Abstract and Keywords

The story of the development of linguistic thought in the major cultures of East Asia centers to a great extent on each country's response to influence from a respected non-related language. In the case of China before the period of European influence, it was Sanskrit, the language of the Buddhist scriptures; for Korea and Japan, it was Chinese, and to a lesser degree also Sanskrit. Both Korean and Japanese borrowed vocabulary extensively from Chinese, as well as the Chinese writing system, and in more recent times, Chinese has borrowed heavily back from Japanese, reflecting the ebb and flow of power between the countries. Korea and Japan later developed their own writing systems, both of which show strong influence from Chinese and Sanskrit. In all three cultures, the formal study of grammar does not begin until the nineteenth century, under European influence.

Keywords: Chinese, Japanese, Korean, history of linguistics, Buddhism

10.1 Background and Context

THIS entry outlines the linguistic traditions of China, Korea and Japan, with special emphasis on China. Genetically, Chinese is a Sino-Tibetan language, unrelated to Korean and Japanese. Korean and Japanese are commonly lumped together in the 'Altaic' phylum ('phylum' being superordinate to 'family') of languages, with some scholars even believing they are distantly related to the Turkic, Mongolic, and Tungusic families, though there is little if any proof for this based on shared native vocabulary with regular phonological correspondences. On the other hand, all the languages in the Altaic phylum share great similarities in their syntactic structure; all are agglutinative.

The story of the development of linguistic thought in East Asia centers to a great extent on each country's response to influence from a respected non-related language. In the case of China before the period of European influence, it was Sanskrit, the language of the Buddhist scriptures; for Korea and Japan, it was Chinese, and to a lesser degree also Sanskrit. Both Korean and Japanese borrowed vocabulary extensively from Chinese, as well as the Chinese writing system, and in more recent times, Chinese has borrowed

heavily back from Japanese (Chung 2001), reflecting the ebb and flow of power between the countries. Korean and Japanese both later developed their own writing systems. This account traces the transmission, acquisition, and subsequent development of writing, both as a tool in its own right and as a medium for spreading literature, religious ideas, and learning. Mention of one example of a philosophical text on language is included in the case of China. Short introductions are provided for China and Japan on the beginnings of Western-style grammar studies, which do not appear in any of the three cultures examined here until the nineteenth century, under European influence. Finally, a number of modern language-related issues, such as the establishment of a national language or writing system, are briefly touched on.

(p. 210) 10.2 China

The history of linguistic thought in China, as in any other culture that has reached a certain level of development, can be approached in two ways: by examining implicit, applied knowledge of language on the one hand, and by surveying records of self-conscious, explicit thought about language on the other. Much of the history of linguistics in East Asia must in fact be constructed from texts and examples of applied language rather than linguistic writings *per se*. And in China, certain subfields of linguistic thought, such as lexicography, poetic prosody, phonology, and character etymology, have from an early period received much more attention than other areas, such as syntax. Rather than trying to force a preset idea of what linguistic thought 'should' be, it perhaps makes more sense to examine whatever traditions have appeared and developed over history, even if some fall more into the area of philology rather than the history of linguistic thought *per se*.

A very rough periodization of the Chinese language, simplified from that of Wang Li, is offered here for reference:

Archaic Chinese: the period prior to the third century AD;
Middle Chinese: from the fourth to twelfth century AD;
Pre-Modern Chinese: thirteenth to nineteenth century;
Modern Chinese: the period from the 1919 May Fourth Movement to the present.

The only truly reliable information we have on China's earliest linguistic thought comes from written records, and the written symbols themselves. The earliest examples we currently have of written symbols in China date to the fifth millennium BC, the period of China's Neolithic Yǎngshào 仰韶 culture. They come to us in the form of engravings on pottery unearthed from the Bànpōcūn 半坡村 archaeological site near Xi'an in northern China (Lin 1981: 3–6, 82–4). Scholars are still unsure of the purpose of the figures and what they mean, if anything. They do in any case seem to be some kind of primitive precursor of the Chinese writing system still in use today.

The next major extant type of early Chinese writing is engravings on tortoise shells and ox bones, called *jiǎgǔwén* 甲骨文 in Chinese, used mainly for divination. These were first discovered in the village of Xiǎotún 小屯 in Anyang, Henan province, starting around 1880

(Chen 1988: 2). Other materials were probably used for writing as well, but being more durable, the shells and bones were better able to survive the millennia. The engraved symbols have largely been deciphered, and clear lines of development have been drawn connecting them with their modern counterparts. These writings are a rich source of information on ancient Chinese religious and superstitious beliefs and rituals, political and social history, and—of special interest to us here—early character forms, variants, and methods of composition, in addition to the syntax of archaic Chinese, covering the period from about 1200 BC to the third century BC. Records of the continued evolution of character forms can be found in inscriptions on bronze vessels from the (p. 211) latter part of this period and later, and in writings on bamboo strips, silk, paper, and other materials after these. Cai Lun 蔡倫 (63–121 AD) is credited with the invention of paper, though he may actually only have improved an existing technology.

10.2.1 Education, Ancient Texts and Philology

The *Yìwénzhì* 藝文志 'Treatise on Arts and Letters' of the *Hànshū* 漢書 'Annals of the Han dynasty', written in the early third century AD, says that at the age of 8, children began their studies at a *Xiǎoxué* 小學 'Little School', where they were trained in the 'Six Arts': (1) the five rituals and rites, (2) the six styles of music, (3) the five styles of archery, (4) the five styles of cart driving, (5) the six principles of character composition, and (6) the nine numbers (Wang 1987: 2–4). So along with learning to read classical texts and to write, children were taught about the methods of character composition. This system of analysis and classification of the characters was a revered cultural achievement, and it is still taught in schools today. The six methods, along with the two examples originally given for each, are:

1. *Zhǐshì* 指事 ideographs: a symbolic depiction of the referent, often used for abstract notions; examples: 'above' *shàng* 上, 'below' *xià* 下.
2. *Xiàngxíng* 象形 pictographs: a drawing of a concrete object, e.g. 'sun' *rì* 日, 'moon' *yuè* 月.
3. *Xíngshēng* 形聲 phono-logographs: composed of a semantic plus a phonetic component. Due to historical sound changes and influence from varying dialects, the pronunciation of the phonetic component itself does not always match the modern pronunciation of the whole character. Whereas the previous two categories use character form as a mnemonic, *xíngshēng* characters have two aids to memory: a general semantic category, e.g. 'plant', 'female,' 'vision,' plus a phonetic reminder. It is notable that approximately 85 per cent of all Chinese characters are constructed according to this model, and it is the default method both for existing characters and for inventing new ones. Examples: '[larger] river' *jiāng* 江 [the semantic category component or 'radical' is 'water,' plus the phonetic: *gōng*, which reflects an earlier pronunciation], '[smaller] river' *hé* 河 ['water' radical plus phonetic *kě*].
4. *Huìyì* 會意: compound ideograph: composed of two or more components from the other character composition types, e.g. 'military arms' *wǔ* 武 [composed of the two elements 'dagger axe' + 'stop'], 'trust' *xìn* 信 ['person' + 'words'].

5. *Zhuǎnzhù* 轉注: Characters which have changed in form over history, e.g. *kǎo*考 is an earlier form of *lǎo*老 'old.' This category has many widely differing interpretations; the one adopted here is the more systematic and convincing one of Lu Shixian魯實先, which he backed up by many examples from the *Shuōwén*.

6. *Jiǎjiè* 假借: A character borrowed purely for its phonetic value, with no semantic connection, as was often done in the Egyptian hieroglyphics. The original two (p. 212) examples were words used in an extended sense rather than as actual phonetic loans; more appropriate examples are *ér* 而, which originally meant 'beard'; the character was later borrowed for a homophone meaning 'and, however'; also the character *yāo*要 'waist' was borrowed to write the word for 'to want' *yào* 要; a 'flesh' radical was then added to the character used for *yāo* 腰 'waist,' which through this process became a *zhuǎnzhù* character. Lu emphasizes that *jiǎjiè* is not a method of character *creation per se*, but a type of character *usage* supplementing the other five methods.

Over time, the term *xiǎoxué* came to mean 'philology,' that is, (1) works on ascertaining character meanings, particularly obscure ones found in ancient texts; (2) character compendia, with an emphasis on the forms of the characters; and (3) rhyme books, primarily intended as a reference in poetry writing, but also doubling as ordinary dictionaries. Interest and work in *xiǎoxué* reach its height in the Qing dynasty (1644–1911 AD).

Literacy education needs were the source of China's first 'word books.' A number of them were in use starting from the Qin dynasty (221–207 BC) and earlier. The only one still extant today, however, is a character primer called the *Jíjiùpiān* 急就篇, *jí* 急 and *jiù* 就 being the first two characters of the work, along with fragments of another, entitled the *Cāng Jié Piān* 倉頡篇 (Cang Jie was China's legendary inventor of written characters), which was in fact a collection of a number of works of this type. Many expansions and similar works followed. These primers were written in a highly stilted style, and consisted mainly of 3-, 4- or 7-syllable rhyming lines. They were also designed to include as many different characters as possible, with little repetition, to maximize the children's exposure to and learning of new words. Literacy education was strongly emphasized since, according to the *Hànshū Yìwénzhì*, a knowledge of 9,000 characters was required to pass the test to become a government official. And once employed, an official was subject to punishment if he wrote a wrong character in a document.

A work following these primers was the *Ěryǎ* 爾雅, dating to about the second century BC. The original author is unknown, and it was expanded by later writers. It is basically a lexicon of words classified by subject. Its format is to list a series of words that are close in meaning, then to end it with a 1–3-character definition or description. It starts off with three chapters of miscellaneous words under the general category of 'language,' followed by such subject areas as family relationships, music, geography, and animals. Its main purpose was to explain unfamiliar characters found in earlier texts. More common characters are generally not treated, since they were assumed to be already known by the reader and their inclusion thus unnecessary.

Dialects were another focus of these early word books. The most representative work is the *Fāngyán*方言 'Dialects' by Han scholar Yang Xiong 揚雄 (53–18 BC). It features comparative lists of vocabulary items from different parts of China, along with the name of the region where each was in use. Language study in Chinese focused above all on the *written* language, even in works that collected dialect data. Again, the absence of (p. 213) a system of phonetic notation is a major drawback to this work, but the data it provides can be used to build up isogloss maps indicating the differences and overlaps in vocabulary use of the various regional dialects of the time.

The primers and related word books declined in importance with the appearance in approximately the year 121 AD of the *Shuōwén Jiězì* 說文解字 etymological dictionary by Han dynasty lexicographer Xu Shen 許慎. Xu established 540 'radicals' or categories based on character components, mainly for the purposes of indexing and easily retrieving individual characters when needed. These often coincided with the semantic marker (rather than the 'phonetic', i.e. phonetic marker) of *xíngshēng*-type or phono-logographic characters. Pronunciations were indicated either by the phonetic marker of the character entry, or with a homophone, preceded by the phrase, *dúruò* 讀若 'read as…'. While the *Shuōwén* contains inevitable errors, many character etymologies are now known thanks to this work, especially when it is examined in conjunction with the bronze inscriptions, much studied during the Song dynasty (960–1279 AD), and the oracle bone and shell inscriptions. The *Shuōwén* is China's first bona fide dictionary, and continues to be a highly valued reference work even today.

