Despite the enormous amount that has been written about China in the few short decades since it re-emerged as a major global power, few non-Chinese are even aware of the country’s long struggle to forge a unified national language, something taken for granted in Western and other nations. Samuel S. Kim reaches deep into Chinese history to trace the country’s long march toward linguistic unification, a march that is still under way.

THE WEIGHT of China’s well-chronicled past has both fascinated and frustrated Western observers concerned with what holds Chinese society together and what it means to be Chinese in a changing world. On the one hand, this rich historical legacy seemed to have provided a combination of pragmatic adaptability and secure identity as Chinese (Zhōngguórén, 中国人). On the other, this proud country has had enormous trouble in its abiding quest for national identity as a rich and powerful state (Fùguó qiángbīng, 富国强兵) as well as for a common national language during what Beijing calls the “modern period” (since the end of the First Opium War in 1841).

To address the why of the relentless quest for linguistic unification, the first section below describes the defining and differentiating features of Chinese from the vantage point of English speakers. The second section tracks the complex and evolving strategies of language reform in the modern period, especially since the turn of the 20th century. The third and final section provides an assessment of progress and future prospects.

1) THE LANGUAGE’S DEFINING AND DIFFERENTIATING FEATURES
The single greatest defining feature of Chinese is its characters: instead of a few dozen Latin alphabetic letters, Chinese has developed thousands of characters that represent morphemes and words. This contrasts with closely related writing systems such as Japanese and Korean, which, while sharing many of the same characters, are able to fully function as purely phonetic scripts. Chinese characters have the longest history of continuous use and are the only ideographs that are still in daily use today.

There is no consensus on the exact number of Chinese characters, since new characters for new brand names or new technological products are being added continuously.1 The dictionary I have been using since 1979 — Hàn-Yīng Cídiǎn, 汉英词典 (The Chinese-English Dictionary) — contains more than 6,000 single-character entries, 50,000 compound-character entries, and 70,000 compounds, set phrases and examples, including a great number of common classical Chinese words, dialect words, four-character idi-
Character Forming

The strokes of Chinese characters have eight main categories:

- horizontal:
- vertical:
- left-falling:
- right-falling:
- rising (lower element):
- dot:
- hook:
- turning:

Chinese they encountered, and they called it Mandarin, from the Portuguese mandarim, meaning “minister” or “official.” The meaning of “Mandarin” was expanded in the 20th century and has now become a common term for modern standard Chinese, basically interchangeable with the term pǔtōgǔn (普通话) or guóyǔ (国语) in Taiwan.

The distinction between language and dialect has been fraught with politically charged connotations, as the government of the People’s Republic of China (PRC) has preferred the convenient fiction of classifying all forms of regional speech as regional dialects, fāngyǔ (方言). The common criterion for the distinction is mutual intelligibility; when different speakers are able to understand each other, the forms of speech are usually regarded as dialects of a single language. If comprehension is difficult or impossible, they are considered different languages, as in the case of the Beijing dialect vs. Cantoneese.

The new Romanization system known as Pinyin (拼音) uses the Latin alphabet, along with four diacritical/tonal marks to represent the sounds of Mandarin Chinese in standard pronunciation. The Pinyin spellings for a few consonant sounds are markedly different from their use in other languages that employ the Latin alphabet; for instance, Pinyin “q” and “x” sound similar to English “ch” and “sh” respectively, e.g., Qing (Qing) dynasty; Xi (Shi) Jinping (习近平). Above is a graphic illustration of the four tonal marks and phonetic renditions of foreign words and names. Pinyin without tone marks is often used to spell Chinese names and words in languages natively written with the Latin alphabet (including English).

Here is a sample of select country names and personal names rendered in characters and Pinyin: Shakespeare (莎士比亚, Shashibiya); Chopin (肖邦, Xiaobang); Mozart (莫扎特, Mozhate); Beethoven (贝多芬, Beiduoifan); Donald J. Trump (唐纳德·特朗普, Tangnade J. Telangpu); Jesus (耶稣, Yeshu); America or United States (美国, Meiguo, “beautiful” country); France (法国, Faguo, “legal”
country); Germany (德國, Deguo, “virtuous” country); Finland (芬兰, Fenlan).

