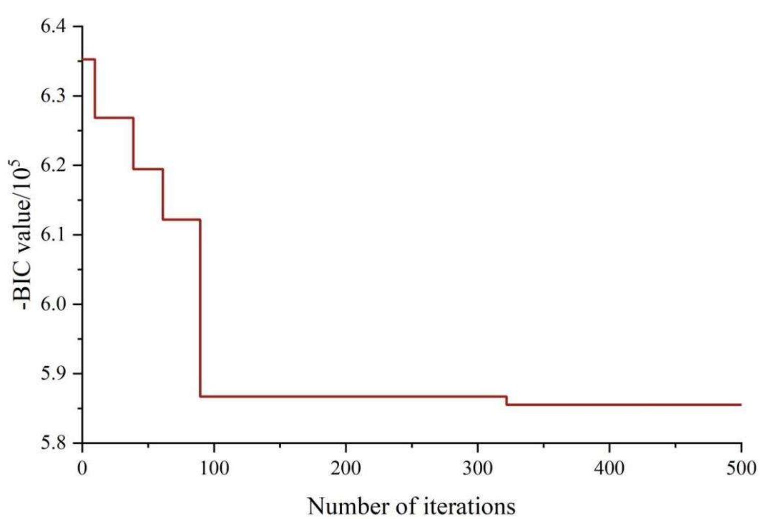

Figure 2 Topology optimization process of model "three"

III. B. Rasch-based unidimensionality test for English listening

Unidimensionality is one of the important characteristics of measurement performance. In order to ensure the unidimensionality of the question bank, this paper carries out a unidimensionality test based on the experimental group's first response data, the specific process is to analyze the students' response data and get the standardized variance data whose detailed results are shown in Table 3.

Rasch model refers to the validation of the original responses in the framework of the standardized principal component analysis, the weights can not be explained by the original total variance of the several comparisons under the original can fully reflect whether it can be explained by the model. If the variance of the first element is greater than 2, it means that the items of the question bank need to be revisited and the reason why this element cannot be explained needs to be found. As can be seen in Table 3, the variance of the first element is 4.5, and a variance greater than 2 indicates that there are unexplained factors that affect the Rasch measurements, and that reasons affecting the explanation of factors such as item difficulty and student ability need to be found. The proportion of variance explained by raw scores was 29.9%, and the results of the analysis of the response data were manageable in terms of statistical significance, so the variance explained was good, and only some of the poorly performing items needed to be modified.

Table 3: Standardized variance

	Empirical situation			Modeling
	Variance	Percent		Percent
The overall original variance of the observations(T)	329.5	100.0%		100.0%
Measure the original variance interpreted(M)	98.4	29.9%		29.9%
The original variance as explained by the student(P)	52.5	15.9%		15.9%
The original variance explained by the item(I)	40.7	12.4%		12.4%
Unexplained original total variance(U)	231.6	70.3%	100.0%	70.3%
Unexplained variance(First comparison)	4.5	1.4%	1.9%	-
Unexplained variance(Second comparison)	3.9	1.2%	1.7%	-
Unexplained variance(3rd comparison)	3.5	1.1%	1.5%	-
Unexplained variance(4th comparison)	3.2	1.0%	1.4%	-
Unexplained variance(5th comparison)	2.9	0.9%	1.3%	-

In order to confirm that the question bank only measured the construct of English listening, this paper further conducted a variance component analysis and presented the possible elements in the form of logarithmic scale and percentage of variance, and the results of the variance component analysis are shown in Figure 3. In the figure, the letters T, U, M, P, and I represent the variance of different elements, and the numbers 1, 2, 3, 4, and 5 represent the elements that may be decomposed, and the percentages of different elements are generally around 3% to 5%. This again confirms that the items in the question bank measure only the construct of English listening.

