

Original Paper

A Study of College English Listening Comprehension Strategies from the Perspective of Multimodal Discourse Analysis

Chenling Lu¹

¹ College of Art and Design, Xi'an FanYi University, Xi'an, Shaanxi, China

Received: February 29, 2025

Accepted: March 29, 2025

Online Published: April 16, 2025

doi:10.22158/wjeh.v7n2p78

URL: <http://dx.doi.org/10.22158/wjeh.v7n2p78>

Abstract

This paper discusses the application of multimodal discourse analysis theory in college English listening comprehension strategies. Through a literature review, it analyzes the current situation of college English listening teaching and points out the potential of multimodal discourse analysis in enhancing the effect of listening comprehension. The article proposes a listening comprehension strategy based on multimodal discourse analysis and demonstrates its implementation effect through a case study. Finally, it summarizes the advantages and challenges of multimodal discourse analysis in listening teaching, and looks forward to future research directions.

Keywords

multimodal discourse analysis, college English listening, listening comprehension strategy, teaching status quo

1. Introduction

With the acceleration of globalization, the importance of English as an important tool for international communication is becoming more and more prominent. In college English teaching, the cultivation of listening comprehension ability is a crucial part. However, the traditional mode of listening teaching is often limited to a single auditory modality, ignoring the role of other modalities in the construction of meaning, resulting in unsatisfactory teaching results. The rise of Multimodal Discourse Analysis (MDA) has provided new perspectives and ideas for college English listening teaching. The purpose of this paper is to discuss the study of university English listening comprehension strategies under the perspective of Multimodal Discourse Analysis, with a view to providing useful references for improving the effect of listening teaching.

2. In-depth Review of the Current Situation of College English Listening Teaching

2.1 *The Curing and Singularity of Teaching Mode*

At present, the problem of the solidification and singularity of teaching mode prevails in the field of university English listening teaching. The traditional teaching mode is teacher-centered and mainly adopts the linear process of "playing recordings-students passive listening-mechanical answers". Although this model follows the basic law of language learning, i.e. "input-output" to a certain extent, it over-emphasizes the authority of the teacher and the one-way transmission of knowledge, and neglects the students' subjective position and initiative in the learning process. Students in this mode are often in a passive acceptance state, lack of active participation and in-depth thinking opportunities, resulting in a dull classroom atmosphere, interaction is seriously inadequate. In the long run, students' learning interest and enthusiasm are greatly inhibited, and it is difficult to form a sustained learning momentum.

From the perspective of educational psychology, a single teaching mode is difficult to meet the diverse learning needs of students. Each student has his or her own unique learning style and cognitive characteristics, and a single teaching mode cannot take these differences into account, resulting in some students not adapting to poor learning results. In addition, the lack of interactivity in the classroom also limits the development of students' oral expression ability and critical thinking, making listening teaching disconnected from the cultivation of other language skills such as speaking and writing teaching, and unable to form an organic whole.

2.2 *Marginalization of Non-Linguistic Factors*

In traditional university English listening teaching, linguistic information is regarded as the core content of teaching, while non-linguistic factors (such as intonation, rate of speech, articulation, body language, etc.) are often marginalized or even ignored. This tendency ignores the complexity and multidimensionality of language communication, making it difficult for students to fully and accurately grasp the content and meaning of the material in the process of listening comprehension.

Phonological features such as intonation and rate of speech are indispensable components of language communication, which not only convey the linguistic information itself, but also contain the speaker's emotional attitude, contextual information and other important content. Ignoring these non-linguistic factors, it is difficult for students to accurately understand the speaker's real intention and contextual background, thus affecting the accuracy and depth of listening comprehension. In addition, body language, as an important aid in intercultural communication, appears frequently in listening materials, but is often neglected by teachers and students. This not only limits students' grasp of the overall meaning of the listening materials, but also affects the cultivation of students' intercultural communication skills.

2.3 *Missing and Inadequate Instruction in Listening Skills*

Listening skills are the key to improving listening comprehension. However, in the current university English listening teaching, listening skill instruction is often neglected or in form. Many students lack

effective guidance on listening skills in the listening process, such as how to grasp key words and phrases, how to make inferential judgments in connection with the context, and how to guess the meaning of words according to the context. The lack of these skills makes it difficult for students to understand the information quickly and accurately when facing complex listening materials, which affects the effectiveness of listening teaching.

The lack of instruction in listening skills is not only reflected in teaching practice, but also in the preparation of textbooks and curriculum. Many university English listening textbooks lack systematic listening skills training content, and the curriculum often separates listening teaching from other language skills teaching, lacking integration and systematicity. This makes it difficult for students to form a complete system of listening skills in the learning process, and they can't apply the learned skills flexibly in actual listening scenes.

