Available Online: 21 October 2026 DOI: 10.54254/3049-7248/2026.36423

Theoretical construction of an embodied approach to Mandarin pronunciation instruction: a case study of alveolar consonants /n/ and /l/

***Qifu Wan*[1], *Zhanyu Wang*[2,*]**

[1]Youth League Committee, Guiyang Institute of Humanities and Technology, Guiyang, China
[2]Broadcasting and Hosting Art, Guizhou University, Guiyang, China

*Corresponding Author. Email: sanqian13708560566@qq.com

Abstract. The confusion between Mandarin alveolar consonants /n/ and /l/ represents a persistent challenge in pronunciation instruction, where conventional approaches rely heavily on auditory discrimination and imitation with inconsistent efficacy. Grounded in embodied cognition theory, this study adopts grounded theory methodology and performs three-level coding of relevant literature to construct an integrative pedagogical model. Analysis reveals that existing studies attribute n/l articulatory errors to inadequate lingual positioning, ambiguous airflow perception, and negative L1 transfer, with the underlying mechanism being a disconnect between perceptual and motor control. Embodied cognition provides a novel explanatory framework that redefines phonetic learning as the acquisition of sensorimotor skills. Accordingly, the proposed model centers on the "embodied transformation of articulatory knowledge", encompassing three core components: Embodied Externalization, Proprioceptive Encoding, and Contextual Transfer, supported by teacher development and resource construction. This model overcomes the limitations of traditional empirical paradigms, offering systematic theoretical grounding and actionable strategies for n/l pedagogy.

Keywords: embodied cognition, articulatory phonetics, alveolar consonants, grounded theory

1. Introduction

1.1. Research background and motivation

General Secretary Xi Jinping has emphasized the need to "comprehensively strengthen education in the national common language and writing system, and continuously enhance the scientific and cultural literacy of people of all ethnic groups". Meanwhile, the *Outline for Building a Leading Country in Education (2024–2035)* proposes "to implement actions for the universal promotion and quality improvement of the national common language, raise the language and cultural literacy of the entire population, and improve the standards system for language standardization and build a new national corpus". The promotion of the national common language is, moreover, a crucial strategy for improving national quality and facilitating social communication [1]. Against this backdrop, the quality of Mandarin pronunciation instruction in basic education—especially

Copyright: © 2026 by the authors. Submitted for possible open access publication under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

pinyin instruction as a tool for literacy and reading—directly bears on the comprehensive development of students' language proficiency [2]. In actual teaching, however, the confusion between the alveolar consonants /n/ and /l/ in Mandarin is a pervasive and persistent difficulty. Existing studies indicate a high detection rate of articulation errors involving alveolar consonants, among which /l/ errors occur most frequently; this confusion seriously impairs speech clarity and communicative efficiency [3]. Confronted with this difficulty, most existing pedagogical methods remain at the empirical level of auditory discrimination and imitation [4]. Teachers widely report a lack of better methods, relying mainly on traditional means such as repeated reading and practice, demonstration, and lip-shape observation. Yet this model not only lacks a scientific explanation of the physiological mechanisms of articulation and a systematic training plan, but also leads to unstable instructional outcomes and large individual differences; after learners disengage from study for a period, regression in pronunciation is commonly observed.

Embodied cognition theory, however, offers a novel perspective for re-understanding this dilemma. This theory holds that cognition is the product of the interaction among body, mind, and environment [5]. At the same time, the successful application of articulatory-phonetics knowledge in the field of speech rehabilitation —where articulatory-organ correction and targeted training can significantly improve articulation clarity— provides valuable potential resources for the scientization of instruction.

In sum, this study aims to bridge the gap between teaching practice and scientific theory. By adopting grounded-theory methodology to systematically integrate relevant literature, it constructs a theoretical model of pedagogy for the alveolar consonants (/n/ and /l/) that integrates the perspective of embodied cognition with articulatory-phonetics knowledge, with a view to providing systematic theoretical guidance for basic pronunciation teaching practice.