“Measure words” in Chinese are special indicators or classifiers (known as liàngcì, 號词) used for counting things, based on what they are — people, flat things, round things, animals, fish, and so on. It is placed between the number and the noun it refers to. English similarly has measure words for mass nouns (e.g. a grain of rice, three pounds of rice, three bowls of rice, a herd of cattle). In Chinese, zhàng (張), for example, is used for flat objects such as paper and tables; táo (桃) is used for long narrow things such as snakes, roads or ships (e.g. 三条船, sāntiáochuán, three ships); gē (个) can be used when one is uncertain about which measure word to use.

Grammatically, Chinese lacks morphological changes in person, tense or gender and also has no definite or indefinite articles. Unlike in most European languages, word order is very important for conveying meaning, and the subject of a sentence is usually placed before the predicate, e.g. 他妈妈好 (“he’s/ she’s mother is good”); wǒ hǎo (我好, “I am fine”). Omissions or deletions are much more common in Chinese. Thus, the answer to the English question, “Do students like him?” must be something like “Yes, students like him,” whereas the Chinese equivalent, 学生喜欢他 (“students like him”), must be something like “him.”

Without historical perspective on change and continuity, it is easy to exaggerate either the possibilities or the constraints against the creation of a new Chinese national language. In its 6,000-year history, the Chinese language has undergone three major transformations. First, during the Shang Dynasty (1766 BCE to 1122 BCE), pictograms (象形, xiàngxíng, “form imitation”) were developed to depict objects (e.g., sun, moon, mountain, water, rain, wood, human). Second, during the Zhou and Han Dynasties (1406 BCE to 220 CE), characters became more complex and stylized for conceptual, philosophical and literary words. The third and final transformation began in the 20th century, especially under Mao Zedong, resulting in the triangulation of Putonghua, the Pinyin Romanization system and character simplification. The interplay of historical, cultural, political, scientific and technological factors reflects the twists and turns in China’s long march toward linguistic unification.

During the Qing Dynasty (1636-1912) there was no language reform. Instead, Sinification was the priority, and the term “Chinese people” (中國之人, Zhōngguó zhī rén) was used to refer to all subjects of the empire, including Han, Manchu, and Mongol. Both “China” and “Qing” were used to refer to the state in official documents, international treaties and maps of the world. The “Chinese language” included the Chinese, Manchu and Mongol languages.

By the mid-19th century, the triumph of Chinese civilization over the Manchu was nearly complete, with the abolition of Manchu even as a secondary official language; Manchus, the second-largest ethnic/national minority, could no longer speak and write their own language. A special version of Chinese called guànhuà (官话, “the speech of officialdom”) was adopted for use among literati and officialdom. Today, Manchu is a severely endangered Tungusic language spoken in Northeast China, and almost all Manchus speak Mandarin Chinese.

Still reeling from the aftermath of the Opium Wars, peasant rebellions and an unending series of rapacious foreign land grabs (祸不单行), Qing China, “the sick man of Asia,” seemed to be continually on life support until the end of the 19th century. The early 20th century brought a series of challenges to the very bedrock of Chinese cultural identity as the ruling elite were forced to call into question virtually every aspect of Chinese traditional culture. In this context, Chinese intellectuals looked to the model of European nation states, for which “one nation, one people, and one language” seemed the natural way of organizing a modern country.

The first line of attack was against the millennium-old language of Classical Chinese, called wényán (文言 or 文言文). Though...
originally parsimonious, wényán gradually fos-
silized into a conventional textual language that
became the scholarly standard. Generations of
literati refined the idiom into an extremely terse
and aphoristic written form that, although an
expressive vehicle for poetry and literature, had
become completely divorced from any of the
spoken vernacular.

There was something of a democratic impe-
tus in the push for an alphabetic script. Beyond
purely linguistic considerations, the promotion of
alphabets was perceived as a move toward sci-
ence, toward liberation of the human spirit and
the key to universal literacy. It is partly this
aspect that attracted the communist Left to sys-
tem-transforming language reforms.

