

A sampler of the world's writing systems

*J. Marshall Unger
Department of East Asian Languages & Literatures
The Ohio State University*

2009

By world standards, writing is uncommon. Most of the languages that have ever been spoken lack writing systems. People improvise small, ad hoc sets of signs or signals whenever necessity demands a way to indicate specific objects or quantities, and no doubt have been doing so for thousands of years. But full writing—a system of marks capable of transcribing any and all utterances of a particular language—requires using signs that stand for speech sounds rather than directly for things or ideas. That two-dimensional signs can be so used—the rebus principle—was independently discovered just three times: once in Mesopotamia about 4,000 to 5,000 years ago; once in China perhaps 3,000 to 4,000 years ago; and once in Mesoamerica just over 2,000 years ago. Today's principal writing systems can all be traced back ultimately to one of the two Asian systems.

This exhibit attempts to convey the graphic diversity of writing systems past and present. Every writing system must strike a balance between logographic and phonographic strategies for representing speech sounds: too much emphasis on words and phrases makes a system unlearnable; too much phonetic detail makes it too cumbersome for daily use. Still, there is a rich variety of ways in which the necessary balance can be achieved. The samples shown here reflect the differences in the structures of the selected languages, their histories, and the cultures that sustain them.



Writing systems of antiquity

Runes are found in inscriptions preserved throughout northern Europe, Scandinavia, the British Isles, and Iceland from about 100 BCE to 1600 CE. A similar-looking Old Turkic script (from which an Old Hungarian script of the Early Middle Ages derives) is sometimes called *runic*, but is unrelated. Though it is now accepted that descendants of the Vikings from Greenland visited Newfoundland, claims of early runic inscriptions in Minnesota and other places where immigrants from Norway and Sweden later settled are not supported by solid evidence.

Linear B was used for writing Mycenaean, an early form of Greek, mostly for making lists and other commercial purposes. Vowels were mostly ignored in this iGreek alphabet by several

centuries. It developed into a syllabary that continued in use on Cyprus alongside the new alphabet until the 2nd century BCE.

Old Chinese refers to the Chinese spoken from the Chinese Bronze Age down to the Qin unification and start of the Former Han Dynasty (206 BCE to 9 CE). The character shapes preserved on early bronzes and so-called oracle bones are quite different from their later developments, and many logograms were pressed into service as phonograms. For instance, 勿 began as a pictogram for 'creature', Old Chinese *mjut. It came to be used for *mjut 'do not!' because that word had the same sound. To keep things straight, Chinese later added 牛 (from 牛 'ox') to 勿 just when it was used to write the syllable for 'creature'. Such innovations enables the inventory of characters to expand greatly.

Egyptian hieroglyphs ("sacred carvings" in Greek) consisted of a combination of logographic and phonographic signs. Vowels were generally not indicated, though some can be inferred from the descendant language Coptic. Egyptians used cursive hieroglyphs on papyrus and wood. The Rosetta Stone contains parallel texts in Egyptian in formal picture-like hieroglyphs, the graphically reduced demotic script, and in Greek. Though partially damaged, the matching portions gave Champollion his starting point for the decipherment of Egyptian.

The *Mayan* glyphs evolved in the pre-Columbian Mayan civilization of Mesoamerica. The earliest inscriptions date to the 3rd century BCE. Writing was in continuous use until shortly after the arrival of the Spanish conquistadors in the 16th century CE, who destroyed almost all texts that provided information on how the writing system worked. The recent decipherment required the inspired guess that the system was essentially a syllabic (with a few logograms thrown in), and the application of modern linguistic analysis of extant Mayan dialects. Mayan glyphs have nothing to do with Egyptian hieroglyphs.

Cuneiform emerged in Sumer in the 3rd millennium BCE. It is derived from a late 4th millennium (Uruk IV period) system of pictograms. The crucial turning point was the discovery of the REBUS PRINCIPLE. The linguistic affinity of Sumerian is unknown, but the Akkadians, who spoke a Semitic language, treated it as a prestige language much as Latin was treated in medieval Europe. They adapted cuneiform to their needs, as did the Late Bronze Age Hittites of northern Anatolia, who spoke an Indo-European language.



Painted wall signs, Piraeus, Greece

Greek

The **Greek** alphabet has been used to write Greek since about 800 BCE. It is the first true alphabet in that it gives vowel and consonant letters equal graphic status. It has been hypothesized that this innovation may have been the invention of a single individual who wished to be able to transcribe dactylic hexameter verse with precision—something not easily done in a system like Linear B. Letters of the Greek alphabet derive from those of the Phoenician alphabet, and are not related to the those of Linear B or the Cypriot syllabary. They provided the model for the alphabets of ancient Italy, but later still the design of most of the Cyrillic letters.



