International Journal of Academic and Applied Research (IJAAR)
ISSN: 2643-9603
Vol. 9 Issue 12 December - 2025, Pages: 119-125

Research on Foreign Language Acquisition from the Perspective of Embodied Cognition

Xiaoquan Pan[1]*, Huijuan Shao[2]

1. Xingzhi College, Zhejiang Normal University, Jinhua, China
pxq@zjnu.cn
2. Xingzhi College, Zhejiang Normal University, Jinhua, China
hjshao@zjnu.cn
* Corresponding author: Xiaoquan Pan, Xingzhi College, Zhejiang Normal University,
688 Yingbin Road, Jinhua, Zhejiang Province, China. E-mail: pxq@zjnu.cn

Abstract: *Based on the theoretical framework of embodied cognition, this research systematically investigates how three core mechanisms (embodiment, image schemas, and conceptual metaphor) influence foreign language acquisition, aiming to address the traditional dilemma of "form-meaning dissociation" in language teaching. The findings demonstrate that language acquisition is inherently a process of cognitive restructuring driven by bodily experience, rather than the memorization of abstract symbols. Through embodied pedagogical practices (e.g., gesture-assisted pronunciation training, multimodal situational tasks, and VR/AR immersive environments), the neuro-coupling between sensorimotor systems and linguistic processing can be significantly activated, enhancing semantic internalization efficiency. This research develops a "situation-task-interaction" closed-loop teaching model to strengthen functional language application skills. However, the paradigm still faces challenges including unclear synergistic mechanisms for higher-order language abilities, insufficient cross-cultural adaptability, and equity concerns arising from technology dependency. Future research should develop multi-level evaluation systems and blended instructional designs to advance the transformation of foreign language education towards a synergistic "bodily-mind-environment" paradigm.*

Keywords—embodied cognition; foreign language acquisition; embodied cognition; image schemas; conceptual metaphor; language learning design

1. Introduction

Despite the abundance of resources in contemporary foreign language teaching, efficiency remains a persistent challenge for learners. Even after acquiring extensive vocabulary and mastering grammatical rules, many individuals still struggle to spontaneously generate and comprehend a foreign language in authentic communication, that is, being unable to do so in a natural, automatic, and embodied manner. Part of this predicament stems from a longstanding overreliance on "form-meaning dissociation" drills in language pedagogy, which neglects a core paradigm in cognitive-linguistic philosophy: Embodied Cognition. This theory posits that the deep foundations of language comprehension and use are rooted in our bodily sensory systems developed through interactions with the world. It is precisely through the repeated internalization of bodily, motor, and sensory experiences that humans construct the conceptual basis for understanding and imposing order on the world (Gibbs, 2006; Lakoff & Johnson, 1999). Consequently, mastering linguistic knowledge is far more than the memorization of discrete symbols; it is fundamentally a process of cognitive restructuring grounded in physical experiences.

The theoretical value of this study lies in its systematic integration of key mechanisms from embodied cognition, namely embodied simulation, image schema transfer, and the cross-domain projection of conceptual metaphors, into the framework of second language acquisition (SLA). This integration helps explain the underlying cognitive dynamics often overlooked by traditional models. Its practical value resides in providing systematically implementable operational paradigms, through embodied instructional designs (such as situated role-playing and strategies integrating multimodal sensory input), to overcome the barrier between learning linguistic forms and applying them functionally. This perspective thus pioneers a novel pedagogical pathway for transforming language instruction from abstract "form training" towards authentic "meaning practice".

2. The Core Theoretical Foundations of Embodied Cognition

As a profound paradigm shift within the philosophy of cognitive science, embodied cognition fundamentally overturns the traditional information processing model based on the computer metaphor. Its core proposition is that language is not an abstract symbolic system isolated from physical and social experiences, but rather an expressive tool of experience that is highly dependent on and mapped onto the sensorimotor systems of the body. This perspective is built upon four key theoretical pillars, which collectively underpin the crucial role of bodily perception in second language acquisition (Wilson & Foglia, 2017).

