

Towards a Comparative Study of Sinographic Writing Strategies in Korean, Japanese, and Vietnamese*

Zev Handel (한 철, 韓哲夫)
University of Washington, Seattle
zhandel@u.washington.edu

1. Introduction

In the first millennium of the common era, Chinese cultural influence led to the introduction of writing, and of the Chinese writing system, to many peripheral areas within and beyond the Chinese polity. Soon after its introduction, the logographic Chinese script was adapted to write the major indigenous languages of the Korean peninsula, the Japanese archipelago, and Vietnam, and within a few hundred years these adaptations had led to fully functional writing systems capable of recording native language texts. The basic methods of adaptation in each region are fairly well documented and understood.¹ Similar techniques were employed in all three locations. If we compare Japanese *man'yōgana*, Korean *hyangchal*, and Vietnamese *chữ nôm*, for example, we find marked similarities in the way that Chinese characters are employed to represent non-Chinese linguistic elements. These result, in part, from potentialities inherent in any logographic script: the capability to extend the use of individual graphs (in this case, Chinese characters) based on either the pronunciation or the meaning of the morpheme(s) that they represent. Making use of these two techniques, logographs originally representing

* I would like to thank William G. Boltz, Richard Salomon, Mark Alves, and Timothy O'Neill for commenting on earlier drafts of this paper. Their suggestions have been extremely helpful. Any errors or misjudgments are of course my sole responsibility.

¹ However, many details remain to be worked out, and recent archeological excavations have unearthed early manuscripts that may shed further light. Such materials are unlikely, however, to significantly alter our understanding of the broad outlines of development.

Chinese morphemes could be employed to represent linguistic elements in other languages.

Developments in Korea and Japan were remarkably parallel. In contrast, developments in Vietnam were strikingly different. In this paper I intend to show that typological differences among the Chinese, Vietnamese, Korean, and Japanese languages were a major factor shaping the direction of writing system developments; these differences constrained and motivated particular kinds of changes and adaptations. In particular, these differences help explain the eventual emergence of simplified Japanese *kana* and Korean *to* (= *gugyeol*) on the one hand, and of complex Vietnamese *chữ nôm* characters on the other. They also may have been a factor in the ultimate adoption of the Hangeul alphabet in Korea and of the modified Roman alphabet (*quốc ngữ*) in Vietnam.

Although this paper will focus on the developments of the Chinese writing system in Asia, the conclusions have broader implications for our understanding of how logographic writing systems change as they are adapted for use with different languages; in particular, they suggest that a teleological evolution from logography to alphabet is an incorrect way of understanding writing system development, and that linguistic typology may play a role at least as significant, if not more so, than cultural and material factors.²

2. Sinographs and the Chinese Writing System

Before beginning a comparative investigation of early writing in Korea, Japan, and Vietnam, it will be helpful to first summarize the features of the Chinese writing system as it was used in China in the first centuries of the common era.³

² The precise definition of writing is a contentious matter. See Sampson 1985, Coulmas 1991, Daniels and Bright 1996, and Rogers 2005 for a sampling of views. Some scholars, such as Sampson, define writing broadly to include both *glottographic* (that is, representing spoken language) and *semasiographic* or *non-glottographic* (that is, representing concepts, ideas, objects, etc. without linguistic mediation). In this article I am concerned only with writing in the narrow (glottographic) sense, while recognizing that non-glottographic subsystems (such as punctuation, or the set of Chinese “radicals”) play a role in many writing systems (see Hyman 2006).

³ For convenience, I will use ‘Korean’, ‘Japanese’, and ‘Vietnamese’ as short-hand terms for the languages spoken on the Korean peninsula (including areas farther to the north than the current polity of Korea), the Japanese archipelago, and the area of modern-day Vietnam, respectively, during the period under discussion here. It must be remembered, however, that the linguistic situation in each of these areas was very different at that time than it is today.

The earliest attested examples of Chinese writing are found in the Shāng 商 oracle bones inscriptions (*jiǎgǔwén* 甲骨文) dating back to approximately 1250 BCE. These texts reveal a fully developed writing system recording a language almost certainly ancestral to the various languages we now identify as Chinese.⁴ That writing system is logographic, and in both structure and usage already boasts all of the features associated with later stages of the writing system, including that of modern written Chinese.

The basic unit of the Chinese writing system is commonly referred to as a Chinese character (Chinese *hànzì* 漢字, Korean *hanja* 한자, Japanese *kanji* 漢字, Vietnamese *chữ Hán* or *Hán tự*). To avoid potential confusion, in this paper I will use the term *Chinese character* only when referring to written forms of the Chinese language; in more general contexts I will instead use the term *sinograph*.

By the Hàn 漢 dynasty (206 BCE – 220 CE), the Chinese writing system can be described as *morphosyllabographic*, as it is today. In other words, the overwhelming majority of Chinese characters represent monosyllabic morphemes of the spoken language. Because a morpheme, by definition, has both phonological shape and semantic content, each Chinese character also had an associated pronunciation and meaning, namely the pronunciation and meaning of the morpheme with which it was conventionally associated.⁵

Chinese characters also have internal structure, and can be categorized accordingly. For our purposes we need identify only three types: unit graphs, semantic-semantic compound graphs, and phonetic-semantic compound graphs.

Unit graphs are essentially indivisible graphic units. Most derive from earlier pictographs or other iconic representations. Examples are 日 *rì* ‘sun’ and 馬 *mǎ* ‘horse’.

There was a great deal more linguistic diversity, and some of the languages spoken were probably not ancestral to the modern Korean, Japanese, and Vietnamese languages. We know very little about some of these languages (such as those spoken in the kingdoms of Goguryeo and Baekche on the Korean peninsula).

⁴ See Keightley 1978, Norman 1988, and Boltz 1994 for more discussion of oracle bone writing. For a general study of the Chinese writing system, see Qiú 2000.

⁵ This statement should not be taken to imply that there was a strict one-to-one relationship between graphs and morphemes. A single graph could be used to write different morphemes, and a single morpheme could, at different times and places, be written with different graphs. However, as the Chinese writing system became more standardized from the Hàn Dynasty on, the ideal of a one-to-one relationship became embedded in the normative use of the Chinese writing system. It is therefore usually possible to specify a single conventionally associated morpheme for each graph.

Semantic-semantic compounds are composed of two graphic elements. In most cases these graphic elements can themselves function as unit graphs, or are abbreviated forms (allographs) of unit graphs. The meaning of the morpheme written by the compound is associated with, or suggested by, the meanings associated with each component graph. An example is 尖 *jiān* ‘sharp’, composed of 小 *xiǎo* ‘small’ and 大 *dà* ‘large’.

Phonetic-semantic compounds also are composed of two graphic elements, but in this case one is conventionally associated with a semantic range related to the morpheme represented by the compound character, and the other is conventionally associated with a pronunciation related to that morpheme. Examples are 芳 *fāng* ‘fragrant’, composed of the semantic element 艸 *cǎo* ‘grass’ (in abbreviated form) and the phonetic element 方 *fāng* ‘square’; and 路 *lù* ‘road’, composed of the semantic element 足 *zú* ‘foot’ and the phonetic element 各 *gè* ‘each’. In the latter example, the modern pronunciation of the phonetic element bears little relation to the pronunciation of the represented morpheme. This is the result of sound changes that have taken place over the last 2000 or more years. The reconstructed Old Chinese pronunciations are **kak* (for the morpheme represented by the phonetic element 各) and **g-raks* (for the morpheme written by the character 路).⁶

The initial development of the Chinese writing system from pre-writing pictographic and iconographic representations depended crucially on the repurposing of graphs to write morphemes other than those they originally represented. A fully functional writing system must be able to represent all elements of spoken language, including such things as grammatical particles that are not amenable to iconic representation. There were two basic techniques for repurposing, or extending the usage, of a graph: *phonetic adaptation* and *semantic adaptation*. In phonetic adaptation, a graph is desemanticized; that is to say, the conventional association with the meaning of the morpheme it writes is ignored. Based only on its conventionally associated pronunciation, the graph is employed to write a second morpheme pronounced similarly to the first. The graph is then resemanticized through its new conventional association with the second morpheme.