The New Culture Movement (新文化运动, Xīn Wénhuà Yuèdòng) of the 1910s sprang from dis-
illusionment with traditional Chinese culture
following the failure of the Chinese Republic,
founded in 1912. However, it was not until after
the May Fourth Movement (1919) and promotion
by scholars and intellectuals — such as pragma-
tist reformer Hu Shi (胡适) and leftist Lu Xun (魯迅) — that vernacular Chinese baihua (白话, lit-
erally “plain speech”) gained widespread accept-
ance. Baihua as a written form was grammatic-
ally patterned on the standard northern Man-
darin dialects. May Fourth intellectuals sought
to release the world of Chinese scholarship from
the stranglehold of Classical Chinese and catalyzed
a movement to publish all books in
baihua so that they could be at least passively understood by a
majority of the Chinese population.

Hu Shi’s passionate appeal for language reform (baihua) in July 1916 is worth quoting: 5

What we need today is a readable, audible, sing-
able, speakable, dictatable language which we can
read aloud without the need to translate into the
spoken language, with the help of which we can
take notes without the need to translate into the lit-	ery language, which we can [use] at the speaker’s
desk as well as on the stage, and which even village
grannies, women and children can understand if
we read it to them. Any language that does not meet
these requirements is not a living language, and can
under no circumstances become the national lan-
guage of our country.

A Madman’s Diary (狂人日記, Kuángrén rìjì), a Chinese short story published in 1918 by Lu Xun,
one of the greatest writers in 20th century Chi-
nese literature, was one of the first and most
influential modern works written in vernacular
Chinese (baihua) and would become a corner-
stone of the New Cultural Movement.

Lu Xun and Mao Zedong (毛泽东) even advan-
ced the eradication of Chinese characters
together, to be replaced with a phonetic system.
Lu Xun saw abolition of Chinese characters as the
only way to expand literacy, a vital goal toward
strengthening the nation. On his deathbed, he is
supposed to have said, “If the Chinese characters
are not eradicated, China is doomed.”

The young Mao was forced to study the Con-
fucian classics, but they only turned him off,
engendering a loathing of the monotonous
memorization and recitation process. Instead,
as a boyish tiger-monkey, Mao searched for
creative ways to continue his subversive read-
ging, “devouring everything I could find, except
the [Chinese Classics],” he told Edgar Snow in 1936. 6 Paradoxically, Mao wrote poetry in clas-
sical Chinese, but he told Snow: 7

We believe Latinization is a good instrument with
which to overcome illiteracy. Chinese characters
are so difficult to learn that even the best system of
rudimentary characters, or simplified teaching, does
not equip the people with a really efficient and rich
vocabulary. Sooner or later, we believe, we will have to
abandon characters altogether if we are to create a new
social culture in which the masses fully participate.

It was not until 1920 that the Chinese Min-
istry of Education ordered that primary and
middle-school texts be changed from wényán
to baihua and mandated that all textbooks be
written in baihua. Since the late 1920s, nearly
all Chinese newspapers, books, and official and
legal documents have been written in baihua.

Classical Chinese is taught primarily by pre-
senting a classical Chinese work and a baihua
version of the meaning of wényán phrases. In the
Gaokao (高考), China’s extremely
competitive college entrance exam, there is a
test on classical Chinese, typically translation
exercises which ask the student to express
the meaning of a paragraph in baihua.

The third major transformation was a long
multi-stage process, starting in the 1910s and
1920s focused on baihua and culminating in the
PRC period. Despite, or perhaps because of,
abandoning the Latinization movement, lan-
guage reform was a low priority between 1949
and 1953; nation-building and China’s entry
into the Korean War in October 1950 became
the top national security and survival chal-
enges. In many respects, the third major trans-
formation occurred after the “end” of the Korean
War in July 1953. Why such a turnaround from
the Latinization movement? As Prime Minister
Zhao Enlai subsequently told a former French
minister of education: “In the 1950s, we tried
to Romanize the writing. But all those who had
received an education, and whose services we
absolutely needed to expand education, were
firmly attached to the ideograms.” 8 In
other words, to create a new Romanized language
could increase China’s literacy rate in the long
run, but it would instantly turn 99 percent of the
Chinese people illiterate!