"Apeklî kazası" haberine açıklama
Turkish newspaper



German magazine



Indonesian tabloid



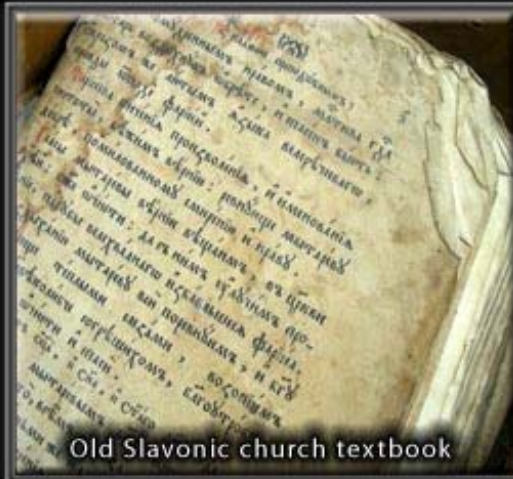
Online Vietnamese newspaper

Latin

The **Latin** or Roman alphabet is the most widely used in the world today. It evolved from a western variety of the Greek alphabet called Cumaeian. During the Middle Ages, it was adapted to the Romance languages, the direct descendants of Latin, and eventually almost all the languages of Europe, including pre-Indo-European Basque and the Uralic languages Hungarian, Finnish, and Estonian. European colonialism spread the Latin alphabet overseas. Modern Turkish, Indonesian, and Vietnamese are now written with Latin-based writing systems, and the International Phonetic Alphabet used by linguists and lexicographers is largely based on the Latin alphabet.



Car with a Bulgarian message,
Plovdiv, Bulgaria



Old Slavic church textbook



McDonald's in Russian, Moscow, Russia



Casino sign, Dnepropetrovsk, Ukraine

Cyrillic

Cyrillic alphabets (there are many variants) have been used in the writing of six Slavic national languages—Belarusian, Bulgarian, Macedonian, Russian, Serbian, and Ukrainian—as well as such non-Slavic languages of the former Soviet Union as Kazakh, Uzbek, and Kyrgyz (Turkic), Tajik (Indo-European), and Mongolian. Letter forms in Cyrillic derive from mostly from the Greek alphabet, though letters from the pre-Cyrillic Glagolitic script, Latin, and even Hebrew also provided models.



Street sign in Yerevan, Armenia

Armenian

The **Armenian** alphabet of 36 letters was devised by the cleric Mesrop Maštoc in 406 or 407 CE. The order of the letters loosely follows the order of the Greek alphabet, but Mesrop was guided by the bi-uniqueness principle: one letter for each sound of Classical Armenian, and only one sound for each letter.



Advertisement for computer classes, T'bilisi, Georgia

Georgian

The **Georgian** alphabet (*mxedruli*) of 33 letters is used to write Modern Georgian, and has been used to write other South Caucasian (Kartvelian) languages, and even the Ossetic, an Eastern Iranian language, in the 1940s. The order of letters and their sounds derive from the Greek alphabet. The earliest texts date from the beginning of the 5th century ce. Letter shapes evolved, reaching something close to their present form in the 11th century.

मुस्लिम मैट्रिमोनियल सर्विसेज
 Book Your Matrimonial Ad. Free & Ad. Your Profile Free For A Short Period

मुस्लिम मैट्रिमोनियल

मासिक हिन्दी / अरबी परिवारिक एवं वैवाहिक मुस्लिम पत्रिका
Muslim Matrimonial Monthly Magazine

शेड: कोर एंजस्ट्रेशन एवं डोम रजिस्ट्रेशन के बिना भी रिजले पर विचार नहीं किया जाय है। रिजले का प्रोटेक्ट-कॉपीरेट सर्वाधिकारी होने पर ही एंजस्ट्रेशन किया जाय है। कोर सर्वाधिकारी एंजस्ट्रेशन के बिना भी रिजले पर विचार नहीं किया जाय।
 वेदा, वेदी, चर्च-मार्ग का रिजले प्रोफाइल की तरह ही तय कराया जाय है।

'मुस्लिम रिजले' संपन्न, मोघ, कुतब, पयन, जंगारी, जर्न, मंगरी, कुंरी, रपरेज, इरासी, सलमान, सुन्नी व शिया धर्मिक उपवासगीय प्रोफाइल से डाक्टर, इन्जीनियर, नर्सिंग, अडमिनिस्ट्रेशन, फार्मिसेट एल्फार्स, उच्च अवस्था, बाल्येड एजुकेशन, लडके/लडकिया एवं विदुर-विदवा सलमानुय उपबन्तगीय प्रोफाइल से रिजले हेतु सम्पर्क करें। **Mob: 9451402293**

मुस्लिम मैट्रिमोनियल सर्विसेज (रजि.)
मुस्लिम मैरिज ब्यूरो

मुस्लिम प्रोफेशनल महिलाओं एवं पुरुषों द्वारा सर्वाधिकारी
 HAS, IFS, IPS, IAS, IIT, IIM, ER, DR, NRI, S/W PROP, BUSINESS, GROOMS & BRIDES
 * HOME VISITS * LIAISONING * HORO. MATCHING * ARRANGING MEETINGS
 ONLINE & OFFLINE MATRIMONIAL SERVICES & ADVERTISER

कृपया समय लेने पर ही मिलें।

Consultant charge Rs. 600/-
 सेवाएं रिजले हेतु सलमान
9451402293

सभी रिजले गोपनीय रखे जाते हैं।
 जिन्दगी की राह में अपना
 इमसफर चुनने का एक खास मौक़ा!!

