Journal of Chinese Language Teaching in Europe 欧洲中文教育 2026/2, pp. 57–69

Balázs Szigethy

(*Eötvös Loránd University Doctoral School of Linguistics Sinology Programme*)

Diachronic Annotation and Network-Graph Structuring in Chinese Character Pedagogy: A Component-Based Framework

Abstract

This article proposes a standards-aligned approach to Chinese character literacy that integrates two complementary components. First, a diachronic annotation file is prepared for each character, grounded in palaeographic and historical lexicographic evidence, and aligned with *bùshǒu* (部首) indexing standards (GF 0011-2009; GB 13000.1). Second, a network-graph approach represents *bùjiàn* (部件) as role-labelled components (semantic, phonetic, or neutral) within a dependency structure. The framework specifies a reproducible decomposition policy, a sequencing logic combining usage frequency with network topology (following Yan et al. 2013), and teacher-facing governance to ensure consistency and prevent folk etymology. The approach is intended to support recognition, pronunciation inference, and transfer, without presupposing measurable gains over existing approaches.

The proposal is intended as a character literacy module that complements, rather than replaces, communicative four-skill CSL instruction, and is most appropriate once learners have established a basic lexical and phonological foundation. An empirical study is outlined to evaluate learning outcomes relative to list-based instruction; claims regarding efficiency remain subject to future validation.

Keywords: Chinese character pedagogy, component-based instruction, network-graph approach, diachronic etymology, standardisation.

szigethybalazs@gmail.com

https://doi.org/10.61637/cltcee.2026.2.4

https://orcid.org/0009-0006-2188-1195

孔茗

汉字教学中的历时注释与网络图结构建构：一种基于部件的教学框架

摘要

本文提出了一种与国家标准对齐的汉字识读教学方法，由两个相互补充的组成部分构成。第一，为每个汉字建立一份历时性注释档案，以古文字学与历史词汇学研究为依据，并与部首索引体系（GF 0011-2009；GB　13000.1）保持一致。第二，构建网络图方法，将部件作为带有语义、语音或中性功能标注的组成单位，置于依存结构中进行呈现。该框架明确了一套可复现的分解策略，以及一种结合使用频率与网络拓扑的排序逻辑（参照闫小勇等，2013），并配套面向教师的治理机制，以确保分析一致性并避免民间字源学的偏差。该方法旨在通过结构化的部件意识，促进汉字识别、读音推断与迁移能力的发展。

本研究所提出的框架定位为一种汉字识读模块，用于补充而非取代以听、说、读、写为核心的综合性对外汉语教学体系，适用于学习者已具备基本词汇与语音基础之后的阶段。文中同时设计了一项实证研究方案，用以比较该方法与传统字表式教学在学习效果上的差异；关于教学效率的相关判断仍有待进一步实证检验。

关键词： 汉字教学法, 部件化教学, 网络图方法, 历时词源学, 标准化

Introduction

Although component-based approaches have been explored in CSL pedagogy (e.g., Taft – Chung 1999; Shen – Ke 2007), these have not resulted in a standardised, interoperable model aligned with both historical scholarship and modern indexing systems.[1] Across contexts, L2 Chinese instruction still leans on list-based rote memorization that treats characters as isolated units. This obscures the system's compositional nature: complex graphs are systematically built

[1] For example: Shen – Ke 2007: 106–109; Tong – Yip 2015: 173–177; Wong 2017: 211–225; Xu – Chang – Perfetti 2014: 789; Takashima – Verhoeven 2019: 77–79.

from recurrent components with relatively stable functional roles. A pedagogy that foregrounds internal structure—components, dependencies, and reusability—may support more efficient learning, better transfer across families, and more transparent assessment.

Prior work has represented characters as nodes in a Directed Acyclic Graph (DAG) where edges encode "is-a-subcomponent-of" relations (e.g., *mù* (木) → *lín* (林) → *sēn* (森)). Using 3500 high-frequency characters and principled decompositions, Yan, Fan, Di, Havlin, and Wu (2013) built a network with a giant component (n = 3687; 7024 links), revealing a hierarchical structure and Zipf-like offspring distributions. They defined learning cost as a function of unlearned subcomponents and proposed *Distributed Node Weight* (DNW) centrality—hybridizing usage frequency with topology (optimal b ≈ 0.35)—which outperformed frequency-only and naïve bottom-up baselines on efficiency: earlier coverage with lower cumulative cost, front-loading reusable components, and surpassing textbook orders.

