Cantonese Dialect-Writing and Korean Goyuhanja: Chinese Characters and Innovative Orthographic Creations*

Marjorie K.M. Chan[1] and Seojin Yang[2]
The Ohio State University[1] and Duke University [2]

Abstract
The Chinese writing system, spanning over three millennia, has enjoyed the longest, uninterrupted literary tradition in world history. At the same time, it had also impacted neighboring countries such as Korea, Japan, and Vietnam. However, the script is not static; time moves on. Some complex characters were later replaced by simpler forms or by forms that add phonetic cues. Some characters became obsolete and new ones were created, reflecting the standard, or national, language, and are entered into authoritative dictionaries for nation-wide distribution and consultation. Orthographic forms that are of regional origin may occasionally be included in such dictionaries but are explicitly marked as dialectal forms. This study analyzes the Sinitic graphs that were created by Cantonese and Korean speakers, focusing on the methods of creating the vernacular graphs in Cantonese and the Goyuhanja (固有漢字) graphs in Korean. Other related issues are also addressed.

Key words
Chinese, Cantonese, Yue dialect, dialect-writing, vernacular graphs, orthography, script, Korean, Hanja, Gukja, Goyuhanja, Hangul, six principles, phonetic loan, semantic loan

* This paper is an extension of research initially presented at the 69th Midwest Conference on Asian Affairs (MCAA 2020) that was held virtually on 16-18 October 2020. Our thanks to feedback from our grad students in the Department of East Asian Languages and Literatures at OSU during our dry run, and from attendees at our MCAA 2020 panel.

1. Introduction

The Chinese writing system has the distinction of spanning over three millennia and has, thus, enjoyed the longest, uninterrupted literary tradition in world history. Nonetheless, despite the rich inventory of Chinese characters—for example, already 47,035 in the *Kangxi Dictionary* of 1716 (Norman 1988:72)—it is still inadequate, or insufficient, for transcribing the vernacular language of regional varieties of Chinese, such as Cantonese.

The script is even less sufficient for transcribing other, non-Sinitic, East Asian languages. A case in point is Korean, which adopted the Chinese writing system (Hanja 漢字) in the decades and centuries after Emperor Wu (r. 141-87 B.C.E.) invaded and conquered the northern part of the Korea Peninsula. The administration of four Chinese commanderies (Hàn Sìjùn 漢四郡) would unquestionably have made use of the Chinese writing system, and from there the eventual introduction of Chinese literary culture and the ultimate growth and development of Sino-Korean vocabulary written in Hanja (漢字).

In this paper, we will explore the Sinitic graphs that were created by Cantonese and Korean speakers, focusing on the methods used by Cantonese speakers for creating vernacular graphs for morphemes that do not exist in the standard Chinese language, and strategies used by Koreans to devise a set of Goyuhanja (固有漢字) graphs that are separate from standard Chinese graphs; that is, the Hanja (漢字) graphs.

Crucial questions addressed here fall under three topics pertaining to form, function, and time. Regarding form, the following are the questions to be addressed: What methods have been used for creating and adding new vocabulary items? And do those methods fall under the rubric of Han dynasty scholar Xu Shen's (許慎) Six Principles (Liù Shū 六書) in analyzing Chinese characters in his *Shuōwén Jiězì* (說文解字)? On function, the crucial question is: What have been the functions, or uses, of these newly created characters for their respective speech communities? And last, regarding time, an important historical question is: How far back can we trace the creation of those Sinitic graphs that are not part of the standard repertoire of Chinese characters?

For the remainder of this paper, the organization is as follows. Section 2 deals with Cantonese dialect-writing. Section 3 explores Korea's adaptation of Chinese characters. And Section 4 closes the paper with some concluding remarks.

2. Cantonese Dialect-Writing

The earliest extant source that we have of dialect-writing for Cantonese is a long love ballad, the *Huājiān Jì* (花箋記, *Romance of the Fancy Note-Paper*. Dating back to the late Ming dynasty (1368-1644),[1] there exists a number of Qing dynasty and later editions, with the earliest, complete extant one being the Jìng Jìng Zhāi (静净齋) edition. Entitled *Dì Bā Cáizǐ Huājiān* (第八才子花箋), this anonymous work has a preface date of the 52nd year of the Kangxi Emperor, or 1713. The *Huājiān Jì* belongs to the *mùyúshū* (木魚書) 'wooden-fish book' (wooden-fish) genre that was written and sung in

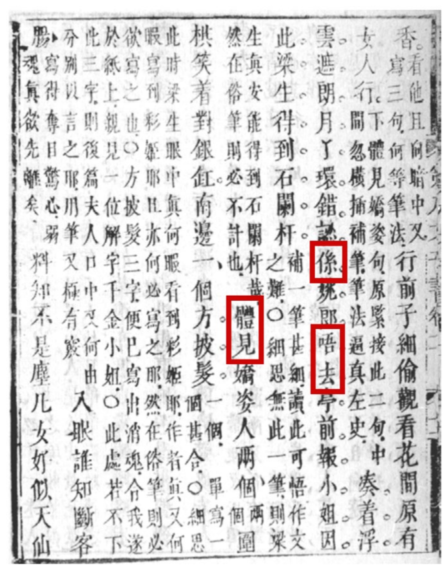

Figure 1. A page from the 1713-prefaced *Huājiān Jì*

[1] Li (1999) reports on the existence of a fragmentary copy of a (late) Ming dynasty edition of this ballad, thus providing evidence of its Ming dynasty origin predating the 1713-prefaced edition. Moreover, Thoms (1824:v), in his English translation of the ballad with bilingual, Chinese-English text, explicitly identifies it as a Ming text.

Cantonese, and more broadly, to the *shuō-chàng* (說唱 'speak-sing'), oral performance tradition in China.[2] Although the work has a great deal of literary language, vernacular Cantonese can still be readily seen, such as 唔去 *m4 heoi3* 'not go' (= *búqù* 不去), shown in Figure 1, from the original page housed in the Bibliothèque Nationale de Paris, where the negative form appears using a vernacular graph, 唔 *m4* 'not, and 冇 *mou5* 'not have.'[3]

In this 1713-prefaced edition of the *Huājiān Jì*, the lexicon that reflects dialect-writing can be analyzed as involving the use of three main strategies, as shown in (1).

(1) Three main dialect-writing strategies in the *Huajian Ji*

A. New compound graphs (*xíngshēng* 形聲 'semantic-phonetic compounds')
B. Stroke modification (*zhuǎnzhù* 轉注 'deflected characters'[4])
C. Borrowing of existing graphs: Phonetic borrowing (*jiǎjiè* 假借 'phonetic-loan characters')

These three categories of the Chinese writing system all fall within the Six Principles. Not included in the *Huājiān Jì* are: 1. *zhǐshì* (指事 'self-explanatory characters, simple indicative characters,' 2. *xiàngxíng* (象形 'pictograph characters,' and 3. *huìyì* 會意 'associative compounds, semantic-semantic compounds.' Some examples for each of the three strategies in (1) are given in Section 2.1 below.

2.1 Dialect-writing strategies and examples from the 1713-prefaced *Huājiān Jì*

Examples from the four strategies in (1) are presented in this subsection, beginning with the creation of new compound graphs, which consists of a semantic component and a phonetic component, given in (2). Cantonese dialect-writing frequently adds the 'mouth' 口 radical to phonetic components to form new graphs for its vernacular, spoken vocabulary, a strategy that has a long history, as can be attested by its use in the *Huājiān Jì* in (2).

