DOI: 10.30050/CLLT.202212_2(2).0001

# Teaching Chinese Character Grammar

James Myers

National Chung Cheng University

## Abstract

Chinese literacy requires being able to recognize and produce thousands of distinct characters. This feat is possible only because Chinese script forms a coherent system: characters are built out of components that come from a set inventory and often have semantic and/or phonetic interpretations, and components are built out of strokes from a set inventory that are written in a relatively consistent order. In this paper, we argue that there is pedagogical value in recognizing this character system as a lexical grammar, with interpretable components as analogous to morphology, regularities in strokes as analogous to phonology, and visual perception and motor control as analogous to phonetics. For each of these three domains, we highlight patterns that teachers and students may benefit from becoming aware of, including visual cues to morphological structure and interpretation, constraints on stroke forms and combinations, and the competing universal forces underlying stroke order.

**Keywords:** Chinese characters, character rules, corpus analysis, psycholinguistics

# 1.Introduction

As the world's sole surviving logographic writing system, Chinese script poses an enormous challenge for learners of Chinese as a second (or third or fourth) language. The very large number of morphemes needing representation means that logographic systems have a very large number of orthographic units, putting a much greater burden on rote memorization than syllabaries, alphabets, or other types of systems.

How then do native Chinese speakers in Taiwan manage to memorize over 5,000 characters by the time they attend college (Hue, 2003)? The essential clue comes from asking another question: how do native English-speaking children manage to memorize up to 10,000 words by the age of 11 (Biemiller, 2005)? They can because the English lexicon, like all lexicons, is highly systematic. In accordance with the principle of duality of patterning (Hockett, 1960), lexicons are systematic at two levels, having both morphology (the grammar of meaningful elements) and phonology (the grammar of meaningless elements). The need for rote memorization is thus limited to the morphemes, phonemes, the rules that describe their combination and modification, and inevitably also a list of exceptions (since "all grammars leak": Sapir, 1921, p. 38). Moreover, articulation and perception are handled by phonetic processes that are, at least in part, hardwired and general-purpose, so they add no extra burden.

Those 5,000 Chinese characters do not form a random set either. Although individual characters may represent whole spoken morphemes or syllables, as visual forms they are built out of components that recur across other characters and that often have at least somewhat predictable meanings and/or pronunciations. As Hue (2003, p. 300) puts it, fluent readers have "a set of strategies to guess the pronunciation and the meaning of an unknown character." Even beyond vocabulary learning, psycholinguists have demonstrated that the decomposability of Chinese characters is an important aspect of mature reading (Lee, 2017) and writing (Chen & Cherng, 2013). Meanwhile, the character components themselves are decomposable into strokes taken from a fixed inventory of stroke types. It is difficult to resist the temptation to see duality of patterning here as well (Ladd, 2014), and Myers (2019) makes this point particularly explicit: the system of character components, being at least potentially

interpretable in meaning and/or pronunciation, form an analogy to morphology, while the system of strokes, being uninterpretable in themselves, form an analogy to phonology (albeit silent, as in sign languages; Brentari, 2019). The psychophysical aspects of reading (visual perception) and handwriting (manual motor control), such as stroke order, are then analogous to phonetics. Myers (2019) argues that this three-part system is the psychologically real grammar of Chinese characters.

Of course, in their amazing feats of memorization, native English-acquiring children and native Chinese-speaking students also benefit from being young, when brains are more flexible. Nevertheless, adults never fully lose their inner child, and so the more teachers are able to build on learner expectations about what makes a plausible human language (whether due to formal universals or general cognitive principles), the easier it is for their students to internalize the new system. Adult learners also have an important advantage of their own: they are much better at conscious reasoning than children. For a learner to become fluent, this grammatical knowledge must eventually be automated, or in psychologist parlance, shift from conscious "system 2" to unconscious "system 1" (Kahneman, 2011), but it is the rare adult who can absorb a new language solely via system 1 alone.

We draw two conclusions from the above discussion. First, overt teaching of the grammar of Chinese characters should be helpful to adult learners. Second, the better grounded this grammar is on general linguistic principles and psychological reality, the more pedagogically effective it should be, both because such grounding makes the grammar more accurate and also because it helps learners build on prior experience with the grammar(s) that they already know. Before future research can test these claims empirically, however, we must first reframe traditionally vague notions about what the Chinese character system is like in clear grammatical terms, in a way that is specifically aimed at teachers and students of Chinese. That is what this paper aims to do.

In section 2 we review traditional views of Chinese characters and recent developments in grammar-oriented views. In section 3 we discuss important patterns in the three domains of character grammar (character morphology, character phonology, character phonetics), many of which have received virtually no attention in the literature despite robust support from corpus-based and experimental evidence, and discuss how explicit knowledge of them

might directly impact Chinese pedagogy. Section 4 summarizes the argument.

# 2. The notion of character grammar

Chinese characters have attracted the attention of theoretical linguists for millennia, and in the modern era, overtly grammatical approaches to orthography have become ever more prominent.

## 2.1. Traditional views

The traditional analysis of Chinese characters goes back thousands of years (Bottéro & Harbsmeier, 2008), but only some aspects affect character teaching. In early classes, a few iconic characters (of the 六書 categories 象形 and 指事) are paraded out (particular favorites include 「山」, 「人」, 「大」, 「水」, 「日」, 「月」, 「上」, and 「下」); while this may help make learning characters more fun early on, it also falsely implies that iconicity is the primary means by which characters are coined. In fact, iconicity is virtually never relevant in modern characters; out of the 213 single-component characters tested by Xiao and Treiman (2012), only 15 were consistently perceived by non-native speakers as iconic, and of course, single-component characters are themselves quite a rare type in the lexicon as a whole. Fortunately, students are also soon taught the most important principle of character grammar, namely the logic of phono-semantic characters (形聲字), which eventually allows them to make reasonable guesses about the meanings and pronunciations of unfamiliar characters.

Interestingly, it turns out that character morphology is also taught indirectly, as a side-effect of how characters are ordered in textbooks. We can illustrate this with two introductory textbooks commonly used in Taiwan: *Practical Audio-Visual Chinese*, Vol. 1, 3rd edition (《新版實用視聽華語》, 第 1 卷; Hsieh, 2017) and *A Course in Contemporary Chinese 1*, 1st edition (《當代中文課程課本 1》; Teng et al., 2016). The very first line of text in *Practical Audio-Visual Chinese* is given in (1a), and the first line in *A Course in Contemporary Chinese* is given in (1b). Note that in (1a), the character 「生」is contained within the character 「姓」, and in (1b), the character pairs 「月」/「明」and 「問」/「嗎」

share components as well, and arguably also the pairs 「請」/「問」, 「明」/「是」, and 「月」/「請」.

(1) a. 李先生：先生，您貴姓？
b. 明華：請問你是陳月美小姐嗎？

In order to test if morphologically related characters cluster together more often than would be expected by chance, we ran a crude quantitative analysis on each textbook. This involved first dividing the text in each textbook into blocks of 20 sequential characters, regardless of whether they came from dialogs, as in (1), or from vocabulary lists, examples in grammar sections, or exercises. We then identified which characters within each block shared at least one component with a later different character within the same block, using the purely visual components cataloged by Wikimedia Commons (2022), which ignores semantics, pronunciation, and etymology. The question was whether related characters tended to cluster more than would be expected if the same texts were randomly scrambled. We quantified clustering through the index of dispersion (variance divided by the mean), where higher dispersion indicates that some blocks have a high concentration of related characters, rather than spreading relatedness thinly across the whole textbook.

The results are shown in Table 1, with *p* values calculated by counting how often dispersion was further from the median for 10,000 random text orders than was the actual text's dispersion. Since both textbooks showed higher dispersion (clustering) than random text (though not to the same degree), it seems that their authors did (perhaps unconsciously) succeed in exposing students to concentrations of characters that share character components.