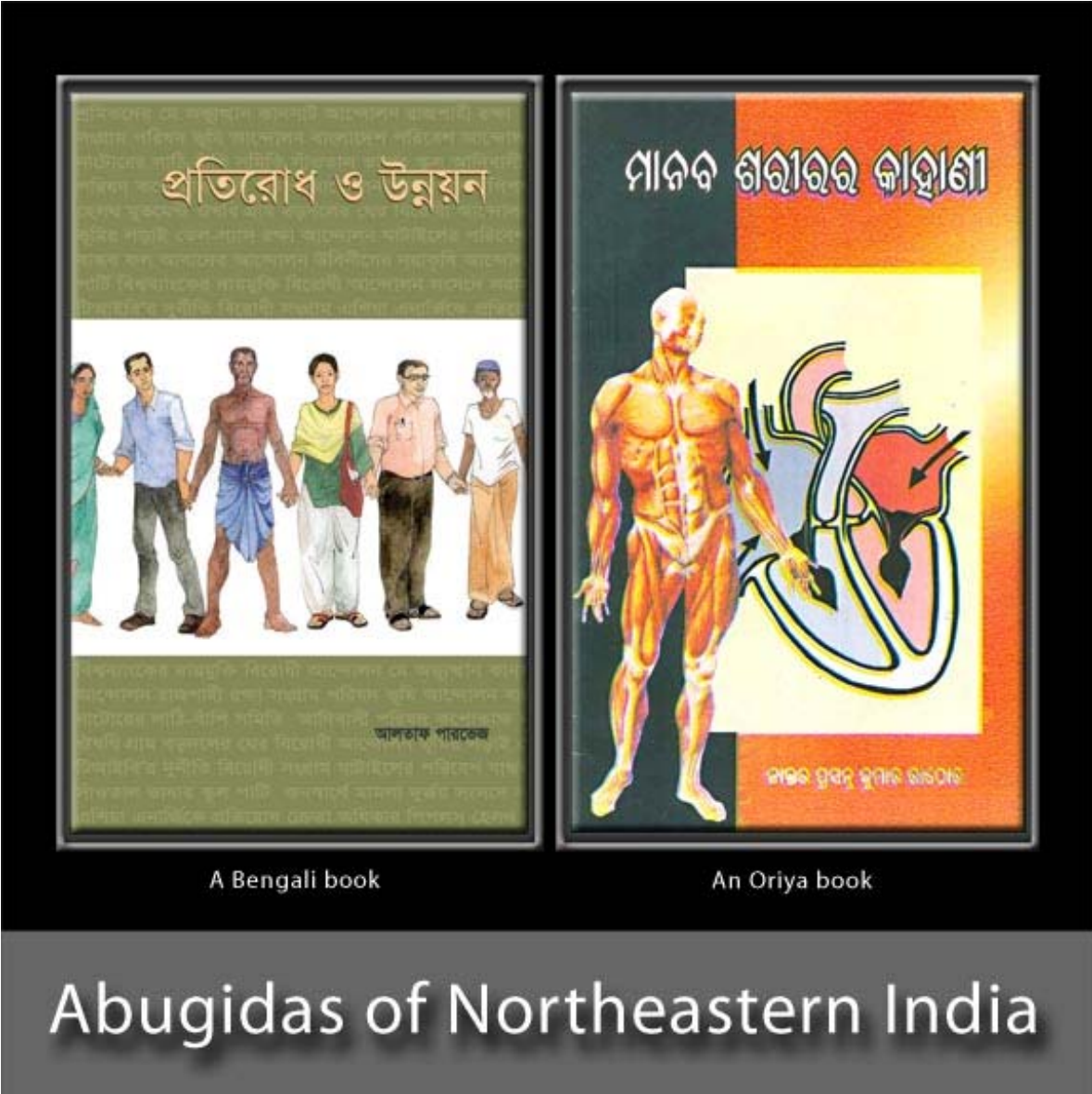
"GHAR BAITHE GHAR BASAO OFFER"
 Form Fee Rs.100/-

