2018

The Challenge of Chinese Character Acquisition: Leveraging Multimodality in Overcoming a Centuries-Old Problem

Justin Olmanson  
*University of Nebraska Lincoln*

Xianquan Liu  
*University of Nebraska Lincoln*

Follow this and additional works at: [https://digitalcommons.montclair.edu/eldj](https://digitalcommons.montclair.edu/eldj)  
Part of the [Instructional Media Design Commons](https://digitalcommons.montclair.edu/eldj)

**Recommended Citation**  
Available at: [https://digitalcommons.montclair.edu/eldj/vol4/iss1/1](https://digitalcommons.montclair.edu/eldj/vol4/iss1/1)
The Challenge of Chinese Character Acquisition: Leveraging Multimodality in Overcoming a Centuries-Old Problem

Justin Olmanson and Xianquan Liu
University of Nebraska Lincoln
1400 R St, Lincoln, NE 68588
jolmanson2@unl.edu
May 23, 2017

ABSTRACT

For learners unfamiliar with character-based or logosyllabic writing systems, the process of developing literacy in written Chinese poses significantly more obstacles than learning to read and write in a second language like Portuguese or Cherokee. In this article we describe the linguistic nature of Chinese characters; we outline traditional and new media approaches to Chinese character acquisition; we unpack how multimodal technologies combined with computational linguistics might be used to provide new types of support for Chinese character learning; and we offer a design that incorporates several of these concepts into a digital writing support tool that could work as a scaffold to enable Chinese language students to leverage their Chinese listening and speaking skills as well as their visual literacies in support of producing and learning Chinese characters.

Keywords: Chinese characters, second language acquisition, design, writing, multimodality, technology

INTRODUCTION

China is historically, culturally, economically, and politically important. With a written history dating back to 2100 BCE, an 18.8% share of the total world population, and the world’s second largest economy as measured by overall GDP, the profile of China and Chinese cultural influence continues to grow (Starr, 2009). For these and other reasons, interest in studying the Chinese language has continued to increase in North America and across Europe.

While learning any language involves a great deal of effort, learning a character-based or logosyllabic language like Chinese can prove especially challenging for learners accustomed to syllabic writing systems such as English and German. In this article we unpack the characteristics of Chinese characters; we describe a range of contemporary approaches to Chinese character acquisition; and we outline how different modal channels could be used in unique ways to make new types of support for Chinese character development possible.

CHINESE CHARACTERS: A BRIEF INTRODUCTION

Over the past 600 years, through 25 dynasties, Chinese characters have evolved to become the dominant form of written communication for 1.3 billion people. Reading and writing in Chinese involves up to 20000 different characters with 7000 characters making up a typical academic vocabulary and knowledge of the 2500 most frequent characters allowing for newspaper reading and everyday written communication. Chinese characters are primarily pictographic and ideographic—meaning they represent real objects and ideas in written form. While some characters or character features offer pronunciation clues there is no one-to-one script-to-sound correspondence.

Being considered literate in written Chinese means being able to successfully navigate abstracted visual representations, combined-character meanings, pronunciation cues, and character stroke order. What is more, even for native Chinese speakers, the pronunciation cues embedded in Chinese characters require character knowledge before they are useful. In the 1950s a phonetic writing system for Chinese called pinyin was developed to support foreign learners of Chinese in the early stages of the language acquisition process. Pinyin was so useful that it was adopted for use with Chinese elementary school students to support their path to literacy as well. Although the phonetic system of pinyin uses nearly the same letters as English, spoken Chinese has additional features. Finals (vowels) are tonal and words have different spelling rules. Pinyin is comprised of three components: initial consonant-like sounds, final vowel-like sounds, and tones that slightly
change the pronunciation of the finals and completely change the meaning of the word (i.e., mā: mother, mā: to bother, mà: horse, mà: to scold, ma: an interrogative particle). Chinese teachers of learners unfamiliar with character-based writing systems often start their language students in pinyin as it allows their ability to express themselves via writing to keep up with their spoken language acquisition. However, as students develop their listening and speaking skills, the eventual and necessary transition from writing and reading via pinyin to using characters becomes more difficult as students struggle to memorize characters and learn the nuances of the writing system while also expanding their spoken vocabulary.

