

○○○○

MULTIMODAL EMBODIED LEARNING DESIGN IN THE SONG "HAWASUL KHAMSA":

Constructing Arabic Vocabulary Meaning for
the Alpha Generation

The 19th CONAPLIN International Conference | Multimodal Discourse Analysis

Presented by Annisa Vadia Haya — Linguistics, Universitas Pendidikan Indonesia

Research Background

Multimodal meaning construction plays a crucial role in language learning. Early Arabic learning often relies on memorization and translation. Songs combine speech, rhythm, repetition, attention, and movement. This study examines how multimodal design constructs meaning in Hawasul Khamsa through gesture and embodied cognition.

Previous Studies

Previous studies reinforce the importance of this issue, but they do not address the focus of this article directly. Research on vocabulary learning through movement shows the benefits of physical activity, gesture, and multisensory enrichment for children (Andrä et al., 2020; Schmidt et al., 2019; Toumpaniari et al., 2015). In Arabic language learning, TPR, lyric videos, singing strategies, and visual-auditory-kinesthetic models have also been reported to support vocabulary acquisition (Afiah & Musyafa'ah, 2024; Kiswanto et al., 2023; Kudsiyah et al., 2021; Maburoh et al., 2024; Raharjo & Siswanto, 2021; Ridwan & Awaluddin, 2019). In early childhood education, movement songs themed around Arabic and the five senses have likewise been discussed as media for language introduction and sensory experience (Nurjanah et al., 2024; Rahma et al., 2025; Riuddani & Pamungkas, 2025). However, most of these studies focus on the effectiveness of media or learning strategies, rather than on a systematic analysis of how lyrics, sound, visuals, and gesture work together to construct meaning. This is where the present study fills the gap.



- The importance of multimodal meaning construction in language learning
- Hawasul Khamsa as a medium for Arabic vocabulary learning
- The role of gesture in multimodal communication and interaction
- Embodied cognition as the theoretical foundation of the study



Research Questions & Objectives

This study focuses on analyzing gesture functions in the multimodal discourse of the song Hawasul Khamisa as a medium for Arabic vocabulary learning. It employs a qualitative descriptive approach with a multimodal discourse analysis framework to understand how gesture supports meaning construction.



How do the four main modes work together in each key scene of Ḥawāsul Khamisa?



What types and functions of gesture appear in the relation between lyric and movement?



How can these cross-modal relations be explained through Embodied Cognition as a basis for Arabic vocabulary meaning-making?

Qualitative Descriptive Approach

Employing a multimodal discourse analysis framework to examine the interaction among modalities in audiovisual texts.

Linguistic & Visual Modalities

Analysis of song lyrics and accompanying visual elements in each verse to understand co-constructed meaning.

Data Source: Video Recording

The primary data consists of a video recording of the song Hawasul Khamisa featuring gestures and vocal expressions.

Auditory & Gestural Modalities

Gesture and auditory element coding conducted integratively to identify multimodal functions and meanings.

Research Method

This study employs a qualitative descriptive approach with a multimodal discourse analysis framework. The data is sourced from a video recording of the song Hawasul Khamisa, systematically analyzed based on four main modalities: linguistic, visual, auditory, and gestural.



Data Analysis Process



01

Transcription and Gesture Coding

The video recording of Hawasul Khamisa was fully transcribed, covering song lyrics, body movements, and visual expressions accompanying each verse.

02

Gesture Type Classification

Gestures were categorized by function: deictic (pointing to body parts), iconic (sensorimotor imitation), and other types according to the multimodal analysis framework.

03

Embodied Cognition Perspective Integration

Each gesture was interpreted through the lens of embodied cognition, examining how sensorimotor experience is manifested in physical movements that support vocabulary comprehension.

04

Integrated Multimodal Analysis

Four main modalities—linguistic, visual, auditory, and gestural—were analyzed integratively to understand holistic meaning construction in Arabic language learning media.



Main Findings

The multimodal analysis of the song Hawasul Khamsa reveals the dominance of two main gesture types: deictic and iconic. Both play a crucial role in constructing meaning and supporting Arabic vocabulary learning through concrete and integrated sensorimotor experience.