निकाह घर आइयो! रिजले लेकर जाइयो! रिजले का प्रोफाइल ऑनलाइन प्रोफाइल के रिजले

Matrimonial advertisement from a Hindi newspaper

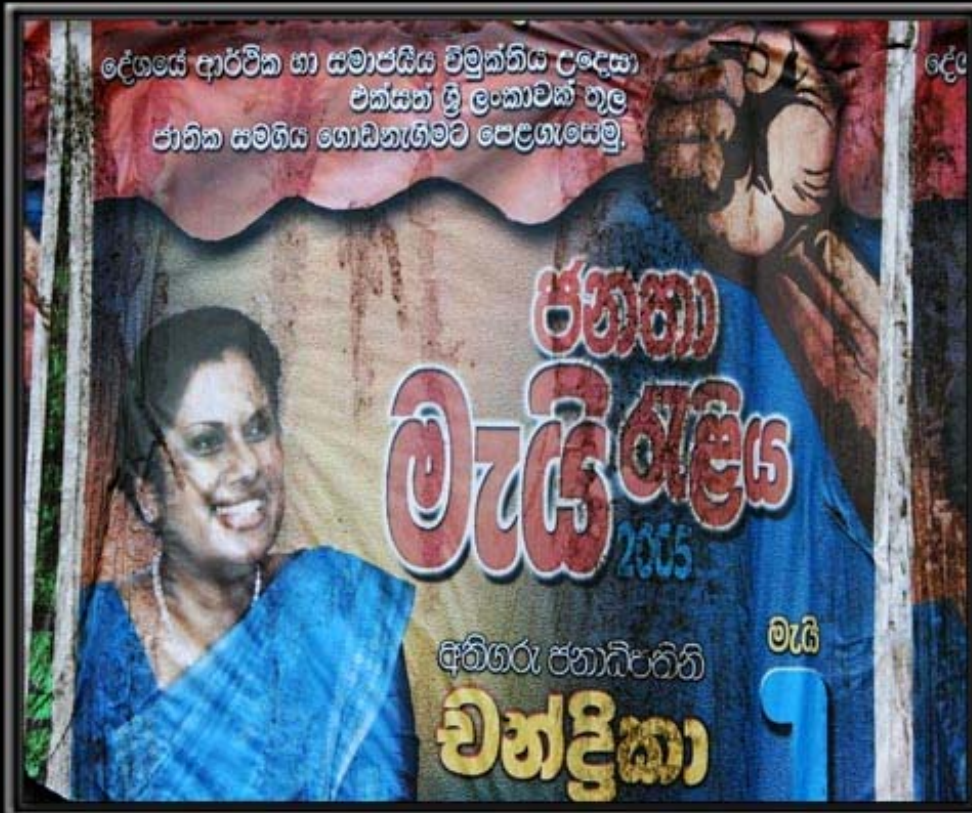
Devanagari

Devanāgarī or **Nāgarī**, is an abugida of India and Nepal. In an abugida, the “unit” letters stand for simple syllables (e.g. प /pa/ = [pə]). Auxiliary letters or diacritics placed above, below, to the left, or to the right of these letters specify a different vowel (पा /pā/ = [pa:], पि/pi/, etc.) or indicate the absence of a vowel (प् /p/). Ligatures (combined letters) are fairly common. Devanāgarī is written left to right. Lines of text appear to have a horizontal line running along the tops of the letters linking them together. Devanāgarī is the principal script for Hindi, Marathi, and Nepali. It has been adapted to the writing of many other languages of the Indian subcontinent, and was formerly used to write Gujarati. Since the 19th century, it has been the most commonly used script for Sanskrit and Pali.



The Bengali (or Bangla) abugida is closely related to Devanāgarī, from which it started to diverge in the 11th century ce. The current printed form of Bengali first appeared in 1778, when Charles Wilkins introduced printing in Bengali. A few archaic letters were modernized during the 19th century.

The Oriya abugida derives from the Kalinga variety of the ancient Brāhmī script. The earliest known inscription in the Oriya language dates from 1051. The curved appearance of letters is a result of writing on palm leaves, which may be torn when making many straight strokes. The script is used for Dravidian and Munda languages spoken in Orissa state, and for Sanskrit there as well.



Sinhalese placard for a political rally for Chandrika Bandaranaike, former president of Sri Lanka

Sinhala

The **Sinhala** (or Sinhalese) abugida, another descendant of Brāhmī, is first seen in inscriptions of the 3rd and 2nd centuries bce. Both the alphabet and the language have changed considerably since then. The earliest surviving literature in Sinhala dates from the 9th century ce. Sinhala is also used to write Pali and Sanskrit Buddhist texts in Sri Lanka.



Dravidian

The **Dravidian** family includes approximately 73 languages, including the four literary languages of Kannada, Malayalam, Tamil, and Telugu. They were probably once spoken throughout the Indian subcontinent—the language of the prehistoric Indus civilization may have been Dravidian—but are now (except for Brahui) most heavily concentrated in the south as the long-term result of the invasions that introduced Indo-European languages many centuries ago. Inscriptions in Dravidian languages date from the 6th century bce.

The *Kannada* abugida developed from the Kadamba and Cālukya scripts, descendants of Brāhmī, which were used between the 5th and 7th centuries ce. Old Kannada script, from about 1500, developed into

modern Kannada and Telugu writing, which became standardized under the influence of Christian missionary organizations at the beginning of the 19th century.

Malayalam first appeared in writing in the *vazhappalli* inscription of about 830 ce. In the early thirteenth century, the Malayalam script began to develop from a script known as *vattezhuthu* (round writing), a descendant of Brāhmī. As a result of the difficulties of printing Malayalam, a simplified version was introduced during the 1970s and 1980s. The main changes involved writing consonants and diacritics separately rather than as complex characters. Muslim speakers of Malayalam in Singapore and Malaysia, and occasionally in the Indian state of Kerala use an Arabic-based script.