Starting in the Han dynasty, a frequently used method of defining an item was *shēngxùn* 聲訓 'use of a homophone or phonetically similar word to define another word.' The belief behind this was that phonetically close words were often also etymologically related and could thus offer insight into the definition of an item (Wang 1987: 10). For example, *zhèn, zhèn yě*, 震 , 振也 'earthquake: vibration.' Other *shēngxùn* are less straightforward and often quite forced, e.g. *rì, shí yě* 日 , 實也 'sun: solid'; *yuè, quē yě* 月, 闕也 'moon: waning'. This approach is frequently drawn on in the *Shuōwén*, and culminated in the compilation of the *Shìmíng* 釋名 'Explaining the Names,' a Han dynasty lexicon of the late second to early third century, attributed to Liu Xi 劉熙, which used this method extensively. In contrast to Western dictionaries, in which entries are arranged alphabetically (roughly, phonetically), and given straightforward semantic definitions, items in the *Shìmíng* and lexica like it are arranged by subject category and usually paired with a phonetically related definition, either a homophone or an alliterative or rhyming gloss. The above examples are found both in the *Shuōwén* and the *Shìmíng*. *Shēngxùn* definitions are generally far from rigorous and are often not even very informative. But they formed a consistent pattern of word definition, and were possibly also useful as a mnemonic.

10.2.2 Central Asia and Buddhism

Credit for opening up contact between China and Central Asia and India goes mainly to Han court envoy Zhang Qian 張騫 (*c*.164–114 BC), who endured extreme duress to visit and return from many countries of Central Asia, and India, between 138 and 115 BC.

(p. 214) Numerous new vocabulary items entered Chinese around this time to name some of the new things he discovered in his travels, particularly names of plants and foods, such as grapes, pomegranates, sesame, and alfalfa; animals, including lions; musical instruments; and other culturally valuable items, such as glass. In general, each syllable in Chinese constitutes a separate morpheme; Central Asian loanwords from this period, on the other hand, are notable because many are disyllabic single morphemes, such as *pútáo* 葡萄 'grapes' from Old Iranian *buduwa*, derived from *buda* 'wine,' and *mùsù* 苜蓿 'alfalfa' from Old Iranian *buksuk* or *buxsux* (Shi 2008: 518–19).

Buddhism was founded in India in the sixth to fifth century BC, and transmitted to other parts of Asia, particularly Central Asia, largely through the efforts of India's Ashoka the Great (*c*.304–232 BC). This is of interest in a linguistic history of China for two major reasons: first, because of the large amount of new vocabulary introduced through Chinese translations of the Buddhist sūtras and second, because the process of translation unavoidably raised awareness of the features of Chinese that differed from Sanskrit and the other source languages.

Translation of Buddhist sūtras from Sanskrit, Pāli, and various Central Asian languages into Chinese began around 58 AD. Translation was typically a team effort. The first step was for a native of the source language, who had come to China specifically for this task —perhaps a Sogdian, Khotanese, Kashmiri, Indian, or Yuezhi monk—to declaim the text, often apparently from memory. Then a bilingual interpreter, usually a native of a Central Asian state or India, or a descendant of same who grew up in China, would interpret the text orally to Chinese scribes, who then transcribed what they heard. Then the texts were carefully edited for accuracy, consistency, conciseness, and style, and proofread. Some translations were also done by Chinese who learned Sanskrit in China. Studies of phonetic equivalents used when going from the source language to Chinese seem to establish that in the early period, the texts were usually translated from a translation into a Central Asian language, called in Chinese *húyǔ* 胡語 'foreigners' language' or *súyǔ* 俗語 'common, popular language,' rather than directly from Sanskrit or Pāli; as time went on, translations were increasingly directly from Sanskrit or Pāli, and this was reflected in different phonetic transliterations. Translations from the early period also tended to be very literal, at the expense of good Chinese style, in an effort to ensure accuracy. The monks also tried to choose transparent descriptive translations in easily understandable Chinese, and to avoid phonetic transliterations of specialized Sanskrit terms as much as possible, though many transliterated Sanskrit and Pāli words did still appear, and are opaque to the uninitiated. Different translators had different styles, and sometimes used different Chinese words for the same term in the source language; and there were many other complicating issues. It is clear from this description that the translation process was highly susceptible to error and omission. Then, as now, linguistically sophisticated, highly skilled translators were in short supply. Translated sūtras in circulation were thus often corrupt and incomplete, so this occasioned some monks, such as Faxian 法顯, between 399 and 412 AD, to undertake arduous and dangerous journeys to India and other countries to learn Sanskrit, and to bring back the original manuscripts (p. 215) for painstaking

translation, thereby deepening and solidifying the connections between Chinese and the languages of Buddhism (Ma 1998: 18–94; Boulnois 2004: 209–34).

Translating religious and philosophical concepts which often had no ready equivalent in Chinese was a challenging task. Many new terms had to be invented, either by coining them from existing Chinese morphosyllables, e.g. *yè* 業 'business, deed' or *yèzhàng* 業障 'business + obstacle' for 'karma'; or often a phonetic rendering of the Sanskrit was used, e.g. *púsà* 菩薩 for 'bodhisattva' and *póluómén* 婆羅門 for 'Brahmin.' These examples are strongly marked as religious loanwords from India. But some of the newly created vocabulary came into general use, and is not easily identified as coming from Buddhism, e.g. *shìjiè* 世界, previously referred to as *tiānxià* 天下 'heaven below' in Chinese, is now the standard word for 'world'; and *xiànzài* 現在, which used to be 見在 *jiànzài* 'what-you-can-see existing' in earlier Chinese, became the normal word for 'now' (Wang 1988: 674–5).

The diverse alphabets, word orders, conjugations, declensions, and agglutinative suffixation of the source languages of the scriptures, i.e. Central Asian languages, which mostly belonged to the Turkic family, and Sanskrit and the Prakrits, of the Indo-Iranian branch of Indo-European, all offered sharp contrasts to the spare, unencumbered morphology of Chinese. Sanskrit and other Indian languages, for example, tend to have SOV word order, though word order can also be flexible, while Chinese is basically an SVO language; Sanskrit is also a highly inflected synthetic language, while Chinese is the classic example of an isolating, analytic one.

