

Reconstructing the Classification System of Modern Chinese Characters: A Pedagogically Oriented Approach

Levente Li, Tufts University

The traditional *liushu* (Six Principles) theory, while effective in explaining the historical formation of Chinese characters, is not readily applicable to the analysis of modern character forms. This study proposes a structure-based interpretation grounded in the present forms of Chinese characters and reconstructs a classification system on this basis.

At the classificatory level, the notion of structural motivation is redefined, and a revised fivefold system is proposed, consisting of pictographic, indicative, associative, phono-semantic, and symbolic characters. At the level of structural mechanisms, a tripartite functional model of components—semantic, phonetic, and symbolic—is introduced.

The concept of “component languaging” is further proposed to account for the transformation of components from graphic units into functional linguistic units capable of participating in subsequent character formation. On this basis, a type–function model is established, integrating classification, structural mechanisms, and structure-based interpretation.

This model provides a unified synchronic framework for the analysis of modern Chinese characters and offers a theoretically grounded and operational approach to their structure-based interpretation.

Keywords: modern Chinese characters; *liushu*; structure-based interpretation; fivefold classification; component languaging; Chinese character pedagogy

Introduction

The *liushu* 六書 “Six Principles” has traditionally been regarded as the foundational framework for analyzing the structural formation of Chinese characters and occupies a central position in the history of Chinese philology.¹ This framework is grounded in evidence from early script forms, with the *xiaozhuan* 小篆 “small seal script” system serving as its primary reference.² While it effectively accounts for ancient characters and certain modern forms that retain traces of their original structural configurations, it does not fully explain the present structures of modern Chinese characters.

¹ Xu Shen 許慎, *Shuowen jiezi* 說文解字 (Beijing: Zhonghua shuju, 1963).

² Qiu Xigui 裘錫圭, *Chinese Writing* (Berkeley: Society for the Study of Early China, 2000).

Through processes such as clericalization, regularization, and modern simplification and standardization, the original structural relationships of many characters have been weakened or even disrupted. As a result, continued reliance on the traditional *liushu* framework and its etymological approach to interpreting contemporary character forms often leads to unstable classifications, inconsistent criteria, and limited pedagogical applicability.

This issue is particularly pronounced in the context of Chinese character pedagogy, where instruction focuses on contemporary character forms rather than their early historical stages. If the structure of a given character can only be explained by reference to ancient script forms—such as oracle bone inscriptions, bronze inscriptions, or *xiaozhuan*—such explanations, while philologically valid, cannot be readily converted into stable, consistent, and operational teaching practices.

It is therefore necessary to distinguish between two different interpretive approaches. One is an etymological approach that reconstructs the historical origins of characters by tracing their earlier forms. The other is a structure-based approach that focuses on the relationships among components within the contemporary form, thereby accounting for the meaning and pronunciation of the whole character. The present study is primarily focused on the latter.

In response to the issues outlined above, this study has two main objectives. First, it seeks to reconstruct a classification system tailored to modern Chinese characters. Second, it aims to account for their fundamental structural mechanisms at the level of components. The former addresses how modern Chinese characters can be categorized into structural types, while the latter concerns the functional principles governing their internal organization.

By integrating whole-character types with component functions, this study establishes a unified framework that is better suited to the analysis and pedagogy of modern Chinese characters. This study aims to provide a unified analytical framework grounded in present forms and applicable to both theoretical analysis and pedagogy.

The Applications and Limitations of the *liushu* Theory

1) The Continuing Validity of the Traditional *liushu* Categories in Modern Chinese Characters

The pictographic, indicative, associative, and phono-semantic categories within the traditional *liushu* framework continue to have explanatory value. In modern Chinese characters, many forms still exhibit features reflecting pictorial representation, indicative relations, semantic combinations, or phonetic-semantic division.

Therefore, the *liushu* framework remains a theoretically justified point of departure for analyzing character structure.³ In particular, phono-semantic compounds remain among the most important structural types in the modern Chinese writing system.

³ Qiu Xigui 裘錫圭, *Chinese Writing* (Berkeley: Society for the Study of Early China, 2000).

However, this validity is not comprehensive. The forms of modern Chinese characters are not identical to their earlier script counterparts, and over the course of long-term development, the original structural relationships of many characters have been weakened or even lost.⁴ As a result, characters that could once be clearly classified within the *liushu* system in earlier stages do not necessarily lend themselves to unambiguous classification based on their present forms. This suggests that while the traditional *liushu* framework still possesses a certain degree of explanatory power for modern Chinese characters, such applicability is conditional rather than universal.

2) From Historical Motivation to Present Structural Motivation

To reconstruct a classification system for modern Chinese characters, it is necessary to first reconsider the notion of “structural basis.” In traditional philology, structural basis is primarily explained with reference to earlier script forms, emphasizing why a character was originally constructed in a particular way.⁵

In contrast, the concept of structural basis as employed in this study does not primarily concern itself with historical origins, but rather with whether a meaningful relationship can still be established, within the present character form, between its components and the character’s meaning and pronunciation. In other words, the focus is not on how a character was originally formed, but on how it can be structurally interpreted in its present form.