2.1 Embodiment

Embodiment constitutes the central pillar of embodied cognition theory, fundamentally rejecting the philosophical premise of Cartesian mind-body dualism prevalent in traditional cognitive science. While the Cartesian "mind as machine" metaphor views cognition as a computational process independent of the body, embodied cognition theory emphasizes that the formation and development of human cognitive abilities are inseparable from the physical structure of the body, the sensorimotor system, and experiences gained through interaction with the environment (Shapiro & Stolz, 2019). Neuroscience provides robust evidence for this view: when we understand the semantics of action verbs like "grasp" or "walk", the brain's motor cortex is activated; similarly, hearing words describing tactile sensations like "smooth" or "rough" triggers synchronized responses in the somatosensory cortex (Pulvermüller, 2013). This indicates that language comprehension is not a process of pure symbolic computation but involves neural simulation that recreates bodily experiences. Essentially, the brain reactivates the perceptual-action systems associated with linguistic symbols through covert imitation, demonstrating that the meaning of language is grounded in the body's experiential interaction with the world (Gallese & Lakoff, 2005).

For foreign language acquisition, this implies that language learning is far more than mere abstract symbol mapping. It is a process through which learners, by means of bodily actions, sensory input, and motor systems, reshape their pre-existing sensorimotor networks. For instance, when L2 learners acquire English action verbs like "kick" or "twist", supplementing learning with actual execution or physical simulation results in significantly deeper semantic encoding compared to groups relying solely on textual memorization (Andrä et al., 2020). This suggests that the efficiency of foreign vocabulary acquisition is positively correlated with the degree of embodied simulation it triggers within the cognitive system. In other words, linguistic "knowing" necessitates its internalization of meaning via bodily "doing".

Within the domain of second language acquisition, the mechanism of embodiment manifests as the reconfiguration of sensorimotor neural pathways. According to the Gesture-for-Conceptualization Hypothesis (Kita et al., 2017), learners "anchor" abstract linguistic forms to the level of bodily experience through gestures, posture, and environmental interaction during foreign language learning, thereby establishing a dual-coding neural network. Take phonetics acquisition as an example: embodied pedagogy significantly improves pronunciation accuracy through techniques like visualizing articulatory movements (e.g., using oral mirrors) and simulating tongue positions with hand gestures (Lin, 2020). Such cases corroborate the core claim of the Embodied Simulation Theory: the perception, production, and comprehension of language rely on the subject's internalized recreation of others' action intentions and sensory states (Rizzolatti & Craighero, 2004).

It is for this very reason that the current paradigm in second language pedagogy is shifting from a disembodied approach towards an embodied one: by designing multimodal perceptual tasks and spatial action games, foreign language forms are directly embedded into the learner's procedural knowledge system. This signifies that the essence of language acquisition is not storing symbols in memory, but rather reconstituting the body as the cognitive medium of language (Wilson & Foglia, 2017). Consequently, embodiment not only reveals the neurobiological foundation of foreign language competence development but also provides physiological grounding for pedagogical reform that moves beyond the "Grammar-Translation Method".

2.2 Image Schemas

Image Schemas are fundamental cognitive frameworks formed through humans' repeated bodily movements and spatial interactions (Johnson, 1987). They are not concrete images but rather abstract, dynamic patterns derived from embodied experiences, constituting the underlying organizational principle of the conceptual system. For instance, the CONTAINER schema originates from the bodily perception of inside/outside spaces; the PATH schema arises from bodily movement trajectories; and the FORCE schema stems from kinesthetic experiences of pushing/pulling objects (Li et al., 2025). These cross-situationally stable schemas possess three key properties: (1) pre-conceptual nature; (2) cross-modality (non-specific to sensory modalities); and (3) spatial topology: preserving relational structures while omitting specific details (Hampe & Grady, 2005). Neuroimaging studies confirm that comprehending language expressions related to these schemas synchronously activates the parietal cortex (spatial processing area) and motor cortex, demonstrating that image schemas serve as neural coupling interfaces connecting language to perceptual and motor experiences (Bajracharya & Peelle, 2023).