(2) A. New compound graphs (*xíngshēng* 形聲): Two examples

1. 唔 *m4* 'not' (口 + 吾)
2. 吤 *go2* 'that' (口 + 个) = 嗰

Note that the graph for 'that' in the second example in (2) is not included in regular Unicode fonts. The graph that is conventionally used in written Cantonese is 嗰, which combines the 'mouth' (口) radical with the phonetic component, *gè* 個. The phonetic components, *gè* 个 and *gè* 個 in (2)

---

[2] See Chan (2022) for a brief study of the *Huājiān Jì,* as precursor to Cantonese opera's role in the spread of vernacular written Cantonese in the 20th century.

[3] Mandarin-based Pinyin romanization is placed before the graphs, while modern Cantonese pronunciation, using Jyutping, is placed after the graphs. In Jyutping, Tone 1 designates Yīnpíng (陰平), Tone 2 Yīnshǎng (陰上), Tone 3 Yīnqù (陰去), and so forth.

[4] This category is controversial (Norman 1988:69), and is difficult to give a precise translation due to the uncertainty of this category in the *Shuōwén Jiězì.* The translation chosen here is 'deflected character,' with other translations include 'transfer character, mutually explanatory character,' etc. Crucially, this category is semantics-based, with modification of the base graph to create the derived graph.

were historically variant graphs that today are treated as belonging to different character sets, 个 as part of simplified characters, and 個 as part of traditional characters.[5]

In (3), only one example is provided for stroke modification, in treating the creation of 冇 *mou5* 'not have' as a case of *zhuǎnzhù* 轉注 'deflected characters.' Its creation involves the deletion of the two horizontal strokes in *you* 有 *jau5* 'have' to form its negative, 冇 *mou5* 'not have.' Here, we find a clever, iconic strategy to represent the opposite meaning of the base graph, *you* 有 'have' by removing two strokes.[6]

(3) B. Stroke modification (*zhuǎnzhù* 轉注): One example
   1. 冇 *mou5* 'to not have' (derived from *yǒu* 有 *jau5* 'to have')

Graphs formed by stroke modification of semantically-related graphs are rare in Chinese. *Xiǎo* 小 *siu2* 'small' and its derived graphs, *shǎo* 少 *siu2* 'few, little, less' and *jié* 尐 *di1* 'few, small amount' may constitute one such set.[7]

The strategy with the largest set of examples is that of phonetic loan, *jiǎjiè* 假借, with ample examples presented in (4). Examples (4-1) and (4-2) are highlighted in Figure 1.

(4) C. Phonetic borrowing (*jiǎjiè* 假借): Many examples
   1. 體 *tai2* 'to look at, to see' e.g. 體見 *tai2 kin3* (= *kànjian* 看見) 'to see'
   2. 係 *hai6* 'to be'
   3. 渠 *keoi5* 's/he' (渠 *keoi4* 'canal, ditch')[8]
   4. 地 *dei6* '(plural suffix)' e.g., 渠地 *keoi5* dei6 'they'
   5. 乜 *mat1* 'what, why; so, all that' 乜 *me1* 'to squint'
   6. 個 *go2* 'that' (個 *go3* '(general classifier)'
   7. 泔 *gam2* 'so'
   8. 哩 *li1* 'this'
   9. 的 *di1* '(plural suffix)' e.g., 個的 *go2 di1* 'those'
   10. 邊 *bin1* 'who, where' e.g., 邊處 *bin1 cyu3* 'where'
   11. 番 *faan1* 'to return' Also: 反 *faan1* 'to return'
   12. 企 *kei2* 'to stand'
   13. 重 *zung6* 'still, yet'

---

[5] Since mainland China uses simplified characters and therefore does not use the traditional graph, 個, one often sees online an alternative form used to display 'that' in Cantonese. To wit, the graph, *guǒ* 果 'fruit,' is used in lieu of 嗰 for 'that' in Cantonese, treating *guǒ* 果 'fruit' as a phonetic loan for spelling 'that' in Cantonese. For speakers whose Cantonese has lost the labialization on velar stops, 嗰 'that' and 果 'fruit' are homophonous: *go2.*

[66] The example cited in the *Shuōwén Jiězì* involves a pair of graphs, *lǎo* 老 'old' and its derived graph, *kǎo* 考 'old' that shares its meaning of 'old.' Hence, *zhuǎnzhù* 轉注 has sometimes been translated as 'mutually explanatory characters.' Here, we interpret this category as involving semantically-related pairs, regardless of whether they are synonyms or antonyms involving some modification of the strokes of the base character. An established, antonymous pair that might typically be analyzed as pictographic could potentially also be categorized as a pair of deflected characters: *āo* 凹 *nap1* 'concave' and *tū* 凸 *dat6* 'convex.'

[7] Written Cantonese today often spells 尐 *di1* using 的 or 啲, or use upper-case letter D.

[8] On *qú* 渠 *keoi4* 'canal, ditch' as a phonetic loan for the third-person pronoun, see Chan (forthcoming).

14. 估 gu2 ‘to guess, to think, to presume’
15. 添 tim1 ‘(emphatic particle)’
16. 埋 maai4 ‘(ASP marker)’

2.2 Semantic-borrowing example in an 18th century Sino-Portuguese glossary

One strategy that is not shown in (1) and is not part of Xu Shen’s Six Principles—but can be illustrated from another early source—is semantic borrowing, *xùndú* (訓讀). This strategy involves the borrowing of an existing graph that is synonymous with the characterless, indigenous morpheme. This is rarely discussed in the Chinese linguistic literature, although its application has been known since at least the 19th century, albeit nameless at that time (e.g., Williams 1856). It was not until mid-20th century that a renowned Chinese linguist, Zhan Bohui (1957), first applied the term, *xùndú* (訓讀), for its usage in Chinese dialects.[9] In contrast to the late recognition of this strategy for writing characterless words in Chinese dialects, the semantic borrowing of a synonym is a crucial part of reading and writing in Japanese. The term *xùndú* (訓讀), in fact, comes from Japanese *kundoku* (訓読), or *kun’yomi* 訓読み), where Chinese characters, *kanji* (漢字), may have at least two readings: a Sino-Japanese reading, such as in Kan-on (漢音) or Go-on (呉音), and a reading based on the pronunciation of a native Japanese morpheme, thus merely borrowing the *kanji* in lieu of creating an indigenous graph for the native vocabulary. An example is *shān* 山 ‘mountain’ from Chinese, with a Sino-Japanese Kan-on reading of 山 as *san*, and a native Japanese reading of *yama*, as in Mount Fuji—Fujisan or Fujiyama—both readings written with the Chinese characters, 富士山.

For Cantonese, a modern Cantonese example of a semantic loan is *wāi* 歪 ‘slanted, askew,’ with two readings: one based on standard Chinese, *waai1*, and one involving semantic loan of the graph for the indigenous morpheme in Cantonese: *me2*. Thus, the same principle as in Japanese is at work here. As to whether semantic loans occurred in the *Huājiān Jì*, it would be nigh impossible to ascertain if semantic loans were used in that early source. We simply do not have recordings to determine if a regular graph was sung based on an indigenous morpheme that is not cognate to the graph displayed in the published source. What we do know, from examples in (4) is that phonetic loans were frequently employed to render an indigenous Cantonese lexical item, even in the use of *tǐ* 體 *tai2* ‘body,’ as in 體見 *tai2 kin3* (= *kànjian* 看見), shown in Figure 1 in the *Huājiān Jì* for the meaning of ‘to look at, to see.’ Later, 體 co-occurs with *dì* 睇 *tai2* (e.g., Thoms 1824), but is subsequently completely replaced by *dì* 睇 in later editions of the *Huājiān Jì* and into the modern era.[10] The ultimate adoption of 睇 *tai2* is understandable, given the suitability of the graph in incorporating the ‘eye’ (*mù* 目) semantic component in *dì* 睇 t*ai2*, with its original meaning of ‘to look askance.’