➔ Dominant deictic gestures

➔ Iconic sensorimotor

➔ Pointing to body parts

➔ Imitating sensory actions

One syntactic pattern, five senses

[organ / instrument phrase] + [1st-person perceptual verb] + [object]

e.g. bi-l-‘aynayni arā al-a‘lām — “with my (two) eyes I see the flags”

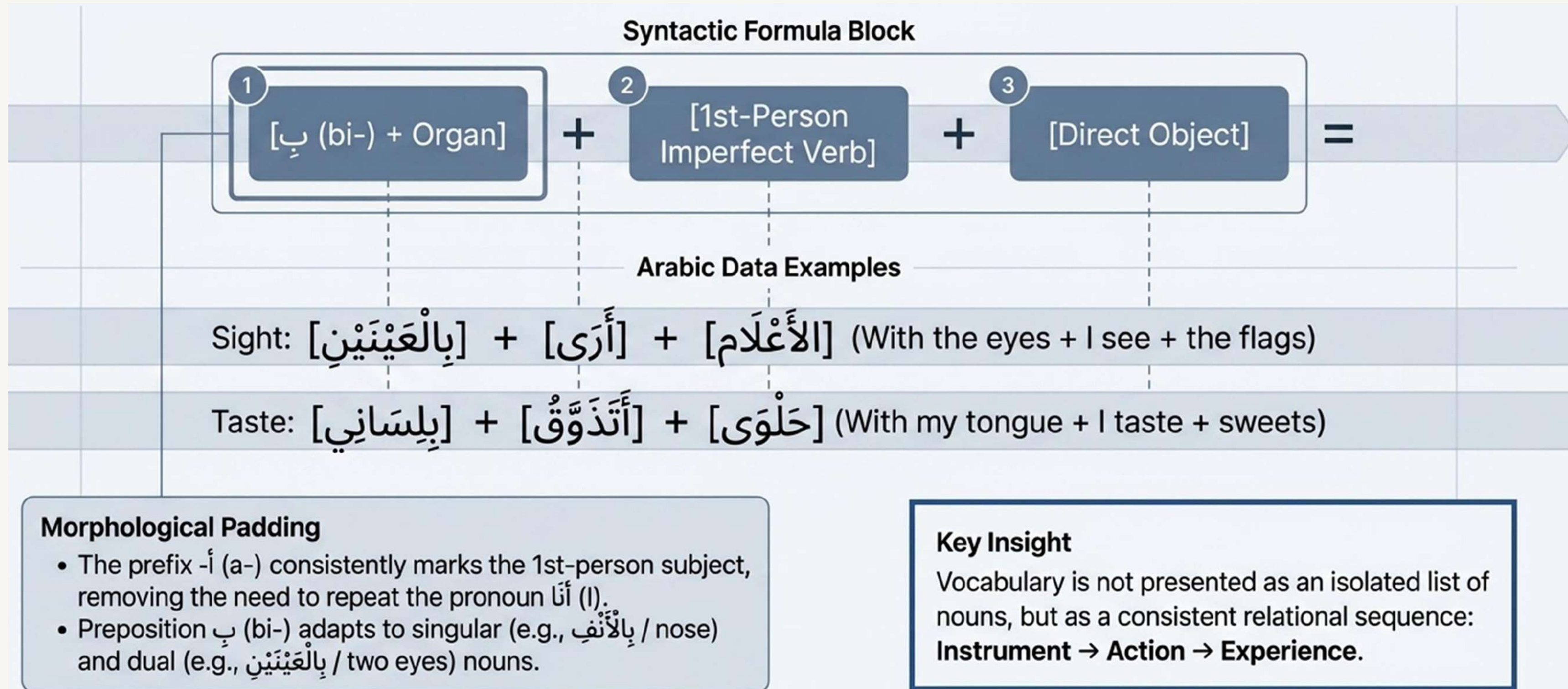
- 10 fixed lyric units recur across 6 thematic scenes
- First-person voice (a-, ‘I’) makes each sensory act personally owned — an interpersonal effect
- Parallel structure across the 5 sense-verses turns vocabulary into a relation — organ → sensory action → object/effect — not a bare noun list
- The flag/bird refrain recurs 10x, working as a textual bridge that keeps verses cyclically connected



Opening scene: five fingers open on “I have five senses”



Linguistic Scaffold: The Syntactic Formula (HK/Ba/B.1–B.6)



Visual Framing:

The Body as Semiotic Center (HK/Vi/B.1–B.6)



Fig. 1 – Opening: Symmetrical group formation. The human performer is immediately established as the subject possessing the senses.



Fig. 2 & 6 – Smell & Taste: Visual focus zooms to specific facial areas creating a direct organ-object pairing.



