Writing Systems in East and Southeast Asia

Francis Bond
based on Goddard Chapter 6

2024

Introduction to Writing Systems

- Language: Primarily spoken, but writing is a stable, visible form.
- Writing is integral to language culture.
- Study of writing systems reveals cultural history and linguistic insights.
- Two basic principles: logographic and phonographic.
- Logographic: Symbols represent words/morphemes.
- Phonographic: Symbols represent sounds (phonemes/syllables).

Types of Writing Systems

- Logographic Systems:
    - Symbols represent morphemes (e.g., Chinese, Hieroglyphics).
- Phonographic Systems:
    - Alphabetic (segmental): Symbols represent phonemes (e.g., Latin script).
    - Syllabic: Symbols represent syllables (e.g., Japanese kana).
    - Featural: Symbols represent phonetic features (e.g., Korean Hangul).
- Mixed Systems:
    - Combination of logographic and phonographic elements (e.g., Japanese).

Alphabetic Systems: Malay/Indonesian === Jawi

- Traditional Jawi script:
  - Arabic origin.
  - Right to left writing.
- Example:

    (1)   السلام عليكم *as-salām 'alaykum* "Peace be upon you."

- Jawi (جاوي) is a shortening of الجزائر الجاوي, *Al-Jaza'ir Al-Jawi* "Java Archipelago"
- *Jawi* is often used to mean 'Malay', for example in *Bahasa Jawi* "Malay Language"
- Jawi has 31 original Arabic letters;
  - 6 new letters constructed to fit phonemes native to Malay: ca (چ /t͡ʃ/), nga (ڠ /ŋ/), pa (ڤ /p/), ga (ݢ /g/), va (ۏ /v/),
  - 1 additional phoneme used in foreign loanwords: nya (ڽ /ɲ/).

- Jawi script is highly stylized for brush writing.
- Letters change form based on their position in the word.
- Example of a Jawi phrase:

(2) بالي جكرتا (Bali Jakarta) - Bali to Jakarta

- Usage has declined but still part of Malaysian education.
- Used in religious documents and decoration

ووردنيت بهاس

Wordnet Bahasa

Alphabetic Systems: Malay/Indonesian — Rumi

- Introduced by colonial powers.
  - Malay spelling followed English
  - Indonesian spelling followed Dutch
- Standardized in 1972 for both Malaysia and Indonesia.
  - *di-buat* → *dibuat*, *di-rumah* → *di rumah*
  - *ch* and *tj* /tʃ/→ *c*: *ketjap/kechap* → *kecap*
  - *njonja/nyonya* → *nyonya*
  - *ke-barat²-an* → *kebarat-baratan* "westernized"

- Derived from Old Khmer script.
  Derived from Mon, originally Indic
- No spaces between words; consonants represented by main letters, vowels by diacritics.
- Example:

  (3) สวัสดี (sawasdee) - Hello

- Complex relationship between letters and tones.

Thai Tones and Consonant Classes

- Consonant letters divided into three classes: high, middle, and low.
- Tone determined by initial consonant, tone marks, and vowel length.
- Thai words with initial /kh/ (Smalley 1994: 184)

Thai Word		Meaning	Thai Word		Meaning
ขา	khǎa	leg	คา	khaa	embedded
ข่า	khàa	galangal	ค่า	khâa	cost
ข้า	khâa	slave	ค้า	kháa	trade
ขาด	khàat	lack	คาด	khâat	to strap
ขัด	khàt	obstruct	คัด	khát	be clogged

- Complicated function of the initial consonant, diacritic and final consonant
- But generally deterministic
- Like most writing systems, sound changes quicker than writing

Thai vowels

อะ	อั	อา	อิ	อี	อึ	อื
a	a	a	i	i	ue	ue
[a]	[a]	[aː]	[i]	[iː]	[ɯ]	[ɯː]

อือ	อุ	อู	เอะ	เอ็	เอ	แอะ
ue	u	u	e	e	e	ae
[ɯː]	[u]	[uː]	[e]	[e]	[eː]	[ɛ]

แอ็	แอ	โอะ	โอ	เอาะ	อ็อ	ออ
ae	ae	o	o	o	o	o
[ɛ]	[ɛː]	[o]	[oː]	[ɔ]	[ɔ]	[ɔː]

อ็	เออะ	เออ	เอิ
o	oe	oe	oe
[ɔː]	[ɤʔ]	[ɤː]	[ɤː]