Table 1. Dispersion (clustering) of related character counts in actual and randomly ordered texts

| Textbook | Actual dispersion | Mean random dispersion | *p* |
|---|---|---|---|
| *Practical Audio-Visual Chinese* | 1.045 | 0.610 | < .0001 |
| *A Course in Contemporary Chinese* | 0.696 | 0.581 | .029 |

In contrast to character morphology, the other two domains of character grammar

receive far less attention in the teaching of Chinese script. Teachers generally set aside traditional stroke analyses and simply have students repeatedly trace over model character forms, perhaps with numbers showing the prescriptively determined stroke order. That is, after character morphology, the next most commonly taught domain is character phonetics, leaving character phonology almost entirely neglected. Moreover, for all of these domains, including character morphology, character teaching is primarily by example, with very little discussion of the underlying principles.

## 2.2. Contemporary views

Modern analyses of Chinese characters are part of the broader study of writing systems in general (Daniels & Bright, 1996; Meletis, 2020). Within this field, grammar-based approaches are not yet mainstream; most linguists still agree with Bloomfield (1933, p. 21) that "Writing is not language, but merely a way of recording language by means of visible marks." Nevertheless, there are two good reasons to seriously consider the proposal that writing systems do have genuine grammars. First, while writing does depend in part on speech, it also has systematic structure of its own. This is particularly obvious in a logographic system like Chinese, with very large numbers of glyphs that have relatively tenuous links to speech; as noted in the introduction, it would probably be impossible to memorize so many of them if they were all completely unrelated to each other. Second, just as with spoken (and signed) languages, children come to know more about their first orthographic system than is explicitly taught in school, whether alphabetic (Pacton et al., 2001) or logographic (Tsai & Nunes, 2003).

These two facts are why the traditional teaching of Chinese script by example has worked so well for millennia: with sufficient time and motivation, young students just "pick up" character grammar, as if the system is pre-adapted to the human brain. The only problem is that traditional character teaching methods only engage the intuitive system 1, leaving the more deliberate system 2, which is particularly important for adult learners, underutilized. Teachers may themselves also benefit from becoming more consciously aware of how character morphology, character phonology, and character phonetics actually work, if for no other reason than to have answers when students ask why characters are written as they are.

Explicit formal grammars of Chinese characters include Rankin (1965), Fujimura and Kagaya (1969), Wang (1983), Stalph (1989), Li and Zhou (2007), Kordek (2013), and Peng (2017), but perhaps the most thorough analysis is that of Myers (2016, 2019, 2021a, 2021b, 2022). Not only does the grammatical framework presented in these works build on well-established linguistic theories of spoken and signed languages, but it has also been tested for psychological reality, wherever possible, through corpus-based or experimental evidence. In the remainder of this paper, then, we explore what this framework may contribute to the teaching of Chinese characters. We illustrate the argument primarily with traditional characters, though simplified characters are also mentioned when relevant (as Myers, 2019, shows, the two systems have virtually identical grammars, despite differing in their lexical inventories).

## 3. Grammatical insights and character teaching

Myers (2019) has a brief discussion about teaching Chinese characters to adults (section 6.3.1.2), but it merely reviews well-known facts that are only tangentially related to grammar. These include the transfer of systematic orthographic knowledge from the learner's first writing system (e.g., Kim et al., 2016) and the benefits of handwriting practice for character recognition in order to build higher-level mental representations that bridge the input and output modalities (e.g., Guan et al., 2011). Regarding the learning of character grammar itself, Myers (2019) unfortunately offers no advice beyond what has already been said by many other studies, such as that students have to learn individual lexical components and their favored positions (e.g., Wong, 2017), the geometry of complex characters (e.g., Loh et al., 2021), stroke types (e.g., Guo, 2008), and prescribed stroke order (e.g., Zhang, 2014).

Yet if the framework of Myers (2019) really does have some degree of psychological reality, then incorporating this framework into the classroom should help students improve the lexical quality that is needed to master Chinese reading and writing (see Perfetti, 2007, for more on lexical quality). Here we go beyond previous work in showing what a detailed character grammar framework actually implies about teaching each of the three domains of character grammar: character morphology, character phonology, and character phonetics. For each domain we begin with a sketch of the theory, and then show how explicit knowledge of

specific generalizations might benefit learning.

## 3.1. Character morphology

### 3.1.1 Theory

Character morphology involves components (部件) that may be interpreted for meaning and/or pronunciation (see Watt, 1975, for a related definition of the morphology of alphabetic systems). Building on traditional character analyses, Myers (2019) takes phono-semantic characters (形聲字) and semantic compounds (會意字) as the two primary types of morphologically complex character. However, he ascribes the latter to two distinct operations. The first is true compounding, which involves the combination of semantic components with relative freedom in how the components are placed and interpreted, as in English compounds (Jackendoff, 2010); for example, a *houseboat* is not a *boathouse*, and a *snowman* is made of snow but a *mailman* delivers mail. This variation in component position and interpretation is illustrated with 「水」 in the semantic compounds in (2).

(2) a. 冰 'ice' (substance)
b. 涉 'wade' (location) (氵 = reduced 水)
c. 盥 'wash (hands/face)' (instrument)

By contrast, reduplicated forms (疊字) are restricted to iconic interpretations like those illustrated in (3). As pointed out by Behr (2006), these semantic restrictions are shared by "true" reduplication (Hurch, 2005), as in spoken Chinese 「人人」 'everyone' (abundance), 「乾乾淨淨」 'very clean' (intensity), and 「看看」 'take a look' (attenuation).

(3) a. 多 'many' (abundance)
b. 晶 'bright' (intensity)
c. 弱 'weak' (attenuation)

Reduplication in character grammar also conform to the fixed arrangements in (4), with the first three by far the most common. Myers (2016) has demonstrated in experimentally collected judgments for fake characters that the restriction to these arrangements is a

productive part of mature reader knowledge. Myers (2019) further argues that this formal constraint is reminiscent of the role of prosodic constituents like syllables and feet in “true” reduplication (McCarthy, 2021).

(4) a. 𠚪: 比 林
b. 吕: 多 炎
c. 品: 森 蟲
d. 翉

Myers (2019) analyzes phono-semantic characters (形聲字) as derived via a third operation called affixation. The essential property of affixation is asymmetry, with affixes “weaker” than their bases. Thus affixes form a closed class, are semantically bleached (in the word *dancer*, the morpheme *dance* has rich semantics but *-er* does not), classify (*-er* marks the class of agents), and are bound, fixed in position, and often reduced (in *dáncer* and *unháppy* the underlined affixes are unstressed).

Like true affixes, semantic components in phono-semantic characters come from a closed class, with only a few dozen widely used across the lexicon and even fewer involved in modern coinages (Handel, 2017). This contrasts with the inventory of phonetic components, which mature readers are quite willing to expand, as shown experimentally by Mattingly and Hsiao (1999). They are also semantically bleached: in the compound in (5a), 「口」 literally means ‘mouth,’ but in the phono-semantic character in (5b), it merely indicates that the whole character sounds like its phonetic element. As argued extensively by Wiebusch (1995), semantic components phono-semantic characters act as semantic classifiers as well, again similar to affixes.

(5) a. 吠 ‘bark’
b. 嗎 (question particle)

Character-forming affixes are also bound: some never appear on their own, like those in (6a-b), and the rest are only historically related to their free cognates due to semantic bleaching, like that in (6c) (as also discussed by Handel, 2017).

(6) a. 冫 宀
b. 氵 [水] ⺾ [艸]
c. 貝 ‘shell’ 賬 ‘account’

The semantic components in phono-semantic characters are not fully fixed in position, but their positions are much less free than semantic components in true compounds: Myers (2019) estimates that almost 85% appear in the same position in over 50% of their phono-semantic characters. Finally, character-forming affixes tend to be reduced in form, not only via lexically idiosyncratic allomorphy as in (6b) above, but also via the regular processes illustrated in (7).

(7) a. Diagonalization: 土 in 地
b. Dotting: 木 in 材
c. Shrinking: 雨 in 雲

Table 2 summarizes the core of character morphology as analyzed in Myers (2019).