In second language acquisition (SLA), image schemas play a central role as cross-linguistic cognitive bridges. Despite surface linguistic differences, all humans share these bodily experience-based schematic structures, providing a deep cognitive pathway for semantic transfer from the first language (L1) to the target language (L2) (Evans, 2007). For example, in acquiring spatial expressions, Chinese learners comprehending the English preposition "over" must activate the PATH+COVERING schema, which originates from bodily experiences shared across cultures.

Empirical evidence shows that SLA proficiency significantly improves when pedagogical methods explicitly invoke schematic associations, such as using line diagrams to illustrate paths or gestures to simulate force actions (Tyler & Evans, 2003). Furthermore, schemas drive grammatical metaphoric extensions: the use of "under" in "under pressure", extending from spatial location to abstract states, constitutes a metaphoric projection of the CONTAINER schema (internal state under external force). This mapping from the spatial to the psychological domain validates the "Schema Projection Hypothesis" (Lakoff, 1990), which posits that the internal logic of abstract grammar is essentially a cognitive transcription of spatial experiences. Therefore, image schemas

International Journal of Academic and Applied Research (IJAAR)
ISSN: 2643-9603
Vol. 9 Issue 12 December - 2025, Pages: 119-125

reveal that the essence of SLA is the re-conceptualization of bodily experiences, not merely the substitution of symbol systems (Hampe & Grady, 2005).

2.3 Conceptual Metaphor

Conceptual metaphor is a meaning-construction mechanism operating across abstract domains: humans utilize concepts formed through embodied experiences to comprehend non-physical, abstract realms (Lakoff & Johnson, 1980). When learners map their L1 metaphorical experiences onto L2 expressions, the linguistic forms gain deeper understanding through cognitive motivation rather than rote memorization. Learning efficiency thus shifts from the superficial level of "translation pairing" to the cognitive level of "conceptual structure transfer".

Conceptual Metaphor Theory (Lakoff & Johnson, 1980) overturned the traditional linguistic view of metaphor as mere rhetorical ornamentation, proposing instead that metaphor constitutes a deep cognitive structure within human conceptual systems. Its core proposition states that humans comprehend abstract, intangible target domains (e.g., time, emotion, morality) through embodied, concrete source domains (e.g., space, motion, force), forming cross-domain mappings. For instance, in the conceptual metaphor LIFE IS A JOURNEY, elements like "starting points", "obstacles", and "forks in the road" from the bodily experience of movement (JOURNEY domain) are systematically projected onto the domain of life, generating natural expressions such as "standing at the start of life's journey" and "overcoming career obstacles".

Cognitive neuroscience research confirms that such metaphors activate the parietal cortex and premotor cortex—brain regions involved in processing spatial actions (Boulenger et al., 2009). The neural representation of abstract concepts is anchored in sensorimotor experiences, proving that metaphorical thinking is fundamentally the cross-domain translation of embodied experiences (Gallese & Lakoff, 2005). Such mappings are systematic and unconscious. Over 70% of abstract expressions in English originate from approximately 200 foundational metaphors (Kövecses, 2010), yet speakers are rarely aware of their metaphorical underpinnings. Metaphor is thus not merely a linguistic phenomenon but the experiential foundation of human abstract thought.

In second language acquisition (SLA), cross-linguistic differences in conceptual metaphors pose deep challenges yet also offer efficient pedagogical pathways. Shared embodied experiences in the L1 (first language) and L2 (target language) provide a cognitive interface for metaphorical transfer. However, cultural models shape divergent mapping preferences: English uses "Time is money" to generate expressions like "save time/waste time", while Chinese conceptualizes "Time as a spatial entity".

