The development of writing and preliterate societies

Author: Cynthia Tung

The development of writing and preliterate societies

Cynthia Tung

Advisor: MJ Connolly
Program in Linguistics
Slavic and Eastern Languages Department
Boston College
April 2015

Abstract

This paper explores the question of script choice for a preliterate society deciding to write their language down for the first time through an exposition on types of writing systems and a brief history of a few writing systems throughout the world. Societies sometimes invented new scripts, sometimes adapted existing ones, and other times used a combination of both these techniques.

Based on the covered scripts ranging from Mesopotamia to Asia to Europe to the Americas, I identify factors that influence the script decision including neighboring scripts, access to technology, and the circumstances of their introduction to writing. Much of the world uses the Roman alphabet and I present the argument that almost all preliterate societies beginning to write will choose to use a version of the Roman alphabet. However, the alphabet does not fit all languages equally well, and the paper closes out with an investigation into some of these inadequacies and how languages might resolve these issues.

Contents

List of Figures

Chapter 1

Introduction

How does a preliterate society choose its writing system? What script to use? Who decides? What factors affect the decision? What constitutes a good or bad writing system? Throughout the world, people have the ability to speak more than one language but most write in only one script. If "one alphabet is as much as most people can be expected to learn in a lifetime" (Lewis, *Turkish*, 34), every script, whether good or bad, has a lasting impact. The questions above pinpoint elements of a practical writing system and help us understand historical precedents. These precedents provide valuable insight to our hypothetical script choice of modern preliterate societies as no matter the time period, societies utilized spoken language prior to writing.

Chapter 2

What is writing?

2.1 Definition

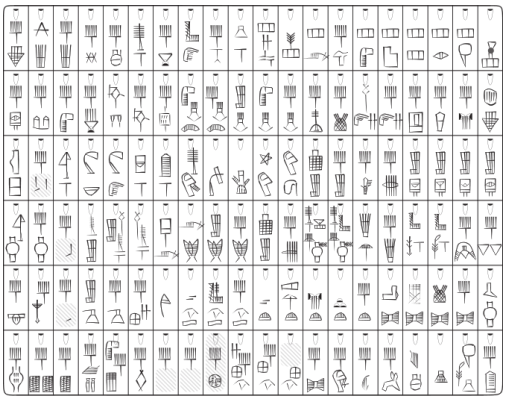
Writing relates to spoken language yet clearly differs from speech, so how do we define it? Writing represents spoken language through graphic signs and symbols (Dobrovolsky, Writing, 545), often transcribes speech (Smalley, Writing, 3), extends language over space and time (Jensen, *Signs*, 17), and outlasts the spoken word (Gnanadesikan, *Revolution*, 3). People may implement their knowledge of language through written symbols (Cook, Writing, 2) but writing does not encompass all of language. Instead, languages might have writing systems (Cook, Writing, 2) and once written down, that particular version of the language becomes fossilized and spreads beyond the original sphere of writing (Gnanadesikan, *Revolution*, 89). Typically, people write more deliberately than they speak (Gnanadesikan, *Revolution*, 5), but with modern technology such as instant messaging and texting, people often exhibit less forethought with writing than they did with a typewriter or pen and paper. Daniels summarizes writing as "a system of more or less permanent marks used to represent an utterance in such a way that it can be recovered more or less exactly without the intervention of the utterer" (Daniels, Writing, 3).

2.2 Types of writing

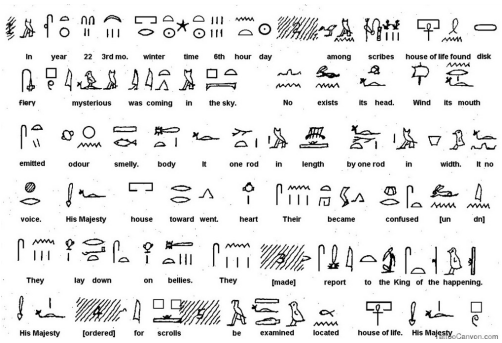
Writing has the difficult task of coding enough information without being overly cumbersome, and scripts meet this challenge differently. Scholars typically separate scripts into categories based on the relationship between the writing system and the corresponding spoken language. From an overarching, global perspective, writing has followed the law of least resistance, transforming from complicated to simple and growing in abstract qualities (Jensen, *Signs*, 22).

