

Concept Paper

Not peer-reviewed version

WuYi. A Three-Level Cascade Architecture for Learning Chinese Radicals Through Sequential Multimodal Encoding, Narrative Chaining, and Mythological Macro-Organization

[Stanislav E. Lauk-Dubitskiy](#)*

Posted Date: 17 March 2026

doi: 10.20944/preprints202603.1303.v1

Keywords: Chinese character learning; multimodal mnemonics; cascade encoding; embodied cognition; Wu Xing; narrative curriculum; radical pedagogy; CFL



Preprints.org is a free multidisciplinary platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This open access article is published under a [Creative Commons CC BY 4.0 license](#), which permit the free download, distribution, and reuse, provided that the author and preprint are cited in any reuse.

Disclaimer/Publisher's Note: The statements, opinions, and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions, or products referred to in the content.

Concept Paper

WuYi. A Three-Level Cascade Architecture for Learning Chinese Radicals Through Sequential Multimodal Encoding, Narrative Chaining, and Mythological Macro-Organization

Stanislav E. Lauk-Dubitskiy

MIREA – Russian Technological University ORCID: 0000-0001-7004-2511 Email: laukdubitskiy@mirea.ru

Abstract

This paper presents WuYi (五仪 "Five Rites"), a methodology for learning Chinese characters based on a three-level cascade architecture that integrates sequential multimodal encoding, inter-item narrative chaining, and a culturally grounded macro-narrative organized according to Wu Xing (五行) philosophy and classical Chinese mythology. At the micro level, five cognitive modalities—mental visualization (Fire), phonological construction (Metal), kinesthetic anchoring (Wood), theatrical episodic simulation (Earth), and graphomotor reconstruction (Water)—are activated in a prescribed sequence with explicit transition criteria, overcoming working memory limitations through temporal unfolding rather than parallel presentation. At the meso level, 2–5 radicals are linked through continuous causal narratives that simultaneously serve discriminative, compositional, and retrieval functions. At the macro level, the entire corpus of 214 Kangxi radicals is distributed across a two-cycle mythological structure—Cosmogonic Cycle (74 radicals) and Legendary Cycle (131 radicals)—each traversing five Wu Xing phases aligned with the canonical mythology of Nüwa, Shennong, Huangdi, and Fuxi. The methodology introduces several novel mechanisms: synthetic narrative calligrams that encode tonal contours through typographic modulation; chimeric tone spirits that bind homophonic morphemes across all four tones into single mnemonic characters; deferred mnemonic anchors that create proactive facilitation through spreading activation; and narrative-aligned primary encoding with autonomous fallback mnemonics activated through self-diagnosis. The theoretical framework integrates dual coding theory, levels of processing, embodied and situated cognition, cognitive load theory, the SPT effect, hierarchical retrieval cues, and narrative transportation theory. A between-subjects experimental protocol (N=90, three groups) for controlled validation is provided. No prior work was found that combines sequential multimodal cascade encoding, inter-item narrative chaining, or mythological macro-organization of character curricula.

Keywords: Chinese character learning; multimodal mnemonics; cascade encoding; embodied cognition; Wu Xing; narrative curriculum; radical pedagogy; CFL

1. The Problem of Traditional Methods

1.1. Limitations of Mechanical Repetition

Traditional approaches to teaching Chinese writing rely on mechanical repetition (rote learning) and focus on the visual-graphic code: repeatedly copying characters, often without explicit work on phonology, radical structure, and bodily experience. Studies show that one-sided, shallow encoding at the level of the symbol's form leads to limited trace durability and high forgetfulness in delayed tests, especially in adults learning Chinese as a foreign language (Shen, 2005a; Kuo & Hooper, 2004; Baddeley, 1986).

Several works emphasize that mastering reading and writing Chinese characters requires hundreds of hours of purposeful practice and meaningful attention to radical structure, not just the holistic form. Passive visual perception creates shallow memory traces localized predominantly in the ventral visual stream, without forming strong associative connections with phonology, motor skills, and episodic memory (Clark & Paivio, 1991; Wang et al., 1992; Shen, 2005b; Taft & Chung, 1999; Barsalou, 2008).

1.2. Disadvantages of Existing Mnemonic Systems

Modern systems such as the Hanzhi Movie Method (Mandarin Blueprint) and classic keyword approaches for vocabulary focus mainly on visual-verbal encoding, often ignoring the learner's bodily experience and the motor aspects of writing. They require mastering an extensive nomenclature of "primitives" and do not always integrate the phonological component and actual gesture execution, while the motor component is reduced to traditional copying treated as a separate task (Heisig, 2007; Cowan, 2001).

In research on computer- and Kinect-based systems for teaching Chinese characters, motor skills are most often used either as an interface (gesture control) or as "air writing" (shukong), but are not included in multi-level narrative mnemonics with phonological and episodic layers. The lack of meaningful physical engagement limits the activation of the premotor cortex and reduces the stability of memory traces (Xu & Ke, 2017; Paivio, 1986; Sadoski & Paivio, 2001; Glenberg, 2008; Fischer & Zwaan, 2008; Ji et al., 2013).

1.3. The Dilemma: Isolated Techniques vs. Cognitive Overload

Modern pedagogy and multimedia learning face a fundamental dilemma. Isolated techniques (only pictures, only sound, only gestures) do not provide sufficient depth of encoding. Attempts to simultaneously engage multiple modalities quickly lead to working memory overload and an increase in extraneous cognitive load (Macedonia & von Kriegstein, 2012; Wilson, 2002; Sadoski & Paivio, 2001; Sweller et al., 2019; Barsalou, 2008; Mayer, 2009).