3. Towards an Embodied Cognition Approach to Foreign Language Teaching: Design, Methodology, and Evidence

3.1 Situated Task-Based Language Teaching

Situated Task-Based Language Teaching (STBLT) serves as the core pathway to deeply integrate embodied cognition into foreign language classroom practice. It not only disrupts traditional pedagogy's overreliance on abstract rules and symbolic representations but also redefines language learning as a process wherein cognitive agents actively construct meaning through embodied activities in authentic/simulated contexts. Its core design principles emphasize language embodiment and situated embeddedness: learning environments must maximally replicate the sociocultural contexts of target language use (Brown et al., 1989), enabling learners' bodies, senses, actions, and minds to deeply engage in accomplishing task goals (Johnson, 1987). For example, when learning restaurant ordering expressions, students participate in a meticulously designed role-play scenario to complete with a visual illustrated menu and communication pressure to "place orders". They activate multisensory inputs (visually identifying food images, auditorily processing server questions, tactilely manipulating prop menus), using oral production accompanied by physical gestures (e.g., pointing to menu items) for genuine interaction. This pedagogy's cognitive advantage lies in its activation of brain regions responsible for spatial processing and social interaction (e.g., mirror neuron systems). Through bodily engagement (e.g., gesturing to indicate dish locations), learners not only process linguistic forms but also internalize the profound connection between language functions and their embodied contexts and behavioral sequences (Nelson, 1980). This facilitates the transformation of stored knowledge into effective retrieval and application, achieving "somatic-linguistic alignment".

Practice design constitutes the vitality of STBLT. Effective situated contexts should focus on three dimensions: (1) Perceptual Affordances: Teachers leverage multi-modal media (images, videos, short skits, VR/AR virtual scenes, or realia) to construct tangible scenario elements, enabling learners to sensorially anchor spatial scenes and action targets. (2) Task-Driven Engagement: Tasks must compel learners to use target language to fulfill authentic communicative needs or solve problems within specific contexts. The evolving task sequence naturally creates "noticing gaps" in language forms, stimulating autonomous learning motivation (Long, 1996). (3) Gradual Scaffolding: To balance task complexity and cognitive load, teachers provide timely support. In Pre-task, teachers activate background knowledge and linguistic input (e.g., presenting frequently used expressions). In Mid-task, teachers offer modeling (e.g., sample dialogues), visual aids (flowcharts, step prompts), and tiered resources. Scaffolds gradually fade as learners transition from imitation to autonomous construction and output (aligning with Vygotsky (1978)'s Zone of Proximal Development). Immediate feedback based on task performance is critical (e.g.,

International Journal of Academic and Applied Research (IJAAR)
ISSN: 2643-9603
Vol. 9 Issue 12 December - 2025, Pages: 119-125

retrospective metacognitive reflection via audio/video recordings).

Robust evidence (Ellis, 2003 meta-analysis; Pica et al., 1993 negotiation interaction study) confirms that STBLT not only enhances oral fluency and pragmatic competence but also strengthens implicit acquisition and long-term retention of complex grammatical structures (e.g., subjunctive mood, passive voice) within functional contexts.

3.2 Embodied Environments of Virtual Reality (VR) and Augmented Reality (AR)

Virtual Reality (VR) and Augmented Reality (AR) technologies serve as revolutionary tools for realizing the ideal paradigm of experiential-cognitive foreign language teaching. They achieve this by creating highly immersive, interactive, multimodal embodied environments. The presence (Schuemie et al., 2001) these technologies foster surpasses traditional classroom settings. Learners participate in spatialized linguistic events either as digital avatars or directly via information layers superimposed onto real physical space. This profoundly simulates the natural ecology of language occurrence (Swain, 2000).