The earliest known *Tamil* inscriptions date back to at least 500 bce. The oldest literary text in Tamil, *Tolkāppiyam*, was composed around 200 bce. The Tamil alphabet is thought to have evolved from the Brāhmī script, though some scholars believe it may have roots in the so far undeciphered Indus script. Originally developed for writing literary Tamil, it poses certain problems for writing the colloquial language, and various attempts to overcome them were made during the 19th century. Nowadays the colloquial written language appears mainly in schoolbooks and in passages of dialogue in fiction.

The *Telugu* alphabet is closely related to that of Kannada. The earliest known inscriptions in Telugu date from the 6th century ce and Telugu poetry begins to appear during the 11th century. Until the 20th century, written Telugu was very different in structure from the vernacular language. During the second half of the 20th century, a new written standard emerged based on the modern vernacular.



Decorative relief wood carving, imperial summer palace, Beijing, China

Chinese

Chinese is actually a language family; its so-called dialects are as diverse as the languages of the Romance, Slavic, or Germanic families of Europe. Most of the words of the modern languages of the Chinese family consist of two or more syllables, but each syllable is written with a single character; many different characters are used for syllables of the same sound, the correct choice depending on the meaning of the word in which the syllables occurs. The largest character dictionaries (such as the multivolume work by the Japanese scholar Morohashi Tetsuji) contain more than 50,000 entries, but many are graphic variants of other characters that have their own headings.

Literary Chinese had an impact on non-Sinitic languages in adjacent cultures, with the result that Chinese characters (*hanzi*) have been adapted to writing in Japan (where they are called *kanji*), Korean (*hanja*), and Vietnamese (*hán tự*).

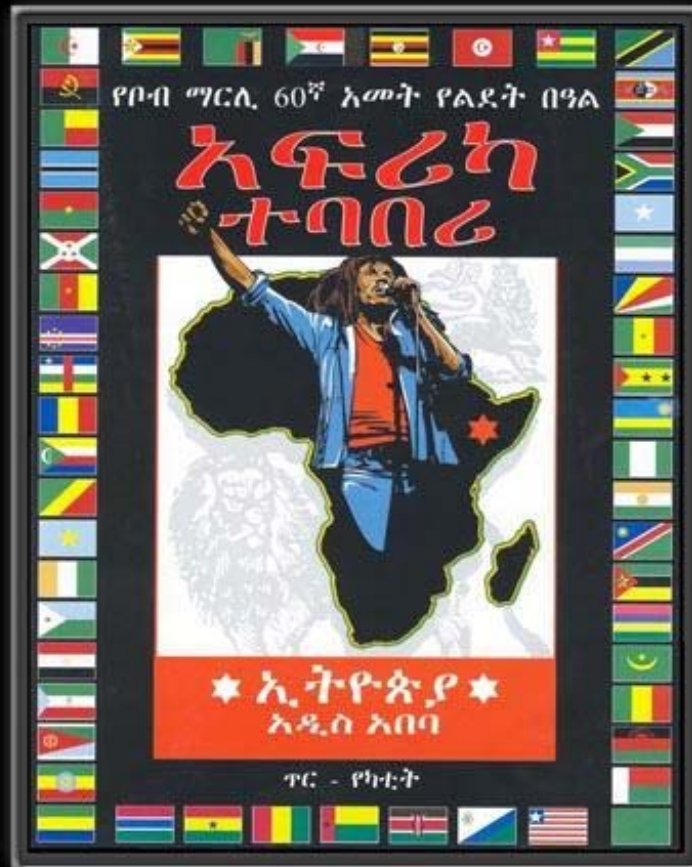


Evening commuters at Shinjuku subway station, Tokyo, Japan

Japanese

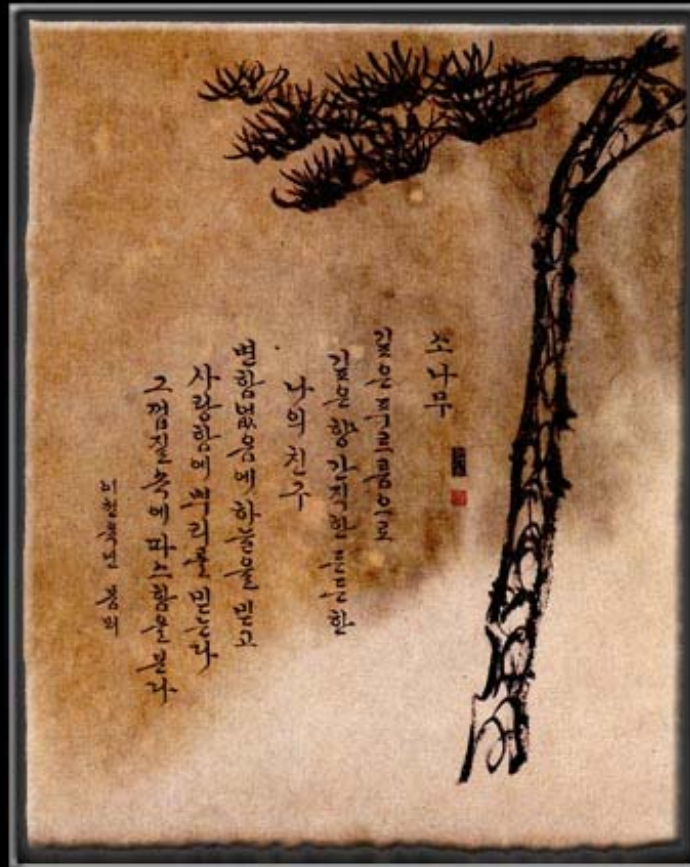
The **Japanese** writing system now comprises three types of symbols, used in somewhat loosely prescribed combinations: (1) *kanji* or Chinese characters; (2) *kana* of two styles, *hiragana* and *katakana*, both anciently derived from Chinese characters; and (3) Western alphanumeric symbols (A-Z, a-z, 0-9, etc.), introduced in the latter half of the 19th century for a variety of special purposes. *Katakana* is used primarily for words borrowed from foreign languages other than Chinese; *hiragana* as the modern default syllabic script; *kanji* represent parts of Sino-Japanese loanwords and coinages and common native roots.

