
SINO-PLATONIC PAPERS

Number 354

August, 2024

Handling Chinese Characters on Computers: Three Recent Studies

by
J. Marshall Unger

Victor H. Mair, Editor
Sino-Platonic Papers
Department of East Asian Languages and Civilizations
University of Pennsylvania
Philadelphia, PA 19104-6305 USA
vmair@sas.upenn.edu
www.sino-platonic.org

SINO-PLATONIC PAPERS

FOUNDED 1986

Editor-in-Chief

VICTOR H. MAIR

Associate Editors

PAULA ROBERTS

MARK SWOFFORD

ISSN

2157-9679 (print) 2157-9687 (online)

SINO-PLATONIC PAPERS is an occasional series dedicated to making available to specialists and the interested public the results of research that, because of its unconventional or controversial nature, might otherwise go unpublished. The editor-in-chief actively encourages younger, not yet well established scholars and independent authors to submit manuscripts for consideration.

Contributions in any of the major scholarly languages of the world, including romanized modern standard Mandarin and Japanese, are acceptable. In special circumstances, papers written in one of the Sinitic topolects (*fangyan*) may be considered for publication.

Although the chief focus of *Sino-Platonic Papers* is on the intercultural relations of China with other peoples, challenging and creative studies on a wide variety of philological subjects will be entertained. This series is *not* the place for safe, sober, and stodgy presentations. *Sino-Platonic Papers* prefers lively work that, while taking reasonable risks to advance the field, capitalizes on brilliant new insights into the development of civilization.

Submissions are regularly sent out for peer review, and extensive editorial suggestions for revision may be offered.

Sino-Platonic Papers emphasizes substance over form. We do, however, strongly recommend that prospective authors consult our style guidelines at www.sino-platonic.org/stylesheet.doc.

Manuscripts should be submitted as electronic files in Microsoft Word format. You may wish to use our sample document template, available here: www.sino-platonic.org/spp.dot.

All issues of *Sino-Platonic Papers* are free in PDF form. Issues 1–170, however, will continue to be available in paper copies until our stock runs out.

Please note: When the editor goes on an expedition or research trip, all operations may cease for up to three months at a time.

Sino-Platonic Papers is licensed under the Creative Commons Attribution-NonCommercial-NoDerivs 2.5 License. To view a copy of this license, visit <http://creativecommons.org/licenses/by-nc-nd/2.5/> or send a letter to Creative Commons, 543 Howard Street, 5th Floor, San Francisco, California, 94105, USA.

Handling Chinese Characters on Computers: Three Recent Studies

J. Marshall Unger

The Ohio State University

ABSTRACT

Writing systems with large character sets pose significant technological challenges, and not all researchers focus on the same aspects of those challenges or of the various attempts that have been made to meet them. A comparative reading of three recent books—*The Chinese Computer* by Thomas Mullaney (2024), *Kingdom of Characters* by Jing Tsu (2022), and *Codes of Modernity* by Uluğ Kuzuoğlu (2023)—makes this abundantly clear.¹ All deal with the ways in which influential users of Chinese characters have responded to the demands of modern technology, but differ from one another considerably in scope and their selection and treatment of relevant information long known to linguists and historians.

Keywords: touch-typing, computerization, Sinitic languages, politics of script reforms, national language

¹ A fourth, earlier book in this genre, Zhong 2019, has already been critically discussed in an excellent review by Shuheng Zhang (2020).

I.

THOMAS S. MULLANEY. *THE CHINESE COMPUTER: A GLOBAL HISTORY OF THE INFORMATION AGE*. CAMBRIDGE, MA & LONDON: MIT PRESS, 2024.

This book, a sequel to Mullaney 2017, tells the story of how the handling of Chinese characters developed from typewriters to computers. Though well-informed by archival research, it is written in the style of a *Time* magazine story, with colorful biographical digressions and mostly decorative details calculated to make readers feel they are being let in on “insider” information. As a work of popular history, it largely succeeds, but Mullaney aims for more, and, in so doing, spoils the book.

The word “global” in the subtitle is not explained until one reaches the end of the work and has learned two concepts that Mullaney develops along the way. One is IMMEDIACY, by which he means the one-to-one relationship of keys and letters like those on an English typewriter keyboard. Typewriting devices that contained hundreds of “keys” or cells, each corresponding to a specific character, were built and used in China and Japan in the twentieth century, but current computer-based input is generally MEDIATED in that one uses a keyboard to type short codes that identify the characters one wishes the computer to register and display or print. The second concept is HYPOGRAPHY, by which Mullaney means the sequence of codes one types to produce a text, as distinct from their visible output.

Given these related concepts, the meaning of Mullaney’s rousing summation is clear: we are, he says, in “the first fifty years of an epochal change in text technology, akin to the opening half-century of the Gutenberg era” (221).