Thai dipthongs

เอียะ	เอีย	เอือะ	เอือ	อัวะ	อัว	อว
ia	ia	uea	uea	ua	ua	ua
[iaʔ]	[ia]	[ɯaʔ]	[ɯa]	[uaʔ]	[ua]	[ua]
อิว	เอ็ว	เอว	แอว	เอา	อาว	เอียว
io	eo	eo	aeo	ao	ao	iao
[iu/iw]	[eu/ew]	[eːu/eːw]	[ɛːu/ɛːw]	[au/aw]	[aːu]	[iau/iaw]
อัย	ใอ	ไอ	ไอย	อาย	อ็อย	ออย
ai	ai	ai	ai	ai	oi	oi
[ai/aj]	[ai/aj]	[ai/aj]	[ai/aj]	[aːi/aːj]	[ɔi/ɔj]	[ɔːi/ɔːj]
โอย	อุย	เอย	อวย	เอือย	อำ	ฤ
oi	ui	oei	uai	ueai	am	rue
[oːi/oːj]	[ui/uj]	[ɤːi/ɤːj]	[uai/uaj]	[ɯai/ɯaj]	[am]	[rɯ/ri]
ฤๅ	ฦ	ฦๅ	อ๋	อ์		
rue	lue	lue	m̥	silences		
[rɯː]	[lɯ]	[lɯː]	[~]	consonant		

Korean Hangul

Image from Wikipedia (2024)

- Invented by King Sejong (15th century)
  - the process of invention is well documented in the Hunmin jeongeum, or "The correct sounds for the instruction of the people"
- Represents phonemes, arranged in syllable blocks.
- Example:

    (4) 안녕하세요 *annyeonghaseyo* "hello"

- It was originally looked down on by scholars (who use hanzi) as used by women, children and the uneducated
- It was not widely adopted until the 19th century

Korean Hangul: Design Principles

- Vowels represented by distinct vertical/horizontal strokes.
- Consonant shapes reflect place of articulation.
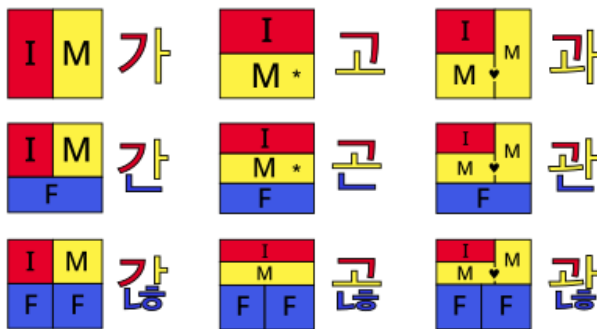- One syllable is one block

Image from Wikipedia (2024), cropped.

Korean Consonants: shape ≈ articulation

ㄱ /k/

ㄴ /n/

ㅅ /s/

ㅇ /ng/

ㅁ /m/

Images from Wikipedia (2024)

Korean Consonants: shape ≈ articulation

	Basic	Simple	Aspirated	Tense
Velar		ㄱ (k)	ㅋ (kʰ)	ㄲ (k̚)
Fricatives		ㅅ (s)		ㅆ (s̚)
Palatal		ㅈ (tɕ)	ㅊ (tɕʰ)	ㅉ (tɕ̚)
Coronal	ㄴ (n)	ㄷ (t)	ㅌ (tʰ)	ㄸ (t̚)
Bilabial	ㅁ (m)	ㅂ (p)	ㅍ (pʰ)	ㅃ (p̚)
Dorsal		(')	(ŋ)	(h)

Summary

- Hangul is a well-motivated alphabetic system, not a direct mapping of phonetic features
- Many sounds change in combination, so technically it is morphophonemic
- It is said that "A wise man can acquaint himself with them before the morning is over; even a stupid man can learn them in the space of ten days."
  Hunmin Jeongeum Haerye, postface of Jeong Inji, p. 27a, translation from Gari K. Ledyard, The Korean Language Reform of 1446, p. 258