Accordingly, this study defines “structure-based interpretation” as an approach that, grounded in the present form of the character, explains its meaning, pronunciation, or both by analyzing the functional relationships among its components within the overall structure.

Any character for which a relatively stable and interpretable relationship can be established within its present structure falls within the scope of structure-based interpretation; conversely, those for which such correspondence cannot be established should not be forcibly subsumed under this approach.

Such a definition preserves the rigor of philological analysis while better aligning with the practical needs of modern Chinese character pedagogy.

3) Limitations of Existing Revision Proposals

Scholarly efforts to revise the traditional *liushu* framework have generally followed two main directions: one seeks to adjust classification criteria within the existing framework, while the other introduces new categories to supplement the traditional system. For example, Chen Mengjia reinterpreted Chinese character

⁴ Qiu Xigui 裘錫圭, *Chinese Writing* (Berkeley: Society for the Study of Early China, 2000).

⁵ Su Peicheng 蘇培成, *Xiandai hanzixue gangyao* 現代漢字學綱要, 3rd ed. (Beijing: Commercial Press, 2014), 103.

categories as stages in the historical development of writing rather than as fixed classificatory principles⁶. These attempts have contributed to a renewed understanding of Chinese character structure.

However, a key limitation remains: despite their formal differences, many of these approaches continue to rely to a large extent on etymological analysis. As long as classification criteria are primarily grounded in historical character formation, problems of interpretive inconsistency are likely to persist when such approaches are applied to contemporary character forms.

More importantly, in modern Chinese characters there exists a group of forms whose internal structural relationships have been significantly weakened, such that no stable correspondence among components, phonology, and meaning can be established on the basis of their present forms. For example, the simplified character 东, derived from the cursive regularization of 東, no longer preserves an interpretable correspondence among its present components, pronunciation, and meaning. It is therefore classified as a symbolic character in the present study.

These characters often lack a clear position within existing frameworks: they are either forcibly assigned to one of the traditional categories or treated as exceptions.

Either approach undermines the overall coherence of the classification system. Therefore, the reconstruction of a classification system for modern Chinese characters does not hinge on making incremental adjustments to existing categories, but rather on reestablishing the underlying principles of classification. Specifically, classification must be based on whether a character can be meaningfully interpreted in terms of its present form, rather than on whether its historical origin can be reconstructed.

Reconstruction of the Classification System of Modern Chinese Characters

1) Principles of Reconstruction

A classification system for modern Chinese characters should be based on two fundamental principles.

First, analysis must be grounded in the present forms of characters. That is, classification should rely on structural relationships observable at the synchronic level, rather than on the reconstruction of earlier script forms.

Second, structural interpretability should serve as the primary criterion for classification. Characters for which a meaningful and interpretable relationship among components can be established within their present structure should be classified as interpretable types. By contrast, those for which such correspondence cannot be established should be assigned to a separate category.

When a character appears to satisfy more than one category, classification is determined according to the principle of present structural interpretability. Priority is

⁶ Chen Mengjia 陳夢家, *Yinxu buci zongshu* 殷墟卜辭綜述 (Beijing: Zhonghua shuju, 1988), 79.

given to the category that provides the most stable and coherent explanation of the relationship among components, pronunciation, and meaning in the present form. Characters for which no stable structural correspondence among components, pronunciation, and meaning can be established in their present forms are classified as symbolic characters.

Based on these two principles, the pictographic, indicative, associative, and phono-semantic categories within the traditional *liushu* framework may be retained, although their definitions should be reinterpreted in light of present character forms. At the same time, a new category *jihaozi* 記號字 “symbolic characters” should be introduced to accommodate those forms for which no meaningful correspondence among components, phonology, and meaning can be established within the present structure.

Accordingly, modern Chinese characters can be classified into five categories—pictographic, indicative, associative, phono-semantic, and symbolic—thereby forming a revised fivefold system.

2) Redefining the Five Categories of Characters

Following the establishment of the fivefold classification system, it is necessary to further redefine these categories on the basis of the present structural forms of modern Chinese characters. The term “redefinition” here does not imply a simple adoption of traditional philological terminology; rather, it involves a rearticulation of each structural type according to the different ways in which relationships among components, phonology, and meaning are manifested in the present structure.

Through this adjustment, the classification of character types can be grounded in a more consistent set of criteria, thereby providing a reliable basis for the detailed analyses that follow.

It should be noted that, in conducting structure-based interpretation, this study does not restrict relevant phonological–semantic relationships to a character’s original meaning or its most common pronunciation. Rather, the analysis is based on whether such relationships can be meaningfully established within the present structural configuration of the character. These relationships may derive from original meanings or standard pronunciations, but may also arise from extended meanings, conventional usages, or less frequent phonological forms.