---

[9] Liu (2007:91) explicitly cites Zhan (1957) as the first to use this term for the semantic borrowing of synonyms in Chinese dialects.

[10] Contra K.C. Leung (1999: 272, fn. 7) who had not consulted early editions of the *Huājiān Jì*, 體 was *not* some careless misprint in Thoms (1824), where both 體 and 睇 appear. In addition to the 1713-prefaced *Huājiān Jì,* 體 was, in fact, consistently used in another *Huājiān Jì* edition, *Xiùxiàng Dì Bā Cáizǐ Zhù* (繡像第八才子註, IHP collection). The presence of the two homophonous graphs in Thoms (1824) and in Morrison’s (1828) three-volume work, *A Vocabulary of the Cantonese Dialect,* suggests that at least up through the early 19th century, the ultimate choice of 睇 for ‘to look at’ in written Cantonese had not yet been firmly established.

An early example in the use of semantic loans in written Cantonese can be found in a mid-18th century, Sino-Portuguese glossary that is placed at the end of the *Àomén Jì Luè* (澳門記略 'Sketch notes on Macao'), written by Yin and Zhang (1751/1968).[11] The morpheme is *zī* 孖 'twins,' which can be found in the Song dynasty rhyme dictionary, *Guǎngyùn* (廣韻). The Cantonese colloquial word for 'twins,' *maa1*, is written using *zī* 孖, a semantic-semantic compound graph formed by the *huìyì* (會意) principle, wherein the graph is composed of the doubling of *zǐ* 子 *zi2* 'child' to form the morpheme for 'twins.' The morpheme *maa1* is clearly not a cognate of *zī* 孖, but is a case of semantic loan, the borrowing a graph with the same meaning for a Cantonese colloquial word. We know that *zī* 孖 was pronounced as [ma] in that Sino-Portuguese glossary, and not [tsi], because of what that graph is used to transliterate, namely, Portuguese /ma/, as exemplified by (5-1), and illustrated in Entry 152 in Figure 2, a page from the 1751-prefaced source.[12] In addition, 孖 is also used for Portuguese /ba/, as shown in (5-2). The use of 孖 [ma] for Portuguese [ba] is explained in Chan (1994) as involving nasal post-stopping; that is, nasal onset with oral release [m^b] in the source dialect, Zhongshan Cantonese. The bilabial nasal is transcribed as [m^b] in (5-2) as a reminder that /m/ may include oral release in Zhongshan Cantonese. The numerical tone values in (5) use Chao's (1980/1930) system of tone numbers, with '1' to '5' for a pitch range from lowest to highest.

(5) An example of semantic borrowing (*xundu* 訓讀) using *zī* 孖 *maa1* 'twins' in the 18th c. Sino-Portuguese glossary

1. Entry 152. 乳 'breast' — 孖 麻
   Modern Portuguese: ['ma ma]
   Modern Zhongshan Cantonese: [ma.55 ma.51]
   Modern Hong Kong Cantonese: [ma.55 ma.21]

2. Entry 287. 煙 'tobacco' — 大 孖 古
   Modern Portuguese: [ta 'ba ko]
   Modern Zhongshan Cantonese: [tai.22 m^ba.55 ku.13]
   Modern Hong Kong Cantonese: [tai.22 ma.55 ku.35]

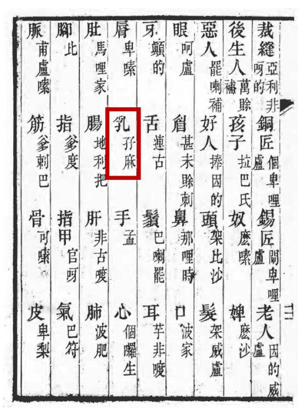

Figure 2. Entry 152: 乳 孖麻

Even today, 孖 *maa1* is used in Cantonese for transliterations. 孖結 *maa1 git3,* for example, has long been used for 'market' in English. As a result, one finds on the internet Cantonese-owned shops in California with such bilingual names as Good Luck Market (廣和孖結 *gwong2 wo4 maa1 git3*), and Man Sung Market (萬生孖結 *maan6 sang1 maa1 git3*).

[11] See Bawden (1954), Boltz (1977), and Chan (1982, 1994) for more information on the *Àomén Jì Luè* (preface date of Qianlong 16 (1751)). Bawden (1954:16) suggests that there is an unnamed third contributor to this glossary, since the two authors came to their government posts from outside Guangdong Province and would not have been native Cantonese speakers. The Sino-Portuguese glossary contains 395 entries covering five different topics (1. celestial and territorial matters (Entries 1-83), 2. human and natural matters (Entries 84-244), 3. clothing and food (Entries 245-296), 4. utensils and numerals (Entries 297-344), and 5. common functions (Entries 345-395). The popularity of the *Àomén Jì Luè* spawned later editions, including one published in 1800 and another in 1880.

[12] The page in Figure 2 is from Yin and Zhang's original woodblock printing of the *Àomén Jì Luè,* with preface date of year 16 of the Qianlong Emperor (1751), republished in 1968 by Guangwen Shuju, in Taipei, Taiwan.

2.3 Post-19th century orthographic innovations

Since the 18th century, many more Cantonese dialect characters have been created, with strategies that fall into four categories, as listed in (6), with the most recent creations involving different ways to mix scripts.

(6) Post-18th century graphic innovations built on four categories of strategies
  A. Compounding (semantic-phonetic and semantic-semantic compounds)
  B. Borrowing (phonetic and semantic loans)
  C. Translanguaging: Kongish (Hong Kong Cantonese-English mix)
  D. Tranßcripted graphs (with mixing of scripts)

Regarding (6-A), we have already encountered compound graphs, be they semantic-phonetic compounds formed by the *xíngshēng* (形聲) principle, or semantic-semantic compounds formed by the *huìyì* (會意) principle. Examples of these two types of compounds are presented in (7), where SFP abbreviates 'sentence-final particle,' ASP abbreviates 'aspect,' and PRT abbreviates 'particle.' In (7-1), graphs with the 'mouth' (*kǒu* 口) or 'person' (rén 亻) semantic components are chosen, with the 'mouth' component the most productive for creating new Cantonese characters. Only a small subset of examples are given here. Because the 'mouth' radical implies speech, it is particularly suitable for creating graphs for the spoken, vernacular language. In (7-2), a more irregular type of semantic-semantic compound is displayed here, involving multiple-component graphs (*hétǐzì* 合體字). Two quite recent, innovative graphs have been created, with neither available in Unicode. Earlier creations in the Chinese language includes *béng* 甭 'to not need to,' from the fusion of *bù* 不 'not' + *yòng* 用 'to use.' The example in (7-2b) is seen on the internet; its actual pronunciation is not certain as to whether it is monosyllabic or disyllabic (i.e., a case of *héwén* 合文 graphs, which are multiply-complex polysyllabic graphs. Perhaps the best known one is the auspicious, four-character phrase seen during Chinese New Year (Spring Festival), namely, *Zhāocáijìnbǎo* 招財進寶 'Bring in wealth and treasure.' That four-character phrase is combined into one complex graph that is, nonetheless, read as a quadrisyllabic phrase.'[13]

(7) A. Compounding
  1. Semantic-phonetic compounds with the 'mouth' (*kǒu* 口) or 'person' (rén 亻) semantic component
    a. 啲 *di1* 'some, (plural suffix)' (口 + 的)
    b. 㗎 *gaa3* '(SFP)' (口 + 架)
    c. 咗 *zo2* '(Perfective ASP)' (口 + 左)
    d. 嘅 *ge3* '(subordinative PRT)' (口 + 既)
    e. 嚟 *lai4* 'to come' (口 + 黎)
    f. 佢 *keoi5* 'he/she' (亻 + 渠)
  2. Semantic-semantic compounds (multiple-component graphs)
    a. 黑警 *hing2* 'gangster cops' (黑 *hak1* 'black' + 警 *ging2* 'cop')[14]

[13] For some other complex graphs, such as those used during the 2019-2020 Hong Kong protests and earlier, see Gomes and Chan (2022).