Modern Japanese may be written horizontally (*yokogaki*), like English, or vertically (*tategaki*), in traditional Chinese format. In practice, a Japanese magazine or newspaper article seldom involves more than 600 or 700 distinct characters.



Amharic

Amharic, the second most widely spoken language of the Semitic family after Arabic, is written with letters that function much like modern Japanese *kana*. Though other speech communities exist in Ethiopia, Amharic is the official working language of the nation.



South Korean calligraphic art

Korean

The **Korean** alphabet was promulgated by King Sejong (1418–1450) in 1446. The alphabet was originally called *Hunmin chōng’ūm*, or “the correct sounds for the instruction of the people,” but is best known by the modern name *han’gŭl*, a term coined by the Korean linguist called Chu Sigyöng (1876–1914). Though the king and the scholars who assisted him were probably influenced by the example of the ‘Phags-pa alphabet designed at the command of Qubilai Khan for use in his multilingual empire, but Sejong’s foresight and originality can hardly be exaggerated.



Mongolin poem attributed to Genghis Khan

Mongolian

The **Mongolian** alphabet has separate letters for consonants and vowels but has its roots in the Syriac script of West Asia. It is written in vertical columns running from top to bottom that span the page from left to right. The letters have a number of different shapes, the choice of which depends on the position of a letter in a word and which letter follows it. The Mongolian language has also been written with Cyrillic and Latin alphabets in modern times.



Manchu text along side Chinese writing on a gate, Inner Mongolia, China

Manchu

The **Manchu** alphabet was commissioned in 1599 by the Manchu leader Nurhachi (1559–1626), the founder of the Manchu state. The letters are based on the Classical Mongolian alphabet; the phonetic value of letters derives from Jurchen, an earlier Manchu script. The alphabet was modified slightly in 1632.

In 1644 the Manchu conquered China and established the Qing dynasty, which lasted until 1911. For roughly the first 150 years, Manchu was the main language of government in China and served as a lingua franca. By the mid 19th century, many of the Manchu had adopted Chinese as their first language, though they continued to produce Manchu versions of Chinese documents until the end of the dynasty and for sometime afterwards.



Road sign in Thai



Rural Military sign posts in Burmese



Traffic signage in Khmer



Guest house sign in Lao

South East Asian writing systems

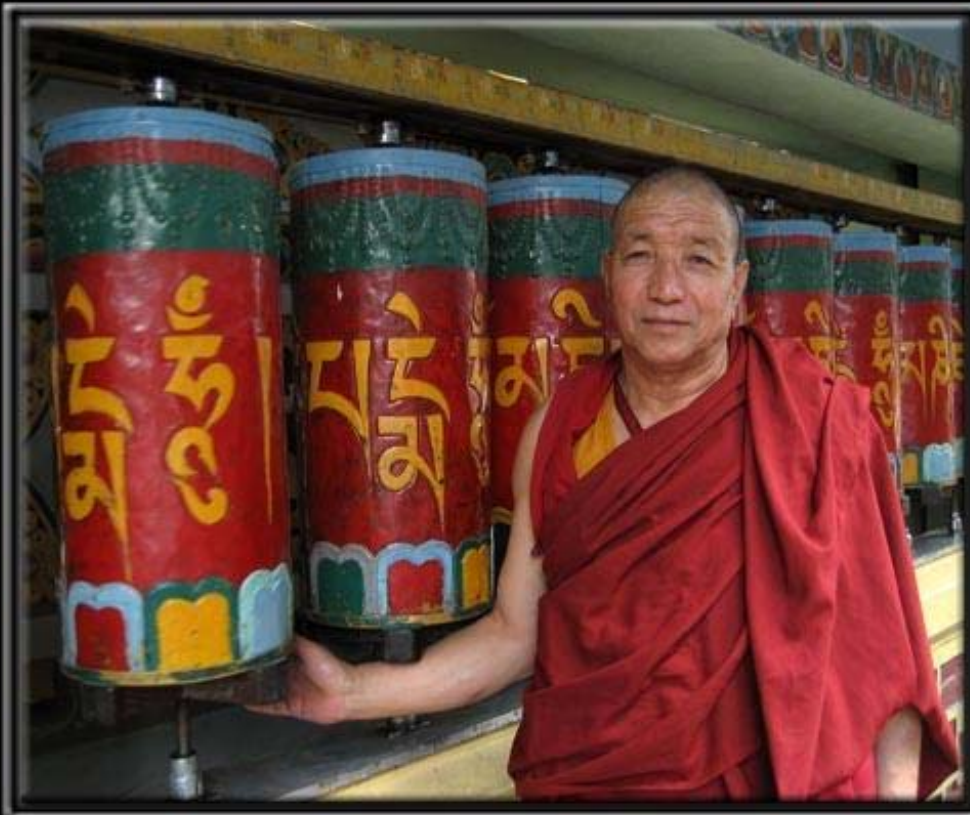
The *Burmese* language (*myanma bhasa*) is the official language of Burma. (“Myanmar” and “Burma” reflect different dialect forms of the same name.) It is the native language of the Bamar and related ethnic subgroups, and is spoken by about 32 million as a first language. The extent prehistoric relationship between the Tibeto-Burman and Sino-Tibetan language families is still a matter of some academic debate. The writing system uses a unique abugida derived from Mon script and ultimately from the Brāhmī script of India.