2.2.1 Pictograms, logograms, and ideograms

Historically, ancient civilizations began writing with pictographs which encode meaning through images and pictures. Pictograms, logograms, and ideograms represent one kind of writing through varying levels of abstraction. Pictograms have the most identifiable images and ideograms the most abstract. Historically, many ancient civilizations including Sumeria and Egypt utilized this type of script, but presently, only Chinese remains a predominately logographic script. This type of writing requires many hundreds to thousands of symbols, making them heftier to learn than other types of scripts. However, two advantages of logographic systems include efficiency with space and flexibility with variable pronunciations (more in 'Sound Changes' 2.6 on page 13) (Gnanadesikan, *Revolution*, 10).



(a) Proto-cuneiform



(b) Egyptian hieroglyphics

Oracle Bone Script	Seal Script	Clerical Script	Semi-Cursive Script	Cursive Script	Regular Script (Traditional)	Meaning
					日	Sun
					月	Moon
					山	Mountain
					水	Water
					雨	Rain
					木	Wood
					禾	Rice Plant
					人	Human

(c) Evolution of Chinese script throughout time

Figure 2.1: Pictograms, ideograms, and logograms

Image Sources: Fig 2.1a. Englund, Robert K. 1998. Wikipedia commons licensed under CC BY-SA 3.0.

http://commons.wikimedia.org/wiki/File:Proto-cuneiform_Vessels_list.svg.

Fig 2.1b. Tattoo canyon. http://www.tattoocanyon.com/36-egyptian-tattoos/2098-foot-tattoos-jefferies-tattoo-egyptian-hieroglyphics/.

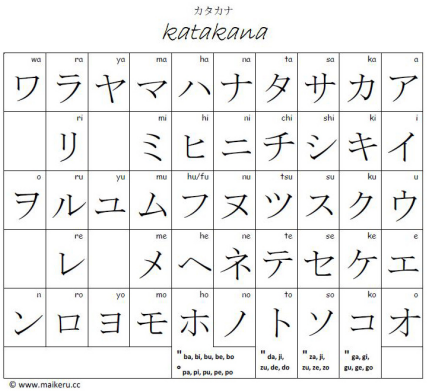
Fig 2.1c. Olivier. 2014. Marketing to China. http://marketingtochina.com/strict-censure-chineses-social-network-weibo/.

2.2.2 Syllabaries

A syllabary represents speech in syllabic units and typically each sign represents one spoken **syllable**[1]. Syllabaries best fit languages which contain primarily **open syllables**[2] (CV) but most syllabaries have systems in place for representing **codas**[3] (CVC) (Gnanadesikan, *Revolution*, 22). Modern examples include Cherokee and Japanese Hiragana and Katakana scripts.

(a) Cherokee syllabary

(b) Japanese Hiragana



(c) Japanese Katakana

Figure 2.2: Syllabaries

Fig 2.2a. First people of the American Southeast. 2002. http://www.tngenweb.org/tnfirst/syllabary.gif.

Fig 2.2b. "Hiragana Script". Maikeru. http://www.maikeru.cc/allthingsjapan_language_hiraganaandkatakanacharts.htm.

Fig 2.2c. "Katakana Script". Maikeru. http://www.maikeru.cc/allthingsjapan_language_hiraganaandkatakanacharts.htm.

[1]**Syllable:** a unit of linguistic structure that consists of a syllabic element and any segments that are associated with it (O'Grady, *Linguistics*, 637) Linguists represent the components of syllables with the letters <C>and <V>, standing for consonant and vowel respectively. Syllables consist of onsets, nuclei, and codas which represent the beginning consonant(s), middle vowel, and final consonant(s) respectively. Onsets and codas may be optional depending on the language and oftentimes, syllables do not contain all three parts.

[2]Syllables classify as either open or closed. **Open syllables** do not contain consonant(s) after the vowel (nucleus) whereas a **closed syllable** does. For example, in English, the word two [tuː] is CV and therefore an open syllable whereas ton [tʌn] is considered a closed syllable because it is CVC.

[3]**Coda**: the element that follows a nucleus (vocalic core of a syllable) in the same syllable (O'Grady, *Linguistics*, 613)

2.2.3 Alphabets

Similar to a syllabary, alphabets correlate phonemes (sounds) with graphemes (forms). Unlike the syllabary, alphabets symbolize individual consonants and vowels, rather than syllabic units. No alphabet fully represents every sound in the language; often one syllable represents **allophones**[4] of a phoneme. In general, people find alphabets the quickest to learn initially because alphabets utilize fewer symbols than any other system. Consider for example, modern alphabets such as English and Cyrillic.