Collectively, Asia, Africa, and the Middle East are now home to some of the most vibrant and lucrative IT markets in the world, with Euro-American firms clamoring to make inroads. For those who cling to the myth of immediacy, such vistas might seem to confirm for them what they already knew to be true: that the magnificence of Western engineering and innovation once again bestowed its gifts upon the “rest” of the world.

But this is a falsehood. It was not the Western-designed computer that saved

China and the non-Western world. *It was China and the non-Western world that saved the Western-designed computer*—saved it, that is, from its foundational limitations, both conceptual and material. Without Input Method Editors, contextual shaping, dynamic ligatures, rendering engines, layout engines, adaptive memory, contextual analysis, autocompletion, predictive text, the “modding” of the BIOS; the hacking of printer drivers, “Chinese-on-a-chip,” and, above all, an embrace of hypography, no Western-built computer could have achieved a meaningful presence in the world beyond the Americas and Europe. Today, hypography is the global norm. Hypography made global computing possible.

What becomes of our understanding of computing and new media, then, when the number of languages requiring “auxiliary” programs outnumber those for which computers work the way they “should”? When hypographic technologies become more widespread than supposedly “normal” orthographic ones, how does our understanding of electronic writing change? (229, emphasis added)

This is the sort of exaggeration one would expect in a puff piece about the Belt and Road Initiative. Mullaney would have done better leaving it out.

Consider first the list Mullaney rattles off after his mention of IMEs: it lumps together things that ought to be kept distinct. “Contextual shaping, dynamic ligatures, rendering engines, [and] layout engines” refer to software that handles graphic output in scripts that use allographs of various kinds (e.g., Arabic abjad scripts, Indic abugida scripts, Korean hankul). Such software has little to do with input per se. By contrast, “adaptive memory, contextual analysis, autocompletion, [and] predictive text” are related to input but by no means exclusively to Chinese input. “The ‘modding’ of the BIOS[,] the hacking of printer drivers, and ‘Chinese-on-a-chip’” are firmware or hardware improvements. They may perhaps have been responses to a perceived need to handle Chinese characters in a computer environment, but, if so, that was a need that existed in other places (e.g., Japan) as well as in China. In fact, there is no evidence that any of the listed items would not have been developed but for Chinese IMEs. As for hypography, as explained below, it doesn’t belong in the list.

It shows up there because Mullaney fails to distinguish between the two sides of the interface

where the user and the computer meet. The computer comes equipped with a keyboard or other device that allows the user to produce a hypographic text and a program that selects characters either autonomously or through a dialog with the user; the user must know the rules for the kind of hypography the machine is programmed to accept and have the skills needed to operate the hardware interface speedily and accurately. Mullaney hardly takes notice of the human side of this dyadic relationship. Computer users must not only be proficient and literate in the language they are inputting but must also have acquired sufficient skill in touch-typing to enable them to inscribe and edit text more easily than they could with pen and paper. In the case of Chinese, which is really a family of related languages, this means that those who do not grow up speaking the now official form of Mandarin (*hanyu*) need to learn it as a second language if they want to use *pinyin*-based input software. Mullaney acknowledges this (213), but only obliquely engages the question of how much additional training and practice is needed to master a hypographic input system.

He also fails to pay adequate attention to the different purposes for which people use computers. What is necessary or sufficient for one task may be neither for another. Hence Mullaney's repeated mentions of high-speed hypographic input, such as the following, ring hollow:

If Huang Zhenyu's mastery of a complex alphanumeric code weren't impressive enough, consider the staggering speed of his performance. He transcribed the first 31 Chinese characters of Hu Jintao's speech in roughly 5 seconds, for an extrapolated speed of 372 Chinese characters per minute. By the close of the grueling 20-minute contest, one extending over thousands of characters, he crossed the finish line with an almost unbelievable speed of 221.9 characters per minute.

That's 3.7 Chinese characters *every second*.

In the context of English, Huang's opening 5 seconds would have been the equivalent of around 375 English words-per-minute, with his overall competition speed easily surpassing 200 WPM—a blistering pace unmatched by anyone in the Anglophone world (using QWERTY, at least). (8, emphasis in original)

Since it takes on average about 1.7 characters to write the average high-frequency Chinese word, 372

characters comes out to about 219 words; in fact, 186 Chinese words may be closer to the truth since bloviating speeches like Hu's are freighted with polysyllabic jargon. More importantly, I can use my touch-typing skills to write and edit an article like this review "on the fly." Does Huang's prowess enable him to do extemporaneous writing and editing at "blistering" speed?² There are no doubt thousands of speakers of Vietnamese, Turkish, Russian, Korean, Japanese, and many other languages quite different from English who make good use of touch-typing skills very similar to mine. I doubt the same can be said for Huang's. In the modern world, copy-typing is an important computer application, but it is just one among many.³