[14] A slightly different pronunciation is given by Hui (2019), as she assigns Tone 5 (Yángshǎng 陽上) to 警; hence, the pronunciation of the fused form that she gives is *hing5*.

b. [illegible] *bye* (?) ‘bye, good-bye’ (再 *zoi3* ‘again’ + 見 *gin3* ‘to see’)[15]

The next set of examples in (8) involves borrowing, be it phonetic loans (8-1) or semantic loans (8-2). The graph in example (8-1d), 唻 *lai4* ‘to come,’ is far less frequently seen than the homophone, 嚟 *lai4* in (7-1e), although it has shown up in subtitles of song lyrics, for example.

(8) B. Borrowing

1. Borrowing via phonetic loans

a.	蚊	*man1*	‘dollar’	蚊	*man1*	‘mosquito’
b.	搵	*wan2*	‘to seek’	搵	*wan3*	‘to press with one’s fingers’
c.	蝦	*haa1*	‘to bully’	蝦	*haa1*	‘shrimp’
d.	唻	*lai4*	‘to come’	唻	lai4	‘(no specific meaning)’
e.	咪	*mai5*	‘don’t’ (= *bié* 别)	咪	*mai1*	‘meow’
f.	咪	*mai1*	‘mike, microphone’			
g.	媽咪	*maa1 mi4*	‘mummy, mommy’			
h.	屋企	*uk1 kei2*	‘home’	*qǐ* 企	*kei2*	‘to stand (on tip toe/up)’

2. Borrowing via semantic loans

a.	凹	*nap1*	‘concave, dent’	*āo* 凹	*aau3*	‘concave, dent’
b.	凸	*dat6*	‘convex, protruding’	*tū* 凸	*gu2*	‘convex, protruding’
c.	來	*lai4*	‘to come’	*lái* 來	*loi2*	‘to come’
c.	甩	*lat1*	‘to fall/come off’	*shuǎi* 甩		‘throw (off), leave behind, fling’
d.	戇	*ngong6*	‘simple, stupid’	*zhuàng* 戇	*zong3*	‘simple, stupid’

The last two categories in (6), translanguaging and tranßcripting, are 20th and 21st century innovations. Both deal with mixing of scripts. Translanguaging in the written context involves the mixing of another script into a Chinese text; such as using the English letter ‘D’ in (9-1), in lieu of 啲 *di1* in (7-1a), as a phonetic loan across orthographic systems. Included in (9) are phonetic loans from English letters of the alphabet, the earliest of which is the use of ‘D’ for 的 *di1* and 啲 *di1* ‘some, (plural suffix),’ already observed and analyzed by Bauer (1982). One of my earliest recollections of seeing the use of the letter ‘D’ (later, the lower-case ‘d’ have also been used) was the informal transcribing of the song title and lyrics of a catchy, wildly popular Cantopop song that uses the melody from the Beatles’ 1964 song, *Can’t Buy Me Love*. That Cantopop song, written and sung in 1965 by Shangguan Liuyun (上官流云), bears the title, 行快的啦 (*Haang4 faai3 di1 laa1* ‘Hurray up and walk faster’).[16] Its popularity was further extended by Samuel Hui (許冠傑) and his brothers’ performance of the song in the 1970s. The song’s popularity continued into the 21st century. Over the years, other orthographic forms of the title appeared, namely, 行快啲啦, 行快 D 啦, and 行快 d 啦. Interestingly, searching using any of the four “spellings” on YouTube retrieves this song. Obviously, the Internet has learned to recognize these orthographic variants, 的, 啲, D, and d! The letter ‘D’ has also shown up later in the lyrics of videorecordings of the lighthearted operatic song, 沙三少情挑銀姐 (*Saa1 Saam1 Siu3 cing4 tiu1 Ngan4-2 Ze2*

[15] We do not know if (2b) is monosyllabic, or if it is read in its full, disyllabic form.
[16] See Chan and Chey (2013) for an analysis of this song.

'Third Master Sha dallies with Miss Silver'), that was originally part of a 1951 Hong Kong Cantonese operatic film.

Tranßcripting in (6-D), in contrast, incorporates foreign elements into the Chinese graph itself, such as using lower-case 'o' for the 'mouth' (*kǒu* 口) semantic component, producing o地 instead of 哋 in (7-1a).

(9) C. Translanguaging (using numbers and English letters)

1. D ~ d *di1* 'some, (plural suffix)'
2. E 卜 *ji1 buk1* 'e-book'
3. E 家 *ji1 gaa1* 'now'
4. JJ *zei1 zei1* 'male body part (young boy)'
5. C 君 *si1 gwan1* 'C Kwan' 鄭<u>詩君</u> *zeng6* <u>*si1 gwan1*</u> <u>(HK hip hop artist)</u>
6. 6 永 *luk6 wing5* 'Luk Wing' <u>陸永</u>權 <u>*luk6 wing5*</u> *kyun4* (HK hip hop artist)
7. Ching → a. C hing
   → b. 師 兄
   → c. *si1* *hing1* 'brother'

Examples (9-5) and (9-6) involve playful use of translanguaging for the names of two Hong Kong hip hop singers who form a duo for their band, Fama (農夫 *nung4 fu1* 'farmer'). Observe that the Cantonese pronunciation of the number '6' and the surname, Lu 陆 *luk6*, are homophonous, thus, a clever case of translanguaging. Examples of translanguaging are given in (9), and those for tranßcripting are given in (10). The example in (9-7) is from Kongish (Hong Kong Cantonese-English mix). Looking only at *Ching*, one is baffled as to exactly what is being conveyed. The mystery is solved when one breaks up that string of letters into two separate units, '*C*' (phonetically [si]) and '*hing*'. Reading those two syllables aloud, one realizes that the Kongish form, *Ching*, hides a homophonous, disyllabic word in Cantonese: 師兄 *si1 hing1* 'brother'!

In (10), three examples are given with 'o' for the 'mouth' semantic component, with (10-3) appearing online with 'o' attached to either the traditional graph, 無, or its simplified form, 无. Example (10-4) is among the examples discussed in Li and Zhu (2019).[17]

(10) D. Tranßcripting

1. o地 *di1* 'some, (plural suffix)' (o + 地)
2. o既 *ge3* '(subordinative PRT)' (o + 既)
3. o無 ~ o无 mou5 'to not have' (o + 無, o + 无)
   wú 無/无 *mou4* 'to not have'
4.  'Hong Kong independence' *dú* 獨 *duk6* + HK
5.  '(an expletive)' (*J* is from *zhuī* 追 *zeoi1* 'chase')
   Graphic variants: J左 ~ J咗 ~ Jo左

In summary, there have been a number of different strategies used over the centuries in rendering spoken Cantonese into written form, including extension to using mixed scripts in the

[17] Li and Zhu (2019:156) observes the association of the tranßcripted graph for 'Hong Kong independence' with "the anti-China political movement." Also see additional tranßcripted graphs in Gomes and Chan (2022).