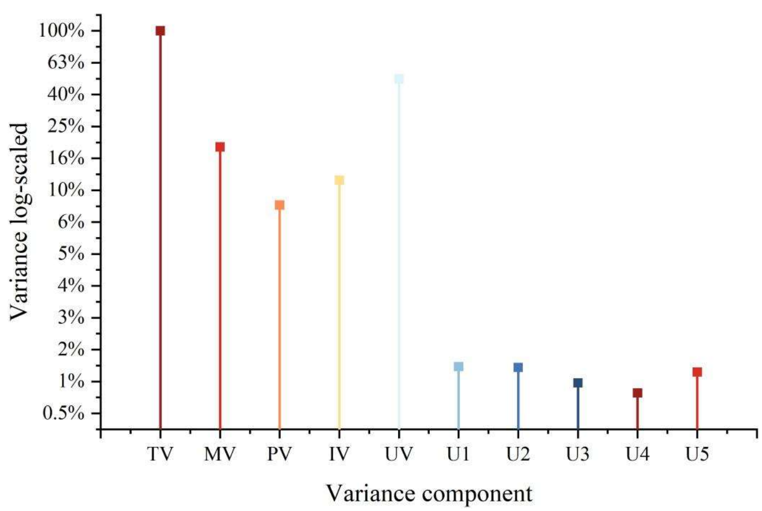

Figure 3: Analysis of variance components

This paper also analyzes the question bank questions in more detail, and finds directions for item improvement by identifying the 26 items that performed poorly.The 26-item standardized variance comparison is shown in Figure 4, which is able to directly demonstrate the correlation between the loading coefficients of the 26 items and the estimated item difficulty values. In this case, each letter indicates a different item, and when the value of the corresponding vertical coordinate exceeds the range between -0.4 and 0.4, it means that the item cannot satisfy the unidimensionality test. According to the results in the figure, it can be found that items A, H, I, J, O, P, and Q are beyond the specified parameter range. Based on the experimental results, this paper eliminates the poorly performing items to ensure the quality of the question bank and formally puts the proposed system into use for teaching experiments.

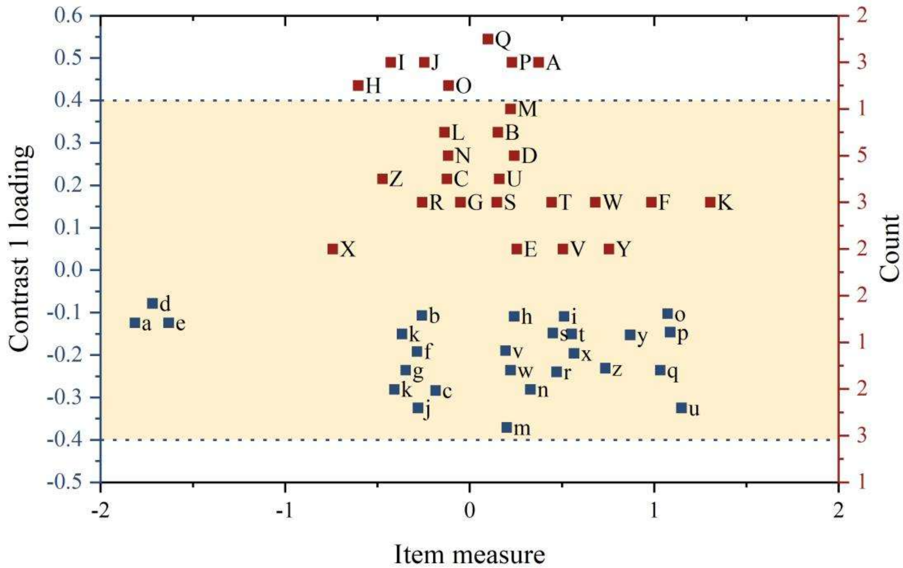

Figure 4: Standard variance comparison

III. C. Analysis of the application effect of English listening ca-MST system

After conducting a semester-long teaching experiment, the data were statistically analyzed for the listening scores of the control and experimental classes before and after the experiment, with the test scored out of 25, and the results of the descriptive statistics of listening ability before and after the experiment are shown in Table 4. According to the data provided, the average English listening performance of the control class after the experiment increased from 18.037 to 18.478, while the average of the experimental class increased from 18.085 to 22.189.Although the scores of the two classes increased, the experimental class improved significantly more than the control class.

Table 4: Descriptive statistical results

Class	Time	Number of cases	Mean value	Standard deviation
Control group	Before the experiment	60	18.037	4.587
	After the experiment	60	18.478	4.603
Experimental group	Before the experiment	60	18.085	4.687
	After the experiment	60	22.189	3.035

Further a paired sample t-test was conducted to verify the students' listening ability and the paired sample t-test for listening ability before and after the experiment is shown in Table 5. The results of the paired samples t-test for the control class show that the t-value of the change in performance before and after the experiment is 0.367, and its significance (two-tailed) is 0.882, which exceeds the significance level of 0.05. This means that in the control class, English listening skills did not show significant improvement after the experiment. However, the paired samples t-test for the experimental class showed that the t-value of the change in performance before and after the experiment was 2.874 and its significance (two-tailed) was 0.003, which was below the 0.05 level of significance. This indicates that in the experimental class, there is a significant improvement in the listening performance after the experiment.