For example, the graph 勿, in origin a pictograph, initially wrote the word ‘creature’, pronounced **mjut* (modern *wù*).⁷ It was then used to write the

⁶ These and all following reconstructions are from the system of Baxter 1992.

⁷ This and following examples are from Boltz 1994. I give the modern forms of the Chinese characters. They are structurally equivalent, but different in appearance, from

homophonous morpheme *mjut ‘do not’ through a process of phonetic adaptation and resemanticization. This first part of this process is technically referred to as *paronomasia* and commonly called *rebus usage*.

It is worth noting that if the graph were not resemanticized—in other words, if the graph became permanently desemanticized and thus was consistently used to represent phonetic content alone rather than specific morphemes—it would no longer be a logograph. It would instead become a phonograph, in this case representing a syllable.

The process of semantic adaptation is similar to phonetic adaptation, but in this case the graph is used to write a second morpheme with meaning, rather than pronunciation, related to the first. Because the second morpheme has a pronunciation, the graph is then rephonologized by virtue of its association with that morpheme.

For example, the graph 月, originally depicting the moon, was used to write the word *ng^wjat ‘moon’ (modern yuè, now written 月). It was also used to write the semantically related word *z(l)jAk ‘night’ (modern xī, now written 夕) by metonymic association. This kind of semantic adaptation only occurred in the earliest stages of the development of Chinese writing, as it depended on the iconic value of graphs to suggest meanings different from those of the morphemes they conventionally wrote. What I refer to here as semantic adaptation is, therefore, a kind of indexical usage (see Boltz 2006).

The phonetic-semantic compound characters described above originally arose through the process of adding disambiguating elements to repurposed characters in order to restore a one-to-one relationship between graphs and their represented morphemes. Following the semantic or phonetic extension of a logograph to represent a second morpheme, a single graph ends up representing two distinct morphemes. The resulting ambiguity could be eliminated by adding a *semantic determinative* or a *phonetic determinative* to the graph, creating a compound graph. Consider the example given earlier of the two morphemes written with the graph 勿, *mjut ‘creature’ and *mjut ‘do not’. The semantic determinative 牛, a modified combining form of 牛 niú ‘cattle’, was combined with 勿 in order to create a new graph, 物, to write the morpheme ‘creature’. This process was based on the perception that the meaning of ‘creature’ is related to the semantics of ‘cattle’. The result of the process is the creation of a phonetic-semantic compound.

By the Hàn Dynasty it had already become common to create new phonetic-semantic compounds without first going through an initial stage of phonetic

earlier forms.

borrowing, by analogy with the many phonetic-semantic compounds already present in the writing system. For example, as far as I know there is no evidence that the simple graph 方, conventionally associated with the morpheme *fāng* ‘square’ was ever used to write the morpheme *fāng* ‘fragrant’. It is possible that the compound character 芳 was created in one step, and was the first and only graph used to write *fāng* ‘fragrant’.

The evidence from other logographic writing systems, such as Sumerian cuneiform, Egyptian hieroglyphs, and Mayan hieroglyphs, suggests that the basic techniques used to extend logographs in order to generate a complete writing system capable of representing any linguistic utterance are an inherent feature of logographs. Or, perhaps it would be more accurate to say that they are inherent in the human cognitive perception and manipulation of logographs. A logograph can be desemanticized and thus used to write other linguistic elements based on pronunciation. If it is resemanticized so that it still conventionally represents only specific morphemes, we can refer to the result as a *phonetically-derived logograph* (PDL). If it is permanently desemanticized it ceases to be a logograph and becomes a phonograph; we can refer to the result as a *phonetically-derived phonograph* (PDP).

The parallel process, by which a logograph is repurposed to write a semantically-related morpheme, results in a *semantically-derived logograph* (SDL).⁸ Within a monolingual context, this process only appears to be possible at the earliest, iconic stages of the writing system. However, as we shall see, SDLs are quite common in cross-linguistic contexts.

The basic techniques outlined above for the development of the logographic Chinese script are precisely the basic techniques that were employed when Chinese writing was adapted to write other languages of Asia; moreover, it seems justified to further claim that these basic techniques will inevitably be employed when any logographic writing system is repurposed to write a second language.

I have just mentioned that if desemanticized graphs are not resemanticized, i.e. are not conventionally associated with specific morphemes, they cease to be logographs and become phonographs; and that this developmental possibility is inherent in any logographic writing system. In the history of China, various forces, both linguistic and societal, ultimately countered any tendency for widespread

⁸ As I have defined them here, PDPs, PDLs, and SDLs are no different in form from the original graph from which they are derived. The nature of the derivation involves a new use of the graph, or to be more precise, a reformulation of the associations holding between the graph and elements of the spoken language.

desemanticization and the development of a phonographic writing system. Generally speaking, desemanticized graphs were resemanticized as they were associated with specific Chinese morphemes.⁹ There is one important exception, however, and that is in the writing of certain layers of borrowed vocabulary, where processes of resemanticization and graphic disambiguation were much more haphazard. In such cases it can be argued that certain characters did double duty, as logographs when writing Chinese vocabulary and as permanently desemanticized phonographs when writing borrowed vocabulary or when used for the transcription of foreign names or other words.

By the late Hàn Dynasty, the use of Chinese characters as phonographs had become somewhat conventionalized in the transcription of Buddhist terminology, and some have argued that what could be called a rough syllabary of phonographs came into common usage. In other words, a small subset—fewer than 100—of the several thousand graphs used in the Chinese writing system came to form a “pool” of desemanticized phonographs regularly employed for the transcription of foreign words (see Bentley 2001).

3. Early Sinograph Usage in Korea, Japan, and Vietnam

The histories of writing in Korea, Japan, and Vietnam are complex and highly specialized topics. In this section I will merely sketch the major developments. I will then show in Section 4 how similarities and differences in these developments can be accounted for by considering the typological characteristics of the languages involved.

Much has been written about early writing in these three locations; I have not yet worked through all of the vast literature on the topic, so my perspective at this point is necessarily somewhat limited. For Vietnamese the problem of accessible scholarship is particularly vexing, since almost nothing of substance has been written in languages other than Vietnamese. The conclusions in this paper must therefore be considered preliminary.

Most of the data presented here is drawn a handful of English-language sources, though others (listed in the bibliography) have been consulted as well. For Korean, the major sources are Lee 1972 and Lee and Ramsey 2000; for Japanese, Seeley 1991 and Lurie 2001; for Vietnamese, Nguyễn 1990.

There does not appear to be an accepted cover term for those writing systems that make use of adapted sinographs to write non-Chinese languages. For

⁹ See Boltz 1994 for one explanation of the cultural forces at work.

convenience I will use the term *sinography*. By sinography I refer to a system for writing a non-Chinese language using an adapted form of the Chinese-character script. Those adaptations might be formal, functional, or both.

3a. Korean

Historical evidence suggests that a variety of languages were spoken on the Korean peninsula at the time when Chinese writing was first introduced there, most likely following the establishment of the Hàn Dynasty commanderies in the northern part of the peninsula in 108 BCE. Unfortunately, there is very little evidence attesting in any detail to the features of those languages. The languages spoken in the Korean kingdoms of Goguryeo 고구려 高句麗 in the north, Baekje 백제 百濟 in the southwest, and Silla 신라 新羅 in the southeast during the Three Kingdoms Period (57 BCE – 668 CE) were probably quite distinct, although they most likely shared common typological characteristics.¹⁰ Modern Korean is descended from the language of Silla, which unified the peninsula under its rule in 668.

A detailed description of what is known about the phonology, morphology, and syntax of the earliest stages of the languages of the Korean peninsula is beyond the scope of this paper. For the purposes of this discussion, it will suffice to point out that there is little reason to think that these languages were not typologically similar to attested Middle Korean and Modern Korean. Some of the major typological features that can be posited for all stages of Korean linguistic history are:

- 1) verb-final word order;
- 2) agglutinating morphology, characterized by verbal suffixation;
- 3) a system of noun-marking case particles;
- 4) significant numbers of polysyllabic morphemes;
- 5) fairly complex syllable structure, including consonant clusters in both onset and coda positions.