The *Khmer* abugida is also descended from the Brāhmī script, but by way of the Pallava script, which was used in southern India and South East Asia during the 5th and 6th centuries ce. The oldest dated inscription in Khmer, found at Angkor

Borei in Takev Province south of Phnom Penh, dates from 611 ce. The Khmer alphabet closely resembles those of Thai and Lao, which evolved from it.

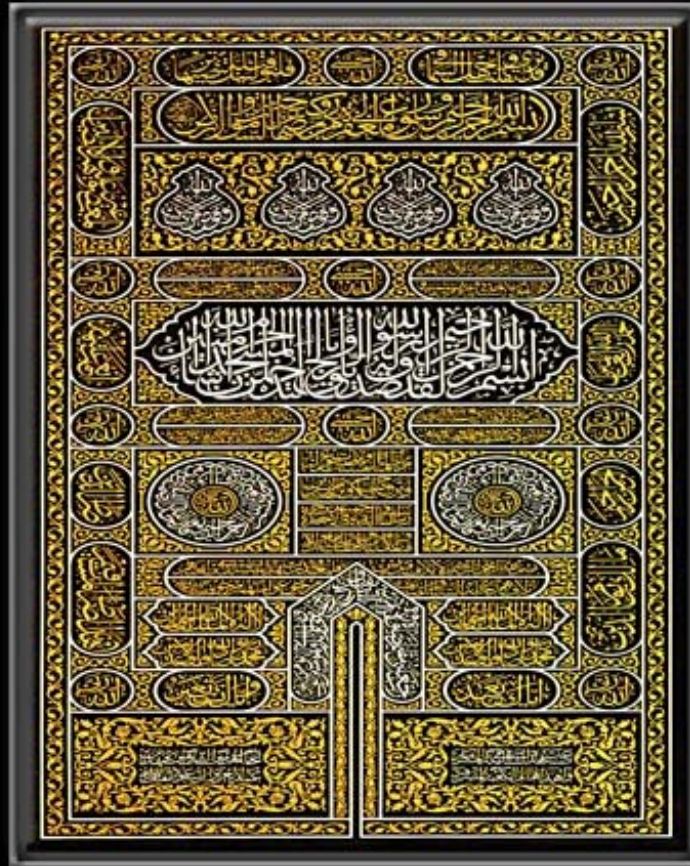
Lao or Laotian is a language of the Tai-Kadai family. It is the official language of Laos, and also spoken in the northeast of Thailand, where it is usually referred to as Isan. The Lao abugida has also been applied to Hmong, a minority language of Laos that is now more often seen in a Latin-based script.

Thai is the official language of Thailand and the native language of Thailand's dominant ethnic group. It too is a member Tai-Kadai language family, believed by some linguists to have had prehistoric links with the Austro-Asiatic, Austronesian, and/or Sino-Tibetan language families. Besides differences in letter shapes, Thai spelling is more complex than that of Lao, which underwent a reform in the 1960s.



Tibetan

Tibetan script is an abugida of Indic origin. The printed form of the script is called *uchen* (orthographically *dbu-can* 'with a head') while the hand-written cursive form used in everyday writing is called *umé* (*dbu-med* 'headless'). The 'Phags-pa alphabet derives from it, but was deployed in Chinese format (top-down vertical columns spanning the page from right to left) while Tibetan texts conform to Indian format (left-to-right lines spanning the page from top to bottom). A vast Buddhist literature has been preserved in Tibetan script; it is also used for related languages in Bhutan and in northern parts of the Indian subcontinent.