Written Chinese has three structural tiers namely stroke, radical, and character. Characters are the smallest meaningful unit in the Chinese writing system, and are made up of combinations of radicals which are made up of strokes (Wong et al., 2013). There are eight basic radicals (Lu, Meng & Tam, 2014) that are used to generate 44 additional radical shapes, which are used to build 7000 frequently used characters—based on relational writing principles, such as stroke order and positioning of radicals (Chang, Xu, Perfetti, Zhang, & Chen, 2014).

Strokes are the basic lines for the writing system. Radicals, or root characters, are created via combinations of strokes based on set rules. Radicals give clues to character meaning and pronunciation. Characters with with only one radical are called simple characters, for example, “口 (mouth)”. Characters made up of more than one radical are called compound characters, for example, “吃 (eat, pronounced as “chi”)” consists of a combination of the root characters “口 (mouth)” and “乞 (beg)”. “口 (mouth)” is used as a radical on the left of the character “吃 (eat)”—cluing the reader in on the character being related to mouth. “乞 (beg)” is pronounced “qi” and is positioned to the right of the character “吃 (eat)” to provide a phonetic cue indicating that “吃 (eat)” has the same final “i” as “乞 (beg)” but different tones. Similar to inflectional affixes in English, radicals can provide meaning or pronunciation clues. For example, the radical “氵” indicates having to do with “water”, which helps semantically explain characters such as “河 (river)” and “海 (sea).” Conversely, some radicals, such as “亻” when used on the right side, indicate that the overall character has an “e” sound in its pronunciation.

In terms of phraseology, most characters are free morphemes, however the majority of them need to be combined to make words and phrases. The ability to use characters in concert with other characters to form words and phrases is central to communicating in written Chinese. For example, months of the year are written based on the characters for 1-12 and 月 (month). January is 一月. February is 二月 and so on. In this way, writing the months of the year is based on a synthesis of understanding how characters can be used together to express ideas via context and juxtaposition and the meaning of the individual characters. Additionally, there are no spaces in Chinese writing—making the recognition of words and phrases more complex without explicit signals for when a word ends and another begins.

Due to the complexity of the writing system, character learning is a complicated and multi-layered process that takes effort, explanation, and time to acquire. In the next section we discuss approaches to learning to read and write in Chinese.

**APPROACHES TO CHINESE CHARACTER ACQUISITION**

There are three dimensions of Chinese character development. The first is the visual presentation of characters, which involves related orthographic knowledge of stroke order as well as phonetic and semantic cues. The second is word formation, and it refers to the ability to use the character in concert with other characters to generate words and phrases to build written texts (Shen, 2013). The third is sound-graph-meaning connections which deals with the relationships between pronunciation, visual presentation, and semantics.

Although there are different pedagogical approaches and learning strategies related to Chinese character instruction and acquisition, most of them rely on rote memorization and mechanical repetition. Thus, Chinese character learning is experienced by students as tedious, time-consuming, and labor-intensive (Tse, Marton, Ki & Ka, 2007). Students are required to hand-copy characters stroke by stroke while saying or thinking each character—repeating this process until they are able to recognize or reproduce characters from memory (Tse, Marton, Ki & Ka, 2007).

Pedagogical strategies to promote character learning involve material organization, presentation, and application (Shen, 2013). Typically, grouping characters takes on either a meaning-centered or character-centered
organization. In the latter, characters of the same graphic form are shown and studied together in order to direct students’ attention to the shared radicals across characters (e.g., 湖 (lake) and 溪 (brook) also share the same water-related radical “氵” as sea and river above). This process supports the development of semantic and orthographic awareness (Chang, Xu, Perfetti, Zhang & Chen, 2014). While meaning-centered approaches present characters in context—something generally seen as pedagogically preferable to decontextualized groupings—this approach alone does not cultivate the phonological and orthographic awareness students need to successfully communicate in written Chinese (Chang, Xu, Perfetti, Zhang & Chen, 2014). Tse and colleagues (Tse, Marton, Ki & Ka, 2009) introduced an integrative perceptual approach that used a combination of meaning-centered and character-centered strategies in presenting characters while emphasizing character structure and form, character meaning, connections between character form and meaning, speech and text associations, and strategies for character usage. Again, despite the power of both meaning and character-centered organizations, rote memorization and automated recall remains the most emphasized approach to teaching and learning Chinese characters.