An official language-reform policy gradually
revealed itself as centering on three mutually
interconnected challenges: (1) promoting Puton-
ghua as an exclusive standard at the expense of
other forms of speech; (2) placing primary empha-
sis on character simplification; and (3) developing
a phonetic system for use primarily in character
annotation while indefinitely deferring considera-
tion of its use as an independent orthography.

Despite flaws in the initial policy, the PRC lan-
guage committees had at least acquiesced to the
enormous amount of expert opinion on the mat-
ter and had sensibly adopted the Beijing dialect
as the basis for the national standard. The Minis-
try of Education also decreed in November 1955
that Putonghua would become the medium of
instruction for Chinese language and literature
courses beginning in the autumn of 1956 and
that teachers of all subjects were to conduct their
classes in Putonghua by 1958.

Promulgation as the spoken language was only
one leg of the post-1953 language reforms. In
China, reform of speech could not be fully imple-
mented without addressing the problem of writ-
ing, and so the government turned its attention
to the written language. At a language planning

---

8 DeFrancis, ibid. p. 258.
conflict between linguistic communities. In 1956, the State Council promulgated the first list of simplified symbols, consisting of 515 simplified characters plus 54 simplified components. About 90 percent were commonly used shorthand forms that had been in use for some time, and others were simplifications replacing complex components with more phonetic ones.

In 1964, further simplification increased the number to 2,238 out of the approximately 7,000 characters in general use. This streamlining yielded a reduction of 12.5 percent in the average number of strokes for the 2,000 most common characters, but even this simplified system still requires years of practice to master. The government attempted a second round of character simplifications in 1977 but these met with strong resistance by the linguistic community, and authorities completely abandoned the additional efforts.

Failure to replace characters with an alphabetic script did not mean that the quest for a scientific Romanization scheme had lost its importance; more than 1,700 phonetic scripts and methods were proposed in the 1950s. Then, in January 1958, Pinyin (Hanyu Pinyin Fāngàn 汉语拼音方案, “Chinese Phonetic Scheme” or “Chinese Phonetic Plan”) was officially promulgated. Pinyin is now the official way of transcribing Chinese personal and place names, ending a long history of chaotic inconsistency—e.g., Pinyin “Zhou Enlai” (周恩来) replaced “Chou Enlai” in English (Wade-Giles system) and “Tchou Enlai” in French. The chief functions of Pinyin have been to indicate the standard pronunciation of Chinese and to aid acquisition of Putonghua via annotated characters. Internationally, Pinyin was generally adopted in the 1980s.

The development of this Romanization system was due, in large part, to the efforts of Zhou Youguang (周有光), who also translated Encyclopaedia Britannica into Chinese, earning himself the nickname “Encyclopedia Zhou.” Even more remarkably, this super-prolific, supercentenarian (1906-2017) managed to publish 10 books between the years 2000 (aged 94) and 2015 (aged 109)!

The spread of Putonghua has been one of the more successful aspects of language reform, becoming popularized beyond large urban areas upon the expansion of broadcast media in the 1980s and 1990s. In 1978, the proposal to establish Radio & TV Universities was approved by Deng Xiaoping. In 1982, Putonghua was given extra clout through its inclusion in the revised PRC Constitution, Article 19: “The state promotes the nationwide use of Putonghua (common speech based on Beijing pronunciation).” This single sentence solidified Putonghua as a constitutionally mandated national language.

In 1994, a national Putonghua proficiency test was established, applying to teachers, television and radio broadcasters, and all movie, television, theatre and soundtrack professionals. Broadcasters can be fined for uttering a non-standard pronunciation or faulty tone.

In the realm of soft power diplomacy, China established the first Confucius Institute (Kǒngzǐ yuàn 孔子学院) on November 21, 2004, in Seoul, South Korea. As of 2014, there were more than 480 Confucius Institutes in dozens of countries in six continents. The main function of these institutes is to improve the understanding of Chinese culture through improved language training facilities, with the aim of promoting greater mutual understanding, as articulated in the idea of a new and harmonious world order. The Ministry of Education estimates that 100 million people overseas may have been learning Chinese by 2010 and currently aims to establish 1,000 Confucius Institutes by 2020.