Yet there is no shared pedagogical standard that reconciles historical analysis with modern orthographic conventions, prescribes a transparent sequencing policy that can be used in curricula and examinations, and provides teacher-facing tooling and quality governance to curb folk etymologies and inconsistent decomposition. The present article proposes a "quasi-Kodály" model (after Zoltán Kodály, 1882–1967, a Hungarian composer–educator whose method builds literacy from systematically sequenced elementary units and graded pattern acquisition) for Chinese characters: a standardised, component-first approach that aims to make aspects of the writing system's generative structure more explicitly teachable and analysable within a pedagogical setting.

The contribution is threefold. Conceptually, it sets out a standards-aligned way to teach components that respects historical scholarship whilst remaining interoperable with contemporary indexing and IT standards. Theoretically, it shows how a network-graph approach operationalises well-established principles in cognitive science and usage-based linguistics. Practically, it details governance, accessibility, and assessment practices that make the approach potentially adoptable at scale.

This proposal does not aim to replace comprehensive CSL curricula that develop speaking, listening, reading, and writing in parallel. Instead, it defines a modular orthographic strand focusing specifically on character literacy. In typical instructional settings, communicative competence—particularly speaking and listening—remains primary at the beginner stage, while structured character analysis is introduced progressively once learners have acquired a minimal working vocabulary. The present proposal is therefore intended as a complemen-

tary system for reading and writing development, not as a standalone approach governing all aspects of language acquisition.

The present article is conceptual in nature and proposes a testable framework rather than reporting empirical results.

Standardised Component Taxonomy

1. *Bùshǒu* (部首) vs. *bùjiàn* (部件): Division of Labour

The proposed standard begins with a clear division of labour between *bùshǒu* (部首) and *bùjiàn* (部件). *Bùshǒu* developed historically as a lexicographic device for finding characters. Xu Shen's *Shuōwén Jiězì* (说文解字) grouped roughly ten thousand graphs under 540 head graphs (indexing key components), a system later consolidated as 214 head graphs in the *Kāngxī Zìdiǎn* (康熙字典) and legally normalised in the PRC by GF 0011-2009 and GB 13000.1 for education, publishing, and information processing. *Bùshǒu* therefore answer the question "where do we look?", and they should remain the basis for interoperable lookup and assessment.

Learning, however, benefits most when instruction targets *bùjiàn* (部件), the synchronically functional subgraphs that serve as semantic or phonetic units, sometimes as both, and sometimes as neutral structural units (as in simplified or abbreviated/omitted forms). The terminological catch-all *piānpáng* (偏旁) is reserved here for historical discussion, since it blurs functional distinctions in the modern script. In teaching practice, each component instance in a character receives a single primary role—semantic, phonetic, phono-semantic or neutral/structural—determined by synchronic evidence and cross-family productivity. Thus, *qīng* (清) 'clear' is analysed as *shuǐ* (氵, the 'three-dots water' form of *shuǐ* 水), which conveys the semantic field of water/cleanliness and at the same time acts as the *bùshǒu* (部首), the indexing component, for lookup and retrievability purposes, together with *qīng* (青) serving as the phonetic component; *qíng* (情) 'feelings, emotions' as *xīn* (忄, the 'standing heart' form of *xīn* 心) plus *qīng* (青); *qǐng* (请) 'to ask, to invite' as *yán* (讠, the 'speech' form of *yán* 言) plus *qīng* (青). By contrast, differentiators without stable contribution (such as the dot distinguishing *yù* (玉) 'jade' from *wáng* (王) 'king') are treated as neutral. This neutral class becomes especially important in simplified forms where graphic residues persist without current semantic or phonetic function, and earlier character forms have to be analysed for semantic clarification. From

a pedagogical stance, it can be said that *bùshǒu* is for interoperable lookup, whilst *bùjiàn* is for decoding and learning.[2]