Models of working memory and cognitive load emphasize that when presenting several uncoordinated elements simultaneously, the learner effectively operates with 4 ± 1 integrated units. Part of the resources is consumed by coordinating modalities rather than the content itself (Sweller, 1988; Cowan, 2001; Macedonia & von Kriegstein, 2012).

1.4. The Problem of Curriculum Organization

Beyond the encoding of individual characters, existing approaches suffer from a fundamental organizational limitation: the 214 Kangxi radicals are conventionally organized by stroke count—a classification that is structurally arbitrary from the learner's perspective. A survey of 42 institutions offering Chinese language instruction found that while 100% agreed radicals should be taught, most simply discuss them as they appear in textbooks, without systematic sequencing (Wang, 2014). No published curriculum was found that organizes the full set of 214 radicals according to a narrative, thematic, or culturally grounded principle. Existing mnemonic methods create independent memory traces for each character, producing "mnemonic islands" with no overarching retrieval structure.

Research on organizational effects in memory demonstrates that hierarchically structured material is recalled significantly better than the same material presented in random order. Bower, Clark, Lesgold, and Winzenz (1969) showed that participants who learned 112 words organized into a conceptual hierarchy recalled 65% of the words, compared to 19% for random order.

2. Cascade Architecture as a Solution

2.1. Principle of Sequential Activation

The proposed methodology solves the encoding dilemma through a cascade architecture: sequential activation where each layer serves as a filter and amplifier for the previous one, and where the step design is subordinate to the principles of the cognitive theory of multimedia learning (limiting simultaneous channels, segmentation, redundancy) and cognitive load theory (minimizing extraneous load, optimizing intrinsic load, enhancing germane load; Mayer, 2009; Sweller, 1988; Sadoski & Paivio, 2001).

The five-phase model includes:

1. **Mental Construction (Element "Fire")** – Presentation of a prototypical image and emotional anchor with expansion potential (visual-semantic core).
2. **Phonological Construction (Element "Metal")** – Sound anchor for pronunciation and tone through auditory symbolism, onomatopoeia, and elements of the keyword approach.
3. **Motor Construction (Element "Wood")** – Kinesthetic anchor through gestures, body movements, and simple interactions with imaginary objects (embodied enactment).
4. **Theatrical Simulation (Element "Earth")** – Narrative episodic anchor uniting all previous levels into a personalized scene (situated gesture / situated action).
5. **Graphomotor Construction (Element "Water")** – Creative practice of writing the character based on the formed narrative, with emphasis on stroke order and spatial structure.

Each step is presented sequentially with a minimal number of active elements, corresponding to principles of segmentation and gradual complexity increase.

2.2. Transitions Between Phases: Completion Criteria

To prevent overload and ensure diagnosability, transition criteria are introduced:

Transition 1 → 2 (image to phonology): The learner can (a) evoke the image and associated "emotional nail" in $\leq 3-5$ seconds; (b) briefly verbalize the meaning.

Transition 2 → 3 (phonology to gesture): (a) Stable reproduction of the syllable and tone ≥ 3 times without hesitation; (b) a simple phonological mediator the learner can describe.

Transition 3 → 4 (gesture to scene): (a) Smooth gesture execution without conscious calculation (1–2 repetitions); (b) ability to describe what the body is doing and how it relates to the meaning.

Transition 4 → 5 (scene to writing): (a) Reconstructing the micro-story in 1–3 phrases from sound or gesture; (b) predicting the general form and stroke order from the story.

Cycle completion: The learner (a) reproduces the character without a sample; (b) can transition from any mediator to the full character in ≤ 10 seconds.

This regulation transforms the cascade into both an encoding protocol and a readiness assessment protocol, aligning with data on the role of successful retrievals and mediators in retrieval practice (Xu & Ke, 2017; Qu et al., 2024).

2.3. Overcoming Working Memory Limitations

Instead of holding 7–8 disparate elements simultaneously, the method distributes them in time and across modalities, ensuring full processing of each component and minimizing competition for resources. Each modal layer: reduces extraneous load by highlighting key elements; enhances germane load by deepening processing (Craik & Lockhart, 1972); and reduces intrinsic load long-term as well-integrated schemas reduce subsequent task complexity (Kalyuga, 2007; Sadoski & Paivio, 2001).

2.4. The Avalanche Effect

Each modal layer screens out inadequate associations and enriches surviving representations with new information. The quality of encoding grows from phase to phase (avalanche effect), aligning with data showing that deeper processing levels provide more stable memory than superficial perceptual rehearsal (Engelkamp, 1998; Sweller et al., 2019; Craik & Lockhart, 1972).

3. Theoretical Rationale

3.1. Dual Coding Theory

Paivio's dual coding theory asserts that information encoded in both verbal and non-verbal (imagery) systems demonstrates better memorization due to two independent access paths (Wang et al., 1992; Shen, 2005b; Taft & Chung, 1999). The cascade model expands dual coding into a multi-level "N-fold" scheme, where visual, phonological, motor, and episodic codes are sequentially layered.

3.2. Levels of Processing

The levels of processing approach (Craik & Lockhart, 1972) interprets memorization as a function of processing depth. Each cascade phase purposefully raises material to a new level: perceptual-visual → semantic → motor → episodic → procedural-graphomotor.

3.3. Embodied and Situated Cognition

Embodied cognition theory asserts that conceptual knowledge is rooted in sensorimotor experience (Chu, 1976; Tan et al., 2005; Guan et al., 2011; Mayer, 2009). Works on embodied learning and L2 gestures show that combining words with iconic gestures improves both immediate and delayed memorization, with neuroimaging demonstrating involvement of motor cortex and premotor areas (Taft & Chung, 1999; Baddeley, 1992; Kalyuga, 2007; Kandel, 2001; He-Zhang et al., 2025).