Cia-Cia

- Cia-Cia (Bahasa Ciacia / 바하사 찌아찌아) is a member of the Celebic branch of the Malayo-Polynesian language family (Austronesian)
- Spoken by about 80,000 people on Buton Island, a part of Indonesia to the south east of Sulawesi
- It used to be written in a Jawi-like script called Gundhul and is now mainly written in Rumi
- In 2007 Professor Chun Tai-hyun, a professor of Malay and Indonesian linguistics at Hankuk University of Foreign Studies in Seoul developed a Hangeul orthography
  - For fricative /v/ it uses the obsolete jamo ㅸ
- It was privately taught since 2009 using a textbook written by Lee Ho-young, a linguistics professor at Seoul National University.
- It was abandoned around 2012, and revived around 2022
- A Cia-Cia dictionary in Hangul was published in December 2021.
- Hangul remains in use in schools and on local signs

Chinese Writing System

- Ancient origins, evolved from pictograms.
- Continually in use for over 3,000 years
- A powerful force for cultural and political unification
- Also adopted by surrounding languages (Japanese, Korean, Vietnamese)
- Characters mainly represent morphemes, often compound structures.
- Example:

  (5) 你好 *nǐ hǎo* "hello 'you good' "

  (6) 山 *shān* "mountain"

- Reformed to simplified characters in the 1950s

Development of Chinese character horse 马

Type of Characters		Descriptions
金文		Bronze script Jin wen 15th - 11th centuries B.C.E.
甲骨文		Oracle-bone script Jia gu wen 12th - 11th centuries B.C.E.
大篆		Large-seal script da chuan c. 8th century B.C.E.
小篆		Small-seal script xiao chuan 2nd century B.C.E.
隶书	馬	Clerical script li shu 2nd century C.E.
楷书	馬	Standard script k'ai shu since c. 4th century C.E.
行书	馬	Running script Xing shu since c. 4th century C.E.
草书	马	Cursive script Chao shu since c. 4th century C.E.

Logograph writing system example: Chinese

Pictographs

女	子	口	日	月	山	川	豕	目	心	雨	田	木	龜
woman	child	mouth	sun	moon	mountain	river	pig	eye	heart	rain	field	tree	turtle

Ideographs

一	二	三	上	下	中	力	凸	凹
one	two	three	above	below	middle	stength (plough)	convex	concave

Compound Pictographs / Compound Ideographs

好	安	明	家	思	牢	雷	男
good (woman + child)	peaceful (woman under a roof)	bright (sun + moon)	home/family (pig under a roof)	thought (heart + field)	prison (cow under a roof)	thunder (rain cloud over a field)	man/male (field + strength)

Semantic-phonetic compounds

Semantic-phonetic compounds

phonetic component

	古 gŭ	扁 biăn	敖 aó	旁 páng	堯 yáo
人 (person)	估 gŭ (to guess)	偏 piān (biased)	傲 ào (proud)	傍 bāng (beside)	僥 jiăo (lucky)
言 (words)	詁 gŭ (commentaries)	論 piàn (to quibble)	謷 aó (to slander)	謗 bàng (to libel)	譊 náo (to argue)
虫 (insect)	蛄 gŭ (mole cricket)	蝙 biăn (bat)	螯 aó ([crab's] nippers)	螃 páng (crab)	蟯 ráo (worm)
金 (metal)	鈷 gŭ (cobalt)		鰲 aó (griddle)	鎊 bàng (pound sterling)	鐃 náo (cymbals)

(semantic component (radical))

97% of Chinese characters are phonetic compounds (Sproat 2010)! Other estimates are lower (Wang, pc 2013)

Structure of Chinese Characters

- Phonetic component gives clues to pronunciation.
- Semantic radical provides meaning hints.
- Characters require memorization due to irregular phonetic-semantic correspondences.

Using Chinese Characters for Non-Mandarin Chinese

- Chinese characters were designed for Mandarin, which poses challenges for other Sinitic languages.
- Problem: Phonetic clues in Mandarin characters often don't apply to languages like Cantonese and Taiwanese.
- For literacy in Cantonese or Taiwanese, familiarity with Mandarin's written form is often assumed.
- Despite differences, some words in Cantonese and Taiwanese can be represented with Mandarin characters.
- Issue: Many words and grammatical elements in other Sinitic languages lack standard Mandarin equivalents.

Writing Strategies for Cantonese

- Cantonese writers use various strategies to adapt characters:
  1. **Unique Cantonese Characters**: Characters created for words without Mandarin equivalents.
  2. **Phonetic Resemblance**: Using Mandarin characters that sound similar to Cantonese words.
  3. **English Letters**: Borrowing letters to represent Cantonese sounds.
- Example of phonetic adaptation: The Cantonese plural suffix -**deih** is represented by combining 'mouth' and 'earth' characters, resembling the pronunciation **deih**: 哋

Workarounds

(7) 你　喺　　嗰喥好　喇,　千　　祈　咪　　搞
　　　你　o係　果度好　la,　千　　祈　咪　　搞
　　　you being there good SFP, thousand pray don't mess with
　　　佢　　　啲嘢　　　　。
　　　佢　　　D 野　　　　。
　　　he/she of things/stuff .