1.2. Research objectives and questions

Building on the above research background and motivation, this study aims to apply grounded-theory methodology to systematically analyze relevant academic literature, so as to construct a theoretical model of pedagogy for Mandarin alveolar consonants '/n/ and /l/' that integrates the perspective of embodied cognition with articulatory-phonetics knowledge.

To achieve this aim, this study poses and answers the following three core questions:

(1) How do existing studies describe the physiological mechanisms and common error types of alveolar consonant (/n/ and /l/) articulation?

(2) How does embodied cognition theory provide new theoretical explanations and design principles for the teaching of alveolar consonants (/n/ and /l/)?

(3) Based on the literature, what core elements should a theoretical model of pedagogy for Mandarin alveolar consonants (/n/ and /l/) comprise?

1.3. Research significance

This study introduces embodied cognition theory into concrete research on pronunciation instruction, extending the theory's application boundary in the field of language education and providing a new analytical framework for Mandarin teaching research. At the same time, by systematically integrating the literature, this study promotes the theoretical transfer and integration of articulatory-phonetics knowledge from the field of speech rehabilitation to that of general education, enriching the theoretical foundation of Mandarin pronunciation instruction.

This study aims to construct a theoretical model of pedagogy for alveolar consonants (/n/ and /l/) that integrates articulatory-phonetics knowledge; the model can provide front-line teachers with teaching-strategy

references that are more scientific, systematic, and actionable than the traditional mode of auditory discrimination and imitation. Moreover, the theoretical model can provide direct theoretical grounding for the subsequent design of teacher-training curricula and the development of school-based teaching resources.

2. Literature review

2.1. Embodied cognition theory

Embodied cognition is a significant theoretical turn in second-generation cognitive science. Its philosophical thinking can be traced to Merleau-Ponty's phenomenological tradition, which emphasizes that cognition is the product of the interaction among body, mind, and environment [6]. The theory was established through the systematic exposition of Lakoff and Johnson; its core claim is that cognition does not occur in isolation in the brain, but is shaped by the organism's bodily experience in its entirety, and is characterized by embodiment, interactivity, and situatedness [7]. In the field of language learning, gesture, as the externalized expression of bodily movement, activates the motor cortex of the brain through an embodied-simulation mechanism, confirming the neural association between language processing and the sensorimotor system [8]. Sensorimotor engagement can significantly improve the efficiency of semantic memory and the speed of concept internalization [9]. Therefore, from the perspective of embodied cognition, instruction should emphasize learning by doing and multisensory participation, constructing situations that integrate bodily activity with perceptual experience.

2.2. An articulatory-phonetics analysis of Mandarin alveolar consonants "/n/" and "/l/"

From the perspective of articulatory phonetics, both /n/ and /l/ are voiced alveolar consonants, with vocal-cord vibration during articulation, but their manners of articulation are essentially different. /n/ is a voiced alveolar nasal: during articulation the tongue tip presses against the upper alveolar ridge to form a blockage, the soft palate descends to open the nasal cavity, and airflow passes through the nose; /l/ is a voiced alveolar lateral: during articulation the tongue tip presses against the rear of the upper alveolar ridge to block the central airflow path, the soft palate rises to close the nasal cavity, and airflow passes through the gap formed between the tongue and the inner sides of the cheeks. Existing studies also indicate that errors in /n/ and /l/ articulation mainly manifest as omission and substitution [10]. Their causes can be summarized as two factors. Physiologically, due to poor tongue-position control and insufficient tongue-tip strength, the tongue tip cannot make precise contact with the upper alveolar ridge during articulation; perceptually, especially in the Guizhou region, because the local dialect makes no clear distinction between nasal and lateral sounds, learners have difficulty with ambiguous airflow perception, finding it hard to distinguish whether airflow passes through the nose or the mouth [11].