Unlike etymological analysis, the present study focuses on the interpretive function of phonological–semantic relationships within the present structure, rather than on their historical development. The primary criterion in analysis is whether such relationships can provide an effective explanation within the present structural framework.

(1) Pictographic Characters

In traditional definitions, pictographic characters are characterized by the principle of “depicting the object,” that is, representing real-world entities through graphic forms. In the modern Chinese writing system, however, most pictographic characters no longer exhibit an immediately recognizable resemblance to their referents. Instead, through long-term evolution, they have developed into structurally stable symbolic units.

Therefore, modern pictographic characters should not be understood as characters that still visually resemble specific objects, but rather as structural units derived from early pictorial representations and recognized holistically within the present system. Their meaning is primarily carried by the character as a whole, and their internal components generally no longer permit effective structure-based interpretation.

The identification of pictographic characters in modern Chinese does not depend on whether their original visual forms can be reconstructed, but rather on whether their overall structure is recognized as a symbolic unit with a pictographic origin within the present system.

(2) Indicative Characters

Indicative characters are formed on the basis of pictographic characters by adding marks or by utilizing positional relationships within the structure to represent abstract concepts, spatial relations, or semantic orientation. Unlike pictographic characters, the defining feature of indicative characters does not lie in visual representation, but in their indexical function within the structure.

The identification of indicative characters in modern Chinese should focus on whether a clear indexical relationship can be established within the present structure, rather than being constrained by the specific forms of earlier script stages.

(3) Associative Characters

Associative characters are traditionally characterized by the principle of “combining meanings through analogy,” whereby new meanings are generated through the combination of multiple components. From the perspective of modern Chinese character structure, however, the key does not lie in the simple accumulation of component meanings, but in the semantic relationships established among the components. The meaning of an associative character arises from the way these components are organized, rather than from a direct summation of their individual meanings.

In specific structures, the semantic roles of components are often not equal, but instead exhibit hierarchical differentiation. Accordingly, the identification of associative characters should not be based on a one-to-one correspondence between component meanings and the overall meaning, but rather on whether a coherent semantic interpretation can be established within the present structure.

Associative characters can be defined as a structural type in which overall meaning

is generated through semantic relationships among components, with identification based on the interpretability of such relationships.

(4) Phono-semantic Characters

Phono-semantic characters constitute the dominant structural type in the modern Chinese writing system. Their defining feature lies in the division of labor between components, whereby one component provides semantic cues while another provides phonological information.

However, within the present system of character forms, this division of labor is not always stable. The phonological correspondence between the phonetic component and the whole character may vary in degree, while the semantic contribution of the semantic component may be weakened due to semantic change over time.

Accordingly, the identification of phono-semantic characters should not be based primarily on formal features, but rather on the recognizability of the division of labor between phonology and meaning within the present structure.

In this sense, phono-semantic characters can be defined as a structural type in which the overall character is generated through the functional differentiation of components with respect to phonology and meaning, with identification based on the interpretive function of phonological–semantic relationships within the structure. From this perspective, phono-semantic structures constitute the most typical form of interpretable structure in modern Chinese characters.

(5) Symbolic Characters

The concept of “symbolic characters” was first proposed by Tang Lan in *Zhongguo wenzixue* 中國文字學 to refer to a class of elements that do not carry explicit phonological or semantic functions and cannot readily serve as a basis for analyzing the overall structure of a character. This concept provides an important theoretical foundation for addressing those structural types in modern Chinese characters that resist straightforward interpretation.

As a result, the present study redefines the scope of symbolic characters and incorporates them as an independent category within the classification system of modern Chinese characters.

From the perspective of structural motivation, symbolic characters correspond to cases in which the structural relationships within Chinese characters have been weakened or disrupted over the course of historical development. As a result, it has become difficult to establish an interpretable correspondence among their internal components, phonology, and meaning within the present structure.

Although some of these characters may have possessed clear structural motivations in earlier stages, such motivations are no longer even minimally

reflected in their present forms under synchronic analysis.

The identification of symbolic characters does not depend on whether a character can be formally decomposed, but rather on whether a meaningful correspondence among its components, phonology, and meaning can be established within the present structure. Characters for which such relationships cannot be established fall outside the scope of structure-based interpretation and should be classified as symbolic characters.

In this sense, symbolic characters can be defined as a structural type that lacks an interpretable component-based mechanism under present-form conditions and instead functions as holistic symbolic units. The introduction of this category enables the classification of modern Chinese characters to achieve comprehensive coverage under a unified set of criteria. This category is not intended as a residual class, but as a theoretically defined type based on the absence of interpretable structural relations.

3) The Theoretical Role of Symbolic Characters

Within the fivefold classification system, the introduction of symbolic characters carries crucial theoretical significance. It may be argued that without this category, the system itself would be difficult to sustain.

This is because there exists, in modern Chinese characters, a group of forms whose historical origins can be traced, yet whose structures can no longer be explained in terms of component relationships within their present forms. If such characters are not treated as a distinct category, they must either be forcibly assigned to existing categories or treated as exceptions, thereby undermining the consistency of the classification criteria.