The core of their cognitive enhancement mechanism lies in the groundbreaking amplification of immersion and embodied action participation: (1) Enhanced Immersion (VR): VR headsets utilize 360° panoramic visuals, spatial audio, and feedback devices to comprehensively block external distractions. They place the learner’s physical senses within highly realistic socio-cultural-physical environments, inducing a powerful sense of spatial presence and emotional arousal (Mikropoulos & Natsis, 2011). This efficiently activates the mirror neuron system, reinforcing the social understanding of linguistic behavior within its context (Kilteni et al., 2012). (2) Driven Embodied Action (AR): AR utilizes smartphones, tablets, or optical see-through headsets to anchor virtual linguistic information precisely to objects (e.g., classroom items) or specific locations (e.g., street shops) within the real environment. This drives learners to perform embodied actions within authentic physical spaces. This immediate coupling of “body-space-language” (Gibson, 1979) significantly optimizes the mechanism of symbolic embodiment mapping. It tightly binds linguistic forms with sensorimotor experiences, forming embodied cognitive schemas (Barsalou, 2008). This lays the neurophysiological foundation for language comprehension, memory, and retrieval. Designing VR/AR teaching applications requires a focus on mapping required scenarios to target language skills and ensuring controllable cognitive load. It also necessitates constructing a “situation-task-interaction-feedback” closed loop. (3) Required Scenario Mapping: This involves carefully selecting high-value scenarios that best exemplify target language functions and skills (e.g., business negotiations, cultural experiences, emergency assistance, complex spatial instructions) for virtual or augmented reconstruction. (4) Controllable Cognitive Load: This is achieved through tiered difficulty mechanisms (e.g., progressing from directive/convergent tasks to open-ended/collaborative tasks) and dynamic interaction support. Key techniques include: System-embedded adaptive scaffolding (e.g., triggering gesture assistance demos or native-language subtitle prompts if speech recognition errors exceed a threshold); Visual focus highlighting (emphasizing crucial interaction objects within the VR/AR environment); and A. I. conversational partner collaboration. This support system aligns with the Zone of Proximal Development (ZPD) principle.

An amount of evidence indicates VR/AR significantly enhances verbal production motivation, reduces anxiety, and improves contextualized vocabulary acquisition and spatial expression ability. However, design must avoid excessive technologization and information overload (preventing “sensory redundancy” from distracting from linguistic input focus). Continuous cultural appropriateness checks during content updates and iterative user interaction optimization are essential. This represents the most promising neuro-situational coupled learning pathway for technology-empowered language teaching under an experiential-cognitive framework.

4. Reflection and Prospects

Although the embodied cognition perspective has infused foreign language acquisition research and practice with transformative momentum, it continues to confront profound challenges at the levels of theoretical construct refinement, empirical breadth, and pedagogical generalizability. The imperative need for finer-grained theoretical integration is paramount: while the embodied cognition framework powerfully elucidates the grounding role of sensorimotor systems in language processing (e.g., Barsalou (2008)’s Perceptual Symbol Systems theory), it has yet to sufficiently incorporate affective motivation, socio-cultural-historical dimensions, and the influence of higher-order cognitive regulatory mechanisms. The neuro-coupling pathways between embodied behaviors (e.g., gesture imitation) and abstract linguistic structures (e.g., subjunctive mood, complex clauses) require clearer modeling (Fogassi & Ferrari, 2007). For instance, experimental evidence demonstrating “gesture-facilitated instruction enhances spatial preposition comprehension” (Tellier, 2008), while confirming an embodied effect, struggles to explain the potential attenuation mechanisms of this effect when fostering complex syntactic awareness or metapragmatic competence.

Simultaneously, significant methodological limitations persist in the empirical domain. Existing research predominantly focuses on short-term, small-sample laboratory tasks (e.g., vocabulary acquisition in VR environments), often neglecting longitudinal tracking of comprehensive language proficiency, particularly in written production and novel cognitive transfer capacities. While neuroimaging techniques (e.g., fMRI, EEG) can reveal neural correlates of language processing within "embodied-situated" contexts (e.g., fronto-parietal network co-activation), their high cost and ecological validity concerns constrain their scalable diagnostic utility in pedagogical applications (Pulvermüller, 2013).