(a) English (b) Cyrillic

Figure 2.3: Alphabets

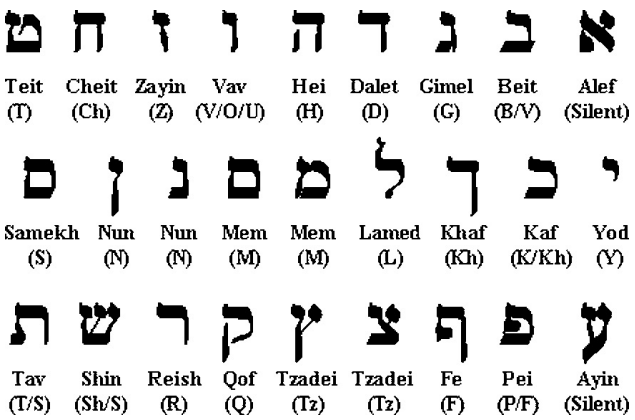
Fig 2.3a. Symbols. http://www.symbols.com/symbol/237.

Fig 2.3b. 1998. WETA. http://www.pbs.org/weta/faceofrussia/reference/cyrillic.html.

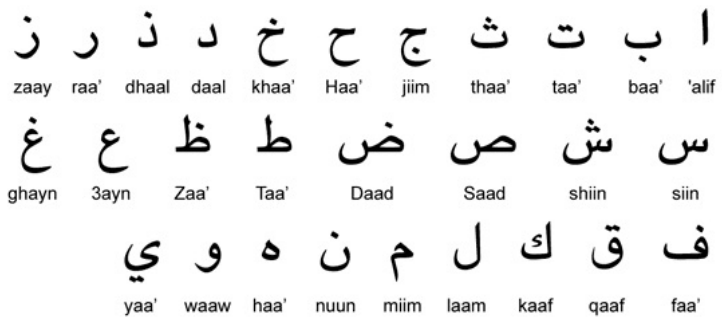
[4]A **phoneme**, defined as a distinctive sounds in a language that contrast with other such that one never appears in the same phonetic context as the other, contain **allophones**, phonetically similar variants of a phoneme.

2.2.4 Abjads

While most scripts classify as pictographic, syllabic, or alphabetic, two additional categories differentiate scripts with characteristics of both syllabaries and alphabets. Abjads only codify the consonants in a language. Languages that use abjad writing systems typically do so because the consonants encompass the meaning of the words and speakers automatically fill in vowels based on patterns of noun declensions, verb conjugations, and other grammatical features. Semitic languages typically use this type of writing including contemporary languages such as Hebrew and Arabic.



(a) Hebrew



(b) Arabic

Figure 2.4: Abjads

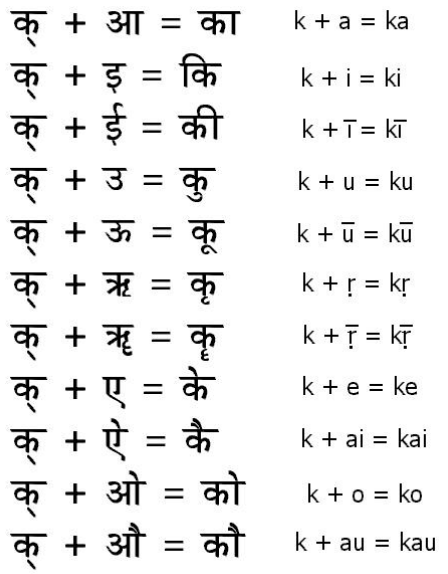
Fig 2.4a. Rich, Tracey R. "The Hebrew Alphabet" 2011. http://www.jewfaq.org/alephbet.htm.

Fig 2.4b. My easy Arabic. http://www.myeasyarabic.com/site/what_is_arabic_alphabet.html.

2.2.5 Abugidas, akṣaras, and alphasyllabaries

The terms abugida, akṣara, and alphasyllabary all describe the same type of script in which the vowel depends on the preceding consonant (Gnanadesikan, *Revolution*, 10). The term **abugida** sounds like the first four consonants of Ethiopic (Daniels, Writing, 4) and akṣara comes from Sanskrit. We consider Sanskrit an alphasyllabary because while each syllable sequence groups together to resemble a syllabary, we can identify the vowel appendages on the consonants, similar to an alphabet (Gnanadesikan, *Revolution*, 174).

(a) Devanagari, Sanskrit orthography



(b) Sanskrit consonant with vowels

(c) Ethiopic script

Figure 2.5: Abugidas, akṣaras, alphasyllabaries

Fig 2.5a. Ancient scripts. http://www.ancientscripts.com/devanagari.html.