When it comes to the language of Silla, from which we get most of our evidence about the early adaptation of sinographs, we can be quite confident about the existence of these typological features.

On the Korean peninsula, as in other non-Chinese-speaking areas, the earliest

¹⁰ Throughout this article I use the official South Korean Revised Romanization system promulgated in 2000.

exposure to, and use of, Chinese characters was in written Chinese (the conventionalized form of which is usually referred to as Literary Chinese). When read aloud, these Chinese characters were pronounced in Korean-accented Chinese; these Korean “readings” of Chinese characters were eventually codified and standardized, and are today referred to as Sino-Korean (SK) pronunciations. The earliest evidence both in Korea and other areas suggests that the first use of sinographs to represent native words was to record local proper names embedded in Literary Chinese texts. This usage can be seen, for example, on the Stele of King Gwanggaeto (광개토대왕비 廣開土大王碑), erected in 414 and containing 1802 Chinese characters.

By the fifth century at the latest, sinographs were being regularly employed in all three kingdoms on the peninsula. The first body of systematic evidence attesting to the various methods employed to represent Korean proper names comes from the place names listed in the geographic section of the *Samguk Sagi* 삼국사기 三國史記 (*Chronicles of the Three Kingdoms*) of 1145. This section lists place names attested in the records of Goguryeo, Baekje, and Silla.

These early Korean uses of sinographs to record native Korean personal and place names reveal that both of the basic methods of repurposing sinographs described earlier, phonetic adaptation and semantic adaptation, are employed. A well-known example is the place name that in modern Korean would be pronounced *Gildong* 길동, composed of two Korean morphemes meaning ‘long piece’ (Lee and Ramsey 2000:47). This place name is sometimes written with the sinographs (a) 吉同 and sometimes with the sinographs (b) 永同. The three graphs involved conventionally write the following Chinese morphemes:

吉 : *jí* ‘auspicious’; SK *gil* 길
 同 : *tóng* ‘together’; SK *dong* 동
 永 : *yǒng* ‘long (time)’; SK *yeong* 영

In transcription (a), both sinographs are PDPs. In other words, the semantics of the Chinese morphemes that they normally represent in written Chinese are ignored, and they are used to represent the homophonous Korean morphemes *gil* ‘long’ and *dong* ‘piece’.

In transcription (b), 同 is a PDP. In contrast, 永 is an SDL, which is to say that the pronunciation of the Chinese morpheme it represents is ignored, and it is used to write the synonymous Korean morpheme *gil* ‘long’. The written form 永同 is thus composed of a logograph and a phonograph.

Semantic adaptation in this context is somewhat different from the type of semantic adaptation that took place within the history of Chinese writing. There, we saw the same graph used to write distinct morphemes whose semantics were related to the iconicity of the graph itself. In other words, that type of semantic adaptation was indexical in nature. When iconicity was lost following conventionalization of the graphic forms, this kind of semantic adaptation fell out of use. In the Korean case, in contrast, semantic adaptation was in effect an act of translation. The scribe “borrows” a graph conventionally used to write a morpheme in one language and employs it to write a synonymous morpheme in the other language. This act depends on a judgment made by a particular person at a particular time and place that is well motivated based on that person’s knowledge of the two languages involved.

The phonetic and semantic adaptations of sinographs to represent Korean morphemes are, as I have said before, capabilities that are inherent in logographs. Such usages are apparently self-evident to literate, bilingual users of logographic scripts. In the following sections we will see these same techniques applied everywhere that sinographs have been employed to write non-Chinese languages. In this most basic respect the developments of writing in Korea, Japan, and Vietnam can be said to be identical.

It is when we move beyond the use of sinographs to represent Korean proper names, and to their eventual use to record entire passages of Korean, that we see more refined and distinctive adaptations of sinographs. It is in these more detailed applications, and the ways that they differ from developments in the writing of Japanese and Vietnamese, that the influence of typological features of the Korean language is evident.

The earliest examples we have of such usage are in the *hyangga* 향가 鄉歌, Silla-era poems (most dating from the 8th century).¹¹ The poems are written in what Korean scholars call *hyangchal* 향찰 鄉札, which simply refers to the particular set of sinographic practices that are exemplified in this body of poetry. In *hyangchal* we see phonographic and logographic uses of sinographs like those

¹¹ We do not have *hyangga* manuscripts dating back to this time, so cannot discount the possibility of textual corruption in the received texts. However, similar writing techniques are seen on datable early manuscripts, such as the Goryeo-era manuscripts known collectively as *Mukseo jipyeon* 목서지편 墨書紙片 (“Paper Fragments with Ink Writing”), the earliest of which is a dedication text written in 1024. They were found in a pagoda at Bulguk Temple (Bulguksa 불국사 佛國寺) during construction work in 1966. See the work of Lee Seungjae 이승재 in National Museum of Korea 2007.

described above. The specific ways that these uses are tailored to the characteristics of the Korean language can be summarized as follows:

- 1) Nominal and verbal roots, i.e. those morphemes with fairly concrete semantics, are usually written logographically, with SDLs.
- 2) Suffixes and grammatical morphemes are generally written phonographically, with PDPs.
- 3) In some cases phonographic usage is only indirectly related to the Sino-Korean pronunciation. This is especially true of conventionalized representations of high-frequency suffixes and case particles.
- 4) A number of sinographs are used proto-alphabetically to represent single consonants, sometimes in disambiguating co-occurrence with logographs.

The four usages described above are exemplified in the poem “Song of Cheoyong” 처용가 處容歌, one of the best understood of the *hyangga*.¹²

At the end of line 3, we find the two sinographs 見昆 writing the verbal stem and inflectional affix *bo-gon* 보곤 “looking”. The first graph, 見 *jiàn* ‘see’ (SK *gyeon* 견), writes the verbal stem *bo-* ‘look, see’ as an SDL, while the second graph, 昆 *kūn* ‘elder brother’ (SK *gon* 곤) writes the suffix *gon* as a PDP. This illustrates usages of types (1) and (2).

In numerous places throughout this and other poems, the sinograph 良 *liáng* ‘good, fine’ (SK *ryang* 량) is used to write the high-frequency inflectional morpheme *ra* 라 or *reo* 러. This phonographic usage is based on a modification of the sinograph’s pronunciation, in which the coda is discarded, and is an example of usage type (3).

In line five of the poem, the phrase 吾下於叱古 *naehayeosko* “are mine” occurs. Here the sinograph 叱 *chì* ‘scold, shout’ (SK *jil* 질) is used to represent the single syllable-final consonant sound *-s*, based on the onset sound of its conventional pronunciation.¹³ It is, in effect, functioning as an alphabetic letter. This exemplifies usage type (4).

Elsewhere in the *hyangga* we see similar alphabetic uses of phonographs to disambiguate or reinforce the pronunciation of other sinographs. For example, the Korean word *bam* 밤 ‘night’ is written with the two-sinograph sequence 夜音.

¹² The Old Korean language of these poems is not fully understood, and many details of interpretation remain controversial. Differences of interpretation do not bear on the description of general sinograph usage presented here. The full text of the poem along with a translation can be found in Lee and Ramsey 2000:49.

¹³ Although the modern Sino-Korean pronunciation begins with *j-*, the usage seen in *hyangchal* is probably based on an older SK reading beginning with *s-*.

These two graphs conventionally write the following Chinese morphemes:

夜 : *yè* ‘night’; SK *ya* 야

音 : *yīn* ‘sound’; SK *eum* 음

The first graph is an SDL writing the native Korean word for ‘night’, *bam*. The second graph, whose Sino-Korean pronunciation lacks an onset and has the ‘neutral’ vowel *eu* [i], is employed phonographically to represent the coda consonant *-m* (Lee and Ramsey 2000:48). Its function is to disambiguate the possible referents of the graph 夜 by specifying that it must represent a morpheme ending in the sound *-m*.¹⁴ Thus the two graphs 夜音 write a single monosyllabic morpheme. One could argue that they form a single unit, a two-component logograph with internal structure indicating, based on the conventional Chinese use of the component graphs, both the meaning and part of the pronunciation of the native morpheme that it represents.