2.3. A review of existing research on 'n/l' teaching methods

Current teaching practices addressing the /n/–/l/ confusion can be characterized as an empirical mode of subjective auditory discrimination overlaid with imitation. Such a teaching model centers on auditory input and lip-shape observation as its core means, and uses rhymes and mnemonic formulas to aid memory; its theoretical assumption is that phonetic acquisition can be achieved through repeated auditory stimulation and articulatory imitation [12]. However, this model lacks a scientific explanation of the physiological mechanisms of articulation. In corrective teaching activities, teachers rely mainly on subjective auditory discrimination and repeated demonstration, and find it difficult to give students precise guidance at the level of the motor

principles of the articulatory organs (lips, tongue, and mandible). The instructional outcomes of this teaching model are unstable and show large individual differences. After experiencing this model, students' learning is characterized by the defect of being effective in class but easily forgotten after class, making it difficult to form stable bodily perception and muscle memory.

3. Research methods

3.1. Research methodology: literature-based grounded theory

Grounded theory is a qualitative research method that builds theory inductively from the bottom up, out of empirical data; it is suited to fields where the existing theoretical system is incomplete and new theory must be explored and generated from data [13]. This study aims to construct a theoretical model of pedagogy for Mandarin alveolar consonants (/n/ and /l/), rather than to verify an existing hypothesis, so grounded theory is the most appropriate fit. This study takes the academic literature of relevant research as its primary data source. Literature is peer-reviewed, systematized theoretical data; its content includes systematic descriptions of articulation mechanisms, teaching strategies, and theoretical frameworks, and can serve as the raw material for grounded coding. The researcher will critically interpret the literature to extract concepts and categories from it.

3.2. Data sources and screening

This study searched the literature in Chinese databases such as CNKI, Wanfang, and VIP, using retrieval terms including "n/l confusion", "alveolar consonants", "articulation disorders", "embodied cognition", and "'pinyin teaching", with the time range set to the most recent five years. Inclusion criteria were: (1) the content concerns the physiological mechanisms of /n/ and /l/ articulation, the characteristics of articulation disorders, and teaching strategies; (2) it involves the application and discussion of embodied cognition theory in language teaching; (3) it involves theoretical or empirical research on Mandarin pronunciation-teaching methods. Exclusion criteria were: (1) literature that merely repeats existing viewpoints without substantive new findings; (2) non-academic articles, conference abstracts, and theses. After layer-by-layer screening, the final set of valid literature included for analysis comprised 33 journal articles and book chapters in total (insert table; for specific data, see Table 1).

3.3. Data analysis process: three-level coding

The literature was read sentence by sentence, and descriptions of the principles of /n/–/l/ articulation, the causes of errors, and teaching methods were conceptualized and labeled to form initial concepts and categories, such as "tongue-position control", "airflow perception", and "tactile guidance". On this basis, the classic model of 'causal conditions → phenomenon → context → intervening conditions → action/interaction strategies → consequences' was applied to relate the categories formed in open coding and develop main categories. Finally, all main categories were systematically analyzed to extract core categories, construct the theoretical relational network among categories, and ultimately form the theoretical model.

3.4. Research reliability, validity, and ethics

During coding, a portion of the literature was set aside and three-level coding was repeated; when new literature no longer generated new important categories or relations, theoretical saturation was judged to be reached. The researcher maintained reflexivity through writing coding memos and conducting peer

discussions, so as to avoid the excessive influence of personal knowledge background on the coding process. All cited literature is annotated with its source in accordance with academic norms, and the intellectual property rights of the original authors are respected.

4. Research finding

4.1. Descriptions in existing literature of the physiological mechanisms and common error types of "n/l" articulation

Through three-level coding, this study systematically sorted out existing literature's descriptions of the physiological mechanisms of Mandarin alveolar consonants /n/ and /l/ articulation and their common error types, and extracted three main categories: "scientific description of the physiological mechanisms of articulation", "types of common errors and their distributional characteristics", and "attribution analysis of error causes".