20th and 21st century. The next section focuses on the adaptation of Chinese orthography to handle the Korean language.

3. Korea's Adaptation of Chinese Characters

Since the introduction of the Chinese script, it has remained indispensable to Koreans as a means of writing and recording. The precise historical period when the Chinese script was introduced to Korean culture cannot be determined from historical records. According to Kim (1990), during the Three Kingdoms period (57 BC to 668 AD), written characters, or writing systems, were not used in Korean society. However, as contact with Chinese culture increased over time, some Koreans began to comprehend the use of Chinese graphs. As the three Korean kingdoms expanded their control over the peninsula, the administration of the four Chinese commanderies came to an end. Even in the absence of direct Chinese cultural influence on the Korean peninsula, there are records indicating that education in the Chinese writing system was already underway in Goguryeo and Baekje during the 3rd and 4th centuries (Kim 1990). When Chinese characters and the Chinese writing system began to be used among the Koreans during the Three Kingdoms period, they were employed in two primary ways based on their application, as shown in (11).

(11) Two primary methods of using Chinese graphs:
  a. 吏頭 Idu[18]
    - 鄕札 Hyangchal, 吏頭 Idu,[19] and 口訣 Gugyeol
  b. 漢文 Hanmun[20]

In (11a), Idu (吏頭) is a term that, in a broad sense, refers to the method of using Chinese graphs to represent the sounds or meanings of Korean words (借字表記). It was primarily used to record Korean personal and place names by borrowing the sounds and meanings of Chinese graphs. Idu can be further categorized into three subtypes: Hyangchal (鄕札), Idu (吏頭), and Gugyeol (口訣). Hyangchal was used to document the popular songs of Silla known as Hyangga (鄕歌). It involves borrowing the sounds and meanings of Chinese graphs to record Korean in accordance with the word order of Korean language. Idu (吏頭), in the narrow sense, involves rearranging texts originally written in classical Chinese. Idu follows Korean grammar rules and incorporates Korean particles and verb endings, in borrowing the sounds and meanings of Chinese graphs. Due to the blending of Korean and Chinese elements, Idu does not represent a pure Chinese literary style, nor can it be considered purely Korean phrases. On the other hand, Gugyeol is a method of adding particles, verb endings, and pronunciation markers to classical Chinese texts to make them more easily understood and recited. While Idu mostly reorganized sentences to adhere to Korean grammar rules, Gugyeol retained the original Chinese writing style, making it understandable and usable only by scholars and the upper class who were literate.

While the use of Idu enabled the representation of meanings and pronunciations of Korean words, this alone was not sufficient as a means of recording. Therefore, intellectuals and the upper

[18] The term Idu (吏頭) used here refers to a broad category encompassing all methods of borrowing Chinese graphs for representation of Korean words（借字表記）. Originally, Idu, along with Hyangchal and Gugyeol, was one of the methods for borrowing Chinese graphs. However, from the early Joseon Dynasty onwards, Idu became used as a broad, general term representing all forms of borrowing of Chinese graphs.

[19] In the narrow sense of Idu (吏頭).

[20] Literary Chinese, Classical Chinese writing.

class used the Chinese writing system to record information in Chinese. This is referred to as Hanmun（漢文）in (11b), which means writing in Chinese graphs. Hanmun was the official means of recording, and was widely used by the government and intellectuals during the Three Kingdoms, Goryeo, and Joseon eras. These two methods of using Chinese graphs—Idu and Hanmun—continued even after King Sejong created *Hangul* in 1443. This was because Hanmun remained the preferred script for intellectuals and the upper echelon of society. Over time, Koreans not only continued to utilize existing Chinese graphs but also created new ones that are unique to Korea. These are known as Goyuhanja（固有漢字）or Gukja（國字）. Similarly, in Japan, also part of the East Asian cultural sphere, there are Waji 和字 characters that had originated in Japan. In Korea, there are three subcategories of unique Goyuhanja（固有漢字）: Gukja（國字）, Guguija（國義字）, and Gugeumja（國音字）.[21]

(12) Goyuhanja（固有漢字）
 a. Gukja （國字）
 b. Guguija （國義字）
 c. Gugeumja（國音字）

The term Gukja（國字) refers to graphs created in Korea that do not exist in China. When meanings unique to Korea are added to existing Chinese graphs, these are called Guguija（國義字）; and when unique pronunciations are added, they are referred to as Gugeumja（國音字）. In the late 19th century and early 20th century, many scholars made efforts to compile and organize Goyuhanja. Currently, there are approximately 200 or more Goyuhanja listed in The Great Chinese-Korean Dictionary（漢韓大辭典）. Korean scholars have continued to research and to add Goyuhanja to the list.

In the remainder of this section, three topics will be discussed: 3.1 deals with Idu, or creative usage, while 3.2 deals with Goyuhanja, in creating new graphs, sounds, or meanings, and 3.3 with modern adaptations that could involve mixed scripts.

3.1 Idu (吏頭): Creative usage

The cases of Idu introduced above primarily correspond to the practice of 'borrowing' Chinese graphs, which means using the meanings or sounds of existing Chinese graphs. Since Idu represents a creative *usage* of existing Chinese graphs, it should be distinguished from Goyuhanja, in which *new* characters, sounds, or meanings are *actually* created. The creative usage of Idu can be categorized into four main types based on how certain elements of Chinese graphs are borrowed and employed. The usage of Idu cannot be fully explained by Xu Shen's Six Principles alone.

---

[21] Among Korean scholars, the term Gukja in the broad sense was generally used in the past; today, however, Goyuhanja is the term that is now widely used, and is the term adopted here.

(13) Four types of Idu

a. Reading by phonetic element 音讀字 (Sino-Korean reading of Chinese graphs)
b. Reading by semantic element 訓讀字 (borrowing of Chinese meaning only)
c. Borrowing phonetic element 音假字 (borrowing of Chinese sound only)
d. Borrowing semantic element 訓假字 (borrowing of Korean sound + meaning)

In (13a), 'Reading by phonetic element'（音讀字）refers to cases where a Chinese graph's given sound is used in conjunction with its given meaning. In this method, the graph is pronounced based on its original sound. For example, the Chinese graph 夜 *ya* meaning 'night' is pronounced as *ya* to represent the concept of 'night.' These are the Sino-Korean readings of Chinese graphs. In (13b), 'Reading by meaning'（義讀字）involves reading the meaning part of a Chinese graph. For example, the character 夜 *ya* meaning 'night' can be used to represent 'night,' pronounced as the Korean word *bam* 'night.' This involves semantic borrowing, using the meaning and discarding the source pronunciation. It is noteworthy that in both (13a) and (13b), the Chinese graphs are used in a way that aligns with their original meanings. These two methods of reading characters by sound or meaning are widely used in countries that are part of the East Asian cultural sphere, such as Korea, Japan, and Vietnam.