Sinographs used phonographically to represent Korean inflectional endings or case-marking particles could not only be used when writing Korean (as in *hyangchal* or other styles of writing Korean words with sinographs, collectively called *idu* 이두), but also as supplemental markers added to written Chinese texts. These markers were reading aids, allowing readers familiar with Literary Chinese to mark the sentential roles and verbal inflections of constituents in the text, and thus facilitate reading comprehension. They can be thought of as signposts, guiding the Korean reader to a successful navigation of the Chinese text. Phonographs used for this marking function are referred to as *gugyeol* 구결 or *to* 토. No doubt because they often had to be squeezed into the marginal spaces between lines of text or between individual sinographs, *gugyeol* graphs were often abbreviated. Abbreviation may also have been a consequence of the practice of jotting the graphs down quickly while listening to an oral rendering of the Chinese text into Korean. Over time these abbreviations became conventionalized, although they were never completely standardized.

Some examples of *gugyeol* graphs, and the characters from which they are

¹⁴ This use of the graph 音 is roughly analogous to the use of “rd” in the English orthographic representation “3rd” for *third*. The logograph “3” can represent words *three* or *third*; “rd” is a phonetic determinative, specifying which reading of “3” is intended and thus redundantly representing the ending consonant sounds of the word. The specialized use of these two letters is indicated typographically by the convention of placing them in superscript.

derived, are:

二 *si* 시, from 示
ㅁ *go* 고, from 古
ㄷ *ni* 니, from 尼

For those familiar with Japanese *katakana*, similarities in both the developmental process and resulting shapes will be apparent, although the sound values of those shapes in Japanese and Korean are quite different.

3b. Japanese

Japanese is typologically similar to Korean. It is a verb-final language, characterized by agglutinating morphology (principally seen in inflectional suffixes on verb stems) and the presence of case-marking grammatical particles. The principle typological difference with Korean lies in the phonology. The Japanese sound system is far simpler. There are fewer consonants and vowels, a much simpler syllable structure, and more restrictive phonotactics. As a result, there are only about 80 possible syllables in modern Japanese. Earlier stages of Japanese were typologically similar.

Both historical evidence and factors observable in early Japanese writing indicate that Chinese writing was introduced to Japan from the Korean peninsula, most likely by scribes from the kingdom of Baekje. Historical sources, though open to interpretation, place this introduction at the beginning of the 5th century.¹⁵ Most likely, techniques of sinographic adaptation for the writing of Japanese were also imported from the Korean peninsula. One question that ultimately must be addressed is the degree to which similarities in sinographic practice in Japan and Korea are attributable to cultural interaction, and the degree to which they are attributable to universal tendencies in script adaptation operating under the constraints of linguistic typology.

Just as in Korea, we find both phonetic and semantic adaptations of Chinese characters to write Japanese words at the very earliest stages, as evidenced by writing found on excavated artifacts (Lurie 2001 chapter 4). By the 8th century, sinography in Japan had reached a level of sophistication capable of fully

¹⁵ This is not to say that occasional artifacts with writing on them did not make their way to Japan prior to this date. See Seeley 1991:4-9 and Lurie 2001:146-154 for more detailed discussion of the historical evidence.

representing Japanese in written form. In such 8th-century texts as the two histories *Kojiki* 古事記 and *Nihon Shoki* 日本書, and the collection of poetry *Man'yōshū* 万葉集, we see not only the semantic and phonetic uses of sinographs familiar from our earlier discussion of Korean, but also an explicit understanding of, and technical vocabulary related to, those uses. Semantic adaptation is potentially ambiguous, because there may be more than one Japanese word viewed as semantically equivalent to the conventional Chinese-based meaning associated with a sinograph. In *Kojiki*, this ambiguity is sometimes eliminated through the use of notes in the text, called *kunchū* 訓注, which use sinographs phonographically in order to specify the pronunciation of a logographically written Japanese word. There are also notes that indicate that preceding graphs should be read phonographically, rather than logographically. These notes are typically of the type “此二字以音” “(read) these two graphs by means of sound”.

In the Japanese-language poems and songs recorded in the *Kojiki* and *Nihon Shoki*, and especially in the poetry of the *Man'yōshū*, we see sinographic usage closely parallel to Korean *hyangchal*. Sinographs are employed both as SDLs and PDPs, with the phonographs used primarily for grammatical particles, inflectional endings, and the transcription of names. Most often the phonographs were used for single Japanese syllables, but sometimes graphs with consonant-coda Chinese pronunciations were used to write two CV syllables of Japanese (Seeley 1991:50).

In parallel with Korean *gugyeol*—sinographs used phonographically to annotate Literary Chinese texts in order to facilitate reading translation into Korean—beginning in the late 8th century we see texts in Japan with *kunten* 訓点. These are markings to aid in the practice of translation-reading from Chinese into Japanese, known as *kanbun kundoku* 漢文訓読 or simply *kundoku*. Among the various types of markings employed are phonographs representing case-marking particles and inflectional endings, and phonographs clarifying the proper reading of sinographs in the main text. As with *gugyeol*, abbreviations of phonographic *kunten* became quite common. These abbreviated forms of sinographs were of two types: those abbreviated through cursivization, and those abbreviated through isolation (the extraction of one part of a graph to represent the whole).¹⁶

In the 10th century, these practices led to the development of precursors to modern *hiragana* ひらがな (平仮名) (phonographs abbreviated from sinographs by cursivization) and *katakana* かたかな (片仮名) (phonographs abbreviated from sinographs by isolation). These scripts were initially neither regular nor consistent, but nevertheless formed cohesive sets of graphs. By the 11th century both scripts

¹⁶ The terms are Seeley's. See 1991:60ff.

can be said to have become fully functional writing systems.

These examples of *katakana* graphs may be compared with the *gugyeol* forms, also derived through the isolating technique of abbreviation, given earlier:

ニ *ni*, from 仁
ロ *ro*, from 呂
ヒ *hi*, from 比

Before we move on to a discussion of Vietnamese sinography, there is one additional development in Japan that warrants description. A small number of sinographs, not found in the Chinese script, were newly created in Japan to represent native morphemes. These *kokuji* 国字 (“national graphs”) were in existence as early as the 9th century. Though the number of *kokuji* was never very large, they are of interest for several reasons. First, they have internal structure analogous to the internal structure of compound Chinese characters. Second, they reveal the existence of perceived gaps in the Japanese sinographic script, which ultimately came to favor the use of logographs for all verbal and nominal roots. If some Japanese roots had no close semantic equivalents in Literary Chinese vocabulary, then there would be no source sinograph appropriate for semantic adaptation to write those roots.

Most *kokuji* are compound graphs whose components are already found in Chinese characters. They tend to be semantic-semantic compounds. This is not surprising, because it would normally not be possible to find a Chinese character whose pronunciation would make it suitable to be a phonetic element in a graph representing a polysyllabic Japanese root.

Two examples of *kokuji*, both still in use in the modern Japanese script, are the following:

<u>Kokuji</u>	<u>Japanese morpheme</u>	<u>Source of components</u>
働	<i>hatarak</i> - ‘work’	人 (‘person’) + 動 (‘move’)
峠	<i>tōge</i> ‘mountain pass’	山 (‘mountain’) + 上 (‘ascend’) + 下 (‘descend’)

3c. Vietnamese

It seems likely that Chinese characters were first used in the area of what is now northern Vietnam as early as the Hàn Dynasty, when Chinese governors first ruled the area known as Nányuè 南越 beginning in 111 BCE. However, the earliest scattered evidence for the use of sinographs to write native Vietnamese is only

found on inscriptions from the 11th to 14th centuries (Nguyễn 1990:395).¹⁷ The largest body of early extant Vietnamese texts written with adapted sinographs is the 15th-century 254 lyric poems of Nguyễn Trãi 阮薦 (1380-1442), although other written poetry, now lost, is said to date back several centuries earlier (Nguyễn 1990). Native Vietnamese writing using adapted sinographs is called *chữ nôm* 字喃, usually translated as “Southern script”. This term is often abbreviated to *nôm*. The *nôm* script was never standardized, so there was considerable variation in the sinographs employed to write a given Vietnamese morpheme.