The translation process also made the Chinese highly aware of the importance and nature of the tones in their language, in contrast to Sanskrit and the other source languages, which were not tone languages. From earliest times, Chinese was in contact with many non-Han languages, which in theory could have led to more insights about Chinese much sooner, but these languages were generally viewed with condescension or even contempt, due to China's sense of her own cultural superiority. (It should be noted that China was in these times most often a collection of independent principalities and states, and not a unified nation.) The position of Sanskrit was entirely different; it was a prestige language because it was the language of the Buddhist scriptures, and this made it worthy of translation and close study.

One outcome of this comparative philological study was the first major work dedicated to a description of the four tones, the *Sìshēngpǔ* 四聲譜 'Tables of the Four Tones,' unfortunately no longer extant, by Shen Yue 沈約 (441–513 AD). From this time on, poetry rhymes were also required to take tone as well as syllable finals into account, although awareness of the tones was already reflected in some earlier works (Wang 1988: 74–5). (It is to be noted that the original four Chinese tones were not the same as the modern Mandarin ones—they went through considerable merging, splitting, and rearrangement before settling into their current form. And other dialects have different numbers of tones: e.g. modern Cantonese, a much more conservative dialect in some aspects, has nine.)

The original four tone categories are described as *píng* 平, 'the level tone'; *shǎng* 上, 'the rising tone'; *qù* 去, 'the falling tone'; and *rù* 入, 'the entering tone,' which is used to refer to syllables ending with a /-p/, /-t/, /-k/ or a glottal consonant stop final. The tones (p. 216) came to have central importance in the rules for Chinese poetic prosody, which were highly evolved by the Tang dynasty. For prosodic purposes, the four tone categories were divided into just two categories, *píng* 平 'level' (the first tone) and *zè* 仄 'oblique' (the remaining three tones lumped together). Various verse forms generally required specific sequences of the two tone types in each line of poetry—for example, two of one tone type, then two of the other, then one of the first type. The following line, often semantically parallel, would have basically the same tone pattern in reverse; there were also some exceptional breaks in the pattern for flexibility and variety. See the poem in Table 10.1, §10.2.4 below, for a sample tone scheme.

In face of the lack of an alphabetic writing system to use for phonetic notation, the *fǎnqiè* 反切 system was developed to indicate character pronunciations. Under the *fǎnqiè* system, two relatively well-known characters, plus the word *fǎn* 反 or later mostly *qiè* 切, were given after a lexical item. The reader needed to take the initial of the first and splice it onto the final rhyme and tone of the second, to derive the pronunciation of the item being looked up. A typical entry is *dōng déhóng qiè* 東 德紅切, i.e. **dé** plus **hóng** in the *qiè* 切 system make *dōng*. (The second tone had not yet separated from the first at this time, thus the difference in tones.) One big advantage of the system is that the *fǎnqiè* characters were already familiar to any literate Chinese, so there was no need to learn a new set of symbols. The disadvantage is that there is no way to know with certainty the actual phonetic realizations of the syllables at the time.

Spurred by his knowledge of alphabetic writing systems, the tenth-century monk Shouwen 守溫, who was possibly not an ethnic Han, developed an 'alphabet' for phonetic notation of Chinese characters for use in the rhyme books. It is interesting that, in spite of having the Sanskrit Devanāgarī alphabet as a model, he did not develop an alphabet or syllabary, but instead chose 30 existing Chinese characters to represent consonant or vowel initials. They were arranged in an order similar to that of the Sanskrit alphabet, according to, for example, whether a sound was voiced, voiceless, or voiceless aspirated. This set was later expanded to 36. The lack of a set of symbols indicating the values of individual segments is a big drawback of the system, but it does give us valuable categorical information on Middle Chinese.

10.2.3 The Rhyme Books and Dictionaries

The rhyme books were the next genre to appear of philological works building on the early primers, lexica, and expanded phonetic awareness of the special features of Chinese. These were primarily intended as reference works for poets checking which character-syllables were permissible as rhymes in the various verse forms; but they also served as dictionaries, since they included *fǎnqiè* pronunciations and simple definitions. Two noteworthy early works are the *Yùpiān* 玉篇 'Book of Jade' by Liang through Chen dynasty (502–89 AD) scholar Gu Yewang 顧野王, and the *Jīngdiǎn* (p. 217) *shìwén* 經典釋文

'Explanations of Words used in the Classics' compiled by Tang scholar Lu Deming 陸德明 (550?–630 AD).

The Song dynasty edition of the *Guǎngyùn* 廣韻 'Expanded Rhymes', compiled by Chen Pengnian 陳彭年 (961–1017 AD) is the most complete extant example of an early rhyme book. Many previous works were combined and re-edited to produce this work. One of the key works on which it was based was the *Qièyùn* 切韻—now lost except for a small number of fragments—compiled by Lu Fayan 陸法言 and other Sui dynasty language scholars in the late sixth century. In the surviving preface of this work, Lu describes how the compilers made a conscious decision to be as inclusive as possible regarding historical and geographical variations:

> In the evenings, after having enjoyed our wine, our discussions always turned to phonology. Differences obtain between the pronunciations of the past and the present, and different principles of selection are followed by the various authors. In the [southern] regions of Wu and Chu the pronunciation is at times too light and shallow. In [the northern regions of] Yan and Zhao it is often too heavy and muted…And so we discussed the right and the wrong of South and North, and the prevailing and the obsolete of past and present. Wishing to present a more refined and precise standard, we discarded all that was ill-defined and lacked precision… A knowledge of phonology is necessary for any literary undertaking…And so, choosing from the various rhyme books and other lexica, old and new, and basing myself on my earlier notes, I organized the material into the *Qieyun* in five volumes, analysing minutiae and making fine distinctions. It is not that I have been the sole judge in these matters. I have merely related the opinions of my worthy colleagues. (Translation from Malmqvist 1994: 11–12).

While the *Guǎngyùn* brings together much valuable data, it is difficult to sort out from it a single, consistent variety of the Chinese of the time.