In contrast to the first two cases, (13c-d) represent somewhat more complex cases. In (13c), 'Borrowing phonetic element'（音假字）refers to cases where the sound component of a Chinese graph is borrowed to represent the pronunciation of a Korean word, a case of Xu Shen's phonetic loan, and not a case of Sino-Korean reading, as in (13a). For example, the Chinese graph 夜 *ya* meaning 'night' is pronounced as *ya,* and only its pronunciation is used, without also including the meaning of 'night.' In (13d), 'Borrowing semantic element' (訓假字）involves borrowing the meaning of a Chinese graph to represent a native Korean word with its native pronunciation, but more may be involved. For example, the Chinese graph 夜 *ya,* meaning 'night' is pronounced *bam*, the native Korean word for night, but it carries a different meaning in Korean than simply 'night.' In (13d), 'Borrowing semantic element' (訓假字) involves borrowing the pronunciation of a Korean word corresponding to the meaning of a Chinese graph to represent the pronunciation in Korean. Similarly, the Chinese graph 夜 *ya* meaning 'night' is pronounced as *bam*, the native Korean word for 'night,' but it carries a different meaning in Korean than simply 'night.'

The examples in (14) and (15) illustrate the complex borrowing practice in Korean, as 'borrowing' of Chinese graphs often involves utilizing a combination of the methods in (13a-d). These methods are used interchangeably to represent Korean names, place names, pronunciations, and more. An example from a Hyangga（鄉歌）of the Silla Dynasty is given in (14).

(14)a. 夜音 *Bam*

	夜	音		夜音
Meaning 訓 :	night (***bam***)	sound (*sori*)	=>	'night'
Sound 音 :	*ya*	*eu**m***	=>	***bam***
	訓讀字	音假字		

b. 卯乙 *Mol*

		卯	乙		卯乙		
Meaning	訓 :	rabbit (tokki)	new (*sae*)	=>	‘secretly’		
Sound	音 :	***myo***	*eu**l***	=>	***mol***	=>	***mollae***
		音假字	音假字				

In (14a), 夜音 *bam* ‘night’ is a monosyllabic word, where the first graph, 夜, involves ‘Reading by semantic element’ (13b, 訓讀字), that is, a case of semantic loan; while the second graph, 音 , meaning ‘sound,’ involves ‘Borrowing phonetic element’ (13c, 音假字), a case of phonetic loan. The word, 夜音 *bam* ‘night,’ involves mixed borrowing, yielding *bam*. The phonetic borrowing of 音 *eum* is only partial, in adding only the nasal coda *-m* to the preceding syllable, *bam*. This additional marking, although cancelled out, may seem unnecessary, but it plays an important role in 夜音 *bam.* It informs the reader that the preceding Chinese character 夜 should be read as *bam*, with a *-m* coda, and not read as Sino-Korean *ya*. Similarly, in (14b) 卯乙 *mol* provides a case of an added *-l* coda. As these two Chinese graphs, 卯 *myo* and 乙 *eul*, are both cases of ‘Borrowing phonetic element’ (音假字, 13c), their meanings are irrelevant. 卯 *myo* and 乙 *eul* are combined to form the first syllable of a native Korean word, *mollae* ‘secretly.’

Idu was commonly used, especially for representing personal and place names. In (15), 蚊火 *mokibeol* is an example that illustrates this usage well (Kim 1990). In (15a), the native Korean word for ‘mosquito,’ *moki*, is borrowed. (15b) shows the simple borrowing of the phonetic element *bul*, which has a similar pronunciation to the native Korean word, *beol* ‘field.’[22] The native Korean toponym, *mokibeol*, is written as 蚊火, but pronounced as *mokibeol.*

(15) 蚊火 mokibeol

		a. 蚊	b. 火		c. 蚊火
Meaning	訓 :	mosquito (***moki***)	fire (***bul***)	=>	mosquito field
Sound	音 :	*mun*	*hua*	=>	***mokibeol***[23]
		訓讀字	訓假字		

We turn next to the study of the new creations, be that the adding of new characters, or the adding of sounds or meanings.

3.2 Goyuhanja (固有漢字): Creation

In contrast to Idu, Goyuhanja should be considered creations because they involve adding of meanings or pronunciations to Chinese graphs that were not present in the original graphs. And Goyuhanja could also involve the creation of entirely new graphs. Among Goyuhanja, Gukja graphs are creations by Koreans and are, therefore, not included in Chinese dictionaries.

Beyond identifying and adding Goyuhanja to lists of graphs, the classification of Goyuhanja has been an important research topic. The classification into *gukja*, *gukuija*, and *gugeumja* is

---

[22] According to Kim (1990), the native Korean word for ‘fire’ *bul*, and ‘field’ *beol* were likely to have had very similar sounds during the Three Kingdom period.

[23] A field/region with a large number of mosquitos.

generally accepted. However, further discussions are needed regarding the structure and principles of creating Goyuhanja.

Shin (2005) suggests that the invention of Goyuhanja is primarily based on *xiesheng* 諧聲 (phonetic-semantic compounds) and *jiajie* 假借 (phonetic loans), two of Xu Shen's Six Principles. Kim (1972) proposes that Goyuhanja are mainly created by borrowing the readings and meanings of existing graphs rather than following Xu Shen's Six Principles. While most research has focused on borrowed notations like personal and place names, Ayukai's (1931) study attempts to classify Goyuhanja's structural types based on Xu Shen's Six Principles.[24] Ayukai's study has attracted recent research in extending discussions and explanations of his classification of Goyuhanja's creation principles. (16-18) are examples of Gukja graphs created using some of these various methods.

In (16), we have an example of a compound graph, where two existing graphs are combined to create a new graph. In this example, the Chinese graph 大, which means 'big,' and the graph 口, which represents 'mouth,' are combined to form a new graph, 夻, which represents a fish with a large mouth. Since 大 is pronounced as *dae* and 口 is pronounced as *gu*, combining these two pronunciations creates the disyllabic Korean word, *daegu,* for a fish with a big mouth.

(16) 合字 Compound graph
夻 cod

		大	口		夻
Meaning	訓 :	big (*keu da*)	mouth (*ip*)	=>	cod **(dae gu)**
		(a fish with a big mouth)			
Sound	音 :	***dae***	***gu***	=>	hua

In (17), 欕 *eom,* meaning 'Kalopanax,' is an example of a semantic-phonetic compound (形聲) graph that is created by combining the meaning of 木 'wood' and the pronunciation of 嚴 *eom*. In Korean, Kalopanax is pronounced as *Eomnamu*, where *namu* means 'tree.' By combining 木 representing 'tree' and 嚴 with the pronunciation *eom*, 欕 (木+嚴) was created to represent the 'Eom tree.'

(17) 形聲
欕 Kalopanax (Eom tree)

		木	嚴		欕
Meaning	訓 :	**tree (*namu*)**	stern (*eom hada*)	=>	Kalopanax (Eom namu) **(Eom tree)**
Sound	音 :	*mok*	***eom***	=>	**Eom**

(18a-b) are examples of creating new characters through phonetic borrowing (假借). (18a) is similar to the example from (14), where a specific final sound in the Korean vocabulary is represented by combining a Chinese graph with a corresponding coda. The graph 石 'stone' and 乙， representing *-l* coda, are combined to create the new graph, 乭. The pronunciation of 乭 is *dol*, which combines the Korean word for 'stone,' *dol*, with the *-l* coda. (18b) demonstrates the

[24] 合字，象形，諧聲，會意，會意 + 諧聲，假借 造字，俗字.

combination of a Chinese graph with a Korean consonant that represents the coda, rather than a combination of two Chinese graphs. The Chinese graph, 巨 *geo*, is combined with the Korean consonant ㄱ, representing 'g/k' as a coda, to create the new graph, 巪 *geok*.