It is no accident that the full flowering of adapted sinographs to write Korean, Japanese, and Vietnamese is seen in poetry, where the faithful representation of the native language is essential, and translation or rendition into Literary Chinese is not an option.

Typologically, Vietnamese is similar to Chinese in many respects. Modern Vietnamese is tonal, has subject-verb-object word order, and is isolating, almost completely lacking in inflection. Morphemes are primarily monosyllabic and invariant. The greatest difference with Chinese is that syllable structure is more complex, especially in the number and type of vowels. There is evidence for the earlier existence of a number of consonant clusters that have disappeared from the modern language over the last several centuries.

A Mon-Khmer language, Vietnamese was once typologically quite different. Many words were sesquisyllabic (that is, composed of a reduced or minor syllable followed by a full or major syllable) and the language was atonal. The change to monosyllabicity and tonality is thought in part to be the result of areal convergence, due to the contact influence of Chinese and perhaps Tai languages. It is significant that the full development of *chữ nôm* seems to have taken place after this major typological shift had occurred probably around the 9th or 10th centuries (Ferlus 1992).

Although the basic techniques of semantic and phonetic adaptation are found in *nôm*, just as they are found in Korean and Japanese sinographic writing, the overall “look and feel” of *nôm* writing is quite different from what we see in Korean and Japanese. Because Vietnamese lacks inflectional endings and case-marking particles, and because of the near uniform monosyllabicity of morphemes, we do not see a bifurcation of form or function in the sinographic representation of word roots on the one hand and grammatical elements on the other. The basic

¹⁷ In fact, Vietnamese legend attributes the invention of *chữ nôm*, the use of sinographs to write Vietnamese, to the Chinese governor Shì Xiè 士燮 (137-226). This story is almost certainly apocryphal.

techniques for representing Vietnamese morphemes with sinographs are essentially uniform across the lexicon.

A major component of *nôm* writing, largely absent from Korean and Japanese sinography, is the use of newly created sinographs not found in the Chinese character script. As we shall see, for the most part the structural principles underlying the creation of these sinographs, or *nôm* characters, closely parallel those seen in Chinese characters as described in Section 2 above.

A number of different Vietnamese scholars have proposed classifications of *nôm* sinographs. These classifications can appear somewhat confusing, because they fail to clearly distinguish structural and functional differences of the graphs themselves from differences in the lexical layers of the morphemes being written. Moreover, the classification categories depend in part on a native view of lexical layers that sometimes seems at odds with the conclusions suggested by linguistic analysis. My classification below is adopted from three different classifications, those of Nguyễn Tài Cẩn and Xtankevich, Lê Văn Quán, and Nguyễn Ngọc San, all listed and described by Nguyễn (1990:397-406). I have made revisions in order to bring the classification into line with the terminology and conceptual framework advanced in this study.

Category I: Borrowed Sinographs

Types 1-3, described below, all involve sinographs found in the Chinese script.

Type 1: Sinograph writes morpheme borrowed from Chinese

The sinograph is employed in its conventional way, representing the Vietnamese morpheme borrowed from the Chinese morpheme that the same graph writes in Literary Chinese.

	<u>Nôm graph</u>	<u>Chinese morpheme</u>	<u>Vietnamese morpheme</u>
1	才	<i>cái</i> ‘talent’	<i>tài</i> ‘talent’
2	頭	<i>tóu</i> ‘head’	<i>đầu</i> ‘head, beginning’
3	山	<i>shān</i> ‘mountain’	<i>son</i> ‘mountain’
4	冊	<i>cè</i> ‘book’	<i>sách</i> ‘book’
5	肝	<i>gān</i> ‘liver’	<i>can</i> ‘liver’
6	買	<i>mǎi</i> ‘buy’	<i>mãi</i> ‘buy’
7	說	<i>shuō</i> ‘talk’	<i>thốt</i> ‘talk’ (SV <i>thuyết</i>)
8	貪	<i>tān</i> ‘greed’	<i>tham</i> ‘greed’

In these examples the Vietnamese morpheme on the right was borrowed from an

earlier form of the Chinese morpheme listed to its left. In most of these cases the Vietnamese morpheme belongs to the Sino-Vietnamese lexical layer, which means that it has the same pronunciation as that assigned to the Chinese character in the Literary Chinese reading tradition of Vietnam. In some cases the morpheme is an earlier borrowing from Chinese.

Type 2: Sinograph is phonetically adapted to represent a native Vietnamese morpheme

The graph writes a Vietnamese morpheme that is homophonous or near-homophonous to the Sino-Vietnamese pronunciation of the Chinese morpheme with which it is conventionally associated.

	<u>Nôm graph</u>	<u>Chinese morpheme (SV pronunciation)</u>	<u>Vietnamese morpheme</u>
9	沒	mò 'sink' (SV <i>một</i>)	<i>một</i> 'one'
10	固	gù 'solid' (SV <i>có</i>)	<i>có</i> 'have'
11	埃	āi 'dust' (SV <i>ai</i>)	<i>ai</i> 'who'
12	別	bié 'separate' (SV <i>biệt</i>)	<i>biết</i> 'know'
13	買	mǎi 'buy' (SV <i>mãi</i>)	<i>mới</i> 'new, recent'
14	甘	gān 'sweet' (SV <i>cam</i>)	<i>cam</i> 'be content, resigned'

Sometimes a diacritic is added to indicate that the sinograph is being used to represent a native Vietnamese morpheme (or a borrowed Chinese morpheme that is not in the SV lexical layer) rather than a borrowed Sino-Vietnamese morpheme. The most common of several diacritics employed for this purpose is a wedge shape <.

Example 13b, below, writes the same morpheme as example 13 above, but the diacritic serves to explicitly differentiate it from example 6.

	<u>Nôm graph</u>	<u>Chinese morpheme (SV pronunciation)</u>	<u>Vietnamese morpheme</u>
6	買	mǎi 'buy' (SV <i>mãi</i>)	<i>mãi</i> 'buy'
13b	買<	mǎi 'buy' (SV <i>mãi</i>)	<i>mới</i> 'new, recent'
15	仍<	réng 'still' (SV <i>nhưng</i>)	<i>những</i> [plural marker]
16	郎<	láng 'young man' (SV <i>lang</i>)	<i>lặng</i> 'quiet'
17	尼<	ní 'Buddhist nun' (SV <i>ni</i>)	<i>này</i> 'this'

Type 3: Sinograph is semantically adapted to represent a native Vietnamese morpheme

This single example is said to be an abbreviated form of 爲, created by isolating the top portion of the character. However, there is some disagreement about its derivation, as I will discuss in Section 4 below.

	<u>Nôm graph</u>	<u>Chinese morpheme (SV pronunciation)</u>	<u>Vietnamese morpheme</u>
18	𠄎	wéi ‘be, do, become’ (SV vì)	làm ‘do’

Category II: Created Sinographs

Types 4-6, described below, all involve sinographs that are not found in the Chinese script. They are Vietnamese inventions.

Type 4: Semantic-Semantic Compound Graphs

Two sinographs are combined into a new compound graph to represent a Vietnamese morpheme with semantics related to the semantics of the two Chinese morphemes conventionally written with the two sinographs.

	<u>Nôm graph</u>	<u>Vietnamese morpheme</u>	<u>Source of components</u>
19	𠄎	trời ‘sky’	天 (‘sky’) + 上 (‘above’)
20	𠄎	trùm ‘village leader’	人 (‘person’) + 上 (‘above’)
21	𠄎	tuổi ‘year of age’ ¹⁸	年 (‘year’) + 歲 (‘year of age’)

Type 5: Phonetic-Semantic Compound Graphs

These newly created graphs are analogous to phonetic-semantic compound Chinese characters. Note that the graph that is desemanticized and employed as a phonetic component need not be a Chinese character; example 26 shows that it may itself be a newly created *nôm* graph.