Table 2. Character morphology

| Operation | Traditional name | Properties |
|---|---|---|
| Affixation | 形聲 | Affix-like semantic component (形符) |
| Compounding | 會意 | Freer organization than in affixation |
| Reduplication | 疊體 | Iconic semantics and constrained arrangements |

Of course there are exceptions to the above points (“all grammars leak”), but Myers (2019) argues that where the patterns have been tested, they are statistically robust in the lexicon and productive in experimental tests on mature readers.

### 3.1.2 Pedagogical implications

Learning grammar, whether via system 1 or system 2, lightens the load on rote memorization. In the case of character morphology, the better the student internalizes the three morphological operations, the easier it will be to learn new characters and to write and read old ones.

Reduplication is the least productive of the three operations, but reduplicated forms are still quite frequently encountered. When students see a whole character formed in this way, it is reasonable to guess that its meaning relates in some way to the universal meanings of reduplication (plurality, abundance, duality, intensity, or attenuation), even if the base component itself is not familiar or is synchronically irrelevant to the modern whole-character meaning. This semantic heuristic works with all of the characters in (8). Often the whole-character meaning also relates to the meanings of the components, as in (8b). Moreover, since reduplication relates to semantics, students can reliably ignore component pronunciation whenever a reduplicated form is encountered. As usual, there are exceptions:「哥」has a pronunciation related to that of its component, and 「單」 'single' contains a doubled「口」(though the fact that its bottom stroke group never appears elsewhere is a clue that this character is not really morphologically decomposable).

(8) a. 多 比 朋 羽 雙 品 弱
b. 林 森 蟲 晶

Even when a reduplicated form is used as a phonetic component, students only need to remember three arrangements: 叠, 吕, 品 (the 叕 arrangement is quite rare and almost always involves this specific character, which can simply be memorized). All of these arrangements involve doubling, even 品, which is doubled along both axes (vertical and horizontal); note that the traditional name 三疊字 confuses the two very different arrangements in (9), with (9a) widely attested and (9b) extremely marginal. Reduplicative structures in the students' native languages are similarly likely to favor doubling (as in English *super-duper*).

(9) a. 品
b. 皿 (cf. 靈 龕)

Knowing the fixed reduplicative arrangements can also help students recognize that an unfamiliar character like that in (10a) must be decomposed as in (10b), since upside down triangles are ungrammatical. Other cases of non-triangular tripling, like those in (11a), and non-square quadrupling, like that in (11b), can also immediately be recognized as not being

morphologically decomposable, and thus neither their arrangement nor their components can be interpreted.

(10) a. 焚
b. 𣏌 + 火

(11) a. 三　川　彡　巛　皿
b. 灬　(as in 黑, 魚, 照 = 昭 + 灬)

Compounding is not a very productive operation either, but as with reduplication, the positions of its components also show some degree of iconicity, which might help some students remember them. For example, in compounds containing a 'hand' component, the hand is sometimes located in a sensible location, as in (12a-c). As noted earlier, however, iconicity is not pervasive enough to be useful very often; other 'hand' characters are iconic but too rare to matter to beginning students, like that in (12c), or else have become completely opaque in their modern forms, like those in (12d).

(12) a. 筆 (holding a pen at the top)
b. 掃 (holding a broom at the side)
c. 盥 (holding water)
d. 寸　弄　丞

The most important thing to learn about compounds is that they are not affixed forms. Given that the former are far less common than the latter, it is tempting for students to adopt the default strategy of treating every character as an affixed form, but of course doing so will lead to overgeneralization. Pronouncing compounds as if they were affixed forms, as in (13), is generally an even worse mistake than pronouncing affixed forms as if spoken Chinese never underwent any sound change, as in (14). Mistaking a compound for an affixed form also throws out potentially useful clues to character meaning, which typically relates, more or less, to the meanings of the components.

(13) a. 法 (*fǎ*, not *qù*)
b. 位 (*wèi*, not *lì*)
c. 宋 (*sòng*, not *mù*)
d. 析 (*xī*, not *jīn*)

(14) a. 媽 (*mā*, not *mǎ*)
b. 精 (*jīng*, not *qīng*)
c. 眼 (*yǎn*, not *gèn*)
d. 絡 (*luò*, not *gè*)

Ironically, one trick to help distinguish the two character types exploits exactly the same thing as the risky default strategy: frequency. Semantic compounds are rarer than phono-semantic characters because compounding is less productive than affixation in character grammar, and by definition, a productive operation is one that is (or has been) more likely to produce novel forms. This means not just that productively generated forms will have high type frequency (many lexical entries) but also that their token frequencies (appearances in text) will skew low, as new forms are created but not yet widely used. This is why English irregular verbs are easier to recite off the top of one's head (*run*, *eat*, *drink*, *hear*) than regular ones (*walk*, *chew*, *gulp*, *listen*); the lexicon has more of the latter, but the former are encountered more often.

Similarly, the more Chinese characters one learns, the rarer the new ones will tend to be, and thus the more likely it is that they are indeed formed via affixation. The flip side of this is that if a student keeps seeing the same not-quite-learned character again and again, its higher token frequency increases the probability that it may be a compound, and so its pronunciation should be looked up rather than guessed via its components.

Another set of heuristics to distinguish character compounding from affixation relate to the position and form of the semantic components. The simplest is that in affixed forms it is usually obvious which component is the affix, since it almost always appears on one of the four character edges, most often the left. This heuristic works even better in the simplified character system, whose creators generalized this pattern, for example replacing 「聽」with 「听」. If the affix is not obvious in a character, that character is more likely to be a compound with an unpredictable pronunciation (or may as well be treated synchronically as such, since

the eye will have trouble finding the phonetic component in fluent reading anyway).

Similarly, because affix positions are more fixed than component positions in compounds, the appearance of a familiar semantic component in an unusual location increases the probability that the character is a semantic compound. For example, when the component in (15a) is used as an affix, it favors the left and bottom, as in (15b), and so when it appears elsewhere, as in (15c), the character is more likely to be a compound (as is in fact the case here). Of course it could also be a phonetic component, as in (15d), but even in these vanishingly rare characters the non-semantic status of 「田」 is made clear by the familiar affixes (「亻」, 「金」, 「氵」) in their familiar location.

(15) a. 田
b. 略　當
c. 畫　男
d. 佃　鈿　沺

In another position-related clue, Myers (2019) observes that dictionaries do not show the left-edge bias when assigning the indexing component (部首) for compounds. Thus seeing an indexing component anywhere than at the left edge also improves the odds that the character might be a compound. Once students have memorized idiosyncratic allomorphs like 「氵」, 「忄」, 「刂」, and 「艹」 (corresponding to 「水」, 「心」, 「刀」, and 「艸」, respectively), they might be able to make use of yet another curious pattern discovered by Myers (2019): affixes on the right and top show idiosyncratic allomorphy about half the time, but in compounds it is more likely that these positions are occupied by components that never reduce. This can be seen in Table 3, which reports idiosyncratic allomorphy rates by position and character type (based on 4,889 traditional characters, some quite rare, that have single-edge indexing components). Thus if a character has a full, never-reduced indexing form on the right or top, the probability increases that it is a compound.

Table 3. Probability of idiosyncratic allomorphy in different character types and positions

| Allomorph position | Affixed form | Compound |
|---|---|---|
| Left | 46% | 41% |
| Right | 43% | 22% |
| Top | 49% | 18% |
| Bottom | 13% | 8% |

Adapted from Table 2.4 in Myers (2019, section 2.3.1.3)

Knowing how affixation works is of course also crucial for interpreting the affixed characters themselves. Because affixes overwhelmingly favor the left edge, it is reasonable (and, here, correct) to guess that the characters in (16) are related to the indicated meanings and pronunciations, with「馬」playing different roles depending on its position. Identifying affixes when they favor positions other than the left edge is a bit harder, but as we show in the next section, affix position is partly predictable from character phonology.