Six cognitive strategies and two metacognitive strategies have been used by non-native beginning-level Chinese learners (Zhao and Jiang, 2002). They are: following stroke order, associating sounds with characters, paying attention to character configuration, understanding radicals, and frequently using learned characters. The metacognitive strategies involve analyzing writing errors and plan-making for character learning. Shen’s study (2013) of Chinese literacy development identifies three factors facilitating Chinese character learning: orthographic-based cognitive strategies, metacognitive beliefs and skills, and strategy training.

As digital media have become increasingly ubiquitous in everyday life, Chinese language learning teachers have integrated education technology applications into the curriculum. In the last few years an increasing number of web and mobile apps have been designed specially to support character acquisition. Such technologies can be classified as primarily supporting: phonological & orthographic awareness, character memorization, semantic association, stroke-by-stroke character creation, and dictionary apps (see Table 1).

Educational technologies often aim to promote motivation and student engagement by transforming tedious tasks like character practice and memorization into more enjoyable experiences. While a few educational technologies like 悟空识字 (wùkōngshízì) provide narrative game-based activities that support

---

**Table 1. Taxonomy of different education technology character acquisition applications.**

<table>
<thead>
<tr>
<th>Categories</th>
<th>App Descriptions</th>
<th>For example...</th>
</tr>
</thead>
<tbody>
<tr>
<td>Phonological &amp; orthographic awareness</td>
<td>Calling out different elements of characters (meaning and pronunciation) and the internal structure of characters (strokes, radicals etc.)</td>
<td>Art of Chinese characters I &amp; II, 写汉字 (xiě hànzi), Chinese alphabet coloring book, 宝宝游戏识汉字 (bāobao yóuxì shí hànzi)</td>
</tr>
<tr>
<td>Character memorization</td>
<td>Providing repetitive experiences for identifying characters supported via (sound-character and/or meaning-character representations)</td>
<td>悟空识字 (wùkōngshízì) monki Chinese class, Fun Chinese, Learn Chinese by minddsnack, Chinese skill</td>
</tr>
<tr>
<td>Semantic association</td>
<td>Associating characters with meaning, with other characters, with the origins and evolution of the character, with related phrases</td>
<td>Art of Chinese characters I &amp; II, 悟空识字 (wùkōngshízì), Linkit</td>
</tr>
<tr>
<td>Stroke-by-stroke character creation</td>
<td>Repetitive writing practice, calligraphic practice of strokes and radicals</td>
<td>Chinese writer, Chinese alphabet coloring book, virtual brush (虚拟毛笔), Live calligraphy</td>
</tr>
<tr>
<td>Dictionary apps</td>
<td>Provide pronunciation with definitions in the learner's first language</td>
<td>Trainchinese: dictionary &amp; Flashcards, Pleco Chinese dictionary</td>
</tr>
</tbody>
</table>
sharing and peer feedback, most of the existing technologies for character acquisition are based on decontextualized rote memorization and mechanical repetition.

The Challenge of Chinese Character Acquisition for Learners Unfamiliar with Logosyllabic Writing Systems

Neuroscience research has found that both the left and right brain hemispheres are active when Mandarin speakers hear Chinese, while only the left of the brain is active when English speakers hear English (Wong, Chai & Gao, 2011). Among all the challenges in learning the Chinese language, Chinese character acquisition has been identified as a primary sticking point for those who have never learned a character-based writing system before (Hu, 2010, Ke et al., 2001, Shei & Hsieh, 2012, Shen, 2015). For these learners, Chinese presents the challenges mentioned above as well as a low number of cognates with Indo-European languages, a tonal element that complicates comprehension and production, as well as other cultural and structural differences unique to the language.