3. Overview of Multimodal Discourse Analysis Theory

3.1 Theoretical Kernel: Collaborative Meaning Construction of Multimodal Symbolic Systems

Multimodal Discourse Analysis (MDA), as an emerging paradigm in social semiotics, focuses on the dynamic interaction mechanism of semiotic systems in the communication process. Its core proposition is that meaning is not generated linearly through a single modality (e.g. language), but is realized through the synergistic action of multiple modalities. This theory breaks through the limitation of traditional linguistics, which takes verbal text as the only object of analysis, and incorporates non-verbal signs (e.g., images, sounds, spatial layout, body gestures, etc.) into the overall framework of meaning production, thus revealing the multidimensionality, dynamics, and context-dependence of human communicative activities.

In multimodal discourse analysis, modality refers to the medium or channel used to convey information, including linguistic modality (spoken language, written language), visual modality (images, videos), auditory modality (music, sound effects), tactile modality (haptic feedback), and so on. Different modalities can complement each other and together express complete information and meaning.

3.2 Theoretical Value and Application Scenarios

3.2.1 Interdisciplinary Research Paradigm

Communication, analyzing the impact of media technology on the use of modality (e.g., algorithmic enhancement of auditory modality in short video platforms). Pedagogy, to optimize the design of multimodal teaching materials (e.g., to reduce the cognitive load of language through visual modality). Artificial Intelligence, to provide theoretical basis for multimodal interaction design (e.g. intonation and semantic correlation in voice assistant).

3.2.2 Practical Application Fields

Advertising and brand communication, analyzing how multimodal symbols construct brand identity (e.g., the minimalist design language of Apple products). Public policy discourse, to study the guidance of modal choice on public cognition in government propaganda films (e.g., visual and auditory

strategies in epidemic information dissemination). Digital Humanities, to develop multimodal corpus tools (e.g., a cross-modal annotation system combining text, image and audio).

3.3 Implications for College English Listening Teaching

The theory of multimodal discourse analysis provides a paradigm shift path from "language decoding" to "meaning negotiation" for listening teaching. Reconstruct teaching materials, develop a multimodal listening corpus, and integrate resources such as video news, audiobooks, and virtual reality scenarios. Design modal comparison tasks (e.g., comparing the information difference between audio-only and synchronized versions). Modal recognition ability, train students to distinguish between linguistic and non-linguistic modal information-carrying functions. Modal Integration, through task design (e.g., drawing a scene from the audio description after listening) to enhance comprehensive comprehension. Modal critique, analyzing the ideology behind modal choices (e.g., power relations of visual symbols in news reports). Develop multimodal listening test tools to comprehensively assess students' ability to interpret verbal and nonverbal modalities in concert. Design modal transcription tasks (e.g., transforming audio content into visual mind maps).

4. The Construction of University English Listening Comprehension Strategy System under the Perspective of Multimodal Discourse Analysis

4.1 Diversified Integration and Utilization of Modal Resources

Under the framework of multimodal discourse analysis, university English listening teaching should break through the traditional single-modal limitations and build a three-dimensional teaching resource system with auditory modality as the core, visual modality as the auxiliary, and multimodal synergy and interaction. Teachers need to systematically integrate multimedia resources and network platform resources to build a multimodal corpus covering image modality (such as theme-related pictures, charts, mind maps and other visual information carriers), video modality (such as film and television clips, news documentaries, academic lectures and other dynamic audio-visual materials), and audio modality (such as standard speech materials, dialect variations, and recordings of cross-cultural communication scenes, etc.). Through the cross-presentation of multimodal resources, we can help students to establish the association network between textual meaning and modal symbols. For example, when explaining culturally loaded words, we can synchronize the display of pictures of objects, videos of cultural practices and audio of standard pronunciation, so as to form a three-dimensional cognitive channel of "image-phonology-semantics" to deepen the understanding of the background knowledge and cultural connotations of the listening materials.