	<u>Nôm graph</u>	<u>Vietnamese morpheme</u>	<u>Source of components</u>
22	𠄎	nhiều ‘many’	堯 (nhiều) + 多 (‘many’)
23	𠄎	ít ‘few’	乙 (ít) + 少 (‘few’)
24	𠄎	ba ‘three’	巴 (ba) + 三 (‘three’)
25	𠄎	cỏ ‘grass’	草 (‘grass’) + 古 (cỏ)
26	𠄎	lời ‘spoken word’	口 (‘mouth’) + 𠄎 (trời ‘sky’)

¹⁸ This Vietnamese morpheme appears to be an early Chinese loan, an etymological doublet of the Sino-Vietnamese pronunciation *tué* for the Chinese character 歲.

Type 6: Phonetic-Phonetic Compound Graphs

This is the only type of *nôm* sinograph that has no parallel within the Chinese writing system. These sinographs were used to represent Vietnamese morphemes with consonant cluster initials.¹⁹ No single Chinese character could serve as an appropriate phonograph for such morphemes, because their Chinese pronunciations at the time (and thus their Sino-Vietnamese pronunciations) could not have had cluster initials.

<u>Nôm graph</u>	<u>Vietnamese morpheme</u>	<u>Source of components</u>
27 巔	<i>trái</i> < <i>blái</i> 'fruit'	巴 (<i>ba</i>) + 賴 (<i>lại</i>)
28 輪	<i>tròn</i> < <i>klon</i> 'round'	巨 (<i>cự</i>) + 倫 (<i>luân</i>)

The first graph is used alphabetically, by which I mean that it is a phonograph representing only the initial consonant sound of the Sino-Vietnamese pronunciation with which it is conventionally associated. This consonant sound is combined with the syllable indicated by the second phonographic element, resulting in a syllable that is homophonous, or nearly homophonous, with the Vietnamese morpheme that the new graph is meant to represent.

The components representing initial consonant sounds were conventionalized and few in number. For example, *b-* is always represented by the graphic element 巴 (Chinese *bā*, name of an ancient state).

It is important to note that every sinograph in the *nôm* writing system is a logograph, employed to write a specific morpheme. There are no PDPs in the *nôm* writing system; phonetic adaptations do not remain permanently desemantized. We have seen exactly the same thing in the development of the Chinese script as described in Section 2.

Abbreviation

Sometimes *nôm* graphs are abbreviated to serve the same purpose as the diacritic described earlier, namely to signal that the graph writes a morpheme other than the Sino-Vietnamese morpheme it is conventionally associated with. For example

¹⁹ These cluster initials are attested in early Vietnamese dictionaries employing Romanized spellings (such as the 17th century *Dictionarium annamiticum Lusitanum, et Latinum ope Sacre Congregationis de Propaganda Fide* by Alexandre de Rhodes), but have disappeared in the modern language.

	<u>Nôm graph</u>	<u>Chinese morpheme (SV pronunciation)</u>	<u>Vietnamese morpheme</u>
9	沒	mò 'sink' (SV <i>một</i>)	<i>một</i> 'one'

also appears in abbreviated form as

9b	𠂇	<i>một</i> 'one'
----	---	------------------

4. Linguistic Typology and the Adaptation of Logographs

I have argued that the basic techniques of adapting a foreign logographic writing system are inherent, and are related to the core processes used in the initial development of the logographic writing system itself. The very process of adaptation will inevitably involve the repurposing of those logographs as SDLs, PDLs and/or PDPs to write elements of the native language.

In the specific case of sinographic writing of Korean, Japanese, and Vietnamese, it is not surprising that phonographic usage should be the earliest and most visible adaptation technique. After all, this is the technique used within the Chinese writing system to represent foreign words, and therefore the technique that was used in Chinese writing to represent proper names in Korea, Japan, and Vietnam, whether that Chinese writing was made by Chinese or non-Chinese, within China or outside China. In other words, the phonographic usage of sinographs was modeled by the Chinese writing system itself, and thus was already available in the first stage of adapting those sinographs to represent indigenous languages.

The logographic use of sinographs through semantic adaptation was no less prevalent, and we have seen that it was widely employed in the earliest Korean, Japanese, and Vietnamese writing systems. As discussed earlier, this type of semantic adaptation differs from that found in the early developmental stages of the Chinese writing system; it is based on judgments of semantic equivalency similar to those applied in translation.

It should be pointed out that all of the adaptation techniques described above require written bilingualism. In the specific case of sinography, initially the techniques must have been carried out by people who could read and write Chinese, and thus who knew both the phonetic and semantic values conventionally associated with each sinograph within the Chinese writing system. (In practice, these Chinese pronunciations were filtered through the native phonological system.) They must of course also have known the indigenous language in question, so that those same graphs could be appropriately applied to represent native morphemes or sound sequences. However, it was not necessary for

subsequent users of the new native writing system to be bilingual; once the writing system was in use, learners had only to acquire the conventional associations of graphs with elements of the native language in order to become literate. Further extensions of the sinographs within the native system could then be made by monolinguals.

The twin techniques of phonetic and semantic adaptation employed when a logographic writing system is repurposed to write another language are likely independent of linguistic typology. Put simply, they are universal. It is in the subsequent application and development of those adapted graphs, and the creation of new graphs through the modification and combination of graphic elements, that the three writing systems under discussion here diverged in ways that I argue were typologically conditioned.

4a. Similarities between Korean and Japanese

Because of the similar typological characteristics of Korean and Japanese, in particular the presence of verbal inflection and case-marking particles, a purely logographic representation of these languages using sinographs would not have been possible, and this accounts for the development of a mixed logographic/phonographic writing system, with logographs representing verbal and nominal roots, and phonographs representing grammatical elements. That these similar developments were dependent as much on typology as on cultural contact is supported by the fact that those aspects in which the writing systems of the two areas did diverge can be persuasively attributed to minor typological differences, as shown in the following sections.

4b. Differences between Korean and Japanese: Alphabetic Spellers in Korean

Phonographic practice in early Japanese and Korean writing differed in one notable respect. In phonographic usage in the *Man'yōshū*, for the most part single phonographs are used to represent individual Japanese syllables (with the occasional phonograph used to represent two Japanese syllables). Korean *hyangchal* practice is similar; however, we also see the use of phonographs to represent single coda consonants. This is directly attributable to the greater complexity of Korean syllable structure compared with both Japanese and Chinese. Unlike Japanese, Korean syllables could end in a number of distinct consonant codas, including some (like *-s*) not found in Chinese. Employing phonographs to represent single consonants was the only way that Koreans could

represent certain inflectional endings.

4c. Differences between Korean and Japanese: Development of a Syllabary

Differences between Korean and Japanese syllable structures also explain why, in Japan, phonographs eventually developed into syllabary scripts. The number of distinct syllables (or, more properly, *moras*) in Japanese was fewer than 100. It was therefore entirely practical to develop a new script based on the conventionalized use of a different graph for each possible moraic element of the language.

The same initial conditions that led to the development of *kana* syllabaries in Japanese were also present in early Korean writing practice, namely, the phonographic use of sinographs, combined with the use of such graphs in abbreviated form to annotate Chinese texts. In the abbreviated *gugyeol* graphs, we see what looks very much like the beginnings of a *kana*-like syllabary. However, this is misleading. While the closed set of syllables needed to represent the inflectional endings of Korean verbs could be represented by a manageable number of phonographs, the very large number of distinct Korean syllables could not be. It was not only that the sheer number of phonographs required would be too large, but more crucially that the inventory of sinographs, with their Chinese-based Sino-Korean pronunciations, were inadequate even for the approximate representation of the more complex native Korean syllable shapes.

In short, the phonology of Korean placed a huge constraint on the development of syllabic phonographs. They could be used for particular purposes (such as a partial syllabary for *gugyeol*), but the process was not generalizable to a full-fledged phonographic script.

4d. Unique developments in Vietnamese: Logography

The most striking and significant difference between Vietnamese sinography on the one hand and Japanese and Korean on the other is that Vietnamese *chữ nôm* was a strictly logographic script, whereas Japanese and Korean were mixed scripts employing many phonographs. Why did Vietnamese remain logographic? Put another way, when sinographs were desemanticized to become phonographs to represent Vietnamese, why were they resemanticized and associated with specific morphemes?