Existing literature has already clearly defined the physiological mechanisms of /n/ and /l/ articulation. Both are voiced alveolar consonants, with vocal-cord vibration during articulation, and in both the place of articulation is contact between the tongue tip and the upper alveolar ridge. The essential difference between them lies in the manner of articulation, manifested specifically in the movement state of the soft palate and the path of airflow. Specifically, /n/, as a voiced alveolar nasal, requires in its articulation a series of actions: the tongue tip rises to press against the upper alveolar ridge to form a blockage, the tongue body retracts, the tongue root exerts force and bulges, the soft palate descends to open the nasal cavity, and airflow passes out through the nose; during articulation, the digastric muscle indirectly affects the structure of the larynx and pharynx by contracting to elevate the hyoid bone, providing a stable support base for the movement of the soft palate. /l/, as a voiced alveolar lateral, requires in its articulation a series of actions: the tongue tip presses against the rear of the upper alveolar ridge to block the central airflow path, the soft palate rises to close the nasal cavity, and the middle of the tongue body exerts force to keep the tongue flat and not in contact with (or only lightly in contact with) the molars on both sides, with airflow passing through the gap formed between the tongue and the inner side of the lower molars; during articulation, the genioglossus contracts to move the tongue forward and downward, and this muscle may assist the tongue tip in protruding forward to press against the rear of the upper alveolar ridge. The precise execution of the above articulatory actions is the physiological basis for the accurate pronunciation of the two sounds. The coding results of this study show the salient features of /n/ and /l/ articulation errors. In terms of error distribution, /l/ errors rank first in frequency among all alveolar consonants, while /n/ errors are relatively less frequent. In terms of error types, /l/ errors are mainly characterized by omission; /n/ errors are mainly characterized by omission coexisting with substitution, but their overall occurrence is lower than that of /l/.

Regarding the causes of the above error phenomena, the literature summarizes two categories of factors: physiological and linguistic-environmental. In terms of physiological factors, poor tongue-position control, insufficient strength of the tongue tip and tongue root to meet articulatory requirements, and ambiguous airflow perception (difficulty in distinguishing whether airflow passes through the nose or the mouth) are the key reasons for inadequate articulatory actions. In terms of linguistic-environmental factors, in some dialect phonologies (such as Southwestern Mandarin) there is no strict phonemic opposition between /n/ and /l/, which leads learners to form negative transfer during language acquisition, confounding the two into a single phoneme and thereby producing systematic articulation errors. The above causes jointly point to a core problem: a coupling misalignment between the articulator's "perception" and "motor control", with no precise articulation model established.

4.2. Theoretical explanations and design principles that embodied cognition theory provides for "n/l" teaching

4.2.1. From symbolic cognition to bodily skill

Embodied cognition theory provides a fundamental theoretical explanation for the teaching of /n/ and /l/. Based on the claim of this theory that cognition is the product of the dynamic interaction of body, mind, and environment, phonetic learning is no longer viewed as a mere process of matching symbols with sounds, but is redefined as the acquisition of a kind of "bodily skill". The core significance of this theoretical turn lies in shifting the teaching focus from external auditory discrimination and imitation to the learner's internal construction of perceptual and motor experience, providing a brand-new epistemological basis for understanding the essence of /n/ and /l/ articulation errors—namely, the mutual misalignment between perception and motor control.

4.2.2. Three principles of instructional design

4.2.2.1. Embodied representation principle (embodied representation)

Based on the principle of embodied cognition that emphasizes multisensory participation, instruction should use visualization tools to transform the invisible physiological mechanisms of articulation into observable external representations. Specifically, visual scaffolds such as dynamic cross-section diagrams, animations, and 3D-printed articulatory-organ models can be used to present intuitively the contact position between the tongue tip and the upper alveolar ridge, the positional changes of the tongue-body force points, the ascending/descending state of the soft palate, and the path of airflow. At the same time, gesture simulation, as the externalized expression of bodily movement, can use the shape of the hand to simulate the changes in tongue position and force-point location, helping learners externalize internal actions into observable images; for details see Figure 1. The above strategies aim to build a cognitive scaffold for learners and reduce the difficulty of understanding abstract articulation essentials.

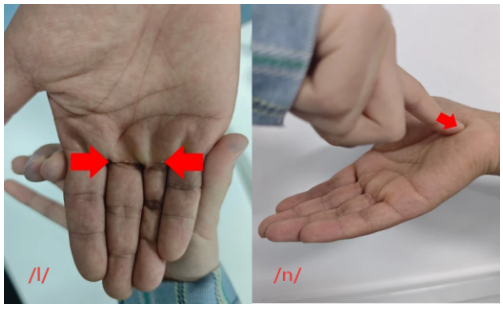

Figure 1. Schematic diagram of hand simulation of tongue-position changes and force-point location changes

Note: This figure was compiled by the authors and photographed on 2026/6/8.