(16) a. 媽 *mā* 'mother':
meaning related to 女 'female', pronunciation related to 馬 *mǎ*
b. 騎 *qí* 'ride':
meaning related to 馬 'horse', pronunciation related to 奇 *qí*

Finally, the high productivity of affixation means that it alone, among the three morphological operations, is recursive, that is, capable of applying to its own output, as in (17). Recognizing this fact should help students memorize what are otherwise forbidding-looking amalgams of strokes. Since the pronunciation of the phonetic component is also passed along recursively, the whole-character pronunciation can be guessed from the most deeply embedded component, a particularly useful trick when the base of the whole-character affix is unfamiliar, as with 「匱」in (17c).

(17) a. 燙 = [[湯]火] = [[氵[昜]]火]
b. 嗨 = [口[海]] = [口[氵[每]]] = [口[氵[𠂉[母]]]]
c. 櫃 = [木[匱]] = [木[匚[貴]]]

## 3.2. Character phonology

### 3.2.1 Theory

Though character morphology dominates traditional character analysis, students also need to master character phonology, which is much richer than generally recognized. Traditional character analysis classifies strokes into fixed types but does not recognize that contrasts in stroke axis, hooking, and curving are readily captured with distinctive features (Wang, 1983; Peng, 2017; see Watt, 1975, and Primus, 2004, for featural analyses of alphabetic letters). In fact, just like features in spoken and signed phonology, they are perceived categorically based on what is lexically distinctive within the system (Yang & Wang, 2018; Myers, 2022). Moreover, as Wang (1983) demonstrated, the same features can also be used to formalize regular stroke alternations (analogous to phonological rules or constraints), like the previously mentioned diagonalization (「土」in 「地」), dotting (「木」in 「村」), and shrinking (「雨」in「雲」).

Character phonology also shows regular distributional patterns, like the restriction of the curved vertical stroke (豎撇「丿」) to the left edge of character components, as in (18a), or less commonly to the left edge of entire characters, as in (18b).

(18) a. 川　介　月　用　飛　明　所　爿　片
　　b. 邦　辣　艸

As Myers (2019) showed experimentally, mature readers' acceptability judgments for invented characters reflect knowledge of this pattern. Wang (1983) also observed informally that curving is only obligatory in narrow components (「月」), with near minimal pairs found only in wider components (e.g.,「角」vs.「甬」). Myers (2019) confirms this generalization statistically in the lexicon, while Myers (2022) shows in a perception experiment that character width really does affect readers' sensitivity to the degree of left stroke curving.

Regarding other stroke types, Wang (1983) suggested (and Myers, 2019 confirmed via corpus analysis and acceptability judgment experiments) that the vertical left-hooked stroke (豎鉤「亅」) is more likely to appear in components that are asymmetrical, as in (19), and/or

have material at the top, as in (20). Note also that the characters (21) become asymmetrical when simplified, which then triggers hooking.

(19) a. 寸
b. 乎 vs. 平

(20) a. 丁
b. 于 vs. 十

(21) a. 東 东
b. 揀 拣

Nevertheless, hooking is less regular than curving, since there are hooked components with neither asymmetry nor topping material, as in (22a), many nonhooked components with one or both properties, as in (22b), and asymmetrical simplified characters without hooking, as in (22c). There are also near minimal pairs for hooking, as in (23).

(22) a. 小
b. 午 耳
c. 车 (車)

(23) a. 于 vs. 干
b. 事 vs. 聿

The most robust generalization in character phonology (again analyzed in Wang, 1983, and experimentally tested in Myers, 2019) is the enlargement of the element at the bottom or right. This element may be an entire component, as in (24a-c), a single stroke, as in (24d), or even a multi-stroke group within a component, as in (24e).

(24) a. 昌 林
b. 巾: 帥 vs. 帖
c. 大: 尖 vs. 奇
d. 三 川 彡 土 未 井
e. 官 飛

Components that violate the stroke-level enlargement pattern are quite rare, with the pairs in (25) the only ones showing lexical contrast. However, as illustrated in (26), in phono-semantic characters, when the affix appears at the right or bottom, it usually does not enlarge, suggesting that its morphological status as a reduced affix outranks this otherwise robust phonological generalization.

(25) a. 士　末　壬
　　b. 土　未　王

(26) a. 戰　歐
　　b. 型　照

Myers (2019) ascribes these enlargement patterns to an orthographic analog of prosody (see Evertz, 2018, for an independent theory of prosody in alphabetic systems). Specifically, Chinese characters and character components have prosodic templates that are something like binary feet, except that they are two-dimensional, with a single strong (S) head at the bottom right, in contrast to weak (W) positions everywhere else. In practice, this prosodic template appears in the four forms sketched in (27), with the full form in (27a) also being symmetrical along the vertical axis.

(27) a. $\begin{bmatrix} W & \\ W & S \end{bmatrix}$　品
　　b. [W　S]　林
　　c. $\begin{bmatrix} W \\ S \end{bmatrix}$　多
　　d. [S]　中

As illustrated above, this prosodic constituent accounts not just for enlargement, but also for component arrangement in reduplication. It also allows us to say that the reason why curving, diagonalization, and dotting apply on the left and shrinking on the top is that these processes are restricted to prosodically weak positions. Myers (2019) argues that curving is less common in wider components because width too is expressed prosodically. Thus the components in (28) have the prosodic analyses as shown, placing the straight leftmost stroke

in the wide character in (28a) in a prosodic head, where curving is not allowed, while the curved stroke in (28b) is in a weak non-head position.

(28) a. 冊 [S][S]
b. 月 [W S]

The prosodic template is applied recursively along with recursive affixation, as in (29).

(29) 燙 $\left[\begin{matrix}[\mathrm{W} & \mathrm{S}]_{\mathrm{W}} \\ \mathrm{S}_{\mathrm{S}} & \end{matrix}\right]$

The link between prosody and morphology explains why the most common affix position is on the left edge: character components tend to be taller than wide, which makes them narrow along the horizontal axis, and thus "light" along this axis, and thus preferentially placed in this axis's weak prosodic position, as in (30). By contrast, intrinsically wide ("heavy") affixes often prefer the right edge, as in (31).

(30) a. 說 [W S]
b. 財 [W S]

(31) a. 鴨 [W S]
b. 翻 [W S]

For similar reasons, the few affixes that are (or have become) flat, like those in (32), generally favor the top edge, since this makes a weak-strong prosodic structure along the vertical axis. The flat yet bottom-favoring affixes 「皿」 and 「灬」 (derived from 「火」) must be treated as lexical exceptions in this analysis, though perhaps their positions were originally motivated iconically (dishes and fire go underneath things).

(32) a. 完 $\left[\begin{matrix}\mathrm{W} \\ \mathrm{S}\end{matrix}\right]$

b. 笨 $\left[\begin{matrix}\mathrm{W} \\ \mathrm{S}\end{matrix}\right]$ (cf. 竹)

Appearing in a weak prosodic position at the top or left can also motivate an affix to become light, as illustrated in (33); note that thinness also triggers curving in (33b). The idiosyncratically reduced affix 「刂」 (derived from 「刀」) is a prominent exception, because it is reduced in the prosodically strong right-edge position.

(33) a. 花 (艸 → ⁺⁺) $\begin{bmatrix} W \\ S \end{bmatrix}$

b. 臉 (肉 → 月) [W S]

As Myers (2019) demonstrates, when left-preferring affixes appear elsewhere, their next-favorite position is usually the bottom, as in (34). This is because their tallness makes them "heavy" along the vertical axis, so in a ⿱ structure they prefer the strong bottom head position.

(34) a. 言 [S]

b. 請 [W S]

c. 警 $\begin{bmatrix} W \\ S \end{bmatrix}$

Together (as Myers, 2019, also points out), the preceding two points explain the otherwise mysterious variation in 「言」 vs. 「讠」 in the simplified characters in (35): the idiosyncratically reduced form only appears in a weak prosodic position. As noted earlier, traditional and simplified characters have virtually identical grammars, despite the obvious lexical differences.