You'd better stay there, and under no circumstances mess with his/her stuff.

- use 'o' instead of 口
- use alphabet: 'D'
- use homophones: 果度 for 嗰喥
- because Cantonese is mainly spoken, there is considerable variation in the orthography

Loan words

Because of the influence of English, Cantonese tends to use more loan words (as does Singapore Mandarin). Note that speakers also freely use Mandarin vocabulary, …

Cantonese	Jyupting	English	Mandarin	Pinyin
BB	bi4 bi1	baby	嬰兒	ying1 er2
巴士	baa1 si2	bus	公交车	gong1 jiao1 che1
拜拜	baai1 baai3	bye bye	再见	zai4 jian4
沙展	saa1 zin2	sergeant	排长	pai2 zhang3

Japanese Writing System

- Combines kanji (logographic) and kana (syllabic).
- Example:

  (8)　こんにちは konnichiwa (hello)
  　　山　　　　　　yama　　　(mountain)

Syllabaries: Hiragana and Katakana

- Hiragana: Used for particles, auxiliary verbs, and suffixes.
- Katakana: Used for loanwords, mimetics, and emphasis.
- Example syllabary table for Hiragana and Katakana.

Hiragana - a true syllabic alphabet

平仮名（ひらがな） hiragana

a	あ	安	i	い	以	u	う	宇	e	え	衣	o	お	於
ka	か	加	ki	き	幾	ku	く	久	ke	け	計	ko	こ	己
sa	さ	左	shi	し	之	su	す	寸	se	せ	世	so	そ	曽
ta	た	太	chi	ち	知	tsu	つ	川	te	て	天	to	と	止
na	な	奈	ni	に	仁	nu	ぬ	奴	ne	ね	祢	no	の	乃
ha	は	波	hi	ひ	比	fu	ふ	不	he	へ	部	ho	ほ	保
ma	ま	末	mi	み	美	mu	む	武	me	め	女	mo	も	毛
ya	や	也				yu	ゆ	由				yo	よ	与
ra	ら	良	ri	り	利	ru	る	留	re	れ	礼	ro	ろ	呂
wa	わ	和	wi	ゐ	為				we	ゑ	恵	wo	を	遠
												n	ん	无

Katkana - a true syllabic alphabet

片仮名（カタカナ） katakana

a	ア	阿	i	イ	伊	u	ウ	宇	e	エ	江	o	オ	於
ka	カ	加	ki	キ	幾	ku	ク	久	ke	ケ	介	ko	コ	己
sa	サ	散	shi	シ	之	su	ス	須	se	セ	世	so	ソ	曽
ta	タ	多	chi	チ	千	tsu	ツ	川	te	テ	天	to	ト	止
na	ナ	奈	ni	ニ	二	nu	ヌ	奴	ne	ネ	祢	no	ノ	乃
ha	ハ	八	hi	ヒ	比	fu	フ	不	he	ヘ	部	ho	ホ	保
ma	マ	万	mi	ミ	三	mu	ム	牟	me	メ	女	mo	モ	毛
ya	ヤ	也				yu	ユ	由				yo	ヨ	與
ra	ラ	良	ri	リ	利	ru	ル	流	re	レ	礼	ro	ロ	呂
wa	ワ	和	wi	ヰ	井				we	ヱ	恵	wo	ヲ	乎
												n	ン	无

Kanji and Mixed System Writing

- Lexical roots written in kanji.
- Grammatical elements in hiragana.
- Foreign words in katakana (or romaji).
- Mixed system allows nuanced text creation.