Traditional analyses of such characters have largely relied on etymological approaches, reconstructing earlier script forms to account for their structural motivations. However, this approach is not grounded in the present forms of characters; its explanatory results depend on specialized historical knowledge and therefore do not readily yield a unified and operational standard. The key issue, therefore, lies not in whether reconstruction is possible, but in whether a character can be meaningfully interpreted within its present structural configuration.

In this sense, the introduction of symbolic characters marks a shift in analytical perspective. Chinese characters are no longer primarily treated as objects requiring historical reconstruction, but rather as structural systems that can be explained at the synchronic level. Accordingly, the basis of classification shifts from historical formation to structural interpretability.

Therefore, symbolic characters should not be regarded as a supplementary addition to the existing system, but as a necessary condition for its coherence. It is precisely through the establishment of this category that the classification of modern Chinese characters can achieve comprehensive coverage under a unified set of criteria.

4) Statistical Analysis of the Fivefold Classification in the Level-1 Character Set

and Its Validating Significance

To evaluate the coverage and applicability of the fivefold classification system proposed in this study, a systematic classification and statistical analysis are conducted based on the characters included in the Level-1 Chinese Character Set. The classification principles follow those established in the preceding sections, namely, an analysis grounded in the present structural forms of characters.

Characters for which a relatively stable and interpretable relationship can be established at the synchronic level are classified into four categories: pictographic, indicative, associative, and phono-semantic. By contrast, those for which such relationships cannot be readily established are assigned to the category of symbolic characters.

It should be noted that the Level-1 Chinese Character Set contains 3,755 simplified characters.⁷ When these characters are converted into their corresponding traditional forms, the total number increases to 3,792, as some simplified characters correspond historically to multiple traditional forms (i.e., a one-to-many relationship between simplified and traditional characters).

The present study conducts classification and statistical analysis on both sets—namely, the 3,755 simplified characters and their corresponding 3,792 traditional forms—and derives two sets of statistical results.

**Table 1. Distribution of the Fivefold Classification in the Simplified Character Set
(n = 3,755)**

Category	Count	Percentage (%)
Pictographic	211	5.61
Indicative	19	0.50
Associative	790	21.03
Phono-semantic	2654	70.67
Symbolic	81	2.15
Total	3755	100.00

**Table 2. Distribution of the Fivefold Classification in the Traditional Character Set
(n = 3,792)**

Category	Count	Percentage (%)
Pictographic	189	4.98

⁷ State Bureau of Standards. *Information Interchange Code for Chinese Characters—Basic Set (GB2312-80)*. Beijing: Standards Press of China, 1981.

Category	Count	Percentage (%)
Indicative	18	0.47
Associative	706	18.61
Phono-semantic	2830	74.63
Symbolic	49	1.29
Total	3792	100.00

From an overall distributional perspective, phono-semantic characters exhibit clear dominance in both the simplified and traditional character systems, accounting for 70.67% and 74.63%, respectively. This result indicates that the primary structural pattern of modern Chinese characters continues to be based on the division of labor between phonology and meaning, whereby components function to provide semantic and phonological cues in the formation of whole characters.

This statistical finding provides corpus-level evidence supporting the earlier claim that phono-semantic structures occupy a dominant position in modern Chinese characters.

Meanwhile, associative characters also account for a relatively high proportion in both systems (21.03% in the simplified set and 18.61% in the traditional set), indicating that structural patterns based on semantic relationships among components continue to play an important role in the modern Chinese writing system.

By contrast, pictographic and indicative characters constitute only a small proportion, suggesting that these types—rooted in early modes of character formation—primarily function as historically retained structural units rather than as productive mechanisms for character formation in the modern system.

Notably, symbolic characters also account for a non-negligible proportion in both systems (2.15% in the simplified set and 1.29% in the traditional set). Although their number is far smaller than that of phono-semantic and associative characters, they do not constitute merely isolated cases.

This indicates that, within the scope of modern commonly used Chinese characters, there indeed exists a group of forms for which stable interpretable relationships among components cannot be established under present structural conditions. It is therefore necessary to recognize them as an independent category within the classification system. In other words, the establishment of symbolic characters is not an ad hoc solution for classification, but a theoretical necessity supported by corpus-based statistical evidence.

A comparison between the simplified and traditional character sets reveals certain differences in their category distributions. Compared with the traditional set, the proportion of phono-semantic characters in the simplified set shows a slight decrease, while the proportions of associative and symbolic characters increase correspondingly.

This pattern suggests that the processes of simplification and standardization have, to some extent, altered the structural interpretability of certain characters and may

affect their classification in synchronic analysis. In particular, under conditions of one-to-many correspondence between simplified and traditional forms, multiple traditional characters with distinct structural origins may be merged into a single simplified form, thereby giving rise to new interpretive configurations in structural analysis.