Fig. 4 & 5 – Sight & Hearing: Frame orientation shifts outward and upward, expanding the spatial context from the body to the external environment.



The visual mode organizes meaning by maintaining the performer's body as the constant point of origin, while secondary objects expand the perceptual context.

Two layers of gesture: point, then perform



Phase 1 — Deictic

Hand raised and open — introduces “hand” as a lexical referent



Phase 2 — Iconic

Physical contact with a peer’s face — directly depicts the action “to touch”

- Layer 1**
Deictic touch/point on the named organ — stable across nearly every occurrence
- Layer 2**
Iconic enactment of the action, direction, or object — more variable across performers and scenes

Visual composition keeps the body as the orientation point in every frame, pairing each organ with its object — nose–flower, ear–bird, hand–grandmother’s face — which extends touch into a relational, kinship meaning rather than a bare body-part label.

Evidence-Based Gesture Typology

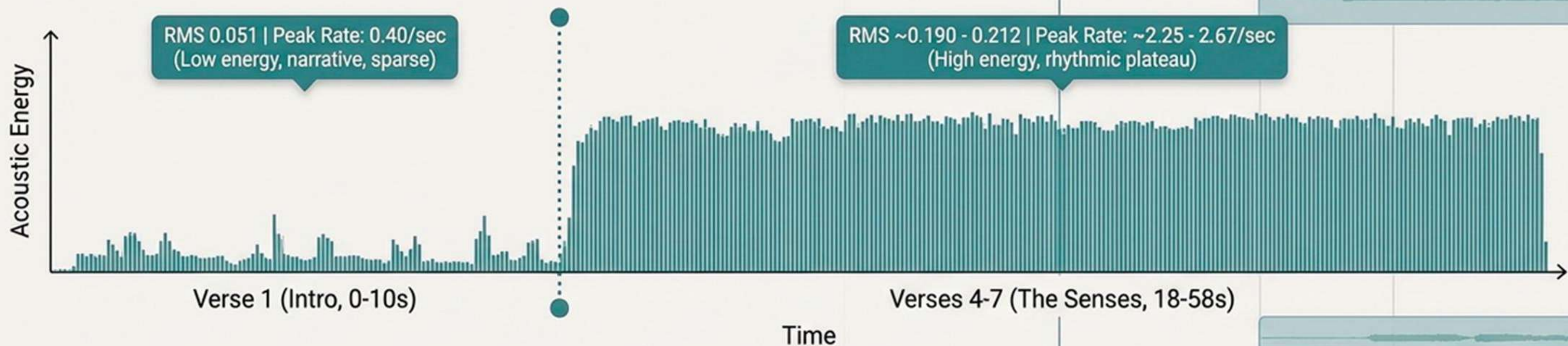
Gesture Typology Table

Data Code	Local Function	Theoretical Classification
HK/Ge/B.1 (Intro)	Pointing to chest, opening 5 fingers.	Self-referential Deictic & Quantitative Iconic .
HK/Ge/B.2 (Smell)	Finger anchored to nose.	Deictic .
HK/Ge/B.4 (Sight)	Finger anchored below eyes.	Deictic .
HK/Ge/B.5 (Hearing)	Touching ear, raising arms like wings.	Deictic (ear) + Iconic (bird wings) .
HK/Ge/B.6 (Taste)	Pointing to mouth, flexing arms.	Deictic (mouth) + Iconic (strength) .

Theoretical Note

Classification is applied after observing local functions, avoiding intuitive guessing. Consistent deictic layers anchor the organ, while variable iconic layers enact the resulting action (Arachchige et al., 2021; Oppici et al., 2023).

Auditory Architecture: The Rhythmic Scaffold



Analytical Correction

- **Initial Assumption:** Music rhythm modulates specifically to match each distinct sense organ.
- **Empirical Reality:** The auditory mode functions as a steady rhythmic scaffold. Once the verse begins, energy stabilizes with minimal inter-verse variation (RMS range only ~0.02).

Key Insight

The auditory mode's primary affordance is textual and interpersonal—synchronizing the emergence of words and gestures while sustaining continuous learner attention, rather than encoding specific lexical meaning.



17

The Semiotic Synthesis Matrix



Linguistic



Visual



Auditory



Gestural



Ideational

(Experience & Representation)

Labels organs, actions, objects, and consequences.

Represents 5 senses as bodily actions within a spatial context.

(N/A - primary function is not representational).