Table 5: Results of paired sample t test

Pairing difference					
Class	Mean value	Standard deviation	t	Degree of freedom	Significance (double tail)
Control group	-0.2864	6.4897	-0.367	62	0.882
Experimental group	-1.9364	4.9274	-2.874	58	0.003

IV. Conclusion

In this paper, we designed the ASR system and English listening ca-MST test model optimized based on topology algorithm, and analyzed its practical application effect through experiments.

At the end of the teaching experiment, the mean value of English listening performance of the control class after the experiment increased from 18.037 to 18.478, while the mean value of the experimental class increased from 18.085 to 22.189. The results of the paired-sample t-test for the control class showed that the t-value of the change in performance before and after the experiment was 0.367, and its significance (two-tailed) was 0.882, which exceeded the significance of 0.05 level. The paired samples t-test for the experimental class showed that the t-value for the change in grades before and after the experiment was 2.874 and its significance (two-tailed) was 0.003, which is below the 0.05 level of significance. This indicates that in the experimental class, there is a significant improvement in the listening performance after the experiment.

The experimental results verify the effectiveness of the topology optimization algorithm-based English listening material design and adaptive testing system for colleges and universities in practical application, and the use of the system proposed in this paper to assist the teaching of English listening in colleges and universities can significantly improve the teaching effect of English listening.

Funding

This work was supported by Henan higher education teaching reform research and practice project in 2021 (2021SJGLX607).

References

[1] Picard, M., & Velautham, L. (2016). Developing Independent Listening Skills for English as an Additional Language Students. International Journal of Teaching and Learning in Higher Education, 28(1), 52-65.

[2] Nurkhamidah, N. (2021). University students' perspective on material and activities in English listening class during pandemic. Elsya: Journal of English Language Studies, 3(2), 94-105.

[3] Ikbar, S., Hermansyah, H., & Rosmiyati, E. (2023). Listening materials in English textbook: Quality Review. Esteem Journal of English Education Study Programme, 6(2), 301-310.

[4] Diachuk, N. (2020). Principles of selection of educational material for developing English listening skills of high school students. Pedagogika. Studia i Rozprawy, (29), 113-126.

[5] Qodir, A., Baehaqi, L., & Miftah, M. Z. (2016). Developing materials of listening comprehension for the English department students. Journal on English as a Foreign Language, 6(1), 1-20.

[6] Liu, L. (2021). Design and implementation of English listening teaching based on a wireless communication microprocessor and virtual environment. Journal of Sensors, 2021(1), 2887302.

[7] Maslikah, K., Widiati, U., Wulyani, A. N., & Sharif, T. I. S. T. (2024). Developing digital literacy application for Extensive Listening materials. Bahasa dan Seni: Jurnal Bahasa, Sastra, Seni, dan Pengajarannya, 51(2), 1.

[8] Ni, D. (2017). Design and research on english listening teaching assisted by computer multimedia. International Journal of Emerging Technologies in Learning, 12(1).

[9] Zaim, M., Refnaldi, R., & Elfiona, E. (2020). Developing mobile based authentic listening materials for senior high school students. Ta'dib, 23(1), 1-10.

[10] Kamarullah, K., Muslem, A., & Manan, A. (2018). Applying English video learning materials in teaching listening. English Education Journal, 9(4), 527-539.

[11] Nurhayati, N., Setiawaty, P. W., & Nur, S. (2024). EFL Teachers Challenges in Designing Assessment Material for Students' Listening Skills. ENGLISH FRANCA: Academic Journal of English Language and Education, 8(2 November), 409-422.

[12] Polat, M., & Erişti, B. (2019). The effects of authentic video materials on foreign language listening skill development and foreign language listening anxiety at different levels of English proficiency. International Journal of Contemporary Educational Research, 6(1), 135-154.

[13] Zaenuri, M. (2015). A model of EFL listening materials development. Indonesian EFL Journal: Journal of ELT, Linguistics, and Literature, 1(2), 119-130.