Rhyme grid books (*yùntú* 韻圖) were a subsequent development of the rhyme book tradition. These also show direct influence from how sounds are combined to form words in the Sanskrit Devanāgarī script. Rather than just giving a *fǎnqiè* pronunciation for each individual syllable, characters were arranged in grids according to tone, and presumably their different main vowel qualities and medials—there is as yet no consensus on how to interpret all of the categories distinguished in the rhyme grids. Perhaps the most representative rhyme grid book is the *Yùnjìng* 韻鏡 'Rhyme Mirror,' author unknown, dating to the tenth century AD.

As the Chinese language evolved, new rhyme books were produced to reflect its changing phonology. The *Zhōngyuán yīnyùn* 中原音韻, compiled in 1324 by Zhou Deqing 周德清, is our best systematic record of the transition between Middle Chinese and modern Mandarin. Rather than attempting to incorporate historical and dialectal material, it focuses only on the actual spoken language of the time, an innovation in this genre (Norman 1988: 49). It documents the disappearance of the entering tone (tones realized with a final stop) from Mandarin, the devoicing of voiced initials, and the split of the even *píng* 平

tone into two new tones, the *yīnpíng* 陰平 and the *yángpíng* (p. 218) 陽平, the first two tones of modern Mandarin. Syllables that originally had a voiced initial ended up in the *yángpíng* tone category, among other changes.

Returning briefly to the dictionary tradition: the comprehensive *Kāngxī Zìdiǎn* 康熙字典 'Dictionary of the Kangxi Emperor' was completed in 1716 after six years of compilation work. Its format and content were based mainly on two earlier lexica, the *Zìhuì* 字彙 'Collection of Characters' of 1615, an innovative dictionary with 33,179 entries that reduced Xu Shen's 540 radicals to the current 214, and arranged characters according to their number of strokes; and the *Zhèngzìtōng* 正字通 'Correct Character Mastery' of 1671, produced as a supplement to the *Zìhuì*. The *Kāngxī Zìdiǎn*, in spite of errors, is still the authority many turn to for obscure characters: it contains 47,035 entries (Liu 1963: 40–56).

Starting from the early 1900s, Swedish linguist Bernard Karlgren (Chinese name: Gāo Běnhàn 高本漢) blazed a trail for studies in Chinese historical phonology based on works such as the *Guǎngyùn* and other written records, together with extensive modern dialect data, which he himself collected. Since then, considerable scholarly energy has been devoted to the phonological reconstruction of Chinese as spoken in various historical periods.

10.2.4 Grammar

It is interesting to contrast the great influence Sanskrit had on Chinese in the area of phonology with the negligible mark it left in the area of Chinese grammar studies. Rather than a division into nouns, verbs, adjectives, and such, the main grammatical categories posited by Chinese for their own language, starting in the Wei-Jin and Northern and Southern Dynasties period (265–589 AD), were content ('solid') words, called *shící* 實詞, vs function ('empty') words, or *xūcí* 虛詞. This simple binary approach was in fact a very reasonable one, since Chinese has very little explicit morphological marking, and was, and continues to be, highly dependent on word order to indicate grammatical relations. Function words are one of the few visible and concrete features of the language on which to build a grammatical analysis. And this approach is a natural outgrowth of the traditional text explication method of studying the written language. Up to the present, entire books are written on the meanings and usages of these 'empty' structural particles, especially those used in Classical Chinese. Such a book would include, for example, explanations of the many functions and uses of the possessive or genitive particle *zhī* 之 in Classical Chinese, and *de* 的 in Modern Chinese. This is probably China's most representative native approach to syntactic analysis.

There is, however, plenty of evidence of *implicit* awareness of the parts of speech as known in the West, reflected in the parallelism found in many types of literary writings, such as poetry. Table 10.1 is an example by Tang dynasty poet Wang Zhihuan 王之渙 (688–742 AD), entitled *Dēng Guànquèlóu* 登鸛鵲樓 'Ascending Stork Tower'; this sample will also

give readers who are unfamiliar with Chinese a good idea of how it works. (Reference translation into English:

(p. 219)

白	**日**	**依**	**山**	**盡**	CHARACTER
bái	*rì*	*yī*	*shān*	*jìn*	PRO-NUNCIATION
white	sun	along/follows	mountains	disappears	ENGLISH GLOSS
adj	noun	prep/verb	noun	verb	PART OF SPEECH
O	O	L	L	O	L[EVEL TONE]
					O[BLIQUE TONE]
黃	**河**	**入**	**海**	**流。**	
huáng	*hé*	*rù*	*hǎi*	*liú*	
yellow	river	into/enters	sea	flows	
adj	noun	prep/verb	noun	verb	
L	L	O	O	L	

Table 10.1 The poem 'Ascending Stork Tower'

欲	窮	千	里	目，	
yù	*qióng*	*qiān*	*lǐ*	*mù*	
if-you-want-to	exhaust	1000	0.3 miles	eye/see	
verb	verb	number	measure	noun	
O	L	L	O	O	
更	上	一	層	樓。	
gèng	*shàng*	*yī*	*céng*	*lóu*	
further	ascend	one	storey	building	
adverb	verb	number	measure	noun	
O	O	O	L	L	

The white sun sinks behind the mountains
The Yellow River flows into the sea;
If you wish to have a broader view
Just climb up one more storey.)