4.2.2.2. Proprioceptive anchoring principle (proprioceptive anchoring)

Aiming at the core difficulties of /n/ and /l/ articulation—tongue-position control, tongue-body force points, and airflow perception—embodied cognition theory emphasizes the intervention of "learning by doing" and bodily experience. At the level of tactile development, the teacher can guide students to perceive the vibration of airflow during nasal sounds by touching the wing of the nose, so as to distinguish the articulation pathways of /n/ and /l/. At the level of kinesthetic guidance, tools such as tongue depressors and cotton swabs can be used to assist in locating the articulation site. Specifically, when producing /l/, a cotton swab can be used to lightly touch the front of the tongue tip (the contact area between the tongue tip and the upper alveolar ridge)

and the two sides of the tongue body, guiding the muscles of the front of the tongue tip and the two sides of the tongue body to form stable tactile and kinesthetic feedback at that point, thereby establishing the muscle memory of the tongue tip rising and airflow passing through the two sides of the tongue—see Figure 2; at the same time, a tongue depressor can be used to press and stimulate the genioglossus from inside the mouth and below the mandible, so that the genioglossus forms a contractile muscular reflex and regularizes the muscular action of the tongue tip protruding forward to press against the rear of the upper alveolar ridge—see Figure 3.

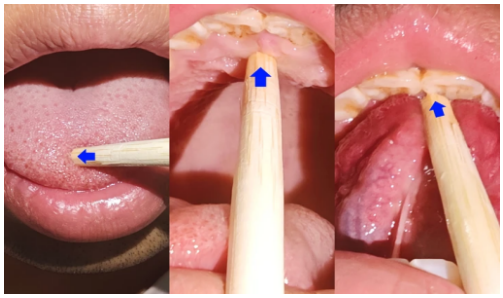

Figure 2. Schematic diagram of tactile and kinesthetic guidance for /l/

Note: This figure was compiled by the authors and photographed on 2026/6/8.

Figure 3. Schematic diagram of tactile and kinesthetic (muscular) guidance for /l/

Note: This figure was compiled by the authors and photographed on 2026/6/12.

When producing /n/, a cotton swab can be used to lightly touch the rear of the tongue surface, guiding the tongue root to rise and close the oral cavity pathway, while strengthening the coordinated action of the front of the tongue tip pressing against the upper alveolar ridge, thereby establishing the muscle memory of the soft palate descending and airflow passing through the nose—see Figure 4. A finger and thumb can also be used to press both sides of the hyoid bone at its junction with the digastric muscle and push upward toward the pharyngeal wall, so that external pressure prompts the digastric muscle to contract and elevate the hyoid bone, completing the action of the tongue root bulging upward toward the soft palate, thereby assisting the internal articulatory action of /n/—see Figure 5.

The core purpose of the above strategies is to transform abstract articulation essentials into tangible, perceptible bodily experience, and to help learners establish the neural connection between articulatory actions

and perceptual experience through direct bodily feedback, thereby achieving precise calibration of articulatory actions.

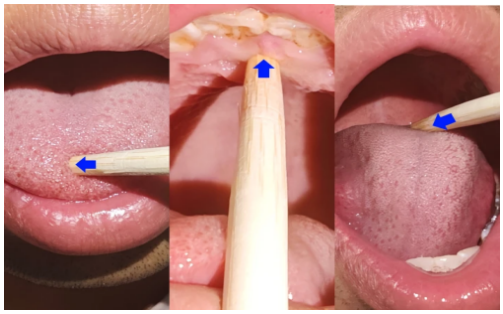

Figure 4. Schematic diagram of tactile and kinesthetic guidance for /n/

Note: This figure was compiled by the authors and photographed on 2026/6/8.

Figure 5. Schematic diagram of tactile and kinesthetic (muscular) guidance for /n/

Note: This figure was compiled by the authors and photographed on 2026/6/12.