In summary, the systematic statistical analysis of the Level-1 Chinese Character Set demonstrates that the fivefold classification system proposed in this study achieves relatively comprehensive coverage of modern commonly used Chinese characters and exhibits strong explanatory power and practical applicability at the scale of a large corpus.

These findings not only provide empirical support for the proposed fivefold system, but also further confirm the practical feasibility of a structure-based interpretive approach grounded in present character forms for the analysis of modern Chinese characters.

The Tripartite Functional Model of Components and Structural Mechanisms in Chinese Characters

1) Whole-Character Types to Component Functions

The fivefold classification system primarily addresses the categorization of Chinese characters at the level of whole-character types. However, an adequate account of their structural mechanisms requires analysis at the level of components. Modern Chinese characters are distinguished not only by differences in overall structural types, but also by the distinct functions that their components perform within these structures.

The analysis of Chinese characters must not only determine their type classification, but also explain the specific roles played by internal components in the generation of meaning and phonological realization.

On this basis, the present study classifies the components of modern Chinese characters into three fundamental functional types: semantic, phonetic, and symbolic. Semantic components provide semantic cues or impose semantic constraints; phonetic components provide phonological information; and symbolic components, although structurally present, no longer carry explicit phonological or semantic functions. Unlike the fivefold classification, which operates at the level of whole characters, this framework takes component functions as its analytical core.

Building on this distinction, the analysis of Chinese character structure can be divided into two levels: the “type level,” represented by the classification of whole-character types in the fivefold system, and the “mechanism level,” represented by the analysis of structural mechanisms based on the tripartite functional model of components. These two levels work in coordination to form the basic analytical framework for the structural analysis of modern Chinese characters.

2) Defining the Three Types of Components

Semantic components are those that function within the structure of a character to provide semantic cues or impose semantic constraints. Their role is to indicate a semantic range or category, rather than to correspond directly to specific lexical meanings.

Phonetic components are those that bear a certain degree of correspondence to the phonological realization of the whole character and provide phonological cues. This correspondence exhibits varying degrees across different characters; however, as long as a component retains a recognizable phonetic function within the present structure, it may be classified as a phonetic component.

Symbolic components are structural elements that participate in the formation of a character but do not carry explicit semantic or phonetic functions. Their primary role is to maintain the structural integrity of the character rather than to contribute directly to the generation of its phonology or meaning.

It should be noted that component functions are not given a priori but are realized within specific structural relationships. The same component may perform different functions in different characters and may, as phonological–semantic relationships weaken, shift into a symbolic role. The identification of component functions is determined not by formal features, but by their actual roles within the structure of the whole character.

3) The Hierarchical Nature of Phonetic Components

Although phonetic components all serve a phonological function, the degree of correspondence between these components and the phonological realization of the whole character is not uniform, but instead exhibits clear hierarchical variation. If they are treated uniformly under the general label of “phonetic elements,” the complexity of the phonological mechanisms in modern Chinese characters may be obscured.

It is therefore necessary, on the basis of component functional analysis, to further differentiate the phonetic roles of these components in terms of their modes of phonological representation.

Starting from the present forms of characters, the phonetic functions of components can be classified hierarchically according to the degree of correspondence between the phonetic component and the phonological realization of the whole character, including aspects such as initial consonants, finals, and tones.

In general, this correspondence exhibits a continuum from strong to weak, ranging from complete correspondence of syllabic elements to partial and then weak correspondence, thereby forming a graded hierarchical system.

In making these distinctions, the present study adopts the phonological system of Modern Chinese as its reference and determines the classification based on correspondences among elements such as initial consonants, finals, and tones. In particular, the classification of rhyme categories follows the phonological system

reflected in *Zhonghua tongyun* 中華通韻,⁸ in order to ensure a relatively consistent basis for phonological categorization across different levels.

It should be noted that the phonological function of phonetic components does not constitute a binary distinction of presence or absence, but rather manifests as varying degrees of phonological cueing. The more complete the correspondence, the stronger the phonetic function; conversely, the weaker the correspondence, the more limited the cueing effect.

Accordingly, the evaluation of phonetic components depends not only on whether they provide phonological information, but also on the strength of the cues they offer.

This hierarchical classification not only facilitates a more precise description of the actual functions of phonetic components in modern Chinese characters, but also provides an operational analytical framework for the teaching of phono-semantic characters, enabling learners to develop a stratified understanding of character forms based on degrees of interpretable relationships among components.

The phonetic functions of components in phono-semantic structures can be classified into the following five types.