Fig 2.5b Figure 3.3. kksongs. http://kksongs.org/language/sanskrit/lesson03.html

Fig 2.5c. Ethiopic script. http://www.tevoortwis.net/reizen/ethio05map/Eth05web/eth053/Alfabet/ethiopic_script.htm.

2.3 Shallow and deep orthographies

Scripts that relate writing to the sounds of the language fall under a spectrum of shallow to deep orthographies. Languages with a direct relationship between sounds and written symbols constitute a **shallow orthography** (e.g. Korean) whereas **deep orthographies** do not employ a direct relationship between sound and symbol (e.g. English). Deep orthographies often exhibit irregular spellings and pronunciations due to changes in the spoken language not paralleled in the written language.

In Korean (shallow orthography), each written symbol stands for a single sound. ㅂ is always [b], ㅁ is always [m] and ㄱ is always [g, k][5]. English (deep orthography) on the other hand, does not employ consistent correspondences. The single letter <s> can sound like [z] as in *dogs*, [s] as in *senior*, and [ʃ] as in *sugar* while the [s] sound can be represented by <s> as in *sound*, <c> as in *cent*, or even represented by both as in *scent*.

Children and second language learners typically find deep orthographies harder to learn initially than shallow orthographies. Anyone who has learned the forms can read Korean writing aloud, even when faced with an unfamiliar word and in fact, one could read entire books aloud without knowing the meaning of a single word. This does not impact long term literacy as once the reader of any language becomes proficient, the depth of the orthography does not impact reading because fluent readers typically do not read one symbol at a time but learn to recognize entire words (Eerie, Theory, 167).

2.4 Different classifications

Linguists (Daniels, Writing, 9) have proposed other categorizations, for example, Viktor Istrin separates alphabetic writing into three categories: phonetic[6], phonlogico-morphological[7], and historico-traditional[8]. Other scholars, Voegelin and Voegelin differentiate "self-sufficient alphabets"[9] as opposed to "alphabet included logographic system" [10] and "alphabet excluded mnemonic systems"[11]. However, scholars generally use the classifications of pictograph, syllabary, alphabet, abjad, or abugida to characterize languages and systems such as Istrin's or Voegelin and Voegelin's only have marginal use.

[5] ㄱ represents both [g] and [k], allophones of the same phoneme, with varying voicing depending on the environment of the phoneme.

[6] **Phonetic:** schemes of phonetic transcription

[7] **Phonologico-morphological:** also known as morphophonemic, unique orthography of phonemes and morphemes, pronunciation may vary in different grammatical forms

[8] **Historico-traditional:** orthography of words correspond with former pronunciation

[9] **Self-sufficient alphabet:** alphabet includes all phonographic scripts including syllabaries, abjads, alphabets, and abugidas

[10] Examples of **alphabet included logographic systems:** Chinese, Egyptian, Maya

[11] Examples of **alphabet excluded mnemonic systems:** pictographic writing, quipus, sign languages

2.5 Form and meaning

Arbitrary sign forms (Jensen, *Signs*, 215) mean that symbols do not inherently stand for one sound or another and no shape of letter is inherently a better choice than another (Smalley, How, 41). In the following Korean and English examples, Korean script contains dots, vertical straight lines, squares, and circles; all of which English also utilizes but for different sounds.

Korean Symbols	Korean Explanations	Shape	English Symbols	English Explanations
ᄌ [dʑ] & ᄎ [tɕʰ]	The dot or dash (depending on font) differentiates the two consonants.	dot or dash	<i>& <j>	English utilizes the dots to differentiate capital and lowercase letters.
ㅡ [ɯ] & ㅣ [i]	Most symbols contain straight lines, but these vowels consist only of one straight line.	straight lines	<l>[l], <I>[I], &<1>	Lowercase <l>, uppercase <I>, and the number <1>all share the same vertical line form.
ㅁ [m]	This consonant is always written as a rectangle.	rectangle	the letter <o>and the number <0 >	Typically English does not use rectangles, but in circumstances where roundness cannot be represented, these symbols are written as rectangles.
ㅎ [h] and ㅇ [ø] word initially and [ŋ] word final	The [h] form contains a circle and the variable ㅇ is purely a circle.	circle	<o>, <O>, <0>, <Q>	The letter <Q>is based off a circle shape and the letter <o>and the number <0>are circular.

Phonetic scripts, despite the importance of matching written forms with the spoken system, do not necessitate a one to one representation for every sound in the language, and the concept of allophones represented as one symbol still applies. Ideally, one symbol should not represent two different phonemes and two symbols should not be used for the same phoneme (Voorhoeve, Tone, 127).