Translates verbal actions into observable physical movements.



Interpersonal

(Engagement & Attitude)

1st-person perspective (I see, I taste) creates immediacy.

Direct body orientation and interactive dyads build relational proximity.

High-energy rhythm builds an active, engaging emotional atmosphere.

Physical contact (touch) introduces effective qualities.



Textual

(Cohesion & Sequence)

Cyclic repetition of the Syntactic Formula and refrain (flag/bird).

Consistent sequencing of organs, ending with the whole body (strength).

Stable RMS plateaus bind scenes together across time.

Temporal synchronization of hand movements with stressed syllables.



Interpretation & Embodied Cognition

The research findings indicate that gestures in the song Hawasul Khamsa are not merely visual supplements, but direct representations of bodily experience. The embodied cognition perspective explains how the body plays an active role in meaning-making and cognitive processes in Arabic language learning.



→ **Gestures embody sensorimotor experience directly**

→ **Multimodal integration strengthens Arabic acquisition**

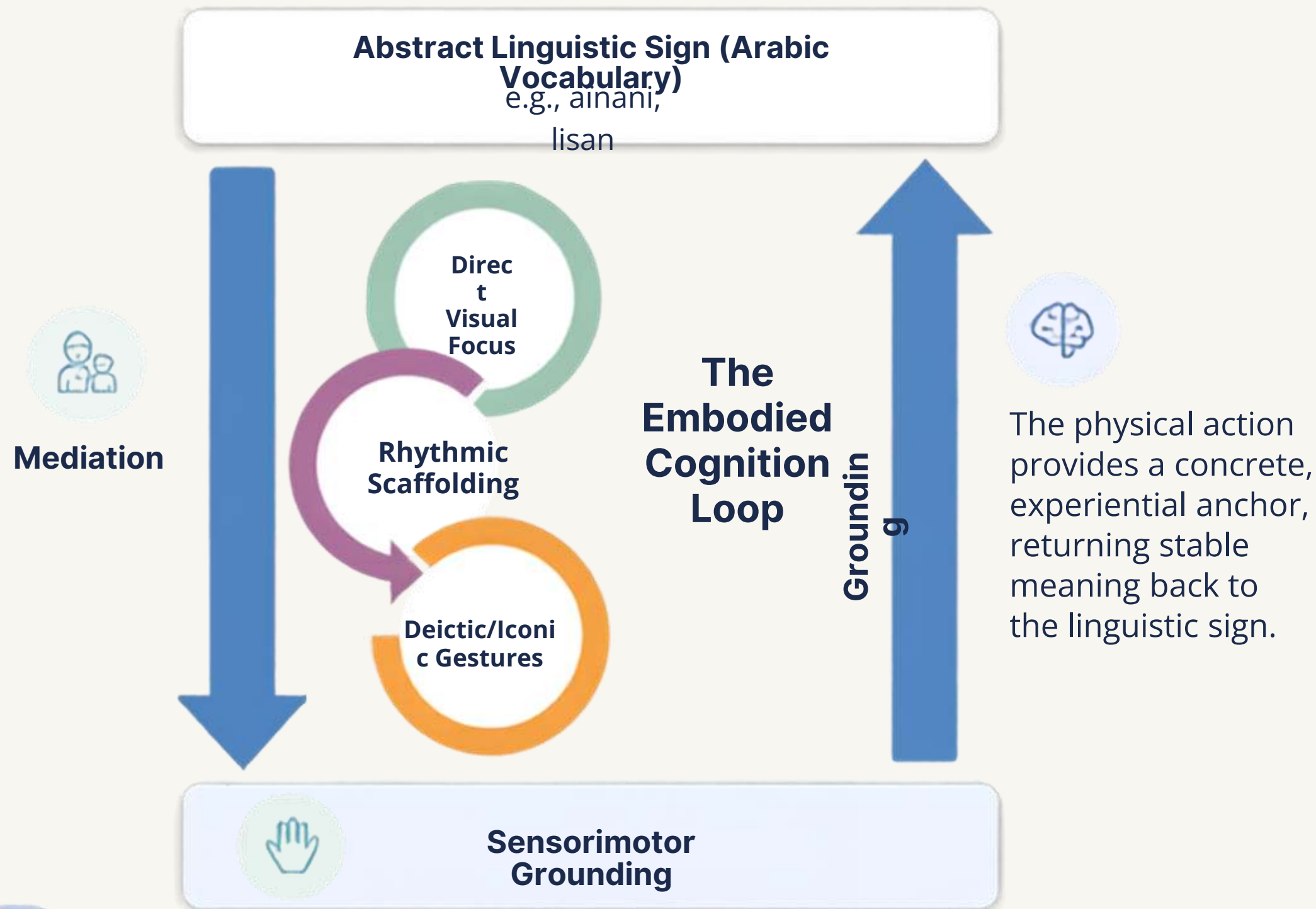
→ **Gesture as a cognitive and communicative tool**