3.4. Multimedia Learning and Cognitive Load

Mayer's (2009) cognitive theory of multimedia learning formalizes conditions under which multimodal encoding improves versus overloads learning. Sweller's cognitive load theory (1988; Sweller et al., 2019) complements this by dividing load into intrinsic, extraneous, and germane. The cascade architecture directly addresses these frameworks: temporal separation reduces channel competition; strict transition criteria ensure automation before progression; theatrical simulation and graphomotor reconstruction maximize germane load.

3.5. SPT Effect and Motor Programs

The SPT effect (Engelkamp, 1998; Engelkamp & Zimmer, 1985) shows that performing actions significantly improves memorization compared to merely observing. Motor programs become part of the semantic representation, creating an additional access path.

3.6. Episodic Memory and the Theatrical Phase

Theatrical simulation relies on episodic memory (Tulving, 1983) — memory of episodes localized in time and space, experienced from the first person. Creating micro-scenes translates characters from abstract symbols into personal episodic experience. Emotional episodes are remembered better than neutral ones, even in L2 (He-Zhang et al., 2025).

4. Five-Phase Cascade Model of Recoding

4.1. Mental Construction ("Fire"): Prototypical Anchor

A rich, multidimensional mental image is formed, integrating visual form, semantic meaning, and initial associations. An "emotional nail" — an unusual, memorable situation — activates the limbic system and enhances memory consolidation.

Example for 八 (bā, "eight/division"): The learner visualizes eight dancers' long legs standing in a row 八八八八; they separate and one character remains. The visual form directly maps onto the radical's diverging strokes.

4.2. Phonological Construction ("Metal"): Sound Anchor

Phonological information is integrated through forced onomatopoeia, where acoustic characteristics are motivatedly correlated with semantic content. The tone is displayed through the movement of sound after primary sounding. Each sound within the character's pronunciation is presented as an association beginning with the same sound in the learner's L1.

Example for 八: The prolonged "bā-a-a-a" — a sound like two spheres striking each other 8 times, "ba-am"; the sound flows like a line to the left and right. The learner may use onomatopoeia, L1/L2 phonetic support, or a hybrid (Xu & Ke, 2017; Qu et al., 2024). This phase generalizes the keyword approach for characters: a phonological mediator connects sound, tone, and semantics into a unified acoustic "scene."

4.3. Motor Construction ("Wood"): Kinesthetic Anchor

Mental representations transform into meaningful bodily motor schemes. The gesture activates the premotor cortex and somatosensory map, creating feedback. Declarative knowledge converts into procedural knowledge based on motor patterns resistant to temporal decay.

Example for 八: Starting position: both hands before the face, index fingers extended. Main gesture: clench fingers into fists except for the index fingers, count how many are hidden = 8! Then spread arms as if pushing apart a curtain — 八. Integration: form, numerical value, and semantics of division are combined.

SPT effect: Motor execution can improve memory by 30–40% compared to verbal encoding (Engelkamp, 1998; Engelkamp & Zimmer, 1985). Empirical data on embodied learning in L2 confirm the critical role of situated gestures: in He-Zhang et al. (2025), 58 novice Chinese learners studied 32 words in four conditions; situated gesture conditions yielded significantly higher immediate reproduction and recognition, especially for emotional words.

4.4. Theatrical Simulation ("Earth"): Narrative Anchor

The culmination of the cascade process: integration of all encoding levels into an emotionally charged episodic narrative with inclusion of the most frequent words containing the current radical. The learner becomes an active participant.

Synthetic Narrative Calligram

The central visual artifact of the "Earth" phase is the **synthetic narrative calligram** — a typographically modulated transcription in which spatial parameters of letters (size, height, weight, trajectory) are isomorphic to the tonal contour of the studied syllable and simultaneously embedded in the narrative scene. Unlike literary calligrams in the tradition of Apollinaire, where text form reproduces the shape of the described object, this technique serves a strictly cognitive task: making the prosodic structure of the language visually perceivable and narratively anchored.

The calligram's placement in the "Earth" phase rather than "Metal" is architecturally determined by the principle of modal isolation. In the "Metal" phase, the learner builds a pure phonological code with eyes closed: the absence of visual input eliminates channel competition and allows the phonological loop to work without interference. By the "Earth" phase, the tonal pattern is already encoded auditorily and kinesthetically; the calligram performs a synthesizing rather than teaching function — it visually seals the already-formed internal acoustic image, creating cross-modal confirmation through pitch-height correspondence (Spence, 2011). The ascending tone is represented by letters rising upward; the falling tone by a sharp descent; the level high tone by a horizontal but graphically tense line.

In digital format, the calligram is realized as kinetic typography: the animation of letters is synchronized with the tempo and prosody of the narrative text, providing multimodal congruence according to Mayer's (2009) multimedia learning principles.

Deferred Mnemonic Anchors and Semantic Network Organization

The narrative scene of the "Earth" phase deliberately includes **lexical deferred anchors** — anticipatory references to the most frequent characters containing the studied radical, which will be formally introduced in subsequent sessions. The learner encounters these characters in the context of a familiar story before their formal study. This mechanism relies on proactive facilitation: preliminary contextual exposure to a node in the semantic network reduces cognitive load during subsequent acquisition through spreading activation (Collins & Loftus, 1975).