When it comes to the ergonomics of touch-typing in relation to the problem of inputting Chinese characters, some of the best research was carried out by Yamada Hisao (1930–2008) and his colleagues at the University of Tōkyō in the 1970s and 1980s. Mullaney does not mention Yamada's work, though he does cite various other Japanese sources, and I think this is one reason he wastes time disparaging the QWERTY keyboard, which he sees as an instrument of "Anglophone" cultural imperialism. It has long been common knowledge that the QWERTY layout—invented in the United States, but with variants for French, German, and many other languages—is not particularly well-designed for touch typing (it was meant to slow down typists so that typebars in mechanical typewriters wouldn't jam). Nevertheless generations of ordinary speakers of all these languages have learned how to touch-type with fair speed and accuracy after about 1,000 hours of training and practice on average. Had Mullaney consulted Yamada's work, he would have known that learning to use a keyboard by touch becomes difficult in direct proportion to the number of keys in the layout. The ANSI standard keyboard

² One is reminded of the acid comment about a speed-demon pianist: "He can play faster than a player piano, but with less expression."

³ Many computer applications could be done successfully in the PRC without the use of characters at all. John DeFrancis, whom Mullaney mentions only *en passant*, advocated for digraphia: the use of whichever writing system is sufficient for a particular task. Most Chinese and Japanese computer users type *pinyin* or *rōmaji* to initiate input anyway: Why not store data in that format? Is character output always needed? (One reason for the problems of Japan's national pension system, which began to come to light in 2007, was that, when information was transferred from traditional paper records to computer-readable form, the readings of *kanji* for names were not retained, making it hard for thousands of paid-up claimants to prove their identities after their names had been input into the system.)

commonly used in the United States, for instance, has 47 keys (26 letters, 10 numerals, and 11 keys for punctuation), not counting function keys such as TAB, SHIFT, BACKSPACE, the space bar, and so on. Beyond 47, accuracy and speed decline: even reasonably good touch-typists, whose fingers hover over the “home row” keys (ASDF and JKL;) may find it hard to be accurate when reaching for keys in the top row or rightward, beyond the tenth diagonal column. This is true no matter how the keys happen to be labelled.

Huang’s “blistering” speed record was achieved using an *wubi* input system, in which keys are mnemonically associated with the shapes of distinctive graphic components of characters. Such a system (Mullaney discusses many in Chapter 4) is a prime example of hypography because “what you type” is not “what you get”: one relies on a program (an “input method editor” in Microsoft jargon) that outputs the desired character when prompted with as few as two keypresses that identify it uniquely, which depends on the structure of the program and the look-up tables it uses. It is up to the typist to know the keypress codes in advance. For that reason, one might think that coding systems based on mnemonic relationships between keypresses and the shapes or readings of individual characters would facilitate recall and would therefore be necessary for rapid input. But Yamada and his colleagues discovered that this intuition is mistaken. They found that a completely non-mnemonic code using two or three keys for each character in a large target inventory produces the fastest touch-typing systems (Unger 1987: 159–63). In the case of Japanese, these characters include *kana*, but that is just as unimportant as the QWERTY (or any other) labels that may appear on a keyboard. Input speed and accuracy are achieved by optimizing the keystroke sequences for characters based on their frequencies in typical texts and the relative ease of finger and hand movements involved in making the keystroke sequences during typing; moreover, mnemonics do not make it easier to learn the system. Yamada and his colleagues demonstrated that *for copy-typing* Japanese subjects could be trained to use their “patternless” keyboard system in about as much time as it takes to train an American subject in English touch-typing.

Instead of Yamada, Mullaney focuses on Chung-Chin Kao, who, in the 1940s, proposed using the old four-digit Chinese telegraphic codebook as the basis of an electric Chinese typewriter. Kao was impressed by the skills of telegraph clerks, tasked with translating from characters to code numbers and vice versa. Yamada’s research corroborated Kao’s empirical observation that a large, arbitrary code

could be learned and used, but, at the same time, highlighted three issues that Kao did not investigate scientifically: the advantages of a non-mnemonic encoding procedure for keyboarding; the time needed to commit an optimized input system to muscle-memory; and the differences among the tasks to which the system might be applied. Kao's codebook happened to qualify as non-mnemonic. As for training time, Mullaney quotes correspondence of 1944 between Kao and Chauncey Griffith of the Mergenthaler Linotype company estimating that an operator would need two months of training to memorize the four-digit codes of 1,000 characters, and another month to become proficient in using them (43). Assuming a rigorous 40-hour work-week, three months comes out to about 520 hours, not a bad guess given Yamada's later finding that training to use a two-keystroke code on a 48-key layout for 2,304 characters takes approximately twice as long.⁴ But this still leaves the crucial question of whether computer users can or should invest the time and effort needed to learn to touch-type hypographic codes.