→ **Body-mind connection in language learning**



Theoretical Implications: The Embodied Cognition Loop

The Embodied Cognition Loop



Theoretical Implication

Hawasul Khamsa does not just teach a list of nouns. It maps unfamiliar linguistic forms directly onto the learner's pre-existing physical schemas (Jusslin et al., 2022; Macedonia, 2019). Words become actions before they become translations.

Conclusion and Research Implications

- Deictic and iconic gestures were found to dominate multimodal meaning construction in the song Hawasul Khamsa, with pointing and sensorimotor imitation functions that reinforce Arabic vocabulary comprehension.
- A strong congruence exists among visual, linguistic, auditory, and gestural modes in conveying meaning, thereby creating an integrated and meaningful learning experience.
- These findings carry significant implications for multimodal-based Arabic language teaching and open opportunities for further research on the role of gesture in language learning and multimodal discourse.



References



Afiah, S., & Musyafa'ah, L. (2024). Penerapan Metode Total Physical Respon dalam Pembelajaran Bahasa Arab. KOLEKTIF: Jurnal Pendidikan, Pengajaran, Dan Pembelajaran. <https://doi.org/10.70078/kolektif.v1i1.14>

Andra, C., Mathias, B., Schwager, A., Macedonia, M., & Kriegstein, K. (2020). Learning Foreign Language Vocabulary with Gestures and Pictures Enhances Vocabulary Memory. *Educational Psychology Review*, 1-36. <https://doi.org/10.1007/S.-Z>

Arachchige, K. K. K., Loureiro, I. S., Blekic, W., Rossignol, M., & Lefebvre, L. (2021). The Role of Iconic Gestures in Speech Comprehension. *Frontiers in Psychology*, 12. <https://doi.org/10.3389/fpsyg.2021.634074>

Bezemer, J., & Jewitt, C. (2018). Multimodality: A Guide for Linguists. *Research Methods in Linguistics*. <https://doi.org/10.5040/9781350043466.ch-012>

Busse, V., Hennies, C., Kreutz, G., & Roden, I. (2021). Learning grammar through singing? An intervention with EFL primary school learners. *Learning and Instruction*, 71, 101372. <https://doi.org/10.1016/j.learninstruc.2020.101372>

Hamilton, C., Schulz, J., Chalmers, H., & Murphy, V. A. (2024). Investigating the substantive linguistic effects of using songs for teaching second or foreign languages. *System*. <https://doi.org/10.1016/j.system.2024.103350>

Jusslin, S., Korpinen, K., Lilja, N., Martin, R., Lehtinen-Schnabel, J., & Anttila, E. (2022). Embodied learning and teaching approaches in language education. *Educational Research Review*. <https://doi.org/10.1016/j.edurev.2022.100480>

Kiswanto, D., Abidin, M., Holis, H., & Mahdi, R. (2023). Exploring the Role of Total Physical Response Method in Arabic Vocabulary Learning. *ALSUNIYAT*. <https://doi.org/10.17509/alsuniyat.v6i2.58293>

Kress, G. (2009). Multimodality: A Social Semiotic Approach to Contemporary Communication. <https://doi.org/10.4324/9780203970034>

Kress, G., & Leeuwen, T. V. (1996). Reading Images: The Grammar of Visual Design. <https://doi.org/10.5860/choice.34-1950>

Kudsiyah, K., Mauludiyah, L., & Murdiono, M. (2021). Arabic Video Lyric Untuk Meningkatkan Pemahaman Kosakata Bahasa Arab Siswa. *JPBA*, 5(1). <https://doi.org/10.15575/JPBA.V5I1.11351>

Liu, D., & Qin, J. (2024). The effectiveness of cognitive linguistics-inspired language pedagogies. *The Modern Language Journal*, 108, 794-814. <https://doi.org/10.1111/modl.12959>