4.2.2.3. Contextual transfer principle (contextual transfer)

Embodied cognition theory emphasizes the situatedness of cognition, and holds that the teaching situation should include a learning community jointly constituted by teachers and students. Instruction should create vivid and realistic situations, integrating pinyin learning into gamified tasks and authentic communicative situations. Gamified pinyin tasks can be designed to let students take the initiative to spell and apply pronunciation while solving problems. For example, in classroom design, the teacher can lead students to make "pinyin + Chinese character" name tags placed on desks; and make and affix "pinyin + Chinese character" name tags to the facilities in the classroom—see Figure 6. By associating the bodily experience acquired in class with life-like situations, the sense of meaning of learning can be effectively enhanced, and

the transfer of the articulation model from the classroom environment to daily communication can be promoted, thereby alleviating the problem of being effective in class but easily forgotten after class.

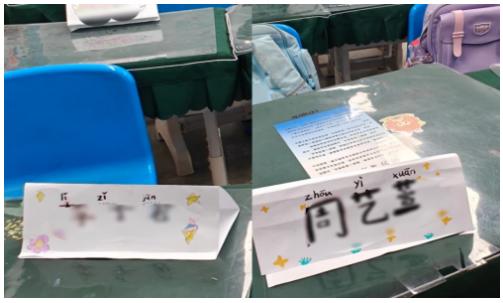

Figure 6. Schematic diagram of desktop "pinyin + Chinese character" name tags

Note: This figure was compiled by the authors and photographed on 2026/6/8.

4.2.3. Extraction of the core category of theoretical explanation and design principles

Integrating the above analysis, this study finds that the core theoretical explanation and design principles that embodied cognition theory provides for /n/ and /l/ teaching can be extracted as follows: the bodily transformation of articulatory knowledge is the construction from conceptual analysis to sensorimotor experience. That is, the core contribution of embodied cognition theory lies in promoting instruction's shift from the conceptual analysis of articulation essentials to the learner's construction of sensorimotor experience; through design principles such as embodied representation, proprioceptive anchoring, and contextual transfer, it transforms abstract articulatory-phonetics knowledge into perceptible and actionable bodily experience, thereby resolving the misalignment between perception and motor control, and ultimately helping learners establish stable articulation schemata and muscle memory at the bodily level.

4.3. Core elements of the embodied Mandarin pronunciation teaching theory (/n/ and /l/)

4.3.1. Theoretical core

The theoretical core of an effective embodied pedagogical theory model for Mandarin alveolar consonants (/n/ and /l/) lies in the organic integration of embodied cognition theory with articulatory-phonetics knowledge. Embodied cognition theory provides an explanatory framework that refuses to treat cognition as the exclusive province of the brain, and holds that the way the body and environment interact is itself part of the operation of mind. It redefines phonetic learning as the acquisition of a kind of bodily skill. Articulatory-phonetics knowledge, for its part, provides the scientific description of the physiological mechanisms of articulation, including the essential differences between /n/ and /l/ in tongue position, tongue-body force points, airflow paths, and soft-palate movement. The integration of the two constitutes the theoretical core of the model—namely, the bodily transformation of articulatory knowledge, which transforms abstract articulatory-phonetics knowledge into bodily experience that learners can perceive, operate, and act upon.

4.3.2. Three core elements

This theory comprises three core elements. First, the Embodied Externalization mechanism (Embodied Externalization): its connotation is that, as implicit procedural knowledge, articulatory actions are difficult for learners to acquire precisely because they cannot directly observe the movement state of their own articulatory organs. This mechanism uses visualization media (dynamic cross-section diagrams, animations, oral-cavity models) and a gestural symbol system to transform invisible physiological processes such as tongue position, soft palate, and airflow into perceptible, operable, and imitable multimodal external representations, realizing

the embodied externalization of implicit bodily procedures into explicit cognitive schemata, and building a cognitive scaffold for learners.