Crucially, deferred anchors differ from disposable mnemonics — associations created solely for memorization that lose function after the goal is achieved. In WuYi, each anchor is a permanent node in the narrative graph: when the corresponding character is subsequently studied, the anchor is activated and strengthened bidirectionally. The resulting mnemonic structure is not a set of isolated associations but a progressively growing semantic graph, where each new element builds on pre-formed network infrastructure. This qualitatively distinguishes WuYi's architecture from classical mnemonic systems (keyword method, Hanzi Movie Method), which build connections between isolated units without establishing systematic inter-level dependencies.

Chimeric Tone Spirits

All high-frequency words sharing the same syllable and tone are presented through a playful **chimeric spirit** — a mythological character whose appearance and actions encode all these words. The learner memorizes them in a single connected contextual series. Using original spirits from Chinese mythology allows combining different concepts more naturally through magical logic and assumptions.

The chimeric tone spirit functions as a **shared contextual node** (hub) in the semantic network, solving the problem of semantic fragmentation when studying homophones. Instead of creating isolated associations for each character, the spirit binds them into a unified narrative.

Example: Bā spirits (first tone). The spirit of Bā is a drummer with large palms, a drum-belly, 8 eyes, and scars, wearing a golden figure-eight chain. He sits in a "rock bar on a rocky hill" (the shared setting for all bā/bá/bǎ/bà spirits), beats his drum-belly with palms 8 times (八 = eight), peels pods and spills 8 peas (巴 = to hope), bakes pea cakes on his reddening belly (疤 = scar), and has buried his bandit past in the bar's courtyard (扒 = to dig). The spirit's tone characteristic: he lies flat on a couch (= level first tone) while drumming.

Corresponding spirits for the second (bá: a tall bartender who pulls corks and picks herbs), third (bǎ: a cash-register bouncer who jumps on the bar counter), and fourth (bà: the bar owner-father who descends from the roof to close up) tones complete the set, with each spirit's posture and trajectory physically encoding the tonal contour.

4.5. Graphomotor Transfer ("Water")

This is not mechanical copying but creative reconstruction. The learner depicts the character through the prism of the created narrative, personalizing the visual form. Neuroimaging studies show that writing practice strengthens connections between visual areas and motor/premotor zones, restructuring the reading network so that graphomotor experience becomes part of the "normal" word processing pathway (Cao et al., 2013; Kandel, 2001; Tan et al., 2005). Experiments with handwriting vs. keyboard input demonstrate the advantage of handwriting for memorizing orientation and configuration (Wilson, 2002).

Example for 八: Draw 八, mirror it downward, then repeat twice — yielding the digit 8 from 8 strokes. Stroke order is embedded in the narrative: "the left stroke falls first (the person walking west), then the right (the person walking east)."

4.6. Narrative-Aligned Encoding with Autonomous Fallbacks

In the full WuYi system, the primary five-phase cascade for each radical is aligned with the macro-narrative: images, sounds, gestures, and theatrical scenes are drawn from the epic arc to which the radical belongs. However, each radical also carries a set of **autonomous reserve mnemonics** —

alternative images, sounds, and gestures independent of the macro-narrative. These are activated through self-diagnosis: if the learner cannot retrieve the primary encoding within 10 seconds during self-testing, the reserve mnemonic is offered. This provides coherence of narrative alignment with flexibility of individual adaptation.

5. Scientific Rationale for Advantages

5.1. Cognitive Offloading Through Temporal Unfolding

The key advantage is overcoming working memory limitations not by ignoring multimodality but by directing it in time. Different brain systems have different processing speeds; the cascade organization accounts for these temporal characteristics.

5.2. Self-Diagnosis Mechanism

Difficulties at one phase signal a problem at a previous level: Problem at Phase 2 → insufficient image at Phase 1; Problem at Phase 3 → weak phonological link; Problem at Phase 4 → insufficient gesture automation. The quality of the mediator is a critical predictor of success in subsequent tests; failure should trigger mediator reconstruction, not repetition increase (Xu & Ke, 2017; Qu et al., 2024).

5.3. Synaptic Multilayering

Each step adds a new neurobiological layer: semantic (hippocampal-cortical), phonological (temporal), motor (striatal), episodic (hippocampal). Neuroimaging data on Chinese radicals confirm the motor component even with purely visual presentation (Taft & Chung, 1999; Hsu et al., 2013).

5.4. Empirical Support

Studies on SPT and motor actions (Engelkamp, 1998; Engelkamp & Zimmer, 1985); gestures and pictures in vocabulary learning (He-Zhang et al., 2025; Andrä et al., 2020; Kalyuga, 2007; Kandel, 2001); situated gestures (He-Zhang et al., 2025; Barsalou, 2008); keyword-mnemonic + retrieval practice (Qu et al., 2024; Xu & Ke, 2017); and Kinect/embodyed systems (Ji et al., 2013; Sadoski & Paivio, 2001; Xu & Ke, 2017) collectively support two key assumptions: (1) bodily and situated encoding increases learning efficiency and retention compared to purely visual-verbal strategies; (2) quality mediators radically increase the return from subsequent testing practice.