In Chapter 6, Mullaney describes improvements in IMEs that use mnemonics based on the readings of words and phrases rather than on the graphic structure of individual characters, noting that "[f]rom the 1990s onward, pinyin Input Method Editors have eclipsed structure-based approaches" (212). But, he concludes, "Insofar as structure-based systems are far more effective at the unambiguous entry of individual characters[,] and because structure-based systems are also able to take advantage of anticipatory technologies such as predictive text[,] the case is very strong that, on the whole, pinyin input represents a step *backward* in terms of the overall speed of Chinese input" (212, emphasis in original). Instead of speculating in this way, Mullaney would have done better to discuss Yamada's non-mnemonic optimization of hypographic codes for touch-typing.

Mullaney immediately goes on to admit that speed isn't everything: "the development of viable pinyin input methods has led to an unprecedented massification of hypography, with far greater numbers of Chinese speakers able to draw upon their knowledge of Hanyu pinyin to employ one or another pinyin input system on their computers or mobile devices" (213). It is at this point that he admits that *pinyin*-based systems put non-native speakers of the standard language at a disadvantage.

⁴ Kao actually proposed using an inventory of 5,278 characters, but most of these occur with very low frequency in typical Chinese texts.

Nevertheless, any hypographic system is fine with him: he has persuaded himself that Chinese-style hypograhya has not only enabled the Sinosphere to guarantee the use of Chinese characters in the future but will do the same for all the world's language communities that use abjads, abugidas, and other non-Western writing systems—QWERTY be damned!

It is therefore bewildering that Mullaney opens his book with a rant against the assertion that Chinese tend to forget how to write characters by hand when they rely heavily on IMEs for their daily writing tasks. As reported in Unger 1987 (to which Mullaney does not refer), such forgetfulness was experienced even by avid proponents of Japanese word-processing systems, such as Kabashima Tadao, long ago. Kabashima cheerfully assured me, when I interviewed him in 1985, that he was not troubled by this discovery. He foresaw a future in which Japanese would be written so ubiquitously by electronic means that one would not need to know how to write characters by hand for everyday purposes. Nothing would matter but the ability to recognize characters at sight and know their readings.

It seems to me that Kabashima, and many later commentators on the computerization of Chinese society, are more honest than Mullaney. He fails repeatedly to draw necessary distinctions, in this case, among the different reasons that people write. Some authors prefer handwriting because, they say, it slows them down and affords them time to think; students of calligraphy obviously must learn how to write by hand. Will writing by hand be confined to such artistic endeavors in the future? Can students be taught to read Chinese and Japanese without learning how to write characters by hand?

These are questions that Mullaney does not address. Speaking for myself, I find it difficult to be optimistic about hypography. As a great historian said of prewar Japanese writing, which made copious use of *furigana* to inform the reader of the writer's intended reading of almost every *kanji*, "One hesitates for an epithet to describe a system of writing which is so complex that it needs the aid of another system to explain it. There is no doubt that it provides for some a fascinating field of study, but as a practical instrument it is surely without inferiors" (Sansom 1928: 44).

II.

JING TSU. *KINGDOM OF CHARACTERS: THE LANGUAGE REVOLUTION THAT MADE MODERN CHINA*. NEW YORK: RIVERHEAD BOOKS, 2022.

Kingdom of Characters became a finalist for the Pulitzer Prize in 2022. It is not hard to understand why. Tsu, even more blatantly than Mullaney, cultivates the *Time* magazine style. She avoids all but the most familiar technical terms used in linguistics and computer science, relying instead on common myths and misconceptions about China and its language (definitively documented in DeFrancis 1984) to lull unwary readers into uncritically accepting her whiggish narrative. This narrative occasionally verges on sinocentric polemic—one even suspects that the choice of “kingdom” rather than “empire” in the title is a kind of nod to CCP political correctness. Though Mullaney occasionally lapses into Chinese boosterism, he offers a more informative and objective history of how the current state of Chinese-script computing came to be. Tsu tries to cover the same ground, but her book is vaguer about the machinery involved, less reliable as history, and weighed down by dubious statements about language and writing in China. For her, input and output, hardware and software, speech and writing are muddled together; she sees nothing wrong in calling Chinese characters ideographs, and she plainly doesn’t know much about the actual history of the Chinese language or the psycholinguistics of reading.