Two related factors account for this. The first is the isolating monosyllabic typology of Vietnamese. Desemanticized sinographs have monosyllabic pronunciations. When used to write Japanese or Korean words, in many cases they

could represent only meaningless parts of polysyllabic morphemes. These graphs would not be associated by Koreans and Japanese with morphemes, and resemanticization into logographs was therefore not possible. In contrast, every time a sinograph was desemanticized to write a Vietnamese syllable, it was also, in each specific instance, writing a Vietnamese morpheme. So the possibility of resemanticization was always present.

Why was this possibility realized? Chinese writing was the model for Vietnamese writing, and because of the typological similarities of the two languages, it was possible to put the ideals of Chinese writing practice—one graph per morpheme—into practice in Vietnamese sinography. The very same disambiguating techniques used to diversify Chinese characters and maintain a one-to-one relationship between graphs and morphemes in Chinese writing were applied by the Vietnamese as well, so that resemanticization of phonographs into logographs was reinforced by the addition of structural elements creating new, unique graphs.

4e. Unique developments in Vietnamese: Lack of Abbreviation

Aside from one notable example, we don't see the kind of drastic abbreviation in Vietnamese sinographs that we see in Korean *gugyeol* or Japanese *kana*. This is because the writing system as a whole remained logographic. There is of course a natural tendency toward abbreviation in all writing systems; it is counteracted by the need to maintain distinctions among the contrasting graphs in the system. Abbreviation can progress farther when applied to a limited set of phonographs; but were abbreviation to obscure the internal structure of compound logographs, whose semantic and phonetic elements are crucial for both the memorization and recognition of large numbers of graphs, the entire system would break down. Logography, made possible by Vietnamese typology, was itself the major constraint on abbreviation, just as it has been in the history of Chinese writing.

It is worth turning to that one example, the sinograph representing the Vietnamese verb *làm* 'do'. This is the most high-frequency verb in Vietnamese, and the original graph 爲 has a large number of strokes. Pressure for abbreviation was probably highest for this one graph, accounting for the forms 𠄎.²⁰

Another type of abbreviation (what Seeley would call isolation) is found in Vietnamese *nôm*. We have seen for example the abbreviation of 沒 to 𠄎. In this case, the abbreviation serves to increase, rather than decrease, the salient

²⁰ See section 4g below for more discussion of this graph.

distinctions among graphs, since the abbreviated form is employed when writing the native Vietnamese morpheme *một* ‘one’, in contrast with the homophonous Sino-Vietnamese morpheme *một* ‘sink’ written with 沒.

4f. Unique developments in Vietnamese: Creation of New Graphs

An adopted logographic system is by necessity going to extend the number of graphs, because the original set of graphs cannot uniquely represent all of the morphemes of the borrowing language. It is natural, in the process of creating new graphs, to employ the techniques already apparent within the existing script, since they present a ready solution. In the case of Vietnamese, that meant creating new phonetic-semantic and semantic-semantic compounds on the model of those character types already found in the Chinese writing system.

Differences between Vietnamese character types and Chinese are due to two factors: 1) the greater pool of semantic and phonetic associations to choose from, since two languages are involved; 2) the need for phonetic-phonetic compounds because of the presence of syllable types too distinct from Chinese to permit the direct phonetic adaptation of Chinese characters based on their Sino-Vietnamese pronunciations.

4g. Unique developments in Vietnamese: Scarcity of Simple SDLs

In Japanese and Korean sinography, semantically-derived logographs were regularly employed to represent verbal and nominal roots. As a result, SDLs were extremely common (and still are in modern Japanese writing). Yet from my survey of the very limited academic literature available to me on Vietnamese *nôm* writing, there appears to be only a single SDL directly borrowed from the Chinese script: 爲 (Chinese *wéi* ‘make, do, serve as’), writing the Vietnamese verb *làm* ‘do’ in an abbreviated form 𠄎.

In fact, even the status of this one example is in some doubt. Nguyễn (1990:408) points out that one scholar has argued that 𠄎 is not an abbreviation of 爲, but of 濫 (Chinese *làn* ‘overflow’; Sino-Vietnamese *lạm*). This would of course mean that the graph writing *làm* ‘do’ is phonetically, not semantically, derived. And that would in turn mean that Vietnamese *nôm* entirely lacks SDLs aside from those compound graphs created within the *nôm* system.

Why should one of the most common sinographic practices in Korean and Japanese be almost entirely lacking in Vietnamese? I believe it is because

Vietnamese lacks inflectional endings and case-marking particles. In Japanese and Korean, these endings and particles help to signal the presence of native verbal and nominal roots, whose morphological behavior differs from that of borrowed Chinese roots.²¹ In this way, based on context, it can be determined whether a sinograph writes the borrowed Chinese morpheme with which it is conventionally associated, or writes a synonymous native morpheme.

Because Vietnamese lacks these morphological clues, SDLs would more often be ambiguous. This same ambiguity does not result from the other techniques of *nôm* character usage and creation for the representation of native morphemes, namely phonetically derived logographs and creation of graphs not found in the Chinese script (whether semantic-semantic compounds, phonetic-semantic compounds, diacritically-marked phonographs, or abbreviated forms).

5. Conclusion

As we have seen, the basic method of adapting Chinese characters was based on two universally available techniques: repurposing individual graphs as phonographs (by disregarding their semantic content) or as logographs (by rephonologizing them to represent synonymous native morphemes). If phonetically-adapted graphs were used to represent specific morphemes or words, they became resemanticized and functioned as logographs within the new writing system. Otherwise they remained phonographs (and, in the case of Japanese, ultimately evolved into members of a syllabary).

Aside from these basic commonalities of development, the specific ways that sinographs were employed to write Korean, Japanese, and Vietnamese varied considerably. Developments for Vietnamese were most markedly different, and I believe that this difference is largely attributable to the very different typological features of the Vietnamese language. At the same time, I think that the similarities found between Korean and Japanese are due to the similarities in their linguistic typologies. Where differences in the development of writing are seen, they correspond well with typological differences, principally the more complex syllable structure of Korean as compared with Japanese.

I do not by any means intend to dismiss the influence and effect of other factors (including cultural contact, political and economic developments, aesthetic

²¹ For example, in both Korean and Japanese, Chinese verbs are borrowed as nouns, to which the verb *hada* 하다 or *suru* する (meaning ‘to do’) is added. With few exceptions, borrowed Chinese morphemes cannot be directly inflected in either language.

sensibilities, etc.) or of historical contingencies in shaping the early development of writing in these three areas of Asia. I do, however, hope that I have demonstrated the likelihood that linguistic typology provides significant constraints within which those other factors must operate.

This hypothesis is testable and generalizable. Korean, Japanese, and Vietnamese are not the only languages that have developed sinographic writing systems. The Altaic Jurchen and Khitan peoples, and the Tai Zhuang people, also developed sinographic writing systems. Though less well studied and less well understood than the developments I have discussed above, these writing systems allow us to test predictions that follow from my hypothesis. Simply put, the hypothesis states that the means of adapting sinographs to represent the agglutinating, inflecting Altaic languages will closely resemble the developments we have seen for Korean and Japanese, while the isolating, monosyllabic Tai language will develop sinographic writing in ways parallel to the Vietnamese model. I intend in the future to investigate those writing systems in more detail in order to test the hypothesis.

More broadly, the notion of typological influence can be applied to our understanding of how writing systems change in general. The primary factor in the development of phonographic writing systems from logographic writing systems may well be typological distance between donor and borrower languages. Conversely, as in the case of Chinese and Vietnamese, typological similarity may well enable logographic writing systems to remain logographic even as they are modified to write a different language.

Similarly, it would be a fallacy to assume that logographic writing systems used continuously to write a particular language will naturally develop into phonographic systems. It may well be that such a change will not take place unless the typology of the underlying language shifts significantly. Thus the most important factor in the persistence of the logographic Chinese writing system may not be cultural, but linguistic.