Mabruroh, H., Nurfadlilah, S., Zarkasyi, A. H., & Nufus, N. H. (2024). Implementation of Visual Auditory Kinesthetic Learning Model in Mastering Arabic Vocabulary. *Al Mahara*. <https://doi.org/10.14421/almahara.2024.0102-02>

Macedonia, M. (2019). Embodied Learning: Why at School the Mind Needs the Body. *Frontiers in Psychology*, 10. <https://doi.org/10.3389/fpsyg.2019.02098>

Macedonia, M. (2025). Your Body as a Tool to Learn Second Language Vocabulary. *Behavioral Sciences*, 15. <https://doi.org/10.3390/bs15080997>

Mahmudah, M. (2025). The Use of Songs and Movement in Teaching Arabic for Early Grade Learners. *EDUKASIA Jurnal Pendidikan Dan Pembelajaran*. <https://doi.org/10.62775/edukasia.v6i1.1427>

Moufarrej, G., & Salameh, C. (2020). The Effects of Songs on Vocabulary Retention in Foreign Language Acquisition: The Case of Arabic. *Al-Arabiyya*, 52, 101-123.

Nguyen, M. (2024). The embodied nature of students' engagement during a total physical response activity. *Classroom Discourse*, 16, 26-51. <https://doi.org/10.1080/19463014.2024.2360412>

Nurjanah, S., Koesmadi, D. P., & Rachman, B. (2024). Peningkatan Kemampuan Bahasa Arab Melalui Gerak Lagu Asmaul Husna. *PELANGI*. <https://doi.org/10.52266/pelangi.v6i1.2754>

Oppici, L., Mathias, B., Narciss, S., & Proske, A. (2023). Benefits of Enacting and Observing Gestures on Foreign Language Vocabulary Learning. *Behavioral Sciences*, 13. <https://doi.org/10.3390/bs13110920>

Raharjo, H., & Siswanto, N. D. (2021). Penerapan Strategi Total Physical Response, Bernyanyi, Dan Permainan Edukatif. *Zuriah*, 2(1). <https://doi.org/10.29240/ZURIAH.V2I1.3363>

Rahma, E., Fitriani, Y., & Arzaqi, R. N. (2025). Analisis Konten Video Klip Lagu Lima Indra sebagai Media Pengenalan Panca Indra Anak Usia Dini. *Aulad*, 8(2). <https://doi.org/10.31004/aulad.v8i2.1056>

Ridwan, R., & Awaluddin, A. (2019). Penerapan Metode Bernyanyi dalam Meningkatkan Penguasaan Mufradat. *DIDAKTIKA*, 13(1). <https://doi.org/10.30863/didaktika.v13i1.252>

Riuddani, A. L., & Pamungkas, J. (2025). Gerak Lagu sebagai Metode Pengenalan Bahasa Arab dan Nilai Tauhid. *Murhum*, 6(2). <https://doi.org/10.37985/murhum.v6i2.1465>

Schmidt, M., Benzing, V., Wallman-Jones, A., Mavilidi, M., Lubans, D., & Paas, F. (2019). Embodied learning in the classroom: Effects on primary school children's attention. *Psychology of Sport and Exercise*. <https://doi.org/10.1016/J.PSYCHSPORT.2018.12.017>

Sun, B., Yang, J., & Liang, Z. (2024). How music facilitates second language (L2) learning. *Innovation in Language Learning and Teaching*, 19, 155-173. <https://doi.org/10.1080/17501229.2024.2350490>

Swanborough-Nilson, C. A. (2020). Actions and words: Meaning making in a bilingual Arabic/English kindergarten. <https://doi.org/10.21954/OU.RO.00011872>

Toumpaniari, K., Loyens, S. M. M., Mavilidi, M., & Paas, F. (2015). Preschool Children's Foreign Language Vocabulary Learning by Embodying Words. *Educational Psychology Review*, 27, 445-456. <https://doi.org/10.1007/s10648-015-9316-4>

Witte, A. (2026). Embodiment, enaction, and the lived body in foreign language learning. *Frontiers in Language Sciences*. <https://doi.org/10.3389/flang.2025.1761548>

Zhang, Y., Baills, F., & Prieto, P. (2023). Singing Songs Facilitates L2 Pronunciation and Vocabulary Learning. *Languages*, 8(3), 219. <https://doi.org/10.3390/languages8030219>

Thank You

Annisa Vadia Haya

Linguistics, Universitas Pendidikan Indonesia

The 19th CONAPLIN International Conference

10 August 2026