Table 3. Hierarchical Classification of the Phonetic Functions of Components in Phono-semantic Structures

Type	Corresponding Level	Criteria	Phonetic Strength	Example
Complete Correspondence	Initial + Final + Tone	The phonetic component fully matches the pronunciation of the whole character	Strong	青 qīng → 清 qīng
Tone Difference	Initial + Final	Only the tone differs	Relatively strong	米 mǐ → 迷 mí
Rime Correspondence	Final	Different initials, identical finals	Moderate	扁 biǎn → 篇 piān
Rime Category Correspondence	Rime Category	Similar finals (same phonological category)	Relatively weak	监 jiān → 籃 lán
Initial Correspondence	Initial	Identical initials, other elements differ	Weak	化 huà → 货 huò

⁸ *Zhonghua shici xuehui* 中華詩詞學會 and National Language Commission 國家語言文字工作委員會, *Zhonghua tongyun* 中華通韻 (Beijing: yuwen chubanshe 語文出版社, 2020).

To further examine the distribution of the above hierarchical classification in modern Chinese characters, the present study conducts a statistical analysis of the phonetic types of phono-semantic characters in the Level-1 Chinese Character Set. The results are presented below.

Table 4. Distribution of Phonological Correspondence Types in the Simplified Character Set (n = 3,755)

Phonological Correspondence Types	Count	Percentage (%)
Complete Correspondence	1568	41.75
Tone Difference	576	15.33
Partial Correspondence	510	13.58
Total (Phono-semantic)	2654	70.67

Table 5. Distribution of Phonological Correspondence Types in the Traditional Character Set (n = 3,792)

Phonological Correspondence Types	Count	Percentage (%)
Complete Correspondence	1730	45.62
Tone Difference	570	15.03
Partial Correspondence	530	13.98
Total (Phono-semantic)	2830	74.63

The statistical results reveal a clearly stratified distribution in the phonetic functions of phonetic components. Strong correspondences, including complete correspondence and tone differences, account for a substantial proportion, while weaker phonological correspondences also constitute a non-negligible share.

This distribution indicates that the phonological mechanisms of modern Chinese characters rely not only on relatively precise syllabic correspondences, but also extensively on varying degrees of phonological similarity, thereby exhibiting a hierarchical structure.

4) Component Languageing

Based on the foregoing analysis of component functions, the operational mechanisms of Chinese character components can be further explained from the perspective of structural generation. The present study conceptualizes this

phenomenon as “component languaging” and regards it as one of the key mechanisms underlying the formation of modern Chinese character structures.

The term “component languaging” refers to the process by which symbols that originally existed as graphic units gradually transform into linguistically functional structural units capable of participating in subsequent character formation, once they establish relatively stable correspondences with phonological or semantic values in the language.

In this process, components no longer rely solely on their original graphic formation, but instead participate in new structural combinations on the basis of the phonological–semantic information they carry.

Under the process of component languaging, the mode of Chinese character formation shifts from an early pictography-based graphic mechanism to one centered on component relationships, involving semantic and phonetic functions. In other words, once a component acquires relatively stable semantic or phonological functions, it can recur across different characters and assume corresponding roles within new structural configurations.

This transformation from “graphic units” to “functional units” constitutes a crucial precondition for the sustained generative capacity of modern Chinese character structures.

Within specific structural configurations, the linguistic properties of components are not simply inherited from their original meanings or pronunciations, but are instead established through their functional roles in new combinatory relationships.

For example, in phono-semantic structures, a given component may function as a phonetic element providing phonological cues, or as a semantic element contributing to semantic constraint. In other structural contexts, the same component may, as phonological–semantic relationships weaken, shift into a symbolic role. It should be noted that component languaging not only accounts for the formation of the division of labor between phonology and meaning in phono-semantic structures, but also provides a mechanistic basis for semantic composition in associative structures. Even in symbolic characters, internal components, once linguisticized, may function as new units participating in subsequent character formation.

Therefore, component languaging is not confined to a particular structural type, but operates as a general mechanism across different types of character structures.

Collectively, the phonological and semantic properties of modern Chinese characters do not derive directly from a single structural origin but are instead gradually generated through structural relationships with the participation of components.

In this sense, component languaging provides an underlying mechanistic basis for structure-based interpretation, enabling modern Chinese characters to maintain relatively stable interpretive pathways within their present forms, even after the loss of their original structural motivations.

5) The Analytical Model of Chinese Character Structure

Based on the preceding analyses, the structural system of modern Chinese characters can be further integrated. From an overall perspective, the fivefold classification system proposed in this study and the tripartite functional model of components correspond to two distinct levels of analysis.

The former operates at the level of whole-character structure, classifying characters into five types: pictographic, indicative, associative, phono-semantic, and symbolic. The latter operates at the level of components, categorizing their functions into three basic types: semantic, phonetic, and symbolic.

These two levels are not independent, but instead operate in coordination within the structural analysis of Chinese characters. The type level provides a framework for classifying whole-character structures, while the functional level reveals the operational roles of components within these structures.

By integrating the two, it becomes possible to establish a unified pathway for structure-based interpretation of modern Chinese characters without reliance on etymological reconstruction.

From the perspective of structural relationships, modern Chinese characters can be understood as a hierarchical system: at the surface level, they are manifested in the classification of whole-character types, while at a deeper level, they are reflected in the combination and distribution of component functions.