2.6 Sound changes

Historically, languages evolved from pictographic to alphabetic systems, but phonetic systems do not improve all aspects of writing. Once societies establish a script, they typically continue using it because changing the system severs historical ties to written traditions (Gnanadesikan, *Revolution*, 144). Spoken language however, constantly changes and phonemic systems eventually will no longer match the sounds of the spoken language. Rigid phonetic scripts that do not adapt to sound changes will eventually fall out of use and an alphabet that fits a language perfectly at one stage will not fit as precisely later on.

Sound change affects logographic systems to a lesser degree since the written symbols represent meaning and not exact sounds. Written Chinese does not codify any one spoken dialect, but all Chinese speakers use the same written forms. Therefore Chinese speakers of Mandarin and Cantonese can pick up the same Chinese book and read it aloud in their own dialect and likewise, a film showing in China can use one set of subtitles regardless of the dialect, but not when a phonetic Chinese script is used, such as pinyin or bopomofo (more about these systems in 'Chinese' 4.1.1 on page 26).

2.7 Decipherment

What is decipherment[12] and why does decipherment matter to our question of preliterate societies and their development of writing systems? As stated in the introduction, studying the historical development of writing helps us identify factors that influence any script choice because each of these societies have undergone a codification process. Decipherment attempts to decode a script and therefore, considering important facets of decipherment can give us valuable insight to the process of encipherment as well.

2.7.1 Process

Certain circumstances lend themselves to an easier decipherment than others (Cleator, *Languages*, 24). Scholars use two primary techniques to decipher new scripts, an analytic method and a probable-word method (Cleator, *Languages*, 26). The analytic studies frequencies of certain forms and experiments with word substitutions to identify the plausibility of different combinations (Cleator, *Languages*, 25). The probable-word process relies on intelligent prediction of likely inscriptions. For example, a tomb or other burial memorial probably contains a prominent or rich person's name and some honorifics exalting them (Cleator, *Languages*, 26).

Context provides a huge clue for decipherers in decoding the script. Useful information includes land or country of origin, time period, and other historical or archeological clues (Cleator, *Languages*, 25). Other helpful components include familiar or similar languages to compare with (Daniels, Decipherment, 143) and proper names (Moorhouse, *Alphabet*, 44). Beyond context markers, decipherers require first and foremost, accurate copies of the inscription (Daniels, Decipherment, 141). Different glyphs occurring in the same context must either stand for the same sound, represent the same meaning, or contrast as members of the same category (Gnanadesikan, *Revolution*, 88). Bilingual inscriptions have benefitted scholars the most but many languages lack such a resource (Daniels, Decipherment, 144). Different languages have different combinations of resources surviving to aid in decipherment, but without any of these tools, scholars have no starting point to decipher an extinct language.

The number of signs can give valuable clues as to the type of script: fewer than 30 signs would likely constitute an alphabet and more than 1,000 signs would indicate an

[12] **Decipherment:** the process of determining the relation between some writing not hitherto understood and the language it represents (Daniels, Decipherment, 141)

ideographic system (Moorhouse, *Alphabet*, 45). Roughly 100 signs imply an abugida or syllabary (Daniels, Decipherment, 142).

2.7.2 Lessons from decipherment

The decipherment process reinforces the importance of context, both societal and linguistic, and that a written language is not required to represent everything about a language. Social and historical markers give decipherers hints as to where to begin. If archaeologists looked at Maya stelae and tried to identify the script as Chinese, they would not get very far. Beyond simply identifying the time and place of the source material, understanding that the Maya had an extensive calendar system helped in the decipherment of Maya numbers. Decipherment requires scholars to look at a source material and identify the function of writing whether it is an accounting list, a memorial set up for a king or other important figure, or prose. The type of document gives context for the decipherers and consciously considering the contexts for which writing will be used in a culture helps to best fit an orthography to the needs of the people. Likewise, a script choice ought to logically fit with the context of the preliterate society.

Scholars have deciphered parts of many languages such as Sumerian, Egyptian, Linear B, and Maya, but we still do not know everything there is to know about these languages. We cannot pronounce the words now as Sumerian cuneiform and Egyptian hieroglyphics did not write vowels, yet speakers of Sumerian and Egyptian were still able to understand the script. The lack of vowels did not detract from the functionality of the orthography, emphasizing that a good script does not need to codify everything about the spoken language.

Chapter 3

Inception of writing

3.1 Prewriting

Not all visual records constitute writing. Scholars do not consider cave paintings as writing because while they likely contained meaning, they lacked the communicative principles pictographs contained. Another example of prewriting[1], Quipus, a system of knots tied into strings used in Peru, enabled the transmission of information but only for record keeping, a limited context.