On the Internet, where China has had a presence since 1989, the Chinese population is the largest national presence, and more than 97 percent of Chinese computer users “write” Chinese by typing Pinyin. Indeed, the Internet and digitalization have become the most powerful force to encourage the spread of Putonghua. Chinese netizens’ reactions to the rise of Xi Jinping (习近平) are surprising and revealing: “At long last, a Chinese leader who can speak Putonghua!”; “Self-confident, young, active, [speaks] Putonghua”; “[let me] say I don’t feel a generational divide while listening to [Xi].”

Most recently, in October 2016, the Chinese government announced its plans to introduce codes for some 3,000 Chinese characters as part of a grand project, known as the China Font Bank, to digitize 500,000 characters previously unavailable in electronic form. The project highlights 100,000 characters from the country’s 56 ethnic minorities, and another 100,000 rare and ancient characters from China’s written corpus. Deploying almost 30 companies, institutions, and universities, it is the largest state-funded digitization project undertaken to date.¹⁰

3) PROGRESS AND PROSPECTS

Despite all these changes, China’s Ministry of Education conceded that language reform had only limited success, stating on its official website

---

GLOBAL ASIA Feature Essay China’s Long Struggle for Linguistic Unification

in September 2014 that fully a third of the Chinese population — 400 million people — can’t speak Putonghua (Mandarin Chinese). But that leaves some 858 million Chinese speaking Putonghua, representing the world’s largest linguistic community. Size matters in assessing progress in China’s quest for linguistic unification. Supposing that Mao had failed to unify mainland China in 1949, what would a territorial-cum-linguistic map of “Mainland China” look like?

It is important to keep in mind that a fair and balanced assessment of progress in linguistic unification first requires answering “compared to when and to whom?” China’s literacy rate should be viewed in historical, contextual and relative terms. The literacy rate (defined by UNESCO as people aged 15 or over who can read and write) was around 20 percent in the 1950s, 78 percent in 1990 and 96.4 percent in 2015, compared to 99 percent for the advanced economies of the West, 100 percent for tiny North Korea, but only 71.2 percent for India, the world’s second largest state. Contextually, the rate for China’s young people (ages 15 to 24) is now 99.6 percent, and the gender gap has been dramatically reduced. In 1990 it was 87 percent for men and 68 percent for women, but by 2010 the gap was reduced to just five percentage points — 98 percent for men and 93 percent for women. Among China’s youth today the gap is virtually non-existent. It is no secret that the literacy rate is higher in urban areas than in rural areas, but from a 19.6 percent urbanization rate between 1949 and 1980, it rose to 51.3 percent in 2011 and is projected to reach 65.40 percent by 2025.

The conventional wisdom espoused by many Western Chinese linguists is that China’s once-great literary and linguistic tradition had in modern times become an albatross around its neck in China’s protracted and torturous march to being a great power (富国强兵). However, a short list of China’s military, economic and technological accomplishments may suffice to argue that this conventional wisdom has no leg to stand on, let alone run:

1) China became a nuclear power in 1964;
2) the World Bank in 1991 singled out post-Mao China as having garnered an all-time global record in doubling per capita output in the shortest period (1977–1987);
3) China surpassed Germany as the world’s largest export power in 2009;
4) China surpassed Japan as the world’s second largest economy in 2010;
5) China became a space power in 2011 when it launched its first orbital space lab; and
6) In 2016, China maintained its No.1 ranking on the 47th edition of the TOP500 list of the world’s top supercomputers with a new system built entirely with processors designed and made in China.

Judging by demographic and digitalization trends, China will most likely reach a 100 percent literacy rate in the not-too-distant future. That said, however, in the coming years, the way China manages its economic reforms, social unrest, environmental pressures, widening inequality and ethno-national pressures from below and within may be more decisive factors in shaping China’s future as a complete and responsible great power. A weak and fragmenting China would be the worst of all possible scenarios, a disaster not only for China but also for peace and stability in East Asia and beyond.

Samuel S. Kim is a Senior Research Scholar at the Weatherhead East Asian Institute at Columbia University. He is series editor of the Asia in World Politics series of Rowman & Littlefield Publishers.