Different types of characters organize component functions in different ways, thereby giving rise to distinct structural patterns.

Within this two-level structure, structure-based interpretation is established as the fundamental analytical method, while symbolic characters define its scope of application from the opposite perspective.

Together, the two operate in coordination to constitute the overall framework for the structural analysis of modern Chinese characters.

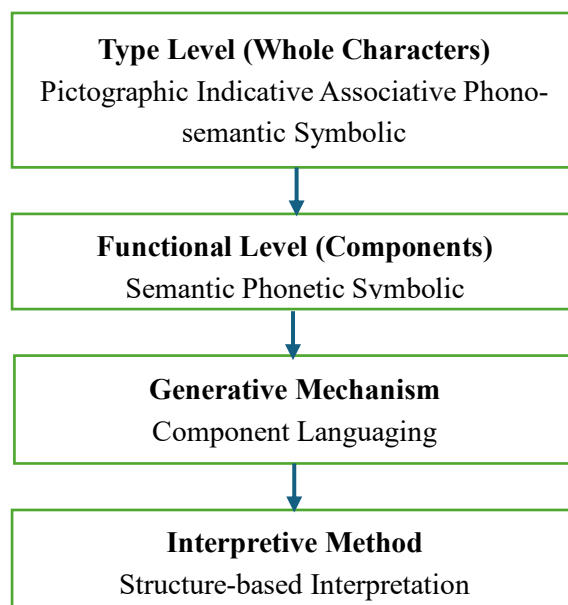
Furthermore, component languaging provides the generative mechanism underlying this two-level structure: once components acquire relatively stable phonological–semantic functions, they can repeatedly participate in structural combinations across different characters, thereby establishing a dynamic relationship between whole-character types and component functions.

Modern Chinese characters are characterized not only by type-based classification, but also by the ongoing operation of components within structural configurations.

Based on the foregoing analysis, the present study conceptualizes the structural analysis of modern Chinese characters as a two-level model comprising a “type level” and a “functional level.” Component languaging is taken as its generative basis, establishing an intrinsic link between structural hierarchy and generative mechanisms.

Grounded in present character forms and guided by the principle of structure-based interpretation, this model integrates classification, structural mechanisms, and interpretive analysis.

A Model for the Structural Analysis of Modern Chinese Characters



Therefore, the theoretical framework of the present study exhibits a clearly structured internal logic. It begins by distinguishing between characters that are interpretable within their present forms and those that are not; it then establishes the fivefold classification on this basis; subsequently, it moves to the component level to explain the functional mechanisms of interpretable character structures; and finally, it invokes component languaging to account for how these functions are sustained within the Chinese character system.

Compared with approaches that rely solely on traditional terminological frameworks, this analytical sequence is better suited to the study of modern Chinese characters and more readily adaptable to pedagogical explanation.

Explanatory Power and Pedagogical Implications of the Classification System

Based on the fivefold classification system and the tripartite functional model of components developed in the preceding sections, a systematic framework for the structure-based interpretation of modern Chinese characters can be established.

This framework not only integrates different structural types of Chinese characters at a theoretical level, but also demonstrates clear practical value in pedagogical contexts.

First, at the level of structure-based interpretation, this system provides a relatively consistent framework for the majority of modern Chinese characters without relying on complex etymological reconstruction.

By taking present character forms as the basis and distinguishing between characters that are amenable to structure-based interpretation and those classified as symbolic characters, the framework avoids inconsistencies arising from dependence on historical reconstruction, thereby enhancing the consistency and stability of analysis.

Second, at the level of pedagogical organization, the fivefold classification system provides a basic framework for categorizing whole characters, enabling instruction to group and generalize character structures at a macro level. The tripartite functional model of components further reveals the internal mechanisms of Chinese characters, allowing learners to understand the specific roles of components within character structures on the basis of recognizing character types.

For example, in teaching phono-semantic characters, learners can be guided to establish phonological–semantic correspondences by distinguishing between semantic and phonetic components. In teaching associative characters, the overall meaning can be understood through analyzing the semantic relationships among components. By contrast, symbolic characters require a different instructional strategy. For example, the character 乐 (traditional 樂) should be introduced as a symbolic character because its present form no longer preserves a stable structural correspondence among its components, pronunciation, and meaning. Rather than forcing a structure-based interpretation, instruction should emphasize holistic recognition together with its different pronunciations (*yuè*, *lè*, and *yào*) and their corresponding meanings and common lexical uses. Such an approach helps reduce unnecessary cognitive burden while improving learning efficiency.

Third, with respect to learning difficulties, the introduction of symbolic characters is particularly significant in accounting for challenges encountered in Chinese character learning. Such characters, lacking stable structure-based interpretive pathways, often require holistic memorization.

By explicitly identifying them within the classification system, instruction can treat them differently from interpretable characters, thereby improving learning efficiency and reducing the cognitive burden associated with over-interpretation.