(35) a. 言 [S]

b. 请 [W S]

c. 警 $\begin{bmatrix} W \\ S \end{bmatrix}$

Exceptions can be formally analyzed as treating weak elements in the "wrong" positions as extra-prosodic, as in (36). This makes psycholinguistic predictions (e.g., that the extra-prosodic elements are indeed "outside" in some processing sense) that have yet to be tested.

(36) a. 刻 [S] W

b. 照 [S] / W

Myers (2021b) extends the prosodic analysis by proposing that the target of enlargement is a unit analogous to the syllable. A key generalization here comes from research on how people draw (van Sommers, 1984) and perceive (Changizi et al., 2006) simple geometric figures, whereby strokes preferentially start rather than end with contact. That is, given the downward and rightward pen movement favored by right-handed writers, the configurations ㇀ and ㇇, as in (37a), are preferred to ㇐ and ㇓, as in (37b). This generalization is statistically true in the Chinese component inventory, and ongoing experimental work suggests that it may also affect the acceptability and handwriting of fake character components.

(37) a. 丁 卜

b. 业 厶

Myers (2021b) sees an analogy here with the articulatory constraints that favor syllable onsets over codas (see Browman & Goldstein, 1988), namely that it is easier to coordinate gestures at the start of their movement than at the end. This analysis implies that the character in (38a) is monosyllabic (analogous to a consonant-vowel structure) but that in (38b) is disyllabic, correctly predicting that the lower stroke is the target of enlargement. Because closed components like that in (38c) are enlarged as wholes, they also form syllables, as illustrated in (38d). For similar reasons, crossed strokes, as in (38e-f), are also treated as syllabic (specifically, analogous to long vowels).

(38) a. 丁: 丁 [S]

b. 工: 丅一 $\begin{bmatrix} W \\ S \end{bmatrix}$

c. 口: 口 [S]

d. 官: 宀口口 $\begin{bmatrix} W \\ [W \\ S] \end{bmatrix}$

e. 十: 十 [S]

f. 爻: 乂乂 $\begin{bmatrix} W \\ S \end{bmatrix}$

A major benefit of applying prosodic theory to Chinese character structure, then, is that it helps capture a wide variety of form patterns that have hitherto only been considered separately, if they are even considered at all.

Table 4 summarizes this section by listing some important patterns of character phonology, according to Myers (2019) and related work.

Table 4. Character phonology according to Myers (2019, 2021b)

| Pattern | Traditional name | Properties |
|---|---|---|
| Prosodic template | None (cf. 表意文字描述字元⿰⿱...) | Head at bottom right |
| Idiosyncratic allomorphy | None (cf. 三點水「氵」, 提手旁「扌」) | Lexically specified |
| Curving | 豎撇/直撇「丿」 | Weak position |
| Diagonalization | 橫 → 提 | Weak position |
| Syllables | None | Includes ㄧ ㄧ 口 |

Even this very long discussion must leave out many character phonology patterns, some of which will be noted below. We feel that character phonology is worth discussing in detail, however, not just because it has been neglected in the traditional literature, but also because it has important implications for the teaching of characters, as we show next.

### 3.2.2 Pedagogical implications

The simplest reason for Chinese teachers to become more conscious of character phonology is to help them answer student questions. Why does the bottom horizontal stroke in 「土」become diagonal in 「地」? Because the left position in a character is prosodically

weak. Why does 「雨」become so tiny in 「電」? Because the top is also a weak position. Why do strange alternations like 三點水 even exist? Because they make affixes fit better into weak prosodic positions at the left or top. What about the reduced semantic components in 「刻」 and 「照」? Those are lexical exceptions, but, like most lexical exceptions, they have high token frequency and thus are easy to remember (which is how they survived grammar's regularizing effect in the first place). Why is there a curved stroke on the left in 「周」 but not in 「同」? Because 「周」 has more material above 「口」 than 「同」, making it taller, hence thinner. Why is there a hook in 「乎」 but not in 「平」? Because the former is asymmetrical due to the topmost stroke. Of course, student questions are of unlimited variety, and answering some may require going beyond the brief summary of character phonology given in the previous section. Nevertheless, attentive teachers (and students) should be able to figure them out for themselves. In fact, in accordance with "flipped learning" (Bergmann & Sams, 2014), some teachers may even want to present sets of characters to their students and ask them to test their own hypotheses about possible formal principles, rather then spoon-feeding the principles to them.

For example, why does 「少」 lack a hook if 「小」 has one, and similarly for the missing hook in the upper component of 「哥」? For the same reason there is no hook in the 刁-shaped stroke in 「口」: hooks cannot make (near) contact. Why does it seem intuitively worse to forget the hook in 「衣」 than in 「同」? Because unlike exception-prone leftward hooking, rightward hooking is obligatory in this context (to the left of crossed strokes), as shown by the set in (39a). The sole exception is the simplified component in (39b), in a rare instance of a grammatical difference between the traditional and simplified systems. Why then is there a rightward hook in the simplified character in (40a)? Because it conforms to the same diagonalization reduction process in (40b-d).

(39) a. 氏 民 長 艮 良 喪 辰 派 畏
b. 长 (長)

(40) a. 请 (請)
b. 土 in 地
c. 七 in 切
d. 匕 in 比

Knowing the distributional patterns of stroke types may also make it easier to parse complex characters into their components, an essential skill for interpreting phono-semantic characters and helpful for remembering the precise details in any type of complex character. Character phonology helps with this because patterns like curving, hooking, and enlargement apply at component edges and thus demarcate them: a curved vertical stroke or a rightward hook implies a component boundary to its left, as in (41a-b), a vertical stroke longer than its neighbors implies a boundary to its right as in (41c), and a horizontal stroke longer than its neighbors implies a boundary just below it as in (41d).

(41) a. 川
b. 低
c. 順
d. 圭

Memorizing the specific form of hundreds of components may also benefit from conscious awareness of the principles of their internal structure, whereby even monstrosities like those in (42a) can be decomposed into the simpler parts in (42b), which are themselves composed of the syllable-like basic stroke groups in (42c). This structure is especially easy to grasp if students literally grasp their pens, since like phonemes in speech, strokes are coordinated with each other in articulatorily motivated ways.

(42) a. 龍　龜
b. 立 亠 竺 月 冫 匕 乚 𠂉 彐 乂 口
c. 丨 一 ㄧ 卜 口 乂

The most important thing for students and teachers to know about character phonology is character prosody. Not only does prosody form the common basis for many disparate patterns (reduplication, curving, and the size of components and strokes), but it also plays the even more fundamental role of making a handwritten character look like a single thing, instead of a jumble of separate components. Prosody is why components cannot be too widely separated from each other and why the size of each component has to harmonize with the

sizes of the rest. For example, when mentally planning out a character with two components standing side to side, the writer should usually leave more room for the one on the right, unless it is an affix, in which case it may be smaller than the one on the left.

Writers also need to take into account the fact that prosody interacts with morphological structure. As Myers (2019, section 1.5.1) points out, it is superficial to treat the characters 「州」, 「琳」, and 「鴻」 as if they all have the same structure ⿲, as is done in the Ideographic Description Characters (IDE) of Unicode; the superficiality of IDE analyses is shown by the fact that mere geometric arrangement only affects the very early visual processing stages of character recognition (Chen & Yeh, 2015). By contrast, the deeper morphological and prosodic structures of these characters, crucial for parsing them in reading and compositing them in writing, are as shown in (43), where A = affix, R = root , S = strong prosodic slot (head), W = weak prosodic slot.