3.2 What constitutes a full writing system?

Considering these precursors to writing, what constitutes a writing system? Writing must represent the grammar of the language and it must be flexible enough to be used in a variety of contexts. Conventionalized pictograms may be used to record specific types of information but not the full language (Gnanadesikan, *Revolution*, 79), such as road signs.

As mentioned earlier, the earliest writing systems were all highly picture-based and later systems became increasingly sound-based. However, no writing utilizes solely phonemic or morphemic representation and these systems still contain some sort of representative units (Gnanadesikan, *Revolution*, 10).

3.3 Sumerian

Sumerian, one of the earliest forms of writing, began as picture writing and transformed into many subsequent writing systems (Diringer, *Writing*, 86). Proto-cuneiform, the source of Sumerian, consisted of 800 signs and 60 to 70 numbers (Michalowski, Mesopotamian, 34) but encoded no grammatical information and little phonological information (Gnanadesikan, *Revolution*, 19), hence the designation as a proto-script.

[1] **Prewriting:** visual symbols that do not contain all the aspects to qualify as writing

Some of the signs in proto-cuneiform included **determinatives**[2], which appeared as unpronounceable aids to eliminate ambiguities and supply phonetic complements (Diringer, *Writing*, 96). In addition, they used **rebus writing**, the practice of using easily pictured objects to stand for an abstract homonym (Gnanadesikan, *Revolution*, 18), a strategy many other logographic and ideographic languages also employed. By 600 BC, the number of signs in proto-cuneiform stabilized at about 600 (Gnanadesikan, *Revolution*, 19) and the number further dropped to 300 (Cooper, Sumerian, 41).

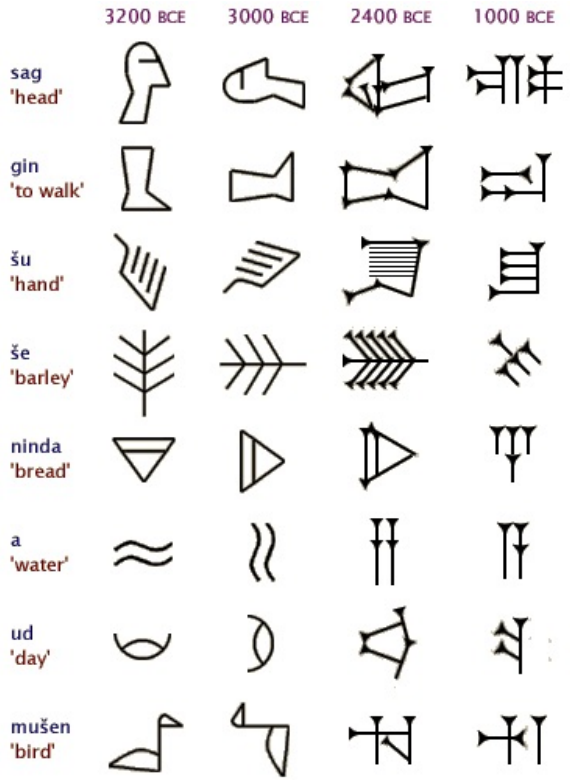


Figure 3.1: Evolution of Sumerian script over time

Fig 3.1. Ancient scripts. http://www.ancientscripts.com/sumerian.html.

Eventually, the Sumerians developed a logosyllabic cuneiform, a wedge-shaped script written left to right within boxes (Gnanadesikan, *Revolution*, 19) of logograms and syllabograms mixed together (Gnanadesikan, *Revolution*, 21). Signs began to represent syllables, forcing an expansion of the writing system as Sumerians spelled out affixes, personal names, and foreign words. By the time affixes were fully represented, Sumerian probably was not a spoken language any longer (Gnanadesikan, *Revolution*, 20). Words spelled with logograms utilized syllabograms as phonetic complements to indicate which of multiple potential pronunciations should be used (Gnanadesikan, *Revolution*, 21). Additionally, the combination of meaningful and phonetic elements of writing contributed to inconsistent spellings (Cooper, Sumerian 42).

An illustration from Gnanadesikan follows: Sumerians pronounced both *reed* and *to render* [gi] and that symbol stood for three things: *reed*, *to render*, and the sound [gi] (Gnanadesikan, *Revolution*, 23). When the Akkadians conquered the Sumerians they continued using the existing cuneiform script for administration, thereby adapting the cuneiform for themselves. When the Akkadians took the logograms, the multiple meanings also carried over. For example, the Akkadian word for *reed* was pronounced [quanuːm] and *to render* was [taːrum] and now that same symbol represented the same meanings but three sounds: [quanuːm], [taːrum], and [gi] (Gnanadesikan, *Revolution*, 21). Eventually, they expanded the syllabary and left the script primarily with syllabic signs and occasional logograms (Cooper, Sumerian, 45).