Shen asserts that the difficulties in Chinese character learning lie in the challenge of retention and retrieval of the three elements of Chinese characters—sound (pronunciation), shape (visual presentation or the written form of the character), and meaning (2013). Hu’s research on the Chinese language learning experiences of students in the UK reported difficulties with character learning: recalling how to write words, recognizing the words, and recalling vocabulary (2010). Lu and colleagues (2014) identified three challenges of Chinese character acquisition for non-Chinese learners. The first is the development of structural awareness of characters. It is very challenging for learners unfamiliar with the Chinese writing system to notice and make sense of information embedded within the Chinese character structure. The second challenge is executing the correct stroke order. There is a positive correlation between following correct stroke-sequence and producing correct characters (Kang, 2011)—which in turn can promote character recognition (Lu, Meng & Tam 2014). However, even after repeated instruction, students constantly struggle in following proper stroke sequencing. The third challenge for students is making connections between characters and their corresponding pronunciation.

As stated above, while understanding character types and radicals can provide supportive insights for individuals learning to read and write Chinese, students often resort to memorizing characters in isolation without the help of radical-embedded phonetic and semantic cues (Chen, Hsu, Chang, Lin, Chang & Sung, 2013). Relying on a repetitive, one-character-at-a-time, write-it-until-you-remember approach creates inefficiencies—wherein the pace of character acquisition remains flat. Relying on repetition, memorization, and stroke practice magnifies the disconnect between the spoken and written systems. This disconnect creates significant differences in the speed of spoken and written language development—resulting in much larger spoken vocabularies than written ones. This disparity can impact learner morale, student persistence, pedagogical options, and overall acquisition rates. Due to these issues, some instructors leave reading and writing up to the students to acquire outside of class (Zhong, 1990, cited in Chang, Perfetti, Zhang & Chen, 2014).

Limited knowledge of characters compared to the more familiar pinyin hinders students from authentic readings and interpersonal writing practice. This limitation in communicative capacity restricts teachers from integrating activities and assignments requiring character writing, which means students are not required to engage in activities involving written output with characters until much later in their course of study—denying them access to an array of authentic texts.

As the computer revolution of the 1980s in the US spread to China in the 1990’s, Chinese character input methods were invented to enable communication over digital mediums. More than a thousand input methods were developed—including methods based on character shape (e.g., Zhengma (郑码) and Wubizixing (五笔字型)) and methods based on phonetic systems such as pinyin. Over time, two-step methods that used pinyin input and character selection became the most popular. With this method, characters are produced by typing initials and finals without tones to generate a list of possible characters. Next, writers select the intended character from the list. While pinyin input methods over digital devices connect pronunciation with characters, they require the user to already be able to identify the correct character from a list of characters that share the same pinyin spelling.

Multimodality and New Media in Chinese Input Methods and Character Acquisition

The use of text and images in combination for educational purposes enjoys a long history—with the
first illustrated language textbook being a Latin text dating to the middle of the seventeenth century (Comenius, 1999). This history exists in part due to the impact of multimodal learning. Words and language represent experiences that are, for most people, formed upon layers of simultaneous multi-sensory input. Visual, aural, haptic, and sometimes olfactory and gustatory experiences are a part of navigating the world and thus, also a part of learning to navigate new cultures and languages. Innovations such as writing, drawing, photography, and print media have offered powerful gateways through which to acquire language and learn cultural nuances (Jewitt, 2005). With the ubiquity of networked devices and new media, the opportunities to use text, sound, and images in concert with one another to support learning and productivity has become increasingly commonplace (Kalantzis & Cope, 2012).

Figure 1. The Sougou pinyin input method uses emojis in place of their corresponding characters.

For example, some pinyin input methods provide visual supports. Sougou pinyin, the most popular pinyin system in China, inserts emojis in place of their corresponding characters (Figure 1). This input method provides visual support for making the process of writing more interesting and potentially speeding up character selection. While this could be useful—allowing Chinese language learners to use their speaking abilities to enter pinyin and click on emoji options to produce characters—the images are only provided for a limited number of high frequency verbs (xihuan 喜欢, like, ♥), nouns (hanbao 汉堡, hamburger, ), and adjectives. Additionally, the images supplant the corresponding characters, meaning that the learners cannot focus on character features when making their selection.