6. Comparative Analysis

No existing method engages more than 3–4 of WuYi's five target modalities, and none prescribes a sequential activation order:

| Method | Vis. | Phon. | Kines. | Epis. | Graph. | Max | Narrative | macro-org. |
|------------------|------|-------|--------|-------|--------|-----|-----------|------------|
| Heisig RTH | ✓ | — | — | ½ | ✓ | 2–3 | — | — |
| Hanzi Movie | ✓ | ✓ | — | ✓ | — | 3 | — | — |
| Chineasy | ✓ | — | — | — | — | 1 | — | — |
| Skritter | ✓ | ½ | — | — | ✓ | 2–3 | — | — |
| Xu & Ke Kinect | ✓ | — | ✓ | — | — | 2 | — | — |
| Keyword Mnemonic | ✓ | ✓ | — | — | — | 2 | — | — |
| SPT/Macedonia | ✓ | ✓ | ✓ | — | — | 3 | — | — |
| He-Zhang 2025 | ✓ | ✓ | ✓ | ½ | — | 3–4 | — | — |
| WuYi | ✓ | ✓ | ✓ | ✓ | ✓ | 5 | ✓ | ✓ |

In Kinect and VR/AR systems, gestures typically serve as control interfaces or visual reinforcement of stroke order but do not integrate phonological encoding and personalized theatrical simulation, limiting the depth of episodic and emotional encoding.

7. Three-Level Cascade Architecture: From Radical to Epic

7.1. Three Levels of the Extended Cascade

Level 1 (Micro): Five-phase encoding of a single radical, as described in Sections 2–4.

Level 2 (Meso): Narrative chaining between 2–5 radicals within a scene. The theatrical scene (Phase 4) of one radical flows directly into the next through narrative causality. Example: 人→入→八→穴. A person stands (人), turns and enters (入), paths diverge (八), and beneath a roof (宀) a cave (穴 = 宀 + 八) is discovered. This simultaneously achieves mnemonic encoding, discriminative training (confusable pairs contrasted in situ), and compositional awareness (the learner sees how 穴 is composed of already-learned components). Meso-level chaining exploits Tulving's (1983) encoding specificity principle, functioning as a narrative equivalent of the method of loci (Bower, 1970).

Level 3 (Macro): A two-cycle mythological epic organized by Wu Xing, distributing all 214 Kangxi radicals across five thematic arcs in two spiraling passes.

7.2. The Two-Cycle Mythological Structure

The macro-narrative is organized as two concentric spirals through the Wu Xing generative cycle, mirroring classical Chinese mythology:

Cosmogonic Cycle (74 radicals): From Nüwa's creation of humanity on Mount Kunlun through the separation of heaven and earth to Gonggong's destruction of Mount Buzhou and Nüwa's repair of the sky with five colored stones.

Legendary Cycle (131 radicals): Post-catastrophe civilization guided by culture heroes Shennong (agriculture), Huangdi (technology, writing), and Fuxi (trigrams, music), through moral temptation and ritual redemption, to spring's return and the closing feast.

Both cycles follow the same Wu Xing generative sequence (Earth → Metal → Water → Wood → Fire):

| Phase | Cosmogonic Cycle | Legendary Cycle |
|-------|--|--|
| Earth | Nüwa creates humans on Kunlun (18 rad.) | Shennong teaches agriculture; cities (46 rad.) |
| Metal | First tools and weapons in caves (20 rad.) | Huangdi's inventions: writing, chariots, war (32 rad.) |
| Water | Shamans provoke divine wrath (14 rad.) | Fuxi's music; temptation; ritual redemption (25 rad.) |
| Wood | World Tree trembles; Gonggong attacks (9 rad.) | Spring returns; harvest; living world (16 rad.) |
| Fire | Nüwa repairs sky with five stones (7 rad.) | Great feast of reconciliation (7+10 rad.) |

The cycle begins with **WUYI-00**: five radicals (木火土金水) presented as stones hanging in the mist of eternity. Their purpose is revealed only at the Cosmogonic climax when Nüwa melts them to repair the sky—creating a narrative loop spanning the entire first cycle.

7.3. Shadow Introductions and Narrative Spaced Retrieval

The two-cycle structure creates two retrieval mechanisms. First, **shadow introductions**: certain important radicals appear in the narrative before their formal five-phase encoding, without cascade scaffolding. The learner encounters the radical in a story context, creating a pre-activation trace that

reduces cognitive load during subsequent formal study (proactive facilitation via spreading activation). Second, approximately 12 radicals appear formally in both cycles, receiving full cascade encoding at first encounter and abbreviated re-encoding (Phases 4–5 only: new narrative + writing) at second encounter, producing re-contextualized retrieval practice at natural intervals of 50–100 radicals (approximately 1–3 weeks), aligned with optimal spacing for long-term retention (Cepeda et al., 2006).

7.4. Alignment with Classical Chinese Mythology

The two-cycle structure maps onto the canonical mythological sequence: Nüwa creates humans → period of harmony → separation of heaven and earth → Gonggong's catastrophe → Nüwa repairs the sky → Shennong teaches agriculture → Huangdi invents technology → Fuxi creates trigrams → Great Flood → restoration. This provides authentic cultural knowledge alongside radical acquisition and functions as an advance organizer (Ausubel, 1968) that will be reactivated when learners later encounter these myths in Chinese texts.

7.5. Canonical Wu Xing Correspondences

Radical distribution follows canonical Wu Xing associations (Huainanzi; Liji): body-related radicals map to Wood (spring, growth); tool and weapon radicals to Metal (autumn, precision); ritual and abstract radicals to Water (winter, wisdom). The square form (方) canonically associated with Earth explains the clustering of visually rectangular radicals (口 日 目) in the Earth epic.

7.6. Structural Design Principles

Four principles govern distribution: **Discriminative priority** (confusable radicals always in the same scene); **Compositional transparency** (component before composite, with causal explanation); **Narrative causality** (each radical arises from the previous through causal logic, per Trabasso & van den Broek, 1985); and **Wu Xing correspondence** (thematic coherence grounded in 2,500-year-old cultural tradition).