Consider, for instance, this cheerful expression of the same thing Kabashima predicted in the 1980s:

One scarcely has to learn Chinese characters the hard way—by memory or by hand—any more, it seems. After a few decades of consolidation and standardization, several major options for Chinese input methods on a computer or mobile keyboard cover just about everyone’s needs. (250)

Evidently Tsu has forgotten that learning to use any input method presupposes knowledge of which characters represent which syllables in a very large set of words. In the case of Chinese, these are words of the modern standard language (*hanyu*). Learning how they are prescriptively written doesn’t come

for free during first-language acquisition (or second-language learning for those who grow up in families that use a different Sinitic language), because language is fundamentally speech. Even children blind from birth acquire it, though they obviously do not learn characters or their uses in the process.

Not only is Tsu's summary of the current state of affairs a kind of fantasy, but her book is full of similarly fantastic statements, e.g., "tones evolved to help differentiate between characters that have the same syllable and are pronounced the same way" (xviii). That Tsu refers to *characters* and not *words*, yet, in the same sentence, also to *syllables* and *pronunciation*, suggests that, in her mind, "characters" and "words" are virtually synonymous, which they most certainly are not. In addition, her statement is historically wrong. To say that two characters "have the same syllable and are [therefore] pronounced the same way" only makes sense if one is looking in a single Sinitic language (such as *hanyu*) and at syllables that happen to be homophonous apart from tone. As historical linguists have shown, tones in Sinitic languages evolved as the result of regular phonemic changes over which no individual speaker had conscious or purposeful control, changes that unfolded over decades if not centuries—indeed, changes that, besides making certain syllables segmentally homophonous in this or that branch of the Sinitic language family, were also responsible for the process of linguistic differentiation itself. Tsu's glib sentence is not a statement of fact; it is just an invitation to the naïve reader to believe a just-so story completely uncoupled from historical reality.

Tsu's knowledge of sociolinguistics seems similarly limited. In discussing a visit to Hanoi, she comments that Vietnamese *chữ nôm* script "looks like an extended application of square-shaped Chinese characters and does not stray far from the character mold" (267).

The attitude in Vietnam toward the Chinese script's thousand-year legacy has gone through its ups and downs. When Vietnam fell to the French in the 1880s, it has been said that one could hear the gentle sound of Sinoscript writing brushes falling to the ground. It was the end of high culture, nurtured by the once-great Chinese Empire. Yet in 1946, Ho Chi Minh, seeing the Chinese influence in light of anticolonial struggles, spoke of the Chinese legacy in a less poetic light: "The last time the Chinese came, they stayed one thousand years.... I prefer to smell French shit for five years, rather than

Chinese shit for the rest of my life." The politics of the Han characters for other countries in East Asia are still real and present. (ibid.)

For this melodramatic synopsis, Tsu cites Woodside 1982: 104, and one page of the Pentagon Papers.

I rather doubt that Woodside would agree that the arrival of the French marked "the end of high culture" in Vietnam or that its high culture had been sustained solely "by the once-great Chinese Empire." The first statement is clearly hyperbolic, and the second plainly wrong because Vietnam was not under Chinese control after 938 CE. As for the quotation from Ho, Tsu apparently chose to rip it out of context because of its shock value. The text of whole typewritten page on which it appears⁵ clearly shows that it was cited to establish "the anti-Chinese odium among the people of North Vietnam" in 1946 and to explain Ho's activities at that time. Tsu's attempt to contextualize the quotation by saying that Ho was looking at "Chinese influence in light of anticolonial struggles" is misleading in two ways. First, the thousand years to which Ho rhetorically gestured lasted from 111 BCE to 938 CE. His immediate concerns were very different, as the document Tsu cites clearly explains: "In famine-wracked North Vietnam, Chinese hordes under booty-minded warlords descended on the DRV [Democratic Republic of Vietnam], supplanting its legal government with committees of their own sponsoring and systematically looting."

Second, Ho was not commenting on *chữ nôm* and literary Chinese, both of which had been in decline even before the founding of the Tonkin Free School (*Đông Kinh Nghĩa Thục*) in 1907, which incorporated romanized Vietnamese into its pedagogy. As explained in DeFrancis 1977 (which Tsu does not cite),⁶ Ho embraced the romanized spelling system for Vietnamese, devised in the seventeenth century by the French Jesuit de Rhodes, primarily because he understood it was essential for educating and mobilizing the masses to expel the French. This is why literacy in the Latin alphabet became

5 (https://en.wikisource.org/wiki/Page:Pentagon-Papers-Part_I.djvu/170)

6 Tsu concludes a lengthy endnote to Chapter 5 by writing, "As is evident in his scholarship, the late DeFrancis was a staunch supporter of Romanizing Chinese. I am grateful for the conversations I had with him in his home at Uluwehi Place, Honolulu, in 2008, the year before he passed away" (299). This in apparently meant to show that Tsu is open-minded, but as "Uluwehi Place" is just the name of a street in residential Mānoa Valley, it strikes me as pretentious.