(18) 假借 造字

a. 乭 stone

	石	乙		乭
Meaning 訓 :	stone ***<u>(dol)</u>***	new (*sae*)	=>	stone (dol)
Sound 音 :	*suk*	***<u>eul</u>***	=>	**<u>dol</u>**

b. 巪 graph for person's name

	巨	ㄱ		巪
Meaning 訓 :	big (keu da)		=>	graph for person's name
Sound 音 :	**<u>geo</u>**	***<u>g/k</u>***	=>	**<u>geok</u>**

The adoption of a Korean-created consonant in (18b), in lieu of using a second Chinese graph to represent a coda in Korean, is clearly a much more efficient and transparent strategy. The example paves the way for studying modern adaptations in the next subsection.

3.3 Modern adaptations.

In the late 1940s, a mixed script known as 국한문혼용체 *Gukanmunhonyongche* (國漢文混用體) was commonly used in Korea. Sino-Korean words were written in Chinese graphs, while the rest of the words were written in Hangul. That continued until the late 1980s when that mixed script was gradually replaced by the sole use of Hangul.

In modern Korean, Chinese characters are no longer used directly as mean of recording the language. Today, Sino-Korean words are written in Hangul, and not in Chinese characters, to represent their pronunciation. Moreover, as Hangul can perfectly represent the pronunciation of all Korean words, including Sino-Korean words, there is no longer the need to represent the pronunciation of Korean using Chinese graphs, such as Idu or Goyuhanja. As a result, the use of Chinese characters in modern Korean is often relegated to wordplay for entertainment. As educated Koreans can easily read and understand simple Chinese characters, usage of the Chinese script is primarily used in cases involving such simple characters. For instance, 無 *mu* 'not have, without,' is often used for emphasis. *Isang mu,* meaning 'no problem, everything is fine,' is frequently written as 이상無, with the mixing of Hangul in the Korean word, *isang* 이상 'problem, issue' with 無 *mu* 'not have, without.' Similar to 이상無, the Chinese character, 無 *mu*, can even be placed before *or* after a noun to emphasize the meaning of 'not, none'!

An example of wordplay using Chinese graphs is given in (19). It involves mock classical Chinese, Hanmun (漢文). At first glance, it appears to classical Chinese that may form a poem. Instead, it is merely a meaningless sequence of Chinese characters. However, when the sequence is read in Sino-Korean, it forms a sentence in Korean. Thus, it is word play where Chinese graphs are used phonetically to mimic Classical Chinese but actually represent a sentence in Korean.

(19)	于里阿其	*wu li a gi*	Our baby
	凱歸蠡魚	*gae gui yeo wo*	(is) so cute.

河鼓風去	*ha go pung geo*	Everything you want to
削多海蘿	*sak da hae la*	do all (you want to)
		'Our baby is so cute. Do everything you want to.'

Another interesting playful usage of Chinese characters is that of letter play, which involves stroke re-arrangement: that is, the strokes of Hangul letters are re-arranged to look like real letters but are not. That is, the strokes of Hangul letters are re-arranged to create letters that look similar to the original word. Many fun neologisms have been created through stroke re-arrangement, especially on the internet. Recently, the letter play has extended to re-arranging Chinese characters as well. Examples of re-arranging just Hangul letters as well as the inter-mixing of Hangul and Chinese characters are given in (20).

(20)a.	멍멍이	댕댕이
b.	팔도비빔면	괄도네넴띤
c.	金長훈	숲튽훈

When the strokes in (20a) 멍멍이 *mungmungyi* 'puppy' are re-arranged as 댕댕이 *Daeng-daengyi*, the result looks like 멍멍이. As a playful trend, young people in Korea have been using such made-up creations as 댕댕이 instead of 멍멍이, or 괄도네넴띤 *Gwaldonenemttin* instead of 팔도비빔면 *paldobibimmyeon*. In fact, a famous Korean singer, Kim Chang-hoon, used Hangul letters to represent the first two characters of his name when he opened a new YouTube channel. In (20c) 숲튽훈, 숲 and 튽 in his name were used to represent the Chinese characters, 金 and 長, respectively. It is interesting that '튽,' which is not a syllable used in any word, is formed by re-arrangement of the consonants ㄴ and ㅈ to mimic the Chinese character, 長.

In summary, Koreans have been using two main strategies in the adaptation of Chinese orthography to record spoken Korean, namely, creative adaptation and the creation of new graphs. Originating in China, Chinese characters hold a historical and linguistic significance that is more than just incorporating some foreign graphs into the Korean language. At the same time, Koreans have not merely adopted Chinese characters, just as they have not simply adopted other aspects of Chinese culture. Rather, they have created their own phonetic script, as well as creatively utilized and adapted the Chinese characters to enhance the efficiency and richness of the Korean language.

4. Concluding Remarks

The Chinese script has, for centuries, been adapted by the Cantonese and Koreans to record—in written form—the oral tradition of the people: their songs, ballads and much, much more. The set of Six Principles has been used, together with further adaptation via semantic loans and the use of mixed scripts. By innovatively adapting the Chinese script, Cantonese and Korean speakers have been able to preserve, for posterity, their rich history and cultural heritage. At the same time, the improvizations made possible through vernacular writing also enabled those with less education, or less opportunities for educational pursuits, such as females, access to inexpensive written materials to read and copy, and perhaps to even compose their own writings.

Even in modern settings, Chinese script is being used by Cantonese and Koreans in various ways. Innovative adaptations and the creation of new graphs serve as inventive means for people to express themselves and communicate, becoming an effective tool for enhancing language and communication, making them more efficient and richer.

References

Ayukai Fusanoshin 鮎貝房之進. (1931) 俗字考・俗文考・借字考 [Korean Vernacular Sino-graphs, Korean Vernacular Sinitic, Korean Borrowed Graphs] (1977). Seoul: Thaehaksa

Bauer, Robert S. (1982) D for two in Cantonese. *Journal of Chinese Linguistics* 10.2, 277-280.

Bawden, C. R. (1954) An eighteenth century Chinese source for the Portuguese dialect of Macao. *Silver Jubilee Volume of the Zinbun Kgaku Kenkyusho*, 12-33. Kyoto: Kyoto University.

Boltz, William G. (1977) Notes on an eighteenth-century Sino-Portuguese glossary. *Romance Philology* 30.3, 442-453.

Chan, Marjorie K.M. (1982) Response to Boltz' notes on Cantonese dentilabialization. *Journal of the American Oriental Society* 102.1, 107-109.

Chan, Marjorie K.M. (1994) Post-stopped nasals and lateral flaps in the Zhongshan (Yue) dialect: A study of a mid-eighteenth century Sino-Portuguese glossary. In Paul Jen-kuei Li, Chu-Ren Huang, and Chih-chen Tang (eds.), *Chinese Languages and Linguistics. Volume II. Historical Linguistics.* [Symposium Series of the Institute of History and Philology, Academia Sinica, Number 2.] 203-250. Taipei: Institute of History and Philology, Academia Sinica.

Chan, Marjorie K.M. (2010) The *Huajian Ji* (花箋記): Glimpses into early 18th century vernacular written Cantonese. The 2010 Midwest Conference on Asian Affairs, 1-3 October 2010, The Ohio State University, Columbus, Ohio.