The application of European-style grammar to Chinese began with Qing dynasty language scholar Ma Jianzhong 馬建忠 (1845–1900), who studied in Paris and knew French and Latin. In his *Mǎshì Wéntōng* 馬氏文通, published in 1898, he analyses Classical Chinese—the spoken vernacular was not yet recognized as worthy of this kind of study—mainly according to Latin-style grammatical categories, though he shows considerable innovation in his system of Chinese syntactic categories. He still adheres to the traditional opposition between content and function words in his treatment of the parts of speech. His list of content words includes nouns, pronouns, verbs, adjectives ('quiet words'), and adverbs ('situation words'); included under 'function words' are prepositions, conjunctions, particles ('helping words'), and exclamations; under parts of a sentence are subject ('beginning words'), direct object ('ending words'), predicate ('speech words'), adjectival, nominal and pronominal predicate ('expression words'), verb complements, including prepositional phrases ('transfer words'), object of a preposition ('operating words'), and appositives ('additive words'). The work is marked by repetition and inconsistencies; yet it is impressive for its time and made a (p. 220) notable contribution to introducing European grammatical traditions to China. Its influence is still felt today. Subsequent grammarians have mostly taken English rather than French and Latin as their model for analysing Chinese.

10.2.5 Philosophy and Language

Reference is made here to just one representative essay as a sample of linguistic thought in Chinese philosophy, the chapter entitled 'Rectifying the Names' *Zhèngmíng* 正名 in the philosophical work *Xún Zǐ* 荀子 (313–238 BC). Presaging Saussure, its main point is that while the forms of words are arbitrary, and differ from culture to culture, names like 'ruler,' 'subject,' 'father,' and 'son' were originally tied to very specific and unambiguous definitions, and can be matched quite precisely with corresponding terms in other languages, using Chinese as the common standard where needed. With the passing of time, however, these terms became blurred in the minds of the people, causing roles to be confused, and eventually leading to an inability to tell right from wrong. The author uses this as a foundation to emphasize the importance of matching words to their intended reality, and of society sticking to firm moral principles.

10.2.6 Choosing a National Language

The need to establish an official national language was felt already in the early Qing dynasty, when the Manchu government launched a number of 'orthoepy [correct pronunciation] institutes' to teach standard Beijing pronunciation, particularly in the Cantonese and Fukienese-speaking southern provinces. These met with very limited success (Chung 1989: 35). In 1913, scholar Wu Zhihui 吳稚暉 was chosen to direct the task of creating a

truly 'national' speech that would transcend locality and dialect. The Beijing dialect was the general foundation of the new National Language, but certain features of various local dialects were also incorporated; in fact, the pronunciation of each individual character was actually voted on by the Commission on the Unification of Pronunciation *Dúyīn Tǒngyīhuì* 讀音統一會 (Wu 1964: 7–30). The newly designated 'standard' pronunciations were recorded by linguist Yuen Ren Chao for public release on a vinyl record. The entire scheme, however, proved impractical, since Chao was the first to point out that he himself was the only one who could actually pronounce everything correctly in running speech (Qin 2011: 25). The Beijing dialect was subsequently adopted more or less wholesale as the basis of the new national language. This was of course easiest for the one-third or so of the country that already spoke a variety of northern Chinese, but the policy encountered more difficulties in areas where more divergent dialects are spoken, particularly in the southern provinces. Philologist Zhang Binglin 章炳麟, better known as Zhang Taiyan 章太炎, came up with a phonetic alphabet for Chinese, which later became the *Zhùyīn fúhào* 注音符號 (p. 221) 'Mandarin Phonetic Symbols' (MPS). The Ministry of Education put out a character list with standard pronunciations, called the *Guóyīn Jiǎnzì* 國音檢字, later used by Wu as a basis to compile a dictionary, in which the characters were arranged according to the 214 radicals.

Another major linguistic event in Republican China was the launching in 1919 of the 'May Fourth' movement, which incorporated writer Hu Shi's 胡適 promotion of the use of the vernacular *báihuàwén* 白話文 rather than the often stilted and unnatural Classical Chinese *wényánwén* 文言文 previously in use for most kinds of writing. The effects were great on literary writing, such as novels, but even more far-reaching for literacy education and language use in society in general.

A branch of the National Language Promotion Commission, originally founded in 1945, was established in Taiwan in 1946, a year after its retrocession to Chinese sovereignty, following 50 years as a Japanese colony. Mandarin was subsequently strongly promoted on both sides of the Taiwan Straits in education and all areas of public life—often to the detriment of local dialects. The Mandarin Phonetic Symbols are still used in Taiwan to teach literacy to children, while the PRC eventually adopted the Hanyu Pinyin Romanization system instead, which in fact closely parallels the MPS structurally, and is now pretty much the world standard for representing Chinese in the Latin alphabet.[1]

10.3 Korean

In the absence of written records from an earlier period, Korea's linguistic history can be said to have begun with the introduction of the Chinese language and writing system into the country, which started perhaps as early as the second century BC. Large-scale importation into Korean of Chinese characters, called *hanja* 漢字, began at the end of the seventh century AD, during the Unified Silla period. Korean previously did not have its own writing system, and in spite of the availability of other writing systems they could have borrowed from, such as Mongolian, Jurchen, and 'Phags-pa, the adoption of Chinese was

all but inevitable because of China's relatively high level of cultural development. Similar to the case of Japan, the Chinese writing system was very poorly matched to the needs of writing Korean. Chinese is an isolating, analytic language in which each syllable generally corresponds to a single morpheme and written character, while Korean is typologically an agglutinative language using cumulative suffixation of words.

(p. 222) In the earliest period, the Koreans learned and adopted the Chinese language itself, together with its writing system, for use in document writing, roughly analogous to how Latin was used in religion and scholarship by speakers of many various languages in Europe during the Middle Ages and beyond. Eventually they tried different approaches to using Chinese characters to write in the Korean vernacular, all variants of what was called the *Idu* script system. It was awkward at best. The three main approaches employed were analogous to the ones adopted for Japanese in the evolution of its own writing systems:

1. *Hyangchal* 鄉札 향찰 'vernacular letters': use of Chinese characters in their original form, mainly to transcribe *Hyangka* 'vernacular poetry.' Basic words in the text were mostly native Korean, Korean word order was used, and each syllable, for both content and function words, was transcribed with a single Chinese character. This system was used into the fourteenth century.