(43) a. 州 [R] [W S]
b. 琳 [A[RR]] [W [W S]$_S$]
c. 鴻 [[AR]A] [[W S]$_W$ S]

Character prosody helps readers as well. Until they have mastered it, Western students may treat component size as superficially as IDE does structure, with their attention unduly attracted by the components at the right and bottom (strong positions that trigger enlargement), unlike Japanese and Korean students with prior character experience (Feng et al., 2005). Knowing that character prosody interacts with character morphology also allows readers to recognize that a small component in one of these strong positions is likely to be an affix, even if it has not yet been learned. The regularity of reduction rules in prosodically weak positions, like diagonalization, dotting, and shrinking, means that their canonical forms can be looked up in memory by simply undoing the rule. Understanding that their reduced forms are not arbitrary but motivated by prosodic weakening can also help beginners view and write components as parts of one whole character rather than as isolated characters (「城」 is quite different from 「土成」). Mastery of character prosody may be particularly crucial when deciphering somebody else's handwriting, which may show variations in size or spacing that would otherwise obscure character form and cross-character boundaries.

## 3.3. Character phonetics

### 3.3.1 Theory

Just as in spoken and signed phonology, all of the patterns of character phonology seem to have their diachronic origins in character phonetics. For example, the hook in 「寸」 presumably arose via the flick of a brush pen while moving from the end of the vertical stroke to the dot, and enlargement probably occurs in the bottom and right in 「土」 and 「川」 because that is where the last-written element goes, and final motoric gestures are universally enhanced (Beckman & Edwards, 1990; Sandler, 1993; Wann & Nimmo-Smith, 1991). However, Myers (2019) argues that character phonology cannot be synchronically reduced to mere character phonetics. There is nothing for the hook to flick towards in 「丁」, for example, and enlargement can occur even where it is not favored by manual articulation (「士」 and 「 末」). In any case, all of the character phonology patterns that we have discussed are also obeyed in mechanical typefaces (like the 明體 used in this paper).

Myers (2019) also argues that conventional stroke order is an aspect of character phonetics rather than character phonology (contradicting Wang, 1983, who places stroke order in the center of his grammar). There are three reasons for thinking this. First, stroke order follows from character form rather than the other way around, as can be seen in the fact that the vertical stroke is written last in the component in (44a), but in its diagonalized allomorph in (44b), the last-written stroke is the diagonal one, in order to shorten the pen distance to the top left of the next component.

(44) a. 牛
b. 物

Second, stroke order is variable in a way that stroke form is not. Not only does conventionalized order differ across different pedagogical traditions, but once free from the teacher's red pen, individual writers generally develop their own stroke order habits. They may even do so while still in school, since, as Zhang (2014) complains, it is typically impossible to determine from a finished character whether the strokes were written in the

prescribed order (see also Katayama et al., 2009, for quantitative data on variability in stroke order among writers of Japanese kanji).

Third, and most importantly, stroke order results from universal articulatory and perceptual factors that help implement character prosody, rather than from an abstract grammar of its own. Regarding articulation, it is mechanically easier to pull than to push the writing instrument towards the writing hand, which favors moving downward and, at least for right-handers, rightward, both within single strokes and across strokes and components. Writers also prefer to minimize total pen movement (Lin, 2014; Myers, 2019), leading to the otherwise counterintuitive stroke order in 「正」: as illustrated in Figure 1, the pen moves a shorter total distance in the conventional order than in an order that maintains stroke contact at every step.

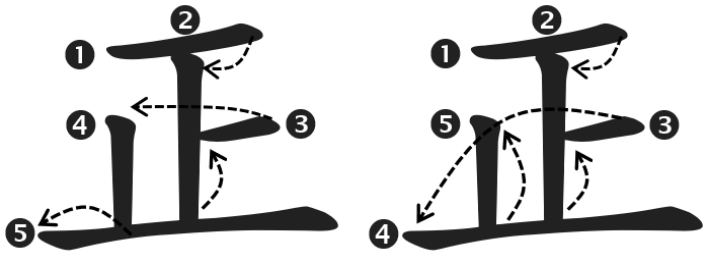


Fig. 1. Conventional (left) and less efficient (right) stroke orders for 「正」 (expanded from Myers, 2019, Figure 3.3)

Regarding perception, since dots have little effect on the overall visual appearance, they are sometimes written last rather than in strict left-to-right/top-to-bottom order (as in 「犬」). It is also often more intuitive to write symmetrical elements together (like the doubled 「糸」 at the top of 「彎」). Moreover, since components are generally prosodic units as well as morphological units, they tend to be written as wholes (so the entire「木」 on the left in 「林」 is finished before the right one is started). Visual considerations also lead some writers to fix large elements first so that the other strokes can be fit in around them, as in (45a), while others follow the articulatorily motivated left-to-right order, as in (45b).

(45) a. 方: 丶一フ丿
b. 方: 丶一丿フ

This last example also illustrates another key point: variation in stroke order arises from competition within and between the articulatory, perceptual, and prosodic principles. When writing a "+" shape, there is no right answer about which stroke to write first. Starting from —, as in Chinese 「十」, requires a disfavored leftward and upward movement to continue to ｜, but starting from ｜, as is conventional in the West for lower case "t" (as in the idiom *crossing ones t's*), requires a disfavored leftward and upward movement to continue to — (see Goodnow & Levine, 1973, for the development of this stroke order convention in young American children). Some Chinese writers write the strokes in 「小」 according to the articulatorily motivated left-to-right principle, while others follow the visually motivated dots-last and symmetry principles. While the whole-component principle is violated in surrounding components like 「口」 in 「因」, where the bottommost stroke is written after the internal component, this is not capricious, but rather is done in order to be consistent with the articulatory principle of writing lower elements later.

### 3.3.2 Pedagogical implications

The notion of character phonetics has less to teach Chinese teachers than other aspects of character grammar because they already know the two essential facts. First, handwriting characters does indeed help improve the accuracy of their mental representations, even for reading (Guan et al., 2011). This conventional wisdom remains true today, despite the fact that even native Chinese speakers tend to type far more often than they write by hand. Myers (2019) points out the reading-writing link reveals the amodal nature of character grammar as it mediates between input (vision) and output (manual gestures); experience with both modalities is thus needed for this central processing system to develop (see Bever, 1975, and Pickering & Garrod, 2013, for related ideas about grammar more generally). Repeated practice, with both reading and writing, is also crucial for the switch from deliberate system 2 to automatic system 1.

Second, despite the emphasis traditionally placed on getting stroke order "right" (Zhang, 2014), many teachers have already learned that there are benefits in deemphasizing stroke order for its own sake (Tamaoka & Yamada, 2000; Chang et al., 2015; Yin, 2016; Zhou, 2019). Much more important is mastering stroke form, an aspect of character phonology, and character components, the core of character morphology. This advice is also consistent with the character grammar framework, which sees stroke order as the natural consequence of

implementing character prosody in the manual modality, and thus it makes no sense to teach it in isolation.

In fact, the naturalness of efficient stroke order follows from its being motivated by the universal but conflicting forces of manual articulation, visual processing, and character prosody. Western students may need to unlearn some of their own (though also universally motivated) stroke order conventions, like starting with the vertical rather than the horizontal stroke in 「十」, but they can build on their prior expectations as well. For example, starting 「女」 with the large L-shaped stroke (rather than with the horizontal stroke as in 「十」) follows the same principle used in visual art: first paint Mona Lisa, then the landscape. Character prosody should have a similarly familiar feel to foreign students: finish each component before starting the next, and if a complex component has multiple subcomponents (like syllables), write them separately as well. Some students might even like to think of character prosody as analogous to music, with the components as phrases and their strokes as notes, all working to build up the tune in a balanced and harmonious way.

## 4. Conclusions

Learning a language means learning a grammar and a lexicon, including exceptions to the grammar, and learning how to use this knowledge in real life. While children seemingly do all this effortlessly, most adults need help to stay motivated, to get sufficient practice, and above all, to understand, consciously, how the grammar actually works, while repeated practice gradually makes this understanding more automatic. If Chinese characters truly do have a psychologically real grammar, then all of the above applies to learning them as well.