7.7. Hierarchical Retrieval Cues

The three-level structure provides hierarchical retrieval analogous to Ericsson and Kintsch's (1995) long-term working memory. When the learner forgets a radical: macro → "Which epic?" → meso → "Which scene?" → micro → "Which phase?" This mirrors expert memory in chess (Chase & Simon, 1973) and medicine (Schmidt & Boshuizen, 1993).

7.8. Novelty Assessment

Extensive review revealed no precedent for: (1) organizing 214 Kangxi radicals into narrative arcs; (2) using Wu Xing philosophy or Chinese mythology as curriculum-organizing principles; (3) applying narrative curriculum design to CFL character pedagogy; (4) combining sequential multimodal encoding, inter-item narrative chaining, and mythological macro-organization; (5) a two-cycle structure where the same Wu Xing sequence is traversed twice with inter-cycle spaced retrieval.

8. Illustrative Example: The Cosmogonic Earth Module (K-1)

To demonstrate how the three levels interact, we present the first teaching module with full cascade examples.

Macro context: The Cosmogonic Cycle, Earth phase. Mount Kunlun, the sacred mountain where gods and humans lived in harmony.

Meso narrative (K-1: The Spark): "In a world of perfect harmony on Mount Kunlun, by a mountain river, sat the goddess-progenitor Nüwa (女). She was lonely. At her feet lay yellow clay (

土). She brought a sacred vessel (皿) and filled it with her blood (血) as a sacrifice. From the fire (火) of life and her heart (心), she shaped small figures from clay..."

Micro cascade for 女 (nǚ, "woman"):

Phase 1 — Fire (Image): Before you stands Nüwa with a serpent's tail. She holds a rope along her body: it slides down her hip and goes behind her back. The body crossed by a rope, and arms embracing — this is 女.

Phase 2 — Metal (Sound): Nüwa lowers the rope into the water and pulls it back with a splash. This is how nǚ sounds: the sound softly descends and returns. Lips like for "u," tongue like for "i." Third tone: down then up, like the splash.

Phase 3 — Wood (Gesture): Spread your arms evenly for an embrace — you are a mother embracing a child, and it embraces you. From above, you resemble this radical.

Phase 4 — Earth (Scene): You stand on the shore before the goddess-progenitor (妈妈, mother). She labored to give beginning (始) to every life. You remember your family. You stand beside the goddess, you are her child, it feels "good" (好)...

Phase 5 — Water (Writing): Draw the curved leftward line, top to bottom — the woman's body. Then the descending stroke that crosses it — the rope. Finally, the horizontal stroke — the arms.

Reserve fallbacks: Fire — mentally double the radical vertically and recall DNA, then the rope and the goddess. Wood — recall a "nude" photograph, a beautiful woman lying on a sheet.

9. The Wu Xing Metaphor: Rationale and Status

The phase-element correspondences follow the Wu Xing overcoming cycle: Fire → Metal → Wood → Earth → Water. Metal overcomes Wood (metal tools cut wood); Wood overcomes Earth (roots penetrate soil); Earth overcomes Water (dams redirect flow); Water overcomes Fire (water extinguishes flame); Fire overcomes Metal (heat melts metal). This cycle ensures each phase checks and balances the previous one.

We do not claim this mapping possesses independent explanatory power in cognitive science. However, the culturally rooted metaphor facilitates memorization of the protocol structure, and in the full three-level architecture, Wu Xing serves as the organizing principle for both the encoding phases and the entire curriculum.

10. Limitations

Time costs: The cascade requires more time per radical than traditional copying, though total time to stable retention (6–12 months) may be lower. **Teacher preparation:** Skills in directing gestures, narratives, and group exercises are required. **Individual differences:** The methodology implies adaptation of phase depth. **Empirical status:** The methodology has not yet undergone full-scale randomized comparison. **Scaling:** Full five-phase cascades are most justified for radicals and high-frequency characters; shortened protocols may be used for low-frequency items.

11. Experimental Protocol

Design: Between-subjects, 3 groups (N=30 each), stratified randomization by digit span.

| Group | Method | Key Feature |
|-------|---------------------------------|---|
| 1 | Cascade encoding | Sequential 5-modality activation with transition criteria |
| 2 | Parallel encoding | Same 5 modalities presented simultaneously |
| 3 | Structured graphomotor practice | Mechanical copying with pronunciation; no mnemonics |

Group 2 receives identical materials to Group 1, isolating the variable "sequence" from "content."

Participants: N=90, age 18–35, HSK 0, L1 Russian, no Korean/Japanese experience.

Stimuli: 35 Kangxi radicals balanced by stroke count, complexity, frequency, position; 4 order variants (Latin square).

Procedure: Offline only, video-recorded. 5 min/radical (fixed). Days 1–3: 10 radicals/day. Day 4: 5 radicals + Test T1. Day 11: Test T2. Day 34: Test T3.

Assessment: Pre-tests (digit span, language experience, expectancy). Battery: Recognition → Recall → Production → Transfer (10 novel characters with learned radicals). Subjective: NASA-TLX daily, Likert after Day 4, strategy questionnaire post-final.

Analysis: LME (lme4 + lmerTest); GLMM for binary data. Fixed: Method, Time, Method×Time, covariates. Random: crossed intercepts for participants and items. Planned contrasts with Bonferroni/Holm. Moderator: Digit span × Method. Bayesian factors (brms). Pre-registered on OSF.

Hypotheses: H1: Cascade > Parallel > Control on Production and Transfer. H2: Gap increases T1→T3 (Method×Time). H3: Subjective ease higher in Cascade (NASA-TLX). H4: Cascade especially effective for low working memory (Digit span × Method interaction).

12. Prospects

Educational integration: HSK 1–6 protocol adaptation; teacher training; validated assessment instruments.