2. Allographs and position

In order to maintain interoperability with indexing and digital infrastructure whilst teaching the forms learners actually see, surface allographs — *shuǐ* (氵), *shǒu* (扌), *xīn* (忄), *yán* (讠), *jīn* (钅), *mì* (纟), *cǎo* (艹) — are mapped to the canonical lemmas *shuǐ* (水), *shǒu* (手), *xīn* (心), *yán* (言), *jīn* (金), *mì* (糸) and *cǎo* (艸) respectively for metadata (GF 0011-2009; GB 13000.1), while instruction uses surface forms.[3] Positional alternations such as *fù* (阝, from *fù* 阜 'mound') when located on the left side and *yì* (阝, from *yì* 邑 'city') when located on the right side of a character are distinguished with typical semantic tendencies.[4] Further terminological confusions arise with expressions such as *zuǒ'ěrpáng* (左耳旁), *yòu'ěrpáng* (右耳旁) and *shuāng'ěrpáng* (双耳旁), 'left ear' key, 'right ear' key and 'double ear' key respectively, which have nothing to do with the character for ear, nor do they have ear-related meanings; the resemblance is merely visual. This approach preserves dictionary skills and functional clarity. Characters are parsed by layout schema — *zuǒyòu* (左右), *shàngxià* (上下), *bāowéi* (包围), *zuǒzhōngyòu* (左中右), *shàngzhōngxià* (上中下), *pǐnzì* (品字), *chuānchā* (穿插) — and split along joints that maximise reuse of attested *bùjiàn* while avoiding ad-hoc fragments. This yields reproducible analyses and stable graph neighbourhoods. Each analysis stores: the *bùshǒu* (部首) in GF 0011-2009, *guībù* (归部)/stroke residuals under GF 0012-2009 for the GB 13000.1 character set, Unicode, and major variants of *Kāngxī* (康熙) characters. It has to be acknowledged that some heads (e.g., *piě* 丿) are convenience categories; indexing convenience should not leak into functional claims.[5]

The cognitive and linguistic rationale for a component-first approach can be stated without resorting to lists. Working memory is tightly constrained (≈4±1 units).[6] Treating characters as unrelated wholes imposes a heavy burden on novices, whereas teaching reusable subparts promotes chunking[7] and allows the same knowledge to be encoded verbally (labels, pinyin, glosses) and visually (layout and edge semantics), which improves retrieval.[8] Moreover, many

[2] Chen 2015: 44–50; Packard 2000: 310–317.
[3] Galambos 2006: 33–61.
[4] Taft – Zhu – Peng 1999: 498–502.
[5] Qiu 2000: 26–58.
[6] Cowan 2001: 88.
[7] Sweller 1988: 257–285.
[8] Clark – Paivio 1991: 166, 170–177.

character families align naturally with embodied image-schemas—verticality, path/goal, container—which help link orthographic form, morphemic meaning, and usage.[9] Category learning itself tends to organise around prototypes with graded extensions[10], a fact reflected in the behaviour of radical families such as *shuǐ* (氵), which prototypically signals liquids yet extends to flow or cleansing/cleanliness. Visual "radial maps" around prototypical cores help learners avoid over-generalisation. At the same time, the modern script is dominated by phono-semantic compounds. Treating shared phonetic elements as families reframes reading as hypothesis-testing: learners form pronunciation expectations for unseen series members and then adjust when they meet exceptions.[11] Finally, frequent form-meaning pairings in the language can be taught as constructions; modelling such patterns as higher-order nodes connected to exemplars supports both comprehension and production.[12]

These theoretical considerations motivate a specific stance on sequencing in the proposed approach. A sequencing policy should reflect both how often learners will encounter an item and how much it unlocks by way of dependents.[13] DNW-type centrality captures that logic by combining text frequency with network topology—reusability and dependency depth. Rather than treating DNW as a one-off optimisation that merely produces a better list, the present proposal treats the network itself as the curricular substrate. Syllabi, pacing guides, assessments, and analytics are all derived from the same graph. In practical terms, prerequisite components precede their dependents; characters that contribute many reusable subparts are introduced early; lesson design respects learnability by limiting the number of similar shapes introduced together; and coverage targets (e.g., HSK/CEFR thresholds) are reached with transparent trade-offs that can be tuned for particular cohorts. The analogy with Kodály is not ornamental: just as musical literacy is built from elemental patterns in graded repertoire, component literacy is built from a small set of high-utility forms whose reuse is then expanded along dependency paths.

While DNW centrality has demonstrated efficiency in network-based simulations, its pedagogical applicability requires empirical validation in classroom settings.