In this paper we have reviewed a wide variety of specific patterns, most of which have been neglected in traditional character analysis, and suggested why teachers and students may benefit from becoming aware of them. The overall pedagogical principles may be called Divide and Conquer (break complex characters, and character grammar itself, into smaller pieces) and Build on What You Know (the grammar of characters is actually not all that different from that of the student's own native language). Even recognizing the bare fact that characters do in fact have a grammar may help students feel more confident about learning individual characters: there is indeed a method in all that madness.

Learning characters means learning three distinct but interrelated sets of patterns: character morphology, character phonology, and character phonetics. Character morphology consists of three morphological operations: reduplication, compounding, and affixation. Identifying which operation is involved in which character allows students to make reasonable guesses about their meanings and pronunciations. It becomes easier to make this identification once one learns some formal regularities that ultimately derive from character phonology; for example, affixes but not compound components tend to reduce even in the prosodically strong right or bottom positions. Learning a bit of character phonology also means that learners need not memorize character components by rote, as if they were bits of abstract art, since many aspects of stroke form and their combinations are predictable (e.g., curving on the left, T-shaped stroke groups). Character prosody can also guide the student's hand and eye to treat characters as wholes and to mentally organize a whole host of other regularities, including relative stroke sizes and reduplication structure. Finally, character phonetics implements the grammar in the manual and visual domains, helping to make things like stroke order make sense.

Admittedly, at this point the benefits claimed for teaching character grammar are speculative, since none have been actually tested in the classroom. It is thus our hope that at least one reader of this paper will want to subject our suggestions, or perhaps even brand new ideas inspired by the character grammar framework, to rigorous empirical tests.

## Acknowledgements

This paper benefited from suggestions and corrections from Jane Tsay (蔡素娟) and Xinhui Chen (陳欣徽), as well as from my research assistants Yangru Chen (陳泱儒) and Meichun Liu (劉美君), who received support from the granting agency formerly known as the Ministry of Science and Technology (MOST 109-2410-H-194-096-MY3). All remaining errors are my own contribution.

# References

Beckman, Mary E. & Jan Edwards. (1990). Lengthening and shortening and the nature of prosodic constituency. In John Kingston & Mary E Beckman. (eds.). *Papers in laboratory phonology I: Between the grammar and the physics of speech*, 152-178. Cambridge: Cambridge University Press.

Behr, Wolfgang. (2006). Homosomatic juxtaposition and the problem of syssemantic (huìyì) characters. In Françoise Bottéro & Redouane Djamouri (eds.). *Écriture chinoise: données, usages et representations*, 75-114. Paris: École des Hautes Études en Sciences Sociales.

Bergmann, Jonathan & Aaron Sams. (2014). *Flipped learning: Gateway to student engagement*. Eugene & Washington, DC: International Society for Technology in Education.

Bever, Thomas G. (1975). Psychologically real grammar emerges because of its role in language acquisition. In Daniel P. Dato (ed.). *Developmental psycholinguistics: Theory and application*, 63-75. Washington, DC: Georgetown University Press.

Biemiller, Andrew. (2005). Size and sequence in vocabulary development: Implications for choosing words for primary grade vocabulary instruction. In Elfrieda H. Hiebert & Michael L. Kamil (eds.). *Teaching and learning vocabulary: Bringing research into practice*, 223–242. Mahwah, NJ: Lawrence Erlbaum Associates.

Bloomfield, Leonard. (1933). *Language*. New York: Holt.

Bottéro, Françoise & Christoph Harbsmeier. (2008). The Shuowen Jiezi dictionary and the human sciences in China. *Asia Major, 21* (1), 249-271.

Brentari, Diane. (2019). *Sign language phonology*. Cambridge: Cambridge University Press.

Browman, Catherine Phebe & Goldstein, Louis. (1988). Some notes on syllable structure in articulatory phonology. *Phonetica, 45*(2-4), 140-155.

Chang, Li-Yun, Joseph Z. Stafura, Ben Rickles, Hsueh-Chih Chen & Charles A. Perfetti. (2015). Incremental learning of Chinese orthography: ERP indicators of animated and static stroke displays on character form and meaning acquisition. *Journal of Neurolinguistics* 33. 78-95.

Changizi, Mark A, Qiong Zhang, Hao Ye & Shinsuke Shimojo. (2006). The structures of

letters and symbols throughout human history are selected to match those found in objects in natural scenes. *The American Naturalist, 167*(5), E117-E139.

Chen, Jenn-Yeu & Rong-Ju Cherng. (2013). The proximate unit in Chinese handwritten character production. *Frontiers in Psychology, 4*(517), 1-7.

Chen, Yi-Chuan, & Su-Ling Yeh. (2015). Binding radicals in Chinese character recognition: Evidence from repetition blindness. *Journal of Memory and Language, 78*, 47-63.

Daniels, Peter T. & William Bright (eds.). (1996). *The world's writing systems*. Oxford & New York: Oxford University Press.

Evertz, Martin. (2018). *Visual prosody: The graphematic foot in English and German*. Berlin: Walter de Gruyter.

Feng Liping [馮麗萍], Lu Huayan [盧華岩], & Xu Caihua [徐彩華]. (2005). 部件位置信息在留學生漢字加工中的作用 [The role of information about radical position in processing]. 語言教學與研究 [*Language Teaching and Linguistic Studies*], *2005*(3), 66-72.

Fujimura, Osamu & Ryohei Kagaya. (1969). Structural patterns of Chinese characters. In *Proceedings of the 1969 conference on Computational linguistics*, 1-17. Stockholm: Association for Computational Linguistics.

Goodnow, Jacqueline J. & Rochelle A. Levine. (1973). The grammar of action: Sequence and syntax in children's copying. *Cognitive Psychology, 4*(1), 82-98.

Guan, Connie Qun, Ying Liu, Derek Ho Leung Chan, Feifei Ye & Charles A. Perfetti. (2011). Writing strengthens orthography and alphabetic-coding strengthens phonology in learning to read Chinese. *Journal of Educational Psychology, 103*(3), 509-522.

Guo, Sheng Lin [郭聖林]. (2008). 漢字的筆畫特點與外國學生漢字筆畫偏誤 [The features of Chinese character stroke and foreign learners' errors of Chinese character stroke]. 暨南大學華文學院學報 [*Journal of College of Chinese Language and Culture of Jinan University*], *2008* (4), 63-69.

Handel, Zev. (2017). The cognitive role of semantic classifiers in modern Chinese writing as reflected in neogram creation. Presented at Seen not heard: composition, iconicity, and the classifier systems of logosyllabic scripts. The Oriental Institute of the University of Chicago.

Hockett, Charles Francis. (1960). The origin of speech. *Scientific American, 203*(3), 88-97.

Hsieh, Chia-ling (ed.). (2017). 新版實用視聽華語, 第 1 卷, 第三版 [*Practical audio-visual Chinese vol. 1*, 3rd ed]. Taipei: Cheng Chung Book Publisher [正中書局].

Hue, Chih-Wei. (2003). Number of characters a college student knows. *Journal of Chinese Linguistics, 31*(2), 300-339.

Hurch, Bernhard (ed.). (2005). *Studies on reduplication*. Berlin: Walter de Gruyter.

Jackendoff, Ray. (2010). The ecology of English noun-noun compounds. In Ray Jackendoff (ed.). *Meaning and the lexicon* (pp. 413-451). Oxford: Oxford University Press.

Kahneman, Daniel. (2011). *Thinking, fast and slow*. New York: Macmillan.

Katayama, Yoshinori, Seiichi Uchida & Hiroaki Sakoe. (2009). Stochastic model of stroke order variation. In Bob Werner (ed.). *10th International Conference on Document Analysis and Recognition* (pp. 803-807). Barcelona: Institute of Electrical and Electronics Engineers.

Kim, Say Young, Ting Qi, Xiaoxia Feng, Guosheng Ding, Li Liu & Fan Cao. (2016). How does language distance between L1 and L2 affect the L2 brain network? An fMRI study of Korean-Chinese-English trilinguals. *NeuroImage, 129*, 25-39.