4.2 Explicit Teaching of Non-verbal Communication Elements

Non-verbal communication elements occupy a key position in multimodal discourse, and their teaching should realize the paradigm shift from implicit perception to explicit cognition. Teachers need to build a framework for analyzing nonverbal elements, including paralinguistic features (intonation pattern, speech rate and rhythm, pause strategy), body language (body movement, facial expression, spatial

distance) and environmental modality (scene setting, props use). In the specific teaching implementation, the four-stage teaching method of "observation-imitation-analysis-application" can be adopted: firstly, through comparative analysis of the differences in nonverbal behaviors in different cultural contexts, the cross-cultural nonverbal communication schema can be established; secondly, immersive teaching activities simulating the real communication scenarios, such as business negotiation simulation, academic conference role-playing, etc., can be designed to guide the students in observing and recording the rules of the use of nonverbal symbols; and finally Through group seminars and case studies, we systematically analyze the mechanism of non-verbal elements in meaning negotiation, attitude expression, identity construction and other levels, and cultivate students' metacognitive ability of multimodal meaning interpretation.

4.3 Systematic Training of Listening Comprehension Strategies

The teaching of listening comprehension strategies should follow the three-dimensional training model of "metacognitive strategies - cognitive strategies - social/emotional strategies". At the metacognitive level, students are guided to monitor the listening process through strategy awareness development activities (e.g., learning journals, reflection reports); cognitive strategy training focuses on the explicit teaching of specific skills.

Prediction strategy, constructing semantic fields based on clues such as titles, pictures, and questions. Keyword localization, using word frequency analysis tools to identify high-frequency semantic nodes and discourse articulation, and identifying means of articulation such as illumination, substitution, and omission.

Pragmatic reasoning, decoding extra-linguistic meanings with contextual factors. Social/emotional strategies emphasize collaborative learning and emotional management, such as peer assessment and self-efficacy assessment to enhance motivation. The teaching implementation can adopt the progressive training mode of "strategy demonstration - decomposition practice - comprehensive application", combined with corpus annotation tools (e.g., ELAN, Praat) to carry out micro-analysis, and improve the automation level of strategy application.

4.4 Ecological Construction of Multimodal Interactive Teaching

Multimodal interactive teaching should build a complete learning chain of "input - internalization - output", and its ecological implementation should focus on three dimensions.

Reconstruction of physical space to create a blended learning environment that integrates smart devices, interactive whiteboards and virtual reality devices.

Design of activity types, developing innovative tasks such as multimodal transcoding (e.g., transforming listening materials into visual narratives) and multimodal creation (e.g., producing theme-related multimodal texts).

Innovative assessment system, establishing performance evaluation criteria covering multiple dimensions such as listening comprehension, multimodal transcription, and collaborative interaction.

In concrete implementation, we can design "multimodal issue debate" activities, requiring students to

use text analysis, image interpretation, physical expression and other skills to complete the argumentation; or carry out "cross-modal narrative" projects, adapting the listening materials into multimodal texts (e.g., graphic novels, animated short films) to deepen the understanding of the modal synergy mechanism in the process of creation. Such activities not only promote the transfer of listening skills, but also cultivate students' multimodal communication literacy in complex communication situations.

5. Case Study

Taking a college English listening course as an example, the teacher innovatively integrates the listening comprehension strategy under the perspective of multimodal discourse analysis into teaching practice, and constructs a "before class - during class - after class" trinity teaching mode, which effectively improves students' listening comprehension and cross-cultural communication literacy.

In the pre-course preparation stage, multimodal resources activate the background schema. Based on the theory of multimodal discourse analysis, the teacher builds an "immersive pre-study scene" through the class WeChat group. In addition to listening texts, they also push theme-related pictures (e.g., cultural scene diagrams, character emoticons), short videos (e.g., news clips, movie clips) and audio clips (e.g., English podcasts, excerpts from speeches), forming a three-dimensional "visual + auditory" pre-reading matrix. For example, when explaining the unit of "Intercultural Business Negotiation", the teacher pushes the video of the comparison between Chinese and Western business etiquette and the illustration of non-verbal communication gestures in advance, guiding students to prejudge the points of cultural conflict in the listening materials by observing the modal features of body language and scene arrangement, activating students' background knowledge schema, and laying the foundation of in-depth understanding in the classroom. Laying the foundation for deep comprehension in the classroom.

In the classroom explanation stage, multimodal analysis of the meaning of the discourse. In the classroom, teachers rely on the multimodal combination strategy of "PPT dynamic demonstration + video contextualization + board logic combing" to deconstruct the listening materials in multiple dimensions. Visual modality strengthens semantic connection, embedding timeline visualization charts and semantic network diagrams of key information in the original listening text in PPT, and helping students to intuitively capture the logical relationship of the discourse through visual symbols such as color marking and arrow pointing; auditory modality trains the decoding ability, and the teacher synchronously pauses and calls up the video clips (e.g., the original TED talk clips) when the listening material is played to guide the students to compare the semantic difference of the pure speech and the multimodal context. Auditory modality training decoding ability, the teacher synchronously pauses and calls the video clip (e.g., TED talk sound clip) when playing the material, guiding students to compare the semantic difference between pure speech and multi-modal contexts to improve the ability of speech

recognition and contextual inference.