While designing and developing emerging technologies in the early 21st century remains a resource-intensive undertaking, access to emoji libraries, or user-generated, creative commons-licensed, meta-tagged images, or text-to-speech audio has expanded the opportunities to offer multimodal representations and supports without having to invent or create each item during the technology development process. While Sougou was not designed for Chinese language learners, their design suggests that supporting expression and interaction via access to high quality images and audio files in real-time affords learning technology designers the opportunity to concentrate on pedagogy, human interface design, and server-side integration of third-party services without being constrained by content generation. In other words, designing learning technologies under conditions wherein multiple types of content are provided externally allows designers to increase the range and quantity of multimodal content in support of learning.

CHINESE CHARACTER HELPER: A DESIGN INSTANTIATION

Despite the wide use and availability of multimodal print and new media supports for Chinese character acquisition, few applications exist that allow students to leverage their spoken language knowledge for character production. This lack of applications that support students in making connections between spoken words and written characters, the prevalence of pinyin character input methods, and the ubiquity of digital devices creates what we see as a design opportunity. Our design, described in this section, supports students in moving from rote memorization toward more authentic contextualized character production and character-sound couplings via a multimodal Chinese writing application.

Chinese Character Helper (Hellwege, Olmanson, & Liu, 2017) has been iteratively designed over the past
year to be a mashup of open-source Chinese input method technologies, character to speech technologies, image search Application Programming Interfaces [APIs], and web-based word processing technologies.

Our design affords learners an opportunity to leverage their Chinese listening and speaking knowledge in the production of texts. By extending the current input method native Chinese speakers use when writing in Chinese we have intended to leverage a decades-old design while adding supports for beginning, intermediate, and advanced Chinese learners.

In the application, students begin by typing the pinyin version of the word or concept they wish to express. Just like the input method used by native speakers, Chinese Character Helper returns a series of options based on how many characters use the same pinyin. For example typing “ma” into the input field returns at least five characters, one each for mà: 媽 (mother), má: 麻 (to bother), mà: 马 (horse), mà: 麻 (to scold), and ma: 麻 (an interrogative particle). As shown in Figure 2 below, “wo” written in the application returns multiple words, many with different tones but some with the same tone—as some words in Chinese are homophones, sharing identical pronunciation but with different meanings and characters.

For advanced Chinese learners and native speakers who already know the characters, scanning the character options is all that is required to identify the intended one created by the input method. For Chinese learners who are not yet able to visually differentiate and identify the correct character, the standard input method described above and shown in Figure 1 is only helpful if the learner means to write hamburger or love or one of the other handful of emoji-linked characters. Our design extends the typical input method in three ways and via two different modalities shown below. First, two seconds after displaying the possible characters, users can access audio pronunciations of each character via a character to speech API by tabbing to it or mouse hovering over it (Figure 3). The use of this type of aural support connects with our interest in enabling the learner to make use of what they know in building new knowledge and understandings. In this way users can anchor new understandings about characters in their developing ability to differentiate tones and identify when two or more character options are homophones. If the learner is able to discern their intended character via this support, they can click the character and enter it into their text and move on to the next word.

On the other hand, if the aural pronunciations prove insufficient in leading the user to make a selection, learners are presented with columns of images intended to be representative of each of the character options (Figure, 4). As with the delay in offering audio support, there is also a three second delay in showing images that correspond with each character. Images are pulled dynamically from Flickr.com, a photo sharing application with over 122 million members and an estimated 10 billion images (Smith, 2015). Flickr offers an API which allows applications to request images based on a number of criteria including keywords, titles, tags, and descriptions. Their API allows for search via text strings made up of Chinese characters. After a waiting period of three additional seconds, the application sends a request for the most relevant 4-10 images that include the characters returned as potential matches. The results are displayed below each character (Figure 4). While we continue to experiment with algorithms to increase the relevance and accuracy of the images displayed, we made the design decision to pull dynamically from Flickr instead of creating a database of images for each character due to gains in speed and scalability—as well as a more modest codebase—outweighing the benefits of offering curated images for a limited number of characters. By including multiple images we aim to offer at least one image that the user can identify as related to their intended word. In other words, if the user wants to insert the character for “strong” into their text they will look through the options first listening for the correct tone pronunciation (zhuàng)
and then for images that most likely represent strength. In this way the image columns do not need to be so precise as to allow the learner to guess the meanings of all the characters but rather offer enough to select images that match the character they need (壮). Flickr allows for images to be returned based on relevance and a “safe” and “moderate” rating signifying they are appropriate for all and most audiences respectively, we use combinations of these designations in our searches.