Second, the Proprioceptive Encoding mechanism (Proprioceptive Encoding): its theoretical connotation is that the essence of phonetic acquisition is the consolidation of the motor memory of articulatory-action programs. This mechanism takes tactile feedback and kinesthetic calibration as dual channels. The former establishes a sensory schema by perceiving airflow vibration and distinguishing the articulation channel; the latter uses tools such as tongue depressors and cotton swabs to locate the articulation site and the muscular force point, encoding the articulatory action as proprioceptive memory, so as to realize the neural connection from conscious control to automatic retrieval, and ultimately form stable muscle memory.

Third, the Contextual Transfer mechanism (Contextual Transfer): its connotation is that phonetic skills need to achieve far transfer in authentic communicative situations. Through a hierarchically progressive structure, this mechanism builds a spiraling upward path from controlled practice, to semi-free application, and then to natural communication. Through this path, gamified tasks create quasi-authentic communicative situations and break through classroom boundaries, transforming articulation experience from the teaching situation to the life situation, thereby realizing social transfer. In this process, gamified tasks and life-like associations jointly create application situations, and progressive training ensures the steady transfer of articulation experience from the classroom to daily communication.

4.3.3. Instructional support system

To ensure the implementation of the theoretical model, the model also includes the instructional support system as a core element. At the level of teacher training, teachers need to be provided with readily usable training in pronunciation techniques and in the use of visual teaching resources. At the level of teaching-resource development, supporting resources such as dynamic tongue-position animations, oral-cavity models, and teaching cards need to be developed. Through systematic training and resource support, teachers' articulatory-phonetics knowledge and embodied-teaching ability are enhanced, providing human and resource guarantees for the practical application of the theoretical model.

4.3.4. Integration of the core elements

Integrating the above analysis, this study finds that an effective pedagogical theory model should comprise the following core elements: with the integration of embodied cognition and articulatory phonetics as the theoretical core, the Embodied Externalization mechanism, the Proprioceptive Encoding mechanism, and the Contextual Transfer mechanism as the three core teaching elements, and teacher training and resource development as the instructional support system. The above elements jointly serve one core objective: transforming abstract articulatory-phonetics knowledge into bodily experience that learners can perceive, operate, and experience, thereby integrating the misalignment between perception and motor control in traditional teaching, and helping learners establish stable articulation schemata and muscle memory at the bodily level.

5. Conclusion

5.1. Theoretical connotation and innovation

The theoretical model constructed in this study has as its core connotation the redefinition of phonetic learning as the acquisition process of a kind of bodily skill. This understanding is essentially different from traditional pedagogy. Traditional pedagogy largely relies on the teacher having students listen and discriminate repeatedly through demonstration and contrastive reading, supplemented by rhymes and mnemonic formulas to aid

memory. This model places the teaching focus on auditory input and external lip-shape observation, lacks a scientific explanation of the physiological mechanisms of articulation, and the teacher's correction of pronunciation largely relies on subjective auditory discrimination and repeated demonstration, making it difficult to give students precise guidance at the level of the motor principles of the articulatory organs. Its instructional outcomes are unstable and show large individual differences, manifested as "effective in class but easily forgotten after class", making it difficult to form stable bodily schemata and muscle memory.

This theory, starting from embodied cognition theory, takes articulatory-phonetics knowledge as the core content of instruction. The model comprises three core teaching elements: the Embodied Externalization mechanism, the Proprioceptive Encoding mechanism, and the Contextual Transfer mechanism. These three elements jointly serve the core objective of "the bodily transformation of articulatory knowledge"—that is, through visualization tools the invisible tongue position and airflow path are transformed into observable external representations, through tactile and kinesthetic guidance abstract articulation essentials are transformed into perceptible bodily experience, and through contextualized tasks the pronunciation ability acquired in class is transferred to authentic communicative situations.

The academic value of the core idea of "the bodily return of articulatory knowledge" lies in the new analytical framework it provides for pronunciation-teaching research. Traditional research mostly proceeds from linguistic or psychological perspectives, focusing on the acoustic features of speech and cognitive processing, whereas this theory incorporates bodily experience into the core mechanism of phonetic learning, emphasizing the construction of the sensorimotor experience of articulatory actions. This perspective expands the application boundary of embodied cognition theory in the field of pronunciation teaching, and promotes the theoretical transfer and integration of articulatory-phonetics knowledge from the field of speech rehabilitation to that of general education.