Technological support: Mobile application; AR gesture visualization; AI personalization of modality profiles; tone feedback through voice analysis.

Research agenda: Comparative studies (cascade vs. parallel, controlled RCT); neuroimaging (fMRI); free-recall tasks testing thematic clustering and hierarchical retrieval; inter-cycle spaced retrieval validation.

Testable predictions of the three-level architecture: (1) Transfer — superior performance on novel characters due to compositional encoding; (2) Clustering — free recall showing thematic clustering by epic arc; (3) Retrieval hierarchy — epic arc names as cues producing higher recall; (4) Discrimination — lower confusion rates for similar radicals taught in the same scene.

13. Conclusion

The systematization of embodied cognition, dual coding theory, levels of processing, and cognitive load in the form of a three-level cascade multimodal mnemonic architecture offers a structurally novel protocol for teaching Chinese writing. At the micro level, five sequential encoding phases produce progressively deeper memory traces through the avalanche effect. At the meso level, narrative chaining exploits discriminative, compositional, and episodic binding between adjacent radicals. At the macro level, a two-cycle mythological epic organized by Wu Xing philosophy and aligned with classical Chinese mythology provides hierarchical retrieval cues, cultural immersion, narrative spaced retrieval, and shadow introductions across the entire 214-radical curriculum.

The methodology introduces several mechanisms without identified precedent: synthetic narrative calligrams for tonal encoding, chimeric tone spirits for homophone binding, deferred mnemonic anchors for semantic network pre-activation, and narrative-aligned primary encoding with autonomous fallbacks for self-diagnosis.

The cascade approach solves the fundamental contradiction of multimodal learning—the striving for encoding completeness with limited working memory—through temporal unfolding at the item level and narrative organization at the curriculum level. Preliminary data from related fields support the key assumptions; the accompanying experimental protocol is designed for direct validation. The methodology is particularly promising for next-generation digital educational products where XR technologies and adaptive AI can enhance, but not replace, the cognitive architecture of learning.

Supplementary Materials: The following supporting information can be downloaded at the website of this paper posted on Preprints.org. Complete two-cycle macro-narrative with all 214 radicals distributed across 28 teaching modules (WUYI-00 through 1-14); full cascade teaching examples for radicals 女, 土, 人, 皿, 血; chimeric tone spirit examples for bā/bá/bǎ/bà.

Funding: This research received no external funding. Institutional Review Board Statement: Not applicable for theoretical framework development. Conflicts of Interest: The author declares no conflicts of interest.

References

- 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. *Educational Psychology Review*, 32, 815–850.
- Ausubel, D. P. (1968). *Educational Psychology: A Cognitive View*. Holt, Rinehart and Winston.
- Avraamidou, L., & Osborne, J. (2009). The role of narrative in communicating science. *International Journal of Science Education*, 31(12), 1683–1707.
- Baddeley, A. D. (1986). *Working Memory*. Oxford University Press.
- Baddeley, A. D. (1992). Working memory. *Science*, 255, 556–559.
- Barsalou, L. W. (2008). Grounded cognition. *Annual Review of Psychology*, 59, 617–645.
- Bower, G. H. (1970). Analysis of a mnemonic device. *American Scientist*, 58, 496–510.
- Bower, G. H., Clark, M. C., Lesgold, A. M., & Winzenz, D. (1969). Hierarchical retrieval schemes in recall of categorized word lists. *Journal of Verbal Learning and Verbal Behavior*, 8, 323–343.
- Cao, F., Vu, M., Chan, D. H. L., et al. (2013). Writing affects the brain network of reading in Chinese. *Human Brain Mapping*, 34(7), 1670–1684.
- Cepeda, N. J., et al. (2006). Distributed practice in verbal recall tasks. *Psychological Bulletin*, 132, 354–380.
- Chase, W. G., & Simon, H. A. (1973). Perception in chess. *Cognitive Psychology*, 4, 55–81.
- Chu, C. (1976). *Chinese Characters: Their Origin, Etymology, History, Classification and Signification*. Dover Publications.
- Clark, J. M., & Paivio, A. (1991). Dual coding theory and education. *Educational Psychology Review*, 3(3), 149–210.
- Collins, A. M., & Loftus, E. F. (1975). A spreading-activation theory of semantic processing. *Psychological Review*, 82(6), 407–428.
- Cowan, N. (2001). The magical number 4 in short-term memory. *Behavioral and Brain Sciences*, 24, 87–114.
- Craik, F. I. M., & Lockhart, R. S. (1972). Levels of processing: A framework for memory research. *Journal of Verbal Learning and Verbal Behavior*, 11(6), 671–684.
- Engelkamp, J. (1998). *Memory for Actions*. Psychology Press.
- Engelkamp, J., & Zimmer, H. D. (1985). Motor programs and their relation to semantic memory. *German Journal of Psychology*, 9, 26–43.
- Ericsson, K. A., & Kintsch, W. (1995). Long-term working memory. *Psychological Review*, 102(2), 211–245.
- Fischer, M. H., & Zwaan, R. A. (2008). Embodied language. *Quarterly Journal of Experimental Psychology*, 61(6), 825–850.
- Glenberg, A. M. (2008). Embodiment for education. In *Handbook of Cognitive Science: An Embodied Approach* (pp. 355–372).
- Graham, A. C. (1986). *Yin-Yang and the Nature of Correlative Thinking*. Institute of East Asian Philosophies.
- Green, M. C., & Brock, T. C. (2000). The role of transportation in the persuasiveness of public narratives. *Journal of Personality and Social Psychology*, 79(5), 701–721.
- Guan, C. Q., Liu, Y., Chan, D. H. L., Ye, F., & Perfetti, C. A. (2011). Writing strengthens orthography and alphabetic-coding strengthens phonology in learning to read Chinese. *Journal of Educational Psychology*, 103(3), 509–522.
- He-Zhang, Y., Duvignau, K., & Huet, N. (2025). The effects of situated gestures on Mandarin Chinese word learning. *Acta Psychologica*, 261, 105806.
- Heisig, J. W. (2007). *Remembering the Hanzi*. University of Hawaii Press.