[2]**Determinative:** a symbol used in non-alphabetic writing systems to assist the reader in knowing the intended meaning or pronunciation of a word (O'Grady, *Linguistics*, 616). Examples of English determinatives include *the* and *a*.

3.4 Egyptian

Sumerian writing may have inspired Egyptian hieroglyphics as the two scripts have similar structures and underwent a series of similar developments (Jensen, *Signs*, 55). Both transformed from idea-writing which had symbolic representation but was not meant to be read aloud (Jensen, *Signs*, 56) to ideographic word picture-writing (Jensen, *Signs*, 57). Eventually, a phonetic word-script developed in Egypt (Jensen, *Signs*, 59), and the script also employed rebus writing ('Sumerian' 3.3 on 16) (Gnanadesikan, *Revolution*, 37).

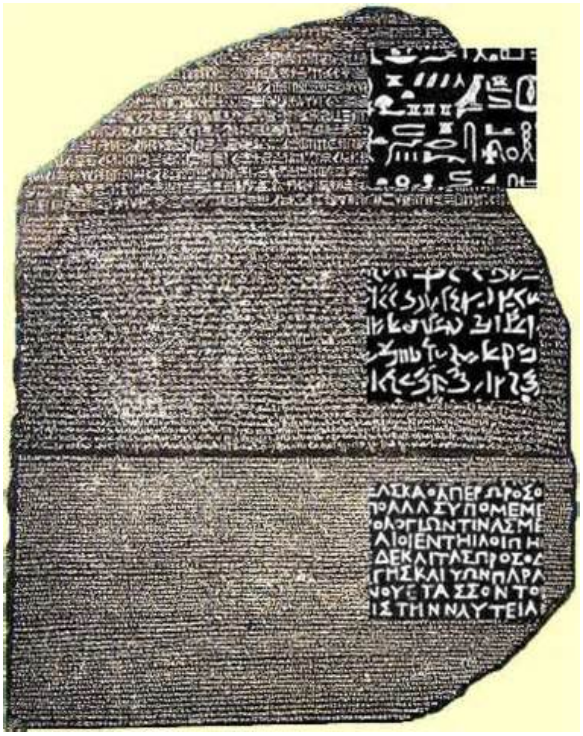


Figure 3.2: The Rosetta Stone containing hieroglyphics, demotic script, and Greek from top to bottom

Fig 3.2. Social studies for kids.

http://www.socialstudiesforkids.com/articles/worldhistory/rosettastone.htm.

Full of redundant features, hieroglyphics (Diringer, *Writing*, 48) contained alphabetic, phonetic, and ideographic writing. Phonetic components accompanied hieroglyphic ideograms, and scribes supplemented logograms with three different kinds of phonograms (Gnanadesikan, *Revolution*, 35). With all these phonograms, Egyptians had the capability to write alphabetically but their writing never became purely phonetic (Moorhouse, *Alphabet*, 75) and retained its pictorial aesthetic (Gnanadesikan, *Revolution*, 35). The script only wrote consonants, like modern day Arabic (see 2.2.4 on page 10) (Ritner, Egyptian, 73). Instead, the phonetic features created a mixed system of logograms and consonants, supplemented by an extensive series of determinatives (Gnanadesikan, *Revolution*, 41).

Egyptian script had a full range of bureaucratic, religious, and literary uses (Gnanadesikan, *Revolution*, 33) but ancient Egyptians never separated formal hieroglyphic representation from art, and the direction of writing varied with the direction people were facing in the carvings (Gnanadesikan, *Revolution*, 44). This type of script required writing too slow for bureaucratic records, literary works, or other everyday purposes, and eventually the script developed the hieratic script, a cursive version of hieroglyphics. Hieratic script developed as a simplified version of writing, always right to left (Ritner, Egyptian 81). Next came demotic script with abbreviations and ligatures, complicating its origins to the original hieroglyphics (Ritner, Egyptian, 82). Demotic remained in use until Coptic, a new script closely associated with Christianity and possibly based on Greek, overtook it in roughly 400 AD (Gnanadesikan, *Revolution*, 49).