widespread in Vietnam within more or less a single generation, while in China and Japan script reform efforts were watered down or blocked.⁷

As already noted, Tsu shows remarkably little interest in the historical reasons that Sinitic languages diverged, mostly from the Middle Chinese of the mid-first millennium, focusing instead on the triumph of Mandarin, which became the *lingua franca* of imperial rule after the collapse of the Yuan dynasty. The large number of people—particularly the well-educated people who could serve the government—who spoke it or closely related dialects made it difficult for the GMD (*Guomindang*) and, later, the CCP (Chinese Communist Party) to choose anything else as the national standard. Yet Tsu tries to persuade the reader (31–38) that the adoption of Mandarin was the result specifically of the heroic efforts of Wang Zhao during a conference held in 1912, soon after the collapse of the Qing dynasty. As for other Sinitic languages, such as Yue, Hakka, and Wu, and the cultures associated with them, Tsu shows no sympathy, which I find particularly ironic given the strongly anticolonialist sentiments she expresses throughout the book. Are CCP efforts to discourage the use of Wu or Yue any less odious than its efforts to compel Uyghurs to speak Mandarin, or, for that matter, the boarding schools for North American “Indian” children where they were punished for speaking their native languages?⁸

If one doesn’t have time for all three books discussed here, Tsu’s is unquestionably the one to skip despite the praise heaped on it by a wide variety of reviewers. Mullaney notably refers to Tsu’s book just once, crediting Tsu with ferreting out the full name of the Chinese man who “slightly adjusted” the inventory of characters in the telegraphic code of Septime Viguiier (1871). In the same footnote, he describes her as “a comparative literature scholar and Winter Olympics commentator” (Mullaney 2024: 254). This would appear to be a well-deserved instance of damning with faint praise.

⁷ Tsu’s uncritical discussion of PRC simplified characters (174–78) as a labor-saving, literacy-boosting development reads like CCP propaganda. One could just as well argue that simplified characters represented a retreat from earlier CCP promises to replace characters with *pinyin*, a betrayal of Lu Xun, and a way to make it easier to identify material not printed in the PRC. It is also curious that Tsu comments (negatively) on attempts to replace the Perso-Arabic scripts of Turkic languages in the Soviet Union with romanizations but doesn’t mention the successful romanization of Turkish proper under Atatürk.

⁸ Tsu does not explicitly invoke the postmodernist authors who disparage “phonocentrism”: she seems simply unaware that the understanding of language as speech emerged from years of empirical work by anthropologists (e.g., Boas) and linguists (e.g., Bloomfield) on languages that had no writing systems (Murray 1994).

III.

ULUĞ KUZUOĞLU. *CODES OF MODERNITY: CHINESE SCRIPTS IN THE GLOBAL INFORMATION AGE*. NEW YORK: COLUMBIA UNIVERSITY PRESS, 2023.

Unlike Mullaney and Tsu, who frequently do not bother to distinguish language from writing, Kuzuoğlu offers an explicit argument for focusing on script and script reform rather than on language. He observes that the adoption of the Latin alphabet in Vietnam, in Turkey, in certain Soviet regions and Mongolia (with a subsequent shift to Cyrillic) as well as proposals for script reform in Japan all occurred during roughly the same period: "There was a global synchronicity in script reforms because they were all precipitated by the metamorphosis of communication practices in the global age of information" (10).⁹ That is, the introduction of the telegraph, Western sciences, and so on disrupted the knowledge economies of these places, created a new need and demand for mass literacy, and made traditional systems of writing look "backward" and inefficient.

Scripts comprised the primary arena of contestation across the world from the late nineteenth century onward because they were the "codes," not metaphorically but literally, for engineering a new society in the operational, technological, and ideological sense of the term." (17)

For this reason, Kuzuoğlu claims that writing systems are political constructs, and that

⁹ "Synchronicity" is a bit strong. Kuzuoğlu mentions the first four decades of the twentieth century, and if Japan is included, one would have to go back to the Nippon-shiki romanization scheme devised by Tanakadate Aikitsu—and much changed technologically in the half century from 1885 to 1935.

Kuzuoğlu also mentions Albania, where at least ten alphabets were used in different regions and periods. Given Albania's small size, the large number of Albanian speakers living across its borders when it gained independence (Treaty of London, 1913), its rocky history after World War I, and its low per capita GDP throughout the twentieth century, one suspects that the precedent of Turkish romanization played a larger role than capitalist development of new information technologies in leading Albanians to settle upon a modified Latin alphabet.