Kordek, Norbert. (2013). *On some quantitative aspects of the componential structure of Chinese characters*. Poznań: Wydawnictwo Rys.

Ladd, Dwight Robert. (2014). *Simultaneous structure in phonology*. Oxford: Oxford University Press.

Lee, Chia-Ying. (2017). Sublexical processes for reading Chinese characters, neurolinguistic studies. In Rint Sybesma, Wolfgang Behr, Yueguo Gu, Zev Handel, C.-T. James Huang & James Myers (eds.). *Encyclopedia of Chinese language and linguistics vol. IV* (pp. 214-218). Leiden: Brill.

Li, Jianyu & Jie Zhou. (2007). Chinese character structure analysis based on complex networks. *Physica A: Statistical Mechanics and its Applications, 380*, 629-638.

Lin, Gui-Sheng. (2014). 漢字筆順規範的理論是路徑法 [Chinese character stroke order theory is the path method]. 黃岡師範學院學報 [*Journal of Huanggang Normal University*]*, 34*(1), 91-96.

Loh, Elizabeth Ka Yee, Xian Liao, Shing On Leung & Loretta Chung Wing Tam. (2021). How do Chinese as a second language (CSL) learners acquire orthographic knowledge: component, structure and position regularity. *Language Awareness, 30*(3), 297-316.

Mattingly, Ignatius G. & Pai-ling Hsiao (1999). Are phonetic elements in Chinese characters drawn from a syllabary? *Psychologia, 42*(4), 281-289.

McCarthy, John J. (2021). Prosodic Morphology. In Carlos Gussenhhoven & Aoju Chen (eds.). *The Oxford Handbook of Language Prosody* (pp. 96–103). Oxford: Oxford Handbooks.

Meletis, Dimitrios. (2020). *The nature of writing: A theory of grapholinguistics*. Paris: Fluxus Editions.

Myers, James. (2016). Knowing Chinese character grammar. *Cognition, 147*, 127-132.

Myers, James. (2019). *The grammar of Chinese characters: Productive knowledge of formal patterns in an orthographic system*. Abingdon, UK: Routledge.

Myers, James. (2021a). Levels of structure within Chinese character constituents. In Yannis Haralambous (ed.). *Grapholinguistics in the 21st Century: From graphemes to knowledge Part II* (pp. 645-582). Paris: Fluxus Editions.

Myers, James. (2021b). Areal script form patterns with Chinese characteristics. *Written Language & Literacy, 24*(2), 259-283.

Myers, James. (2022). Thin & curvy: Unconscious knowledge of a subtle Chinese character stroke pattern. Presented online at the 34th North American Conference on Chinese Linguistics.

Pacton, Sébastien, Pierre Perruchet, Michel Fayol & Axel Cleeremans. (2001). Implicit learning out of the lab: the case of orthographic regularities. *Journal of Experimental Psychology: General, 130*(3), 401-426

Peng, Xuanwei. (2017). Stroke systems in Chinese characters: A systemic functional perspective on simplified regular script. *Semiotica, 218*, 1-19.

Perfetti, Charles. (2007). Reading ability: Lexical quality to comprehension. *Scientific Studies of Reading, 11*(4), 357-383.

Pickering, Martin J. & Simon Garrod. (2013). An integrated theory of language production and comprehension. *Behavioral and Brain Sciences, 36*(4), 329-347.

Primus, Beatrice. (2004). A featural analysis of the Modern Roman Alphabet. *Written Language & Literacy, 7*(2), 235–274.

Rankin, Bunyan Kirk. (1965). *A linguistic study of the formation of Chinese characters*. Philadelphia: University of Pennsylvania Ph.D. thesis

Sandler, Wendy. (1993). A sonority cycle in American Sign Language. *Phonology, 10*(2), 243-279.

Sapir, Edward. (1921). *Language: An introduction to the study of speech*. New York: Harcourt, Brace and Company.

Stalph, Jürgen. (1989). *Grundlagen einer Grammatik der sinojapanischen Schrift* [Fundamentals of a grammar of Sino-Japanese script]. Wiesbaden: Harrasowitz Verlag.

Tamaoka, Katsuo, & Hiroyuki Yamada. (2000). The effects of stroke order and radicals on the knowledge of Japanese kanji orthography, phonology and semantics. *Psychologia, 43*(3), 199-210.

Teng Shou-hsin [鄧守信] (ed.). (2016). 當代中文課程課本 1 [*A course in contemporary Chinese* 1]. Taipei: Linking Publishing Company [聯經出版事業公司].

Tsai, Kuan-Chun & Terezinha Nunes. (2003). The role of character schema in learning novel Chinese characters. In Catherine McBride-Chang & Hsuan-Chih Chen (eds.). *Reading development in Chinese children* (pp. l09-125). London: Praeger.

Van Sommers, Peter. (1984). *Drawing and cognition: Descriptive and experimental studies of graphic production processes*. Cambridge: Cambridge University Press.

Wang, Jason Chia-Sheng. (1983). *Toward a generative grammar of Chinese character structure and stroke order*. Madison, Wisconsin: University of Wisconsin-Madison Ph.D. thesis.

Wann, John & Ian Nimmo-Smith. (1991). The control of pen pressure in handwriting: A subtle point. *Human Movement Science, 10*(2-3), 223-246.

Watt, William C. (1975). What is the proper characterization of the alphabet? I: Desiderata. *Visible Language, 9*(4), 293-327.

Wiebusch, Thekla. (1995). Quantification and qualification: Two competing functions of numeral classifiers in the light of the radical system of the Chinese script. *Journal of Chinese Linguistics, 23*(2), 1-41.

Wikimedia Commons. (2022). *Chinese character decomposition*/. https://commons.wikimedia.org/-wiki/ Commons:Chinese_characters_decomposition (accessed 3 October 2022)

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.

Xiao, Wen & Rebecca Treiman. (2012). Iconicity of simple Chinese characters. *Behavior Research Methods, 44*(4), 954-960.

Yang, Ruoxiao, & William Shi Yuan Wang. (2018). Categorical perception of Chinese characters by simplified and traditional Chinese readers. *Reading and Writing, 31*(5), 1133-1154.

Yin, John Jing-hua. (2016). Chinese characters. In Sin-Wai Chan (ed.). *The Routledge encyclopedia of the Chinese language* (pp. 51-63). Abingdon, UK: Routledge.

Zhang, Haiwei. (2014). A review of stroke order in hanzi handwriting. *Language Learning in Higher Education, 4*(2), 423-440.

Zhou, Jianjiao. (2019). Discussion on the correct direction of stroke order teaching. In *Advances in Social Science, Education and Humanities Research 341* (Proceedings of the 5th international conference on arts, design and contemporary education advances) (pp. 612-615). Atlantis Press.

〔收稿：2022.10.24 修訂：2022.11.27〕

麥傑 James Myers
中正大學語言學研究所
Graduate Institute of Linguistics
National Chung Cheng University
lngmyers@ccu.edu.tw

# 漢字的語法教學

麥傑

中正大學

## 摘要

中文的讀寫能力需要能夠認得與書寫數千個不同的字符。中文使用者能掌握此技能，是因為漢字有一個有條理的系統：漢字由固定的部件所構成，這些部件通常帶有語義或表示語音；而部件則是由一組筆畫依照某種固定順序所組成。在本文中，我們將漢字字符系統視為書寫系統的語法，稱為漢字語法，並認為此語法具有教學價值。漢字中有意義的部件可類比於構詞學，筆劃的規律性可類比於音韻學，視覺感知和動作控制則類比於語音學。我們詳述漢字語法的構詞、音韻與語音三方面的規律。若教師和學生能掌握這些漢字的語法規則，可能會因此受益。這些規律包括漢字的字形結構與釋義的視覺線索、筆劃樣式與組合的音韻限制，以及筆順背後幾個互相競爭的普遍原則。

**關鍵詞：**漢字、漢字規律、語料庫為本的分析、心理語言學