There is, nonetheless, a perennial tension between the mnemonic appeal of historical etymology and the demands of synchronic accuracy. Historical notes can vividly motivate forms such as *běn* (本), *xiū* (休), or *ān* (安), but not all

---

[9] Johnson 1987: 90–100; Lakoff – Johnson 1980: 14–21.
[10] Rosch 1975: 192–233; Lakoff 1987: 58–67.
[11] Perfetti – Tan 1998: 114; Taft – Chung 1999: 249–250.
[12] Goldberg 2009: 201–224; Ellis 2003: 82–95.
[13] Bybee 2006: 724.

historical motivations remain productive for modern learners. For this reason, every etymological comment in the Diachronic Annotation File is paired with a plain "Today it functions as…" line that states the current contribution of each component—semantic, phonetic, or ornamental—with no ambiguity. The result is a pedagogy that uses history responsibly whilst keeping present-day inference on firm ground. It also acknowledges that diagrams function as external memory: when relations are made persistent in the environment, novices can off-load some of the burden and rely on perception to guide retrieval.[14] Under realistic classroom time budgets this typically yields better near-transfer and delayed retention than linear lists, a pattern anticipated by the theory and borne out in prior studies grouping characters by component families.[15]

These requirements imply that the model is not immediately applicable in all instructional contexts and may be more suitable for digitally supported or research-oriented programmes.

From Visual Networks to Semantic–Phonetic Truth

A network-graph is only as good as the decompositions that license its edges. If the parsing is driven solely by modern regular-script silhouettes (or even graphic variants), the resulting graph inherits late graphic biases and the logic of traditional indexing rather than the historically correct semantic–phonetic relations that motivated character formation.[16] This produces precisely the sort of false generalisations that a principled component method is meant to avoid.

The case of *qiáng* (强 / 強 / 彊)[17] 'strong' is instructive. Diachronically, *qiáng* (彊) is analysable as *gōng* (弓) 'bow' contributing the semantic field of 'tension/force' (or the force needed to draw a bowstring) and *jiāng* (畺) functioning as the phonetic component. Later simplifications compact the right-side complex into an opaque block.[18] A visual parser that relies only on modern contours may mislabel the right side as a free-standing semantic element or dismiss it as an ornament, erasing the phonetic cue learners most need. The present account centres on visual reanalysis rather than phonological development; the phonetic trajectory of that component in *qiáng* (彊 / 強) as reconstructed for Old Chinese[19] is acknowledged but not explored here. Modern forms also contain many

[14] Risko – Gilbert 2016: 681–685.
[15] Xu – Chang – Perfetti 2014: 789; Takashima – Verhoeven 2019: 77–79.
[16] Qiu 2000: 208; Boltz 1994: 158–167.
[17] Schuessler 2007: 427.
[18] Su 2009: 34–35.
[19] cf. Baxter – Sagart 2014. Their work systematically reconstructs Old Chinese phonology and examines phonetic series, sound change, and related semantic shifts—demonstrating how pho-

box-like components that resemble *kǒu* (口) 'mouth/opening' but are unrelated historically. Examples such as *wéi* (韦 / 韋), *chū* (出), *gè* (各), or *zhèng* (正) easily tempt learners (and automated tools) to create spurious "mouth" connections, when in fact the shapes descend from enclosures, frames, places, cities or other complexes. Without appropriate diachronic anchoring, such mislabelling proliferates.

For an ideal standardised syllabus, the registry behind the graph must therefore carry diachronic metadata for each component (oracle/bronze → small seal → clerical → regular) with role tags at each stage; inheritance rules that preserve phonetic tagging when simplification compresses a complex into an opaque block; disambiguation of look-alikes that only mimic *kǒu* (口); and explicit alignment with contemporary indexing standards that also notes where indexing convenience diverges from historical function. In classroom tools, a guiding principle is that visual edges are licensed by etymology. The interface may show modern components for recognition and lookup, but every node must be able to reveal its period-specific role and series membership on demand. In practice, a learner hovering over the right side of *qiáng* (强) would see that the element corresponds to *jiāng* (畺) with phonetic value in earlier stages, and the network would connect *qiáng* (强) to its phonetic series siblings rather than to arbitrary rectangles. Where a phonetic element also carries meaning, the interface should render both roles explicitly.

A standards-aligned approach requires infrastructure. On the authoring side, teachers need assisted decomposition checks, graph builders that enforce role labelling, and exportable network sheets for classroom use. On the programme side, governance must include versioning, audit trails, peer review of proposed changes, and explicit citation of palaeographic claims, so that the standard remains both stable and improvable. Accessibility is designed in from the start: colour-independent encodings, screen-reader semantics, keyboard navigation, and functional rather than decorative visuals. These choices are not cosmetic. They determine whether institutions can adopt the approach at scale without producing new inequities.