(9) 彼女　　は　　T シャーツ を　　着.て いる
　　　kanojo ha　　T sha-tu　　wo　ki te iru
　　　she　　　TOP T shirt　　　ACC wear PRG

　　　She is wearing a T-shirt

Homophones and Context Dependence

- Japanese has many homophones due to the limited set of sounds in the language.
- Kanji helps distinguish words with the same pronunciation but different meanings, yet:
    - Words with identical readings may use different kanji.
    - Context often determines which kanji is appropriate, adding complexity.
- Example: こうしょう (kōshō) can mean "negotiation" (交渉) or "factory" (鉱床), depending on kanji and context. (WWWJDIC has 44 possible entries)

Why Kanji Have Multiple Readings

- Kanji have multiple readings due to their integration from Chinese into Japanese over centuries.
- Japanese and Chinese are linguistically different, so adaptations were necessary.
- Kanji readings can be divided into two main types:
  - **On'yomi** (音読み): Chinese-derived readings.
  - **Kun'yomi** (訓読み): Native Japanese readings.
- This dual system has led to multiple ways of reading kanji, depending on context.

Historical Borrowing from Chinese

- Kanji were imported from China starting in the 5th century, with Chinese pronunciations.
- **On'yomi** (音読み) readings were derived from Chinese pronunciations.
- Different periods and regions of Chinese influence led to varied on'yomi.
- For example, 山 has the reading さん ("san") in compounds, like in 火山 ("kazan," meaning "volcano").
- The Chinese-derived readings are often used in compound words.

Native Japanese Words (Kun'yomi)

- Many kanji represented concepts that already had native Japanese words.
- To adapt, kanji were assigned **kun'yomi** (訓読み), or native Japanese readings.
- Example: 山 can also be read as やま ("yama"), the native Japanese word for "mountain."
- Kun'yomi is often used when kanji stand alone or in names.
- This creates another layer of readings that are essential for native vocabulary.

Multiple Chinese Influences: Go-on, Kan-on, Tō-on

- Kanji were borrowed in waves, reflecting different Chinese dynasties and regions:
  - **Go-on** (呉音): Earlier readings from the Wu region.
  - **Kan-on** (漢音): Pronunciations from the Tang Dynasty.
  - **Tō-on** (唐音): Later readings from the Song and Ming Dynasties.
- Some kanji have multiple on'yomi due to these varied influences, used in different contexts.
- Example: 明 has readings like めい ("mei") and みょう ("myō") based on context.

Character Complexity and Memory Load

- Japanese kanji characters are often complex and require memorization of many details:
  - A standard-educated adult in Japan knows about 2,000 kanji.
  - Each kanji has multiple readings (pronunciations), which change based on context and word combinations.
- Learning kanji is time-intensive, requiring years of study.
- Even native speakers may struggle with less common kanji or readings, leading to a heavy memory load.

Historical Reforms and Modern Challenges

- Orthographic reforms have been implemented to simplify kanji, but challenges remain:
  - The **Tōyō Kanji** reform (1946) and **Jōyō Kanji** list (1981) aimed to standardize kanji usage.
  - However, non-standard or rare kanji still appear in literature and names, requiring additional study.
- Modern technology poses new issues:
  - Typing kanji requires selecting the correct character from multiple homophones.
  - Many younger people may rely on digital aids, impacting kanji writing skills.
- These factors continue to make Japanese orthography a challenging system to master.

Calligraphy in East Asia

- Highly valued art form, more esteemed than painting.
- Reflects cultural aesthetics and individuality.
- Styles include Seal Script, Clerical Script, Standard Script, Running Script, and Grass Script.
- Chinese calligraphy emphasizes brush techniques and stroke order.
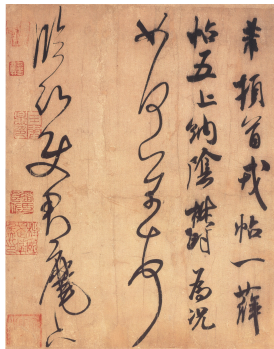
Mi Fu (1051–1107)   Emperor Huizong of Song (1082–1135)
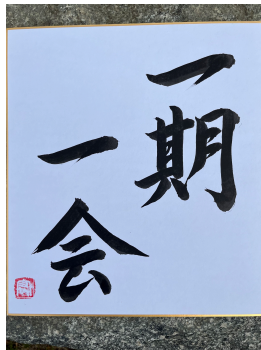
Japanese Calligraphy

- Emphasis on the beauty of handwriting.
- Grass style and Zen calligraphy notable for their fluidity and abstraction.

別無工夫
betsu-naku kuufuu
nothing else but focus

一期一会
ichi-go ichi-e
one lifetime one chance

Indonesian Calligraphy

- Jawi calligraphy is based on Arabic calligraphy
- Often very ornate

Acknowledgements

- Images taken from Omniglot (`https://omniglot.com/`) unless otherwise stated