![Figure 4](image)

**Figure 4.** After a three second delay students are supported via several images corresponding to the meaning of the word or idea embodied in the character.

We draw on several scholars and constructs in making the design determination of when to include support, and the sequencing of that support, in a way that intentionally includes delays in support levels beginning with no support, then offering character pronunciations, and then images that correspond to each character. By building in delays of three or more seconds, research on wait time suggests that this affords students the opportunity to pause, notice, and think. In the case of Chinese Character Helper the chance to pause, to notice, and think about the characters, their pronunciations, and potential graphic equivalents (Nunan, 2000; Tobin, 1987).

Additionally, these supports enable students to accomplish tasks that otherwise would be impossible, tedious, and/or dissimilar to the authentic practices of early 21st Century writing. Thus, this design rationale fits within both Vygotsky’s construct of the Zone of Proximal Development [ZPD] as well as Wood, Bruner, and Ross’s notion of scaffolding (Vygotsky, 1978; Wood, Bruner, & Ross, 1976). While ZPD establishes the potential for support to expand student learning, the construct of scaffolding, specifically adaptive scaffolding, focuses on how, when, and for how long help might be offered within learning environments (Azevedo & Hadwin, 2005; Wood & Wood, 1996). Chinese Character Helper’s scaffolds appear after time for user contemplation. Moreover, they appear in progressively more explicit forms—adapting to the user’s need for guidance in a way that fades support for characters learners can identify by sight or pronunciation.

Finally, providing opportunities for meaningful and purposeful language output is essential for language learning. Swain states that learners should be provided with opportunities to produce language (1995). However non-Chinese learners often prefer to express themselves in spoken Chinese rather than in written form due to the challenges of learning and producing Chinese characters. Chinese Character Helper aims to make writing accessible to novice-level Chinese learners who are typically not given many opportunities for meaningful written output.

**DISCUSSION AND CONCLUSION**

Chinese characters are part of a writing system that does not operate with one-to-one symbol-sound correspondence. Thus there are no explicit phonetic clues embedded in Chinese characters for learners to follow to fully decode Chinese texts. The presence of tones, homophones, and homographs, the contrast between character and syllabic writing systems add to the complexity of learning to write Chinese. These issues, coupled with the complex nature of stroke sequence with regard to writing Chinese by hand leaves speakers of syllabic languages faced with a daunting challenge, one that proves insurmountable for some learners.

While multimodal supports have been used in language teaching and learning for centuries, recent developments have emerged that make technology-supported multimodal solutions possible in ways that scale to offer support for all ability levels. The utilization of audio and images in relation to character form and pronunciation is meant to facilitate binding—a way of connecting meaning to symbolic forms (Terrell, 1986). Additionally, offering this meaningful input via aural and visual modalities supports students in associating Chinese characters with their pronunciations and semantic values in graphic form instead of associating characters with the equivalent words in their first
language (Shrum & Glisan, 2016). One form of association is noticing. Although noticing does not guarantee language learning, it is the starting point to acquiring language knowledge since learners are not able to acquire language unless they are able to consciously or unconsciously notice them (Schmidt, 1990; Venkatagiri & Levis, 2007). Our application provides opportunities and wait time for learners to notice features of the written language.

In the design of this application we seek to encourage meaningful, authentic, contextual interaction and expression with characters. We have worked to create a design that uses student understanding of spoken Chinese as a gateway to writing. We aligned our supports with the strategies of non-native Chinese learners: associating sounds with characters, paying attention to character configuration, and frequent exposure to learned characters (Zhao & Jiang, 2002). We have sought to offer scaffolds that address some of the struggles non-Chinese learners face when learning to write, namely recalling how to write characters, recognizing characters, building an awareness of the features of characters, and making connections between characters and their pronunciation (Hu, 2010; Lu, 2014).

The next steps in this line of research include a) completing alpha testing with Chinese language students and teachers, b) integrating the application into several beginning and intermediate level Chinese language courses, and c) investigating the ways in which students and teachers use the design in support of Chinese language development.

REFERENCES