2. *Idu* proper 吏讀 이두 'clerk reading': Chinese characters were employed for basic words, Korean word order was used, and grammatical elements written in *Idu* characters were inserted. The main purpose of *Idu* was to clarify government documents written in Chinese for average Korean readers, and it was also used in teaching the Chinese language. It lasted into the nineteenth century.

3. *To* 吐 토 'particle': the original Chinese word order was retained, and additional *itwu* characters, often abbreviated from standard Chinese characters, were added to indicate Korean particles, verbal suffixes, and basic verbs not in the Chinese texts (Sohn 1999: 125–8).

These were clearly makeshift expedients. Korea badly needed an efficient writing system adapted to its own needs—and finally got one, in the year 1446.

The story of the invention and popularization of the Korean alphabet, known as the *Hanguel* 한글 (also written *Hangul* or *Hankul*), is the *pièce de résistance* of Korea's linguistic history. It has been called 'the only alphabet completely native to East Asia' (Kim-Renaud 1997: ix). It was originally intended as a writing system for the masses who were not literate in Chinese.

The *Hanguel*, originally called the *Hunmin Jeongum* 'Correct Sounds for the Instruction of the People', was promulgated in 1446 by King Sejong (born in 1397, reigned 1418–50) of the Joseon (Chosun) kingdom or Yi dynasty. The original system contained 28 letters; of these, only 24 are still in use, due to phonological change. The scholarly commentary that originally accompanied the *Hunmin Jeongum*, not discovered until 1940, describes the linguistic and philosophical underpinnings of the alphabet and its usage.

The system shows clear influence from the Sanskrit Devanāgarī script, which reached Korea via China through the transmission of Buddhism. Like the various Japanese scripts, the *Hanguel* were also clearly influenced by Chinese. However, unlike the Japanese syllabaries, in which syllables are added continuously in linear fashion, the *Hanguel* system chunks Korean into individual syllables and represents each one as a (p. 223) separate, integrated unit of initial (optional), vowel, and final (optional), presented as a square shape similar to that occupied by a single Chinese character on a page.

What is truly remarkable about the *Hanguel*, however, is that the forms chosen to represent the sounds are basically conceptual diagrams of the position of the articulatory organs used to pronounce each sound. For example, the letter for velar /k/ is ㄱ, which indicates the back of the tongue being raised to touch the soft palate; /n/ is written ㄴ, indicating the alveolar ridge, to be touched by the tongue tip. This represents a highly original and innovative concept in alphabet design that is unique among the alphabets of the world.

The new system was not universally embraced; it was viewed as too simple to be taken very seriously. Chinese held onto its cachet as the language and writing of prestige. The *Hanguel* thus came to be used only by 'women and commoners' up through the end of the nineteenth century.

Chinese characters continued to be an integral part of Korean up until very recently, in spite of the brief rise of nationalism in response to aggression by colonial powers, including Japan and Russia (Song 2005: 164). Their use has been entirely phased out in North Korea, where the *Hanguel* have been used exclusively since 1949; Chinese characters are taught only as a separate class in schools in the North. Most writing in South Korea is in *Hanguel* as well, though the study of Chinese character writing continues to be cultivated, and *hanja* are sometimes used in abbreviations or for disambiguation. The future role of the *hanja* in South Korea is as yet unclear.

Like Japanese, Korean has ended up with an enormous number of Chinese loanwords in its vocabulary. *The Comprehensive Dictionary of Korean* places the number of Sino-Korean words at about 52.1 per cent of its total 164,125 entries, pure Korean words at 45.5 per cent, and other loanwords, 1.4 per cent. Since the Second World War, however, English has been the main source of foreign loans into South Korean, accounting for about 90 per cent of the total. Many of these came into Korean in Japanized form, which the Koreans have tried to purge from their language since the end of Japanese colonial rule in 1945, either by nativizing them or by replacing them with direct loans from English. Some English loans are in fact homemade Korean coinages, as is also the case for Japanese (Song 2005: 19).

Finally, a note on Romanization. South Korea, which previously used a modification of the McCune–Reischauer (MR) Romanization system, promulgated the Revised Romanization of Korean in 2000. Opinions vary on the new system, which is more phonemic than the previous one, and eliminates diacritics which many readers didn't understand; but the

values of some of the new letters, like *c*, may not be immediately obvious to the uninitiated, and the concurrent use of competing systems can be confusing.[2]

(p. 224) 10.4 Japanese

Japan and Korea, in addition to the strong structural affinity between their spoken languages, also have roughly parallel histories as regards the introduction and adoption of Chinese writing, the subsequent tweaking of written Chinese for their own use, and ultimately, the development of their own scripts well adapted to their respective languages. And contact with western European ideas led in the cases of all three East Asian cultures to the development of grammatical studies, first only of earlier forms of their respective languages, and closely following European models, but eventually of the modern spoken language as well, and showing an increasingly greater independent development.

There is no record of a native Japanese writing system before the introduction of Chinese around 400 AD. In order to read early Chinese texts, the Japanese developed the *kanbun* or *kambun* 漢文 'Chinese texts/literature' system of markings around written Chinese characters to indicate Japanese word order, grammatical particles, suffixes, and inflectional endings, reminiscent of the various *itwu* schemes for Korean. Japan's own earliest literary works were generally composed directly in Chinese, with some concessions to the local language. Japan's earliest surviving written record, the *Kojiki* 'Record of Ancient Matters,' completed in 713 AD during the Nara period, is basically written in Chinese, but occasionally a sentence would be rendered phonetically in Japanese to preserve a personal or place name, a particular phrase, or a poem in its native form. The language is more precisely described thus: 'It is not Japanese, and at the same time it is not Chinese, but a quasi-Chinese which (to quote Chamberlain) "breaks down every now and then, to be helped up again by a few Japanese words written phonetically, and is surely the first clumsy attempt at combining two divergent elements" ' (Sansom 1928: 19). This was the beginning of a mixed phonetic and semantic use of Chinese characters to represent Japanese. A single character, like the Chinese possessive particle 之 *zhī*, was borrowed both for its meaning—which would be の *no* in Japanese—and for its sound, probably *shi*, useful for example in verb endings. Later, native Japanese morphemes and pronunciation were paired with Chinese characters, resulting in sequences of characters that looked odd in Chinese—because the text was in fact representing spoken Japanese.