Expanding these insights into pedagogical practice urgently necessitates addressing three major bottlenecks: technological dependency vs. accessibility chasm, teacher professional reorientation, and equity risks. Despite providing revolutionary contextualization, immersive technologies like VR/AR present implementation barriers, such as high costs, operational complexity, and tendencies to induce "sensory distraction" (e.g., excessive focus on the virtual scene at the expense of linguistic form)—that exacerbate digital disparities in resource-constrained regions (Chun et al., 2016). More critically, implementing an "embodied classroom" mandates a fundamental shift in the teacher's role, from knowledge transmitter to orchestrator of context, facilitator of cognition, and diagnostic observer of embodied interactions. This necessitates developing dual-space pedagogical expertise: proficiency in both the eco-design of physical/digital environments (e.g., spatial anchoring of AR tasks, social scripts for VR scenarios) and dynamic assessment literacy capable of interpreting emergent language-cognitive patterns during student embodied interactions (e.g., potential links between gesture morphology and grammatical errors. However, current professional development frameworks lack established pathways for cultivating these competencies. Furthermore, the high-frequency interaction patterns inherent in embodied group tasks may inadvertently amplify individual differences, demanding inclusive design strategies to mitigate embodied inequality (Alibali & Nathan, 2012).

Future research should prioritize: (1) Constructing multi-layered embodied assessment frameworks that integrate eye-tracking (visual focus), motion capture, verbal transcripts, and analytical tests to quantify the transition from explicit knowledge to implicit competence; (2) Exploring cost-efficient, ecologically robust technological alternatives, such as smartphone AR or "hybrid enactment" blending physical spaces; (3) Deepening cross-cultural adaptability research to examine the transferability of embodied strategies within non-Indo-European languages and high-context cultures; and (4) Innovating "teacher-technology-learner" tripartite synergies, for example, developing AI agent-based systems providing real-time interaction feedback to reduce teachers' cognitive load in monitoring multimodal input. Only by confronting these challenges comprehensively can the embodied cognition paradigm genuinely usher foreign language education towards a sustainable future characterized by the resonant integration of "body, mind, and environment".

5. Conclusion

This study, through a systematic integration of embodied cognition theory, unveils the cognitive underpinnings of the persistent form-meaning dissociation problematic in traditional foreign language pedagogy. It substantiates that the core drivers for reconfiguring second language acquisition (SLA) pathways are the tripartite mechanisms of embodiment, image schemas, and conceptual metaphor. Repositioning pedagogy on the sensorimotor system, which is effectively implemented through embodied instructional techniques such as gestural scaffolding, multisensory engagement tasks, and VR/AR-based contextual simulation, proves efficacious in activating neural coupling mechanisms fundamental to language processing. This activation demonstrably enhances both the depth of linguistic form internalization and its functional application. Empirical evidence confirms that experiential classroom design, via multimodal interaction between the body and environment, provides crucial embodied anchors for abstract linguistic knowledge (Lei & Zhang, 2024). Consequently, this approach facilitates a fundamental shift in learner cognition from mere symbolic storage towards meaningful practice.

Nevertheless, advancing this paradigm necessitates addressing three salient challenges: firstly, the synergistic model linking embodied behaviors to higher-order linguistic competencies requires further theoretical and empirical elaboration; secondly, the pedagogical applicability and efficacy across diverse cultural contexts demand stronger empirical corroboration; thirdly, heavy reliance on technology introduces the potential risk of exacerbating educational inequalities. Future research should be directed towards: (1) developing inclusive assessment frameworks for embodied learning; (2) exploring robust low-cost, high-ecological-validity blended learning models; and (3) fostering teacher professional development in transitioning toward roles as "architects of situated scaffolding". The ultimate objective is realizing the dynamic convergence of the "body-mind-environment" continuum within SLA.