Embroidery on the door of the Kabbah, Mecca, Saudi Arabia

Arabic

The letters of **Arabic** constitute an abjad, a system for notating consonants with optional diacritics for vowels. Systems of this kind arose naturally in the writing of Semitic languages because their basic words have stable consonantal skeletons, with different vowel combinations inserted to form grammatically inflected forms. Classical Arabic is the language of the Qur`an, and the Arabic abjad followed in the wake of Islam as it spread. Since pictorial religious painting and sculpture was prohibited, a calligraphic tradition of remarkable diversity and beauty, based decorative uses of the abjad, flourished throughout the Islamic world.

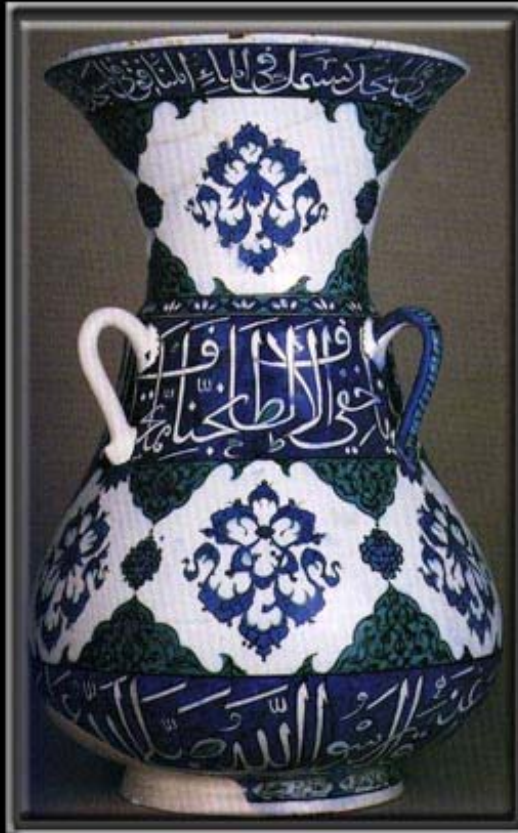
Modified forms of Arabic writing evolved for a wide variety of languages including Persian, Urdu, Pashto, Baloch, Kurdish, Punjabi, and Sindhi (Indo-European), Balti (Tibeto-Burman), Brahui (Dravidian), Ottoman Turkish, Kazakh, Azeri, and Uighur (Turkic), Malay (Austronesian), and the African languages Hausa, Mandinka, and Swahili. Many of these systems are still in use.



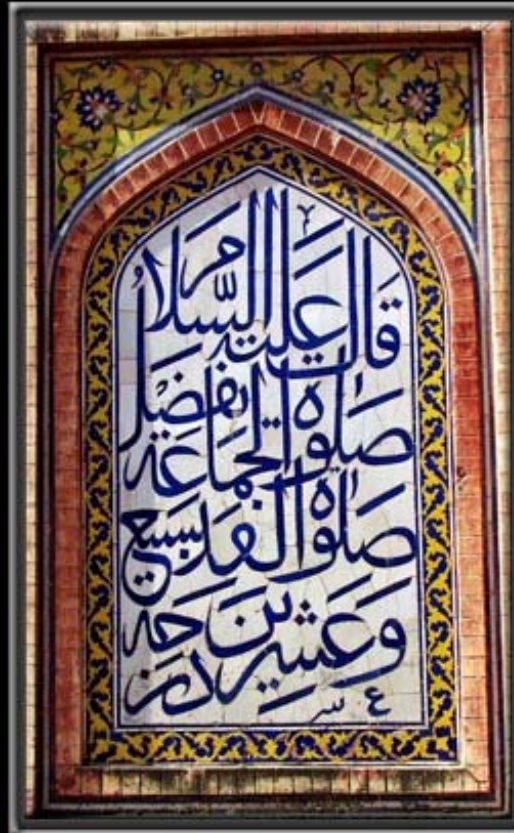
Hairdressing salon, Gaza street, Jerusalem, Israel

Hebrew

The earliest known writing in **Hebrew**, another Semitic language, dates from the 11th century bce. The language had become the liturgical and scholarly language of the Jews by the middle of the first millennium bce, with Aramaic replacing it as the everyday vernacular. It was revived and modernized by Zionists, and is now the national language of Israel. The Hebrew abjad is also used for writing Yiddish, a Germanic language augmented by numerous loanwords from Hebrew and languages of Central and Eastern Europe. Though for most scientific notation Latin and Greek letters suffice, denoting the cardinality of the real numbers as \aleph_0 *aleph-null*, with the first letter of Hebrew abjad, has become a well-established convention since the seminal work on set theory by Georg Cantor.



Iranian porcelain lusterware
with Farsi inscriptions



Urdu faience tile work, the Wazir
Khan Mosque in Lahore, Pakistan

Perso-Arabic

Writing **Persian** with the Arabic abjad following the Islamic conquest of Iran required adding four letters to represent sounds not found in Arabic, namely پ [p], چ [tʃ], ژ [ʒ], and گ [g]. Many other writing systems emerged from this Perso-Arabic script, including those for Kurdish Sorani, Lurish (Luri), Urdu, Balochi, Punjabi, Ottoman Turkish, Shahmukhi, Tatar, and Azeri. Perso-Arabic scripts typically contain many ligatures (joined letters), which complicate text formatting in computer-based environments.