Chan, Marjorie K.M. (2014) Orthographic variants in the Cantonese love ballad, *Romance of the Fancy Notepaper* (*Huajian Ji* 花箋記): 1713 to Present-day editions. The Second Workshop on Innovations in Cantonese Linguistics (WICL 2). 7 March 2014, University of Chicago, Chicago, Illinois.

Chan, Marjorie K.M. (2022) Vernacular written Cantonese in the twentieth century: The role of Cantonese opera in its growth and spread. In Richard VanNess Simmons (ed.), *Studies in Colloquial Chinese.* 36-58. Hong Kong: University of Hong Kong Press.

Chan, Marjorie K.M. (2023) Challenges in D2 and D3 acquisition: Dr. Sun Yat-Sen's 1924 Cantonese and Mandarin audio recordings. In Chien-Jer Charles Lin, Alex Cherici, and Bihua Chen (eds.), *Proceedings of the 34th North American Conference on Chinese Linguistics (NACCL-34)*, Bloomington, IN: Department of East Asian Languages and Cultures, Indiana University Bloomington. (No pagination; 34 pages)

Chan, Marjorie K.M. and Jocelyn Chey. (2013) 'Love you to the bone' and other songs: Humour and *rusheng* 入聲 rhymes in early Cantopop. In Jessica Milner Davis and Jocelyn Chey (eds.), *Humour in Chinese Life and Culture: Resistance and Control in Modern Times.* Volume Two. 103-130. Hong Kong: Hong Kong University Press.

Chao, Yuen Ren (趙元任). (1980) A system of "Tone-letters." *Fangyan* (方言) (1980) 2, 81-83. (Originally published in 1930 in *Le Maître Phonétique*, troisième série 3, 24-27.)

Gomes, Skylor E. and Marjorie K.M. Chan. (2022) The 2019-2020 Hong Kong protests: Dualling messages of the authorities and the protesters. In Mineharu Nakayama, Richard Torrance, Zhiguo Xie, John Bundschuh, Jennifer Nunes, and Lindsey Stirek (eds.), *Buckeye East Asia: Occasional Papers.* Volume 1. *Special Edition: DEALL 50th Anniversary Celebration Edition:*

Issues in East Asian Languages and Literatures. 71-82. Columbus: OSU Knowledge Bank. https://kb.osu.edu/handle/1811/101064

Hui, Mary. (2019) Cantonese is Hong Kong protesters' power tool of satire and identity. (20 June 2019) https://qz.com/1647631/extradition-law-hong-kong-protests-deploy-cantonese-as-satire-tool

Kim, Jonghun (김종훈). (1972) Go-yu-han-ja go (고유한자（固有漢字（國字））고（攷）（一）) [A Research of Go you han ja] Gug-eo-gug-mun-hag (국어국문학 *Korean Language and Literature*) 55, 175-188.

Kim, Sanggeun (김상근). (1990) Han-gug-ui han-ja-su-yong yeog-sa-wa bang-beob (한국의 한자수용 역사와 방법) [How to accept and take use of the Chinese Characters in Korea.] Jung-gug-eo-mun-hag (중국어문학 *Journal of Chinese Language and Literature 17*), 281-311.

Lee, Gunsick (이건식). (2009) Han-gug go-yu-han-ja-ui bal-dal (한국（韓國）고유한자（固有漢字）의 발달（發達）) [A Study on the History of Hanzi Characters Designed in Korea] Gu-gyeol-yeon-gu (구결연구 *Journal of the society Kugyol*) 22, 219-259.

Lee, Gunsick (이건식). (2012). Han-gug go-yu-han-ja ja-hyeong gu-seong bang-beob yeon-gu i-je (한국 고유한자 자형（字形）구성（構成）방법（方法）연구 二題-기존 연구의 비판적 검토와 형성（形聲）으로 만들어진 한국 고유한자의 몇 가지 사례) [A Study of Composition Methods of Korean Go you han ja] Dong-yang-hag (동양학 *Oriental Studies*) 52, 187-211.

Lee, Gunsick (이건식). (2015). Tong-hab han-han-dae-sa-jeon-ui gug-ja cheo-li-e dae-ha-yeo (통합 한한대사전（漢韓大辭典）의 국자（國字）처리에 대하여) [On the Treatments of Korean-originated Chinese Characters in the Unified Great Chinese-Korean Dictionary] Dong-yang-hag (동양학 *Oriental Studies*) 59, 143-167.

Leung, K.C. (1999) Chinese courtship: The *Huajian Ji* 花箋記 in English translation. *Chinoperl Papers* 20-22, 269-288.

Liu, Xinzhong (刘新中). (2007) Hanyu-zhong xundu-de Xingzhi, Yuanyin ji qi Yingxiang (汉语中训读的性质、原因及其影响 [The nature, origin and influence of semantic borrowing in the Chinese language] *Yuyan Jiaoxue yu Yanjiu* (语言教学与研究, *Language Teaching and Linguistic Studies*) 2007.2, 91-96.

Nam, Poonghyun（南豊鉉）. (1975) Han-ja cha-yong-pyo-gi-beob-ui bal-dal(漢字借用表記法의 發達) [The Development of Chinese Character Borrowing Notation] Gug-mun-hag-non-jib (國文學論集 *Essays on the Korean Language and Literature*) 8, 3-47.

Nam, Poonghyun (남풍현). (1981) Cha-ja pyo-gi-beob yeon-gu (차자 표기법 연구) [A Research of Adaptation of Chinese Characters]. Dan-gug-dae-hag-gyo chul-pan-bu (단국대학교 출판부 Danguk University Press, Seoul)

Norman, Jerry. (1988) *Chinese*. Cambridge: Cambridge University Press.

Morrison, Robert. (1828) *A Vocabulary of the Canton Dialect.* Three volumes. Macao: Printed at the Honorable East India company's press, by G.J. Steyn, and brother.

Shin, Sang Hyun. (2005) Han-gug go-yu-han-ja jo-sa yeon-gu (韓國 固有漢字 조사 연구 – 人名用 固有漢字를 중심으로) [Investigating Study on the Goyu-hanja (固有漢字, Chinese Character of Korean Origin) used in the names of people] Min-jog-mun-hwa-yeon-gu (민족문화연구 *Korean Cultural Studies*) 2005, vol.43, 155-202

Thoms, Peter Perring. (1824) *Hua Jian* 花箋 *Chinese Courtship. In Verse.* London: Parbury, Allen, and Kingsbury.

Yin, Guangren (印光任) and Rulin Zhang (張汝霖). (1751/1968). *Aomen Ji Lue* 澳門記略 [Sketch Notes on Macao]. Reprint of an edition with preface date of Qianlong 16 (1751). *Shiliao Cong Bian* (史料叢編). Taipei: Guangwen Book Company (廣文書局). (Also readily available is an 1800 edition that was also reprinted in 1968 in Taiwan, in Volume 109 of the *Zhongguo Fangzhi Congshu* (中國方志叢書). Taipei: Cheng Wen Publishing Co. (成文出版社).

Zhan, Bohui (詹伯慧). (1957) Hainan Fangyan-zhong Tongyizi-de "Xundu" Xianxiang (海南方言中同義字的 "訓讀" 現象) [The phenomenon of semantic borrowing of synonyms in the Hainan dialect]. *Zhongguo Yuwen* (中國語文) 60, 50.

Website

한국민족문화대백과사전 Encyclopedia of Korean Culture
(한국학중앙연구원 The Academy of Korean Studies, http://encykorea.aks.ac.kr)

chan.9@osu.edu
sy307@duke.edu