Position within Four-Skill CSL Instruction

A central concern in CSL pedagogy is the balance between communicative competence and orthographic knowledge. The present framework is designed to operate within this balance by targeting character literacy as a distinct, but

netic evolution can reshape meanings, relations, and category boundaries over time.

integrated, component of instruction; furthermore, it does not assume that character structure determines communicative competence, nor that orthographic structure is a primary driver of overall language acquisition; rather, it targets specific aspects of literacy development.

At the beginner level, learning priorities centre on pronunciation, basic vocabulary, and functional communication. For initial, high-frequency expressions such as *nǐ hǎo* (你好), learning is optimally achieved through communicative practice; structural decomposition is introduced only once learners possess sufficient familiarity with the writing system to benefit from it.

The network-graph approach becomes relevant at a later stage, when learners encounter a growing number of characters and require strategies for managing orthographic complexity. At this point, component awareness supports recognition, generalisation across character families, and informed pronunciation inference.

The diachronic annotation system is primarily teacher-facing in early stages, with only selectively disclosed information presented to learners. This principle of progressive disclosure ensures that historical and structural knowledge supports learning without introducing unnecessary cognitive load.

In this sense, the framework aligns with communicative CSL instruction by strengthening literacy outcomes while remaining subordinate to broader language-learning objectives.

The model does not claim that orthographic structure is a primary driver of overall language acquisition, but rather that it can provide targeted support within the domain of literacy.

Methodology

The study evaluates character-level outcomes (recognition, retention, and inference) as one component of overall language development, without claiming to measure full four-skill proficiency. The empirical study is designed to evaluate the core claims of the proposal under realistic classroom conditions. It asks the following questions:

RQ1: *Does network-graph instruction improve immediate and delayed recall and recognition compared to traditional list-based teaching?*

RQ2: *Do the effects of network-graph instruction generalise across different L1 backgrounds and levels of prior exposure?*

RQ3: *What cognitive-load and usability trade-offs arise in network-graph instruction, and does progressive disclosure help to reduce overload?*

RQ4: *Does network-graph instruction enhance learners' ability to infer the pronunciation of previously unseen characters that share phonetic components with known families?*

RQ5: *Does network-graph instruction enhance learners' ability to infer semantic categories or approximate meanings of unfamiliar characters by recognising shared semantic components?*

Participants include L2 learners with varied L1s, stratified by prior exposure. Materials consist of matched character sets balanced for frequency, stroke count, and family membership; all characters in the network condition are supported by brief Diachronic Annotation Files pairing historical notes with a clear statement of the current function of each component. Instruction in the network condition proceeds by decomposition and reassembly, practice over phonetic and semantic families, and the use of construction nodes, with ordering guided by DNW-style centrality and lessons staged to avoid visual interference. The comparison group studies frequency-ordered lists with the same examples but without graphs. Measures include a brief cognitive-load instrument, immediate and delayed recognition tests, pronunciation prediction for unseen series members, sense classification tasks, short construction tasks, and near/far transfer probes in which students place unseen items into an existing network. Analysis uses mixed-effects models with participant and item as random factors, preregistered contrasts, and time-on-task controls.

The study is designed to test whether network-graph instruction leads to differences in subjective load, retention, pronunciation inference, and sense classification relative to list-based approaches.

For each taught character the programme produces two artefacts. The *Diachronic Annotation File* is a concise, referenced, standards-aligned note that links the historical dossier to modern function. The *Network Sheet* presents the character's local neighbourhood with role-labelled nodes and family links; it is designed for teaching, practice, and assessment. Roll-out aligns with HSK/CEFR levels; teacher training focuses on reproducible decomposition and on how to read and build network sheets; and a public repository tracks issues and proposed changes. Known limits are acknowledged: phonetic families contain irregulars; some simplified forms obscure earlier cues; and diachronic certainty varies. The standard addresses these with family-level exception tags, explicit neutral labels, and transparent confidence notes.