[1]language politics, counterintuitive as it may sound, followed script reforms.... Language politics were obviously important in the fabrication of a national identity, but it was at the material level of the script that battles over the politics of linguistic representation were waged. (58)

He goes on to compare proposals for script reform to bridges over the Long Island Parkway (58–59), which Robert Moses allegedly designed to keep buses (and hence the lower classes) out of Jones Beach (Winner 1980). (Kuzuoğlu knows that there are problems with this analogy (250n3), but presents it anyway.)

I imagine Kuzuoğlu takes this unnecessarily extreme position to show that he is suitably postmodernist. The design of artifacts, including information-handling technologies, may, of course, reflect political biases, but not all their intensional structures or extensional relationships to speech are politically determined. It is a gross exaggeration to say

the Chinese writing system thrived precisely because it could not be reduced to the representation of any spoken language. As a writing system, it was able to accommodate linguistic differences, hence its adoption across Japan, Korea, and Vietnam. Within the longer history of writing, Chinese was neither an exception nor an anomaly. In fact, if there is an anomaly to be addressed, it would be historical emergence of the alphabet itself... [N]one of the earliest writing systems in Mesopotamia, Mesoamerica, the Nile River and the Yellow River were invented to record speech. (11–12)

If Kuzuoğlu had consulted DeFrancis 1989, he would have known that countless instances of pictographic code-like partial writing (visible tokens) have languished; in the few loci where such partial writing became full writing (capable of transcribing any utterance), it was because someone had the bright idea of using symbols to represent the names of things rather than the things themselves (what DeFrancis calls the rebus principle), and, by extension, the language sounds in those names. It was independent implementations of this idea that turned the loci Kuzuoğlu lists into cradles of full

writing, even though, in each, some characters continued to be deployed logographically. In this regard, it was Chinese writing that was exceptional: almost all its phonographicity came to be embedded in the graphic structure of the ever-growing number of its phonetic/signific characters. What is anomalous in the larger picture is that Chinese writing continued to add hundreds of phonetic/signific characters to its inventory, not that abjads, abugidas, well-tempered syllabaries, or alphabets developed elsewhere.¹⁰ Moreover, the claim that mostly logographic uses of Chinese characters facilitated their adoption "across a diverse linguistic landscape in East and Inner Asia" (12) is doubtful. *Pace* Handel (2019), whom Kuzuoğlu cites, Tibetans, Mongols, and Manchus did not drop their writing systems, and Japan, Korea, and Vietnam, which had no previous contact with a literate society before making contact with China, each developed graphic innovations that lessened the difficulty of using Chinese script to write their respective languages (Unger 1990: 396–97; Unger 2021: 138–44).

Despite these weaknesses in Kuzuoğlu's thesis, his forthrightness is commendable and his premise has the merit of liberating him from Mullaney's and Tsu's sinocentrism and preoccupation with the latest technology, allowing him to focus on the role of politics, to which they pay relatively little attention in their surveys of roughly the same historical periods. Indeed, Kuzuoğlu implicitly criticizes both Mullaney and Tsu near the end of his Introduction:

Is it meaningful to dismiss alphabetization as an outcome of Orientalism and argue for a celebratory history that solely relies on the technologization of Chinese characters?
At what point does celebration morph into fetishization?

.....

I find it difficult to reduce modern Chinese information history to a technical history of adapting Chinese characters to Western technologies like the typewriter, not only because it disregards the hybridity of the information regime in China in which characters and Pinyin eventually came to coexist, but also because the ideological issues that surrounded scripts in a modernizing knowledge economy went well beyond

¹⁰ Indeed, if writing systems *are* political constructs, one wonders what political circumstances of ancient China favored the development of a writing system that relegated phonographic mnemonics to a secondary role compared with the role they came to play in Egypt and the Levant.

technical tinkering. In the twentieth century, there were many different scripts in China for Chinese languages. Some of them were native inventions, like the National Alphabet or simplified characters, and others were imported from the outside world, such as the Chinese Latin Alphabet. But regardless of the differences, they were all central to ideologically driven programs that were forged through transnational networks in the global information age. (19)

Chapter 3 is a prime example of how Kuzuoğlu's approach differs. He concentrates on the 1936 proposal of Hong Shen for a "Basic Chinese" inspired by the "Basic English" of C. K. Ogden and I. A. Richards. Neither Mullaney nor Tsu discuss Hong, whose plan for developing mass literacy was to limit the number of characters in general use to 1,100.¹¹ In other chapters, Kuzuoğlu introduces a host of other interesting people and their ideas, all of which Mullaney and Tsu skip over in their celebratory histories.