References

[1] Alibali, M. W., & Nathan, M. J. (2012). Embodiment in mathematics teaching and learning: Evidence from learners' and teachers' gestures. *Journal of the Learning Sciences*, 21(2), 247-286. https://doi.org/10.1080/10508406.2011.611446

[2] Andrä, C., Mathias, B., Schwager, A., Macedonia, M., & von Kriegstein , K. (2020). Learning Foreign Language Vocabulary with Gestures and Pictures Enhances Vocabulary Memory for Several Months Post-Learning in Eight-Year-Old School Children. *Educational Psychology Review*, 32, 815-850. https://doi.org/10.1007/s10648-020-09527-z

[3] Bajracharya, A., & Peelle, J. E. (2023). A systematic review of neuroimaging approaches to mapping language in individuals. *Journal of Neurolinguistics*, 68: 101163. https://doi.org/10.1016/j.jneuroling.2023.101163

[4] Barsalou, L. W. (2008). Grounded Cognition. *Annual Review of Psychology*, 59, 617-645. https://doi.org/10.1146/annurev.psych.59.103006.093639

[5] Boulenger, V., Hauk, O., & Pulvermüller, F. (2009). Grasping ideas with the motor system: Semantic somatotopy in idiom comprehension. *Cerebral Cortex*, 19(8), 1905-1914. https://doi.org/10.1093/cercor/bhn217

[6] Brown, J. S., Collins, A., & Duguid, P. (1989). Situated Cognition and the Culture of Learning. *Educational Research*, 18, 32-42. http://dx.doi.org/10.3102/0013189X018001032

[7] Chun, D., Smith, B. and Kern, R. (2016) Technology in Language Use, Language Teaching, and Language Learning. *The Modern Language Journal*, 100, 64-80. https://doi.org/10.1111/modl.12302

[8] Ellis, R. (2003). *Task-Based Language Learning and Teaching*. Oxford: Oxford University Press.

[9] Evans, J. S. B. T. (2007). *Hypothetical thinking: Dual processes in reasoning and judgement*. Psychology Press.

[10] Fogassi, L., & Ferrari, P. F. (2007). Mirror neurons and the evolution of embodied language. *Current Directions in Psychological Science*, 16(3), 136-141. https://doi.org/10.1111/j.1467-8721.2007.00491.x

[11] Gallese, V., & Lakoff, G. (2005). The Brain's Concepts: The Role of the Sensory-Motor System in Conceptual Knowledge. *Cognitive Neuropsychology*, 21, 1-26. https://doi.org/10.1080/02643290442000310

[12] Gibbs, R. W. (2006). *Embodiment and cognitive science*. Cambridge University Press.

[13] Gibson, J. J. (1979). *The ecological approach to visual perception*. Houghton, Mifflin and Company.

[14] Hampe, B., & Grady, J. (2005). *From perception to meaning: Image schemas in cognitive linguistics*. Berlin/New York: Mouton de Gruyter.

[15] Johnson, M. (1987). *The body in the mind: The bodily basis of meaning, imagination, and reason*. University of Chicago Press.

[16] Kilteni, K., Normand, J. M., Sanchez-Vives, M.V., & Slater, M. (2012). Extending body space in immersive virtual reality: a very long arm illusion. *PLoS One*, 7(7): e40867. https://doi.org/10.1371/journal.pone.0040867

[17] Kita, S., Alibali, M. W., & Chu, M. (2017). How do gestures influence thinking and speaking? The gesture-for-conceptualization hypothesis. *Psychological Review*, 124(3), 245-266. https://doi.org/10.1037/rev0000059

[18] Kövecses, Z. (2010). *Metaphor: A Practical Introduction* (2nd ed.). Oxford University Press.

[19] Lakoff, R. T. (1990). *Talking Power: The Politics of Language in Our Lives*. Glasgow: HarperCollins.