It is my hope that the very preliminary work done here on sinography will ultimately take us closer to a comprehensive theory about the causes of specific kinds of changes that occur during transfers of writing systems.

6. Coda

Given the venue of this paper presentation, it is appropriate to conclude with some words about the development of the Korean alphabet, originally called Hunminjeongeum 훈민정음 訓民正音 (“Proper Sounds for Enlightening the

People”) and now known as Hangeul 한글. The alphabet was invented in 1443 and promulgated by King Sejong 세종대왕 in 1446. After several centuries of relative obscurity, the alphabet became an indispensable element of Korean writing toward the end of the 20th century, and today has almost completely displaced sinographs as the exclusive Korean script in daily life.

I have noted the significant parallelisms in the development of writing in Japan and Korea, and argued that the similar typologies of those two languages were at least as much of a factor as cultural contact in driving those parallel developments. Yet unlike the modern Japanese writing system, the Korean alphabet is not derived from sinographs. It represents a clean break with all that had come before in the history of Korean writing. To what can we attribute its primacy in Korea today, and what explains this radical and growing divergence between Korean and Japanese writing over the last 500 years?

I would argue that the success of the Korean alphabet can be attributed in part to the complexity of the language’s syllable structure. While the sinographic writing system known as *idu* 이두 was perfectly adequate for most purposes in the pre-modern era, it was ultimately not as flexible as Japanese writing.²² This was because the set of sinographs employed phonographically to render the sounds of native Korean words and grammatical elements was too limited to adequately represent Korean phonology. Moreover, it always would be; the phonological system of Sino-Korean pronunciations was too impoverished to supply the raw materials even for a Korean alphabet, let alone a syllabary. This was in marked contrast to Japanese, where the simple phonological structure of the language meant that phonographs could be systematized into a functional script. Once Korean society had developed to the point where widespread literacy had become both feasible and desirable, yet where it was impractical to expect all citizens to master Literary Chinese in order to achieve literacy, *idu* could no longer serve as the primary writing system. Koreans turned instead to Hangeul.

King Sejong was, arguably, ahead of his time in creating and promoting an alphabet that was opposed by many literati of the era, and which was largely neglected for centuries afterward. His recognition of the value of such an alphabet, however, was not misguided, and has been justified in modern times by its

²² *Idu* was a writing system that used sinographs both logographically and phonographically to represent word roots and grammatical elements, respectively. However, the lexicon and aspects of the syntax of the written language were highly sinicized; its use required a fair degree of mastery of Literary Chinese and was inadequate for the written representation of less sinicized forms of Korean.

widespread adoption and demonstrated utility.

Bibliography

- Baxter, William H. 1992. *A handbook of Old Chinese phonology*. Berlin, New York: Mouton de Gruyter.
- Bentley, John R. 2001. "The origin of Man'yōgana." *Bulletin of the School of Oriental and African Studies* 64.1: 59-73.
- Boltz, William G. 1994. *The origin and early development of the Chinese writing system*. New Haven: The American Oriental Society.
- Boltz, William G. 2006. "Pictographic myths". *Bochumer Jahrbuch zur Ostasienforschung* 30:39-54.
- Buzo, Adrian. 1980. "Early Korean writing systems." *Transactions of the Korea Branch of the Royal Asiatic Society* 55:35-62.
- Cho, Seung-Bog. 1990. "On Idu". *Rocznik Orientalistyczny* 46.2:23-32.
- Coulmas, Florian. 1991. *The writing systems of the world*. Oxford: Blackwell.
- Daniels, Peter T. and William Bright, eds. 1996. *The world's writing systems*. New York: Oxford University Press.
- DeFrancis, John. 1989. *Visible speech: the diverse oneness of writing systems*. Honolulu: University of Hawai'i Press.
- Ferlus, Michel. 1992. "Histoire abrégée de l'évolution des consonnes initiales du vietnamien". *Mon-Khmer Studies* 20:111-125.
- Habein, Yaeko Sato. 1984. *The history of the Japanese written language*. Tokyo: University of Tokyo Press.
- Hé, Chéng 何成 et al. 1960. *Yuè-Hàn cídiǎn 越漢辭典 = Từ điển Việt Hán [Vietnamese-Chinese dictionary]*. Beijing: Shangwu yinshuguan.
- Hannas, William C. 1997. *Asia's orthographic dilemma*. Honolulu: University of Hawai'i Press.
- Hyman, Malcolm D. 2006. "Of glyphs and glottography". *Language & Communication* 26:231-249.
- Jīn Lǐxīn 金理新. 2006. *Shàngǔ Hànyǔ xíngtài yánjiū 上古漢語形態研究 [A study of Old Chinese morphology]*. Hefei: Huangshan.
- Keightley, David. 1978. *Sources of Shang history: the oracle-bone inscriptions of bronze-age China*. Berkeley: University of California Press.
- Kim, Yongduk. 1999. "On I-du: writing the Korean language in Chinese characters." *Transactions of the Korea Branch of the Royal Asiatic Society* 74:49-57.
- Lee, Iksop and S. Robert Ramsey. 2000. *The Korean language*. Albany: State University of New York Press.
- Lee, Kimun 이기문. 1972. *Gugeosa gaeseol 국어사개설 國語史概說 [Introduction to the*

- history of Korean*]. Seoul: Minjung Seogwan.
- Lurie, David B. 2001. *The origins of writing in Early Japan: from the 1st to the 8th century C.E.* Columbia University doctoral dissertation.
- Miller, Roy Andrew. 1987. "Chinese script in Korea and Japan". *Asian and Pacific Quarterly of Cultural and Social Affairs* 19.2:1-18.
- Nam, Punghyun. 1994. "On the Relations Between Hyangchal and Kwukyel". In Young-Key Kim-Renaud, ed. *Theoretical issues in Korean linguistics*, Stanford: CSLI Publications, 419-423.
- National Museum of Korea. 2007. *Seokgatap balgyeon yumul josa junggan bogo* 석가탑 발견 유물 조사 중간 보고 [Interim report on the investigation of the relics unearthed from the Seokga Pagoda]. Seoul: National Museum of Korea.
- Nguyễn, Đình-Hoà. 1959. "Chữ nôm, the demotic system of writing in Vietnam". *Journal of the American Oriental Society* 79.4:270-274.
- Nguyễn, Đình-Hoà. 1990. "Graphemic borrowings from Chinese: the case of chữ nôm--Vietnam's demotic script". *Bulletin of the Institute of History and Philology* 61.2:383-432.
- Norman, Jerry. 1988. *Chinese*. Cambridge: Cambridge University Press.
- Qiú, Xīguī 裘錫圭. 2000. *Chinese writing*. Berkeley: The Society for the Study of Early China and The Institute of East Asian Studies. [English translation by Gilbert L. Mattos and Jerry Norman of *Wénzìxué Gàiyào* 文字學概要, Shangwu, 1988.]
- Rogers, Henry. 2005. *Writing systems: a linguistic approach*. Oxford: Blackwell.
- Sagart, Laurent. 1999. *The roots of Old Chinese*. Amsterdam, Philadelphia: John Benjamins.
- Salomon, Richard. In press. "Some principles and patterns of script change". Proceedings of The Shape of Script conference, School for Advanced Research, Santa Fe NM, 2007.
- Sampson, Geoffrey. 1985. *Writing systems: a linguistic introduction*. London: Hutchinson.
- Seeley, Christopher. 1991. *A history of writing in Japan*. Leiden, New York, København, Köln: E.J. Brill.
- Senner, Wayne, ed. 1989. *The origins of writing*. Lincoln, Nebraska: University of Nebraska Press.
- Shibatani, Masayoshi. 1990. *The languages of Japan*. Cambridge, New York: Cambridge University Press.
- Sohn, Ho-Min. 1999. *The Korean language*. Cambridge, New York: University of Cambridge Press.
- Trần, Văn Chánh. 1999. *Từ điển Hán Việt: Hán ngữ cổ đại và hiện đại = Gǔjīn Hàn-Yuèyǔ cídiǎn* 古今漢越與辭典 [Historical Chinese-Vietnamese dictionary]. Hồ Chí Minh City: Trẻ.