In addition, the hierarchical classification of phonetic components provides a more refined basis for the teaching of phono-semantic characters. By distinguishing degrees of phonological correspondence, the phonetic functions of these characters can be explained in a graded manner, enabling learners to progress from stronger to weaker patterns and thereby develop a more comprehensive understanding of phonological structures in Chinese characters.

Furthermore, the concept of component languaging contributes to a holistic understanding of the generative mechanisms underlying Chinese character structure. By emphasizing that components, once endowed with phonological–semantic functions, can participate in subsequent character formation, it helps learners recognize that Chinese characters are not isolated symbols, but rather a system continuously generated through reusable components and structural relationships.

Such an understanding enhances learners' ability to analyze character structures and deepens their awareness of the systematic nature of the Chinese writing system.

In conclusion, by taking the fivefold classification system as the typological foundation, the tripartite functional model of components as the mechanistic support, and integrating both the hierarchical classification of phonetic components and the mechanism of component languaging, this study establishes an analytical framework for modern Chinese characters that combines strong explanatory power with practical applicability.

This framework not only enhances the systematic nature of structural analysis of Chinese characters, but also provides an effective analytical tool for Chinese character instruction.

Conclusion

This study takes modern Chinese characters as its object of investigation and reexamines the applicability of the traditional *liushu* framework within the present character system. On this basis, it proposes a classification framework grounded in structure-based interpretation.

It argues that the classification of modern Chinese characters should no longer rely primarily on historical reconstruction, but should instead be based on whether relatively stable phonological–semantic correspondences can be established between components and whole characters within their present forms.

According to this principle, modern Chinese characters are classified into five types—pictographic, indicative, associative, phono-semantic, and symbolic—forming a fivefold classification system.

On this basis, the study further proposes a tripartite functional classification of components—semantic, phonetic, and symbolic—and employs the concept of component languaging to explain how components are transformed into functional linguistic units capable of participating in subsequent character formation.

As a result, whole-character types and component functions are integrated into a two-level analytical model comprising a “type level” and a “functional level,” thereby enabling the classification, structural mechanisms, and structure-based interpretation of modern Chinese characters to form a coherent theoretical framework.

Overall, the aim of this study is not to reject the traditional *liushu* framework, but to reconstruct its interpretive pathway in terms of present character forms, shifting from an explanation based on structural origins to one grounded in structural relationships.

Following this shift, modern Chinese characters can be analyzed at the synchronic level within a more unified analytical framework. The significance of this approach lies not only in enhancing the consistency of structural analysis of modern Chinese characters, but also in providing a clearer theoretical foundation for Chinese character instruction and related research. This study thus contributes to the ongoing reexamination of Chinese character theory by establishing a structurally grounded analytical paradigm for modern Chinese characters with clear

pedagogical applicability.

Reference List

- Boltz, William G. *The Origin and Early Development of the Chinese Writing System*. New Haven: American Oriental Society, 1994.
- Chen Mengjia 陳夢家. *Yinxu buci zongshu* 殷墟卜辭綜述. Beijing: Zhonghua shuju, 1988.
- DeFrancis, John. *The Chinese Language: Fact and Fantasy*. Honolulu: University of Hawai'i Press, 1984.
- Duan Yucai 段玉裁. *Shuowen jiezi zhu* 說文解字注. Shanghai: Shanghai guji chubanshe 上海古籍出版社, 1981.
- Norman, Jerry. *Chinese*. Cambridge: Cambridge University Press, 1988.
- Qiu Xigui 裘錫圭. *Wenzixue gaiyao* 文字學概要. Beijing: Commercial Press, 1988.
- . Translated by Gilbert L. Mattos and Jerry Norman. *Chinese Writing*. Berkeley: Society for the Study of Early China, 2000.
- State Bureau of Standards 國家標準局. *Information Interchange Code for Chinese Characters—Basic Set (GB2312-80)* 《信息交換用漢字編碼字符集·基本集 (GB2312-80)》. Beijing: Standards Press of China 中國標準出版社, 1981.
- State Council 國務院. *Jianhuazi zongbiao* 簡化字總表. Beijing: Zhongguo wenzi gaige weiyuanhui 中國文字改革委員會, 1964.
- Su Peicheng 蘇培成. *Xiandai hanzixue gangyao* 現代漢字學綱要. 3rd ed. Beijing: Commercial Press 商務印書館, 2014.
- Tang Lan 唐蘭. *Zhongguo Wenzixue* 中國文字學. Shanghai: Shanghai Guji Chubanshe 上海古籍出版社, 2005.
- Xu Shen 許慎. *Shuowen jiezi* 說文解字. Beijing: Zhonghua shuju, 1963.
- Zhonghua shici xuehui 中華詩詞學會 and National Language Commission 國家語言文字工作委員會. *Zhonghua tongyun* 中華通韻. Beijing: Yuwen Chubanshe, 2020.
- Zhou Jian 周健. *Hanzi jiaoxue lilun yu fangfa* 漢字教學理論與方法. Beijing: Peking University Press, 2007.