Perhaps the clearest differences between the three studies are seen in their respective treatments of the latinization (*ladinghua*) movement of the 1930s, the Chinese-speaking Dungan minority in the Soviet Union, and the ascendancy of *pinyin* in the 1950s. In Mullaney's book, Zhou Youguang, often called the father of *pinyin*, shows up only in the references and some footnotes; Qu Qiubai does not appear at all, and the Dungan are not mentioned. Tsu takes note of all three, but her discussion of the Dungan script (in which she mentions Qu *en passant*) is not nearly as rich as Kuzuoğlu's, and her account of Zhou Youguang fails to note the irony of the ascendancy of *pinyin*. Kuzuoğlu stresses that latinization did not refer to a specific system of romanization for a single Sinitic language but rather to finding ways to use the Latin alphabet for the representation of all the languages of China. His assessment of *pinyin* is therefore quite different from Mullaney's or Tsu's. Because "latinization's power had stemmed from its commitment to linguistic pluralism," Kuzuoğlu concludes that

¹¹ As Kuzuoğlu by no means ignores Japan, why does he not note that post-surrender Japan, under pressure from both outside and within to do something to increase literacy, chose to limit the number of characters for general use to 1,850 (the *tōyō kanji*) in November 1946?

Pinyin was in fact a betrayal ... it stood ideologically closer to the GMD's Phonetic Symbols [often called *bopomofo*] or Yuen Ren Chao's Gwoyeu Romatzyh than the CLA [Chinese Latin Alphabet]. Mandarin was finally the sovereign language of the nation enshrined in the letters of Pinyin. (229)

Moreover, in 1958, the CCP decided that *pinyin* would only be an alternative to characters, not a replacement for them (thus betraying the hope of the revered writer Lu Xun, who predicted that either characters would be abolished or China would perish). At the same time, the CCP decided to push ahead with "pinyinizing" the scripts of non-Mandarin and non-Sinitic languages under its control. This "frontier pinyinization," as Kuzuoğlu calls it, was not unlike Stalin's efforts to cyrillicize the scripts of non-Russian languages in the USSR and make Russian the language of the realm.

Precisely because Kuzuoğlu avoids a narrow focus on computer technology and looks beyond the PRC, his book is more rewarding than either Mullaney's or Tsu's. Theirs seem to gloat over the way the current state of technology makes Lu Xun's dire prediction seem vaguely ridiculous; Kuzuoğlu, who mentions Lu Xun in just a single footnote, is more cool-headed. It may be, as Kuzuoğlu says, that language politics generally follow script politics, but one is left wondering how long the PRC can put off a reckoning with the politics of language differences, computer technology notwithstanding.

REFERENCES

- DeFrancis, John. 1977. *Colonialism and Language Policy in Viet Nam*. The Hague: Mouton.
- DeFrancis, John. 1984. *The Chinese Language: Fact and Fantasy*. Honolulu: University of Hawai'i Press.
- DeFrancis, John. 1989. *Visible Speech: The Diverse Oneness of Writing Systems*. Honolulu: University of Hawai'i Press.
- Handel, Zev. 2019. *Sinography: The Borrowing and Adaptation of the Chinese Script*. Leiden: Brill.
- Mullaney, Thomas S. 2017. *The Chinese Typewriter: A History*. Cambridge, MA: MIT Press.
- Murray, Stephen O. 1994. *Theory Groups and the Study of Language in North America. A Social History*. Amsterdam Studies in the Theory and History of Linguistic Science, series III. Studies in the History of the Language Sciences, volume 69. Amsterdam/Philadelphia: John Benjamins.
- Sansom, [Sir] George Bailey. 1928. *An Historical Grammar of Japanese*. Oxford: Clarendon Press.
- Unger, J. Marshall. 1987. *The Fifth Generation Fallacy: Why Japan Is Betting its Future on Artificial Intelligence*. New York: Oxford University Press.
- Unger, J. Marshall. 1990. "The Very Idea: The Notion of Ideogram in China and Japan." *Monumenta Nipponica* 45:4. 391–411.
- Unger, J. Marshall. 2021. "The Cambridge Spelling Study Hoax and Writing System Typology." *Scripta* 12: 129–47.
- Winner, Langdon. 1980. "Do Artifacts Have Politics?" *Daedalus* 109:1. 121–36.
- Woodside, Alexander. 1982. "Conceptions of Change and Human Responsibility for Change in Late Traditional Vietnam." In D. K. Wyatt and A. B. Woodside, eds., *Moral Order and the Question of Change: Essays on Southeast Asian Thought*. New Haven: Yale University Press.
- Zhang, Shuheng (Diana). 2020. *Review of Zhong 2019. Modern Chinese Literature and Culture*. Accessible at <https://u.osu.edu/mclc/book-reviews/shuheng-zhang/>.
- Zhong, Yurou. 2019. *Chinese Grammatology: Script Revolution and Literary Modernity, 1916–1958*. New York: Columbia University Press.

All issues of *Sino-Platonic Papers* are accessible to readers at no charge via our website.

To see the complete catalog of *Sino-Platonic Papers*, visit

www.sino-platonic.org