For example, when parsing "airport broadcasting delay notice", the teacher combines the video of the information scrolling on the big screen of the airport to comparatively analyze the correlation between the speed and tone of the broadcasting and the anxiety of passengers, so as to deepen the students' understanding of the functionality of the service class listening discourse.

In the post-lesson consolidation stage, the multimodal platform promotes the internalization of knowledge. In the post-course session, teachers rely on platforms such as "Learning Channel + WeChat + B Station" to build a "blended learning community". Multimodal homework design, "listening text dubbing show" (audio + subtitles), "cultural differences Vlog" (video + narration) and other tasks, requiring students to use the multimodal discourse analysis strategies learned in the classroom, independent creation of learning outcomes; peer peer evaluation Peer mutual assessment mechanism, through the WeChat group to carry out "modal comparison discussion", such as comparing the difference in comprehension difficulty of the same listening material in three modes: audio-only, subtitled video, and non-subtitled video, to promote students' metacognitive reflection; Teacher's feedback innovation, the use of "voice comments + emoticon incentive Teacher feedback is innovative, adopting the feedback form of "voice comment + panda motivation", which not only corrects pronunciation and intonation problems through voice annotation, but also enhances the fun and affinity of feedback with the help of emoticons (e.g., "Academic Cat Kudos" and "Scholar Panda Doubts").

6. Conclusion and Prospect

Based on the theoretical framework of Multimodal Discourse Analysis (MDA), this study systematically explores the innovative path of university English listening comprehension strategies and reveals the synergistic mechanism of multimodal resources in listening teaching.

The synergistic effect of modal resources, the dynamic interaction between linguistic modalities (phonology, vocabulary, grammar) and non-linguistic modalities (visual, auditory, spatial) can significantly enhance the efficiency of listening comprehension. For example, by combining multimodal resources such as subtitles, body language illustrations, background sound effects, etc., students can effectively reduce the cognitive load of a single linguistic modality and enhance the ability to decode information. Explicit training of non-linguistic factors, such as intonation changes, pause patterns, contextual background and other non-linguistic elements, can help students build up their Modal Sensitivity, enabling them to actively capture the implicit semantic clues in the multimodal information and make semantic inferences accordingly.

Modal Reconstruction of Listening Skills, Traditional listening skills (e.g., prediction, inference, summarization) need to be integrated with modal analysis strategies in a multimodal framework. For example, by observing a speaker's facial expressions and gestures, students can more accurately predict his/her viewpoints; by analyzing the spatial layout of a video scene, students can infer character relationships and plot development. Modal Empowerment for Interactive Teaching, Multimodal

interactive platforms (e.g., online collaborative whiteboards, virtual reality scenarios) can facilitate modal negotiation among students and enhance the ability to build meaning. For example, in group discussions, students can share visual mind maps and label modal nodes in real time to realize the collaborative construction of multimodal knowledge.

The application of multimodal discourse analysis in listening teaching also faces some challenges, such as how to accurately identify and parse the relationships between different modalities and how to rationally allocate modal resources. Future research can further explore these issues and try to apply the theory of multimodal discourse analysis to the teaching of other English skills, such as speaking, reading and writing. Meanwhile, the applicability and effect differences of multimodal teaching in different cultural backgrounds and learning environments can also be studied to provide more effective teaching strategies and methods for English education.

References

- CHEN, T., WANG, J., & ZHOU, G. L. (2020). Study on Learning Strategies of College English Listening Comprehension from the Perspective of Discourse Reasoning. *English Square*, (20), 86-88.
- Li, X. X. (2019). Research on English Teaching Mode of Art Universities under the Perspective of Multimodal Discourse. *Journal of Ningbo Radio and TV University*, 17(03), 48-52.
- Sun, J. (2022). The use of multimodal discourse analysis in college English classroom teaching. *Overseas English*, (20), 89-90.
- Wang, S. Y. (2024). *Research on English Teaching Based on Multimodal Discourse Theory*. Cultural Development Press:202407.176.
- You, Q. D. M., & Liu, Z. X. (2019). A survey study on the status quo of university English listening teaching from the perspective of multimodal discourse analysis theory—Taking Hainan Normal University as an example. *Education Informatization Forum*, 3(08), 85-86.
- Zhang, X. (2023). Research on college English teaching under the perspective of multimodal discourse analysis theory. *Journal of Lianyungang Normal College*, 40(03), 38-41.