[20] Lakoff, G., & Johnson, M. (1980). The metaphorical structure of the human conceptual system. *Cognitive Science*, 4(2), 195-208. https://doi.org/10.1207/s15516709cog0402_4

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

[22] Lei, Q., & Zhang, C. (2024). Using multimodal resources to design EFL classroom lead-ins: A multimodal pedagogical stylistics perspective. *Linguistics and Education*, 83: 101338. https://doi.org/10.1016/j.linged.2024.101338

[23] Li, Z., Buruk, O. O., Bujić, M., Kerouš, B., & Hamari, J. (2025). Natural manipulation of virtual reality: The effects of kinesthetic gloves with force feedback on user performance and experience for accurate 3D object manipulation. *International Journal of Human-Computer Studies*, 199: 103507. https://doi.org/10.1016/j.ijhcs.2025.103507

[24] Lin, Y.-L. (2020). A helping hand for thinking and speaking: Effects of gesturing and task planning on second language narrative discourse. *System*, 91: 102243. https://doi.org/10.1016/j.system.2020.102243

[25] Long, M. H. (1996). The role of the linguistic environment in second language acquisition. In W. C. Ritchie, & T. K. Bhatia (Eds.), *Handbook of second language acquisition* (pp. 413-468). New York: Academic Press.

[26] Mikropoulos, T. and Natsis, A. (2011) Educational Virtual Environments: A Ten-Year Review of Empirical Research (1999-2009). *Computers & Education*, 56, 769-780. http://dx.doi.org/10.1016/j.compedu.2010.10.020

[27] Nelson, S. A. (1980). Factors influencing young children's use of motives and outcomes as moral criteria. *Child Development*, 51(3), 823-829. https://doi.org/10.2307/1129470

[28] Pica, T., Kanagy, R., & Falodun, J. (1993). Choosing and Using Communicative Tasks for Second Language Instruction. In G. Crookes & S. Gass (Eds.), *Tasks and Language Learning: Integrating Theory and Practice* (pp.9-34). Multilingual Matters, Clevedon.

[29] Pulvermüller, F. (2013) How Neurons Make Meaning: Brain Mechanisms for Embodied and Abstract-Symbolic Systems. *Trends in Cognitive Sciences*, 17, 458-470. http://dx.doi.org/10.1016/j.tics.2013.06.004

[30] Rizzolatti, G., & Craighero, L. (2004). The mirror-neuron system. *Annual Review of Neuroscience*, 27, 169-92. https://doi.org/10.1146/annurev.neuro.27.070203.144230

[31] Schuemie, M. J., Van Der Straaten, P., Krijn, M., & Van Der Mast, C. A. (2001). Research on Presence in Virtual Reality: A Survey. *CyberPsychology & Behavior*, 4, 183-201. https://doi.org/10.1089/109493101300117884

[32] Shapiro, L., & Stolz, S. (2019). Embodied Cognition and Its Significance for Education. *Theory and Research in Education*, 17(1), 19-39. https://doi.org/10.1177/1477878518822149

[33] Swain, M. (2000). The Output Hypothesis and beyond: Mediating Acquisition through Collaborative Dialogue. In J. P. Lantolf (Ed.), *Sociocultural Theory and Second Language Learning* (pp. 97-114). Oxford: Oxford University Press.

[34] Tellier, M. (2008). The effect of gestures on second language memorisation by young children. *Gesture*, 8(2), 219-235. https://doi.org/10.1075/gest.8.2.06tel

[35] Tyler, A., & Evans, V. (2003). *The Semantics of English Prepositions: Spatial Scenes, Embodied Meaning and Cognition*. Cambridge: Cambridge University Press.

[36] Vygotsky, L. S. (1978). *Mind in Society: The Development of Higher Psychological Processes*. Cambridge, MA: Harvard University Press.

[37] Wilson, R. A., & Foglia, L. (2017). Embodied Cognition. In E. N. Zalta & U. Nodelman (Eds.), *The Stanford Encyclopedia of Philosophy* (p. 1). Stanford, CA: Metaphysics Research Lab, Stanford University.