- Hsu, C.-H., et al. (2013). Processing Chinese hand-radicals activates the medial frontal gyrus. *Scientific Reports*, 3, 2759.
- James, K. H., & Engelhardt, L. (2012). The effects of handwriting experience on functional brain development in pre-literate children. *Trends in Neuroscience and Education*, 1(1), 32–42.
- Ji, J., Yu, H., Li, B., Shen, Z., & Miao, C. (2013). Learning Chinese characters with gestures. *International Journal of Information Technology*, 19(1), 1–11.
- Kalyuga, S. (2007). Expertise reversal effect and its implications for learner-tailored instruction. *Educational Psychology Review*, 19(4), 509–539.
- Kandel, E. R. (2001). The molecular biology of memory storage. *Science*, 294, 1030–1038.
- Kuo, M.-L. A., & Hooper, S. (2004). The effects of visual and verbal coding mnemonics on learning Chinese characters in computer-based instruction. *ETR&D*, 52(3), 23–38.
- Longcamp, M., Boucard, C., Gilhodes, J. C., & Velay, J. L. (2006). Remembering the orientation of newly learned characters depends on the associated writing knowledge. *Human Movement Science*, 25(4–5), 646–656.
- Macedonia, M., & von Kriegstein, K. (2012). Gestures enhance foreign language learning. *Biological Letters*, 8, 393–396.
- Macedonia, M., Müller, K., & Friederici, A. D. (2011). The impact of iconic gestures on foreign language word learning and its neural substrate. *Human Brain Mapping*, 32(6), 982–998.
- Mayer, R. E. (2009). *Multimedia Learning* (2nd ed.). Cambridge University Press.
- McQuiggan, S. W., Rowe, J. P., Lee, S., & Lester, J. C. (2008). Story-based learning: The impact of narrative on learning experiences and outcomes. In *ITS 2008, LNCS 5091* (pp. 530–539).
- McNamara, D. S., & Healy, A. F. (2000). A procedural explanation of the generation effect. *Journal of Memory and Language*, 43(4), 672–685.
- Needham, J. (1956). *Science and Civilisation in China, Vol. 2*. Cambridge University Press.
- Paivio, A. (1986). *Mental Representations: A Dual Coding Approach*. Oxford University Press.
- Qu, K., Liu, T., Qiao, Y., & Wang, P. (2024). The facilitative effect of the keyword mnemonic on L2 vocabulary retrieval practice. *Heliyon*, 10, e25212.
- Repetto, C., Pedroli, E., & Macedonia, M. (2017). Enrichment effects of gestures and pictures on abstract words in a second language. *Frontiers in Psychology*, 8, 2136.
- Sadoski, M., & Paivio, A. (2001). *Imagery and Text: A Dual Coding Theory of Reading and Writing*. Lawrence Erlbaum.
- Schmidt, H. G., & Boshuizen, H. P. A. (1993). On acquiring expertise in medicine. *Educational Psychology Review*, 5(3), 205–221.
- Schmidt, M., et al. (2019). Embodied learning in the classroom. *Psychology of Sport and Exercise*, 43, 45–54.
- Shen, H. H. (2005a). An investigation of Chinese-character learning strategies. *System*, 33(1), 49–68.
- Shen, H. H. (2005b). Level of cognitive processing: Effects on character learning. *Language and Education*, 19(2), 167–182.
- Spence, C. (2011). Crossmodal correspondences: A tutorial review. *Attention, Perception, & Psychophysics*, 73, 971–995.
- Su, D., Zhong, Y., Zeng, H., & Ye, H. (2013). Embodied semantic processing of Chinese action idioms. *Acta Psychologica Sinica*, 45(11), 1187–1199.
- Sweller, J. (1988). Cognitive load during problem solving. *Cognitive Science*, 12(2), 257–285.
- Sweller, J., van Merriënboer, J. J. G., & Paas, F. (2019). Cognitive architecture and instructional design: 20 years later. *Educational Psychology Review*, 31, 261–292.
- Taft, M., & Chung, K. (1999). Using radicals in teaching Chinese characters to second language learners. *Psychologia*, 42, 243–251.
- Tan, L. H., Spinks, J. A., Eden, G. F., Perfetti, C. A., & Siok, W. T. (2005). Reading depends on writing, in Chinese. *PNAS*, 102(24), 8781–8785.
- Trabasso, T., & van den Broek, P. (1985). Causal thinking and the representation of narrative events. *Journal of Memory and Language*, 24, 612–630.
- Tulving, E. (1983). *Elements of Episodic Memory*. Oxford University Press.

- Wang, J., Thomas, M., & Ouellette, M. A. (1992). The importance of character recognition. *Journal of Chinese Language Teachers Association*, 27(1), 1–20.
- Willingham, D. T. (2004). The privileged status of story. *American Educator*, 28(2), 43–45.
- Wilson, M. (2002). Six views of embodied cognition. *Psychonomic Bulletin & Review*, 9(4), 625–636.
- Xu, X., & Ke, F. (2017). Learning Chinese characters through body movements. *Language Learning & Technology*, 21(3), 138–147.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.