Conclusion

This article proposes a two-part, standards-aligned approach to Chinese character literacy that integrates diachronic annotation with network-graph instruction. The framework is designed to support structured inference and component awareness while remaining compatible with existing CSL curricula. Part I—*diachronic annotation*—preserves *bùshǒu* (部首) indexing for interoperability with dictionaries and examinations while correcting common misanalyses through palaeographic evidence. Part II—*network-graph instruction*—teaches *bùjiàn* (部件) as role-labelled components within a dependency graph, operationalising findings on chunking, dual coding, and phono-semantic regularities.

The proposed framework offers three properties that make it adoptable at scale:

1. Strict alignment with national standards;
2. Reproducible decomposition under peer-reviewed governance;
3. Graph-aware sequencing that integrates usage frequency with reusability.

An empirical study has been outlined to test recall, retention, cognitive load, and transfer, including learners' ability to infer pronunciation and meaning of unfamiliar characters.

In sum, the approach reframes character learning from rote memorisation to structured inference, while preserving the librarian's path of traditional lookup. Implemented as paired *Diachronic Annotation Files* and *Network Sheets*, it offers a potential basis for a more systematic descriptive framework—historically accurate, pedagogically motivated, and accessible across contexts.

The proposal is deliberately bounded. It is not intended as a comprehensive CSL approach, nor is it designed for absolute beginners, and its effectiveness will depend on appropriate sequencing, teacher training, and integration with communicative instruction. It is therefore best understood as a theoretically motivated and operationalised proposal, whose pedagogical value remains to be established through empirical study under classroom conditions.

Bibliography

Baxter, William H. – Sagart, Laurent. 2014. *Old Chinese: A New Reconstruction*. Oxford: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199945375.001.0001

Boltz, William G. 1994. *The Origin and Early Development of the Chinese Writing System*. American Oriental Series, vol. 78. New Haven, CT: American Oriental Society.

Bybee, Joan. 2006. 'From Usage to Grammar: The Mind's Response to Repetition,' *Language* 82 (4): 711–733. https://doi.org/10.1353/lan.2006.0186

Chen, Chunlei. 2015. '现代汉语教材中偏旁、部首和部件的定义问题 [Definitional Issues of *Piānpáng*, *Bùshǒu*, and *Bùjiàn* in Modern Chinese Textbooks],' *Academic research on language and writing* (5): 44–50. https://doi.org/10.14014/j.cnki.cn11-2597/g2.2015.05.013.

Clark, James M. – Paivio, Allan 1991. 'Dual Coding Theory and Education,' *Educational Psychology Review* 3: 149–210. https://doi.org/10.1007/BF01320076

Cowan, Nelson. 2001. 'The Magical Number 4 in Short-Term Memory: A Reconsideration of Mental Storage Capacity,' *Behavioral and Brain Sciences* 24 (1): 87–114. https://doi.org/10.1017/S0140525X01003922

Ellis, Nick C. 2003. 'Constructions, Chunking, and Connectionism: The Emergence of Second Language Structure', in Catherine J. Doughty – Michael H. Long (eds.). *The Handbook of Second Language Acquisition*. Malden, MA: Blackwell, 63–103. Reprinted 2008. https://doi.org/10.1002/9780470756492.ch4

Galambos, Imre. 2006. *Orthography of Early Chinese Writing: Evidence from Newly Excavated Manuscripts*. Budapest: Department of East Asian Studies, Eötvös Loránd University, (Budapest Monographs in East Asian Studies 1): 33–61.

Goldberg, Adele E. 2009. 'Constructions work,' *Cognitive Linguistics*, vol. 20 (1): 201–224. https://doi.org/10.1515/COGL.2009.013

Johnson, Mark. 1987. *The Body in the Mind: The Bodily Basis of Meaning, Imagination, and Reason.* Chicago: University of Chicago Press. https://doi.org/10.7208/chicago/9780226177847.001.0001

Lakoff, George – Mark Johnson. 1980. *Metaphors We Live By*. Chicago, IL: University of Chicago Press.

Lakoff, George. 1987. *Women, Fire and Dangerous Things*: *What Categories Reveal about the Mind.* Chicago, IL: University of Chicago Press. https://doi.org/10.7208/chicago/9780226471013.001.0001

Ministry of Education of the People's Republic of China [中华人民共和国教育部] and National Language Commission [国家语言文字工作委员会] 2009. *Specification for Identifying Indexing Components of the GB 13000.1 Chinese Character Set* [*GB13000.1*字符集汉字部首归部规范]. Beijing: Chinese Language and Literature Publishing House [语文出版社]. Standard GF 0012–2009. http://www.moe.gov.cn/jyb_sjzl/ziliao/A19/200901/t20090102_186104.html (last accessed: 02.10.2025.)