Phonetic use of the characters was the rule in poetry, in which it was essential to maintain syllable count and rhythm. Japan's earliest collection of poetry, the *Man'yōshū* 万葉集 'Myriad Leaves Collection,' uses characters mainly as phonetic symbols, and it is fairly easy to reconstruct the Japanese of the time on this basis; but characters were also sometimes used for their meaning. Using Chinese characters for the meaning of content words, and for the pronunciation of grammatical endings, in theory should have put the Japanese well on their way to creating the *kana*, their own native script. However, in Japan as in Korea, reverence for the Chinese culture and language kept scholars more focused on studying Classical Chinese rather than developing a native (p. 225) script for a

vernacular literature. This meant that the Japanese had in the meantime to get by as best they could with a script full of pitfalls and gross inefficiencies, such as assignments of multiple Japanese equivalents to the same Chinese character (Sansom 1928: 7–40).

Parallel to how Chinese characters were often simplified in the Korean *to* system, cumbersome Chinese characters used phonetically in Japanese writing also began to be abbreviated, based on the 'grass' or cursive *cǎoshū* 草書 calligraphic style. According to Japanese tradition, Kobo Daishi 弘法大師, a famous priest who lived from 774 to 835 AD, chose 47 of these phonetic characters to form a syllabary called the *hiragana* 平仮名 'easy kana,' which went a long way to standardizing phonetic writing in Japanese. At about the same time, a second system came into use, called the *katakana* 片仮名 'side kana,' simplified from the square style of Chinese calligraphy. Chinese characters continued to be used, with either Chinese-based pronunciations, called *onyomi* 音読み, or Japanese pronunciations, called *kunyomi* 訓読み, mostly for content words, and Japanese grammatical syllables were written in *kana*, called *okurigana* 送り仮名, when used in combination with Chinese characters in the same word, e.g. the *mi* み in *kunyomi* 訓読み. Small *kana* written on one side, called *furigana* 振り仮名, were and still are used to indicate the pronunciation of the characters where needed. Chinese characters were often re-adopted in different historical periods, with different pronunciations, due to phonological change in Chinese; the three main layers of these varying pronunciations are (1) *go-on* 呉音, the older pronunciation of Chinese characters of the Nara and Heian periods used in particular for Buddhist and legal terms; (2) *kan-on* 漢音, readings of the Chang'an area borrowed from the seventh to ninth centuries, in the Nara period; and (3) *tōsō-on* 唐宋音, dating approximately to China's Song dynasty (960–1279 AD) and the mid-Heian to Edo periods in Japan. While the layers of inconsistencies made Japanese writing more and more complex, this in fact increased the attraction of the system for the elite who mastered it, as was also the case in Korea (Sansom 1928: 41–6).

Similar to China, though starting several decades earlier, grammatical study of Japanese began under the influence of intense contact with western European ideas in the period leading up to the Meiji Restoration (1868–1912). Japan maintained connections to the West via a Dutch outpost near Nagasaki, and in this way gained access to European learning, for which it had a ravenous hunger. Fujibayashi Fuzan 藤林普山 published a verbatim translation into Japanese of a Dutch grammar in 1815. Shinto priest and scholar Tsurumine Shigenobu 鶴峰戊申 applied the categories of Dutch to early Japanese in his *Gogaku Shinsho* 語学新書 'New Book of Language Study,' published in 1833, which described, for example, governing (nominative) and governed (local) cases. As with Chinese, the earlier language is what was considered most worthy of study. Toward the end of the nineteenth century, however, grammarians had begun focusing their attention on the modern vernacular rather than just the language of the ancient *waka* 和歌 poetry, and started incorporating the notions of parts of speech and other European grammatical concepts into their grammars. The first grammar of spoken Japanese written by a Japanese author is Baba Tatsui's 馬場辰 (p. 226) 猪 *Elementary Grammar of the Japanese Language*, published in English in London in 1873. This was followed by Tanaka Yoshikado's 田中義廉 *Shōgaku Nihon Bunten* 小学日本文典 'Elementary Japanese Grammar,' published in 1875. Ōtsuki Fumi-

hiko 大槻文彦 based his *Kō Nihon Bunten*広日本文典 'A Comprehensive Grammar of Japanese,' published in 1897, on both early indigenous studies and Western grammar; and many more similar works followed (Frawley 2003: 195–6). After the Meiji Restoration, Japanese grammar became a required subject for students at all levels, from elementary school through university (Liu 1993: 11).[3]

Notes:

(1) Additional sources: Chao (1965), Chung (1995), Itkonen (1991), Karlgren (1954), Li (1975), Pu (1990), Shao (2005), Wang (1974), Zhou (1988). A very useful list of influential Chinese linguists, particularly more recent ones, with links to articles on each: en.wikipedia.org/wiki/Category:Chinese_linguists. On Chinese etymology www.chineseetymology.org/CharacterEtymology.aspx

(2) Additional sources: Kim (1987), en.wikipedia.org/wiki/Korean_romanization; en.wikipedia.org/wikiRevised/_Romanization_of_Korean.

(3) Additional sources: Frellesvig (2011), Heinrich (2012), Shibatani (1987, 1990).

Karen Steffen Chung

Karen Steffen Chung (史嘉琳 Sh_ Jialín), originally from St. Paul, Minnesota,USA , has taught English and linguistics in the Department of Foreign Languages and Literatures of National Taiwan University since 1990; current rank: Associate Professor. Education: BA, East Asian Languages, University of Minnesota, 1976; MA, East Asian Studies, Princeton University, 1981; PhD, Linguistics, Universiteit Leiden, 2004. Dissertation title: Mandarin Compound Verbs.