For front-line teachers, this theoretical model provides teaching-strategy references that are more scientific, systematic, and actionable than traditional auditory discrimination and imitation. Teachers no longer judge students' pronunciation right or wrong merely by experience, but can, based on articulatory-phonetics principles, help students establish correct articulation models through visualization tools, tactile guidance, and other means. For teacher training and teaching-resource development, this theoretical model provides direct theoretical grounding for the design of training curricula and the development of supporting resources such as dynamic tongue-position animations, helping to enhance teachers' teaching confidence and professional competence.

5.2. Research limitations and future prospects

This study has the following limitations. First, the theoretical construction is entirely based on literature analysis, lacking empirical testing from real classrooms. The validity and operability of the model have not yet been verified through teaching practice. Second, the literature sample has limitations in source and quantity. Although the study included journal articles and monograph materials, the sample is concentrated in specific fields and periods, and may miss some important research results.

Based on the above limitations, future research can proceed in the following directions. First, quasi-experimental research can be used to apply this model to real teaching scenarios, test its actual effect on improving students' accuracy in /n/ and /l/ pronunciation, and revise the model according to feedback. Second, the scope of the literature sample can be expanded to include research results from more dialect regions and different school stages, so as to enhance the model's generalizability. Third, supporting teaching resources can be developed based on this model—such as dynamic tongue-position animations, teaching cards, and teacher

guides—and systematic teacher-training curricula can be designed to help front-line teachers master embodied teaching methods, thereby promoting the transformation of theoretical results into teaching practice.

References

[1] Wang, H. L. (2025). Current situation and high-quality development path of mandarin popularization in remote ethnic mountainous areas of Guizhou. *Journal of Liupanshui Normal University, 37*(4), 59–68.

[2] Wang, Y. F. (2024). Teaching dilemmas of Chinese phonetic alphabet, textbook interpretation and countermeasures. *Primary Chinese*, (12), 14–17.

[3] Xu, L. N., Li, F., Gao, N., Zhang, Y. Y., & Lü, Z. Y. (2016). Phonetic characteristics and rehabilitation training of alveolar sounds disorders in patients with functional articulation disorders. *Journal of Audiology and Speech Pathology, 24*(4), 321–326.

[4] Wang, Y., & Wan, Q. F. (2026). Application of articulatory phonetics principles in primary-school Pinyin teaching: Taking retroflex consonants in Guizhou as an example. *Guizhou Education*, (3), 92–95.

[5] Xu, X. W. (2010). Principles and theoretical foundations of embodied language cognition. *Yuejiang Academic Journal*, (4), 107–114.

[6] Muller, R. M. (2018). Merleau-Ponty and the radical sciences of mind. *Synthese, 195*(5), 2063–2083.

[7] Lakoff, G., & Johnson, M. (1999). *Philosophy in the flesh: The embodied mind and its challenge to Western thought*. Basic Books.

[8] Garcia, A. M., & Ibáñez, A. (2024). Hand gestures as embodied cognition: Neurophysiological signatures of listening to speech accompanied by gestures. *Language, Cognition and Neuroscience, 39*(2), 167–185.

[9] Qi, W. J., Sui, L., Gong, J., Chen, H., & Zhang, R. (2024). Research progress on internal neural mechanisms and engineering applications of embodied cognition. *Progress in Biomedical Engineering*(1), 45.

[10] Wan, Q. F., & Qiu, C. (2021). On vocal language art teaching under the background of Guizhou dialect: Mandarin phonetics teaching. *Guizhou Education*(16), 36–38.

[11] Tan, J. S. (2024). A review of studies on English pronunciation learning status of Sichuan dialect speakers. *International Journal of Education and Humanities, 54*, 1689.

[12] Chan, M. J. (2023). Using chopsticks to distinguish /l/ and /n/. *Pronunciation in Second Language Learning and Teaching Proceedings, 12*(1).

[13] Chu, R. L., & Ma, X. Y. (2026). Application of grounded theory in college English teaching within second-language acquisition. *Modern English*, (3), 4–6.