Ministry of Education of the People's Republic of China [中华人民共和国教育部] and National Language Commission [国家语言文字工作委员会] 2009. *The Table of Indexing Chinese Character Components* [汉字部首表]. Beijing: Chinese Language and Literature Publishing House [语文出版社]. Standard GF 0011–2009. Released 2009-01-12; implemented 2009-05-01. https://archive.org/details/GF0011-2009/page/8/mode/2up (last accessed: 02.10.2025.)

Packard, Jerome. L. 2000. *The Morphology of Chinese: A Linguistic and Cognitive Approach*. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9780511486821

Perfetti, Charles A. – Li-Hai Tan. 1998. 'The Time Course of Graphic, Phonological, and Semantic Activation in Chinese Character Identification,' *Journal of Experimental Psychology: Learning, Memory, and Cognition* 24(1): 101–118. https://doi.org/10.1037/0278-7393.24.1.101

Qiu Xigui [Mattos, Gilbert L. – Norman, Jerry (trans.)] 2000. *Chinese Writing. (Early China Special Monograph Series No. 4)*. Berkeley: The Society for the Study of Early China – The Institute of East Asian Studies, University of California.

Risko, Evan F. – Sam J. Gilbert. 2016. 'Cognitive Offloading'. *Trends in Cognitive Sciences* 20 (9): 676–88. https://doi.org/10.1016/j.tics.2016.07.002

Rosch, Eleanor. 1975. 'Cognitive Representations of Semantic Categories.' *Journal of Experimental Psychology: General, 104(3):* 192–233. *https://doi.org/10.1037/0096-3445.104.3.192*

Schuessler, Axel. 2007. *ABC Etymological Dictionary of Old Chinese*. Honolulu: University of Hawai'i Press.

Shen, Helen H. – Cheng-Chao Ke. 2007. 'Radical Awareness and Character Acquisition,' *The Modern Language Journal* 91 (1): 97–111. https://doi.org/10.1111/j.1540-4781.2007.00511.x

Su, Peicheng. 2009. '简化汉字60年 [Sixty Years of Simplified Chinese Characters]'. 语言文字应用 [*Applied Linguistics*], 4: 26–35. https://doi.org/10.16499/j.cnki.1003-5397.2009.04.004

Sweller, John. 1988. 'Cognitive load during problem solving: effects on learning,' *Cognitive Science, 12(2):* 257–285. https://doi.org/10.1207/s15516709cog1202_4

Taft, Marcus – Kevin Chung. 1999. 'Using Radicals in Teaching Chinese Characters to Second Language Learners,' *Psychologia* 42 (4): 243–251.

Taft, Marcus – Xiaoping Zhu – Danling Peng. 1999. 'Positional Specificity of Radicals in Chinese Character Recognition,' *Journal of Memory and Language* 40 (4): 498–519. https://doi.org/10.1006/jmla.1998.2625

Takashima, Atsuko – Ludo Verhoeven. 2019. 'Radical Repetition Effects in Beginning Learners of Chinese as a Foreign Language Reading,' *Journal of Neurolinguistics* 50: 71–81. https://doi.org/10.1016/j.jneuroling.2018.03.001

Tong, X. – Yip, J.H.Y. 2015. 'Cracking the Chinese character: radical sensitivity in learners of Chinese as a foreign language and its relationship to Chinese word reading,' *Reading and Writing* 28: 159–181. https://doi.org/10.1007/s11145-014-9519-y

Wong, Yu Ka. 2017. 'The Role of Radical Awareness in Chinese-as-a-Second-Language Learners' Chinese Character Reading Development,' *Language Awareness* 26 (3): 211–225. https://doi.org/10.1080/09658416.2017.1400039

Xu, Y. – Chang, L. – Perfetti, C. A. 2014. 'The effect of radical-based grouping in character learning in Chinese as a foreign language,' *The Modern Language Journal, 98(3):* 773–793. https://doi.org/10.1111/modl.12122

Yan, Xiaoyong – Ying Fan – Zengru Di – Shlomo Havlin – Jinshan Wu. 2013. 'Efficient Learning Strategy of Chinese Characters Based on Network Approach.' *PLoS ONE* 8 (8): e69745. https://doi.org/10.1371/journal.pone.0069745