Overall, hieroglyphics contained roughly 500 common signs throughout its history. By

Roman times, the script split into two, and only priests were still using the older hieroglyphics, the number of which they expanded to around 6,000 (Gnanadesikan, *Revolution*, 43) while demotic hieroglyphics on the other hand, became phonetic and contained an estimated thirty-two signs (Gnanadesikan, *Revolution*, 48).

In 1799, Napoleon Bonaparte and his soldiers discovered the Rosetta Stone (Gnanadesikan, *Revolution*, 51) on which King Ptolemy V had inscribed with the same message in hieroglyphics, demotic, and Greek. This discovery led scholars such as Thomas Young and Jean-François Champollion to a full decipherment of Egyptian hieroglyphics (Jensen, *Signs*, 76).

3.5 Chinese

We cannot ascertain the exact timeline of the development of Chinese script (Jensen, *Signs*, 165), but modern day Chinese descended from the Chinese script of the Shang period (Boltz, Chinese, 191). Chinese also employed **rebus principle** in the script, but unlike other pictographic or logographic scripts, Chinese combined the determinative, rebus, and actual character together into a single character with variable positions to the point where often one cannot positively identify the function of part of the character (Gnanadesikan, *Revolution*, 59).

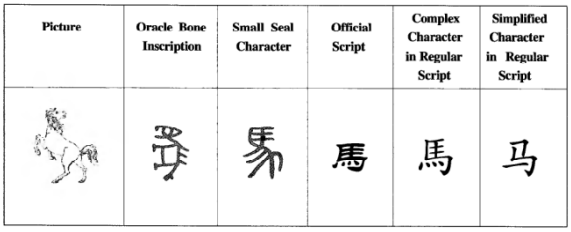


Figure 3.3: The historical evolution of Chinese script.

Fig 3.3 Asians Mind.

http://asiansmind.com/en/4-good-reasons-to-learn-chinese/

As the spoken language changed, the language did not evolve to match. The writing remained, having shifted away from phonological components of the language but the writing remained, further solidifying Chinese's logographic identity (Gnanadesikan, *Revolution*, 67). Chinese has no inflections and few affixes, creating many monosyllabic morphemes and in order to combat ambiguity, Chinese developed **compound characters** by adding a secondary graph to an ambiguous primary graph (Boltz, Chinese, 194). Despite each character pronouncing as one syllable, Chinese does not classify as a syllabary as each character still contains meaning (Gnanadesikan, *Revolution*, 61).

Scholars acknowledge 214 semantic determinatives in Chinese, called **radicals** (Boltz, Chinese, 199). Without an alphabet, Chinese does not have an alphabetical order with which to order words. Instead, the Chinese developed a dictionary utilizing radicals and stroke order. Children learn the shape and stroke of each character, which helped to standardize the written language. The first dictionary consisted of roughly 3,300 characters (Gnanadesikan, *Revolution*, 62).

As mentioned earlier, many languages once employed pictograms and logograms in their writing. However, many languages have evolved their script or adopted an alphabetic script. Yet Chinese remains primarily logographic and we will investigate modern strategies Chinese uses to stay recent in 4.1.1 'Chinese' on page 26.

3.6 Maya

Away from the rest of the world, the Maya created their own writing system in the South American Yucàtan. The symbols look complicated initially but upon closer inspection, each block contains multiple compounds of simpler signs compressed together (Jensen *Signs* 234).

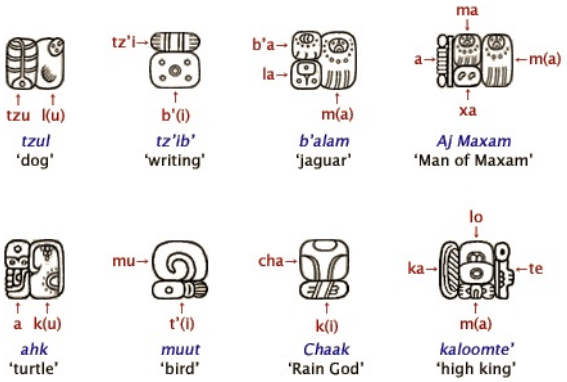


Figure 3.4: Sample Maya script

Fig 3.4 Ancientscripts.com. http://www.ancientscripts.com/maya.html.

With the death of the last remaining Maya scribes in the early 18th century, the world lost access to the written script (Gnanadesikan, *Revolution*, 84). Luckily, Diego de Landa, a Franciscan priest held a genuine interest in Maya culture and wrote *Account of the Affairs of the Yucàtan*, in which he recorded Maya glyphs with corresponding Spanish letters (Jensen, *Signs*, 236). Despite his good intentions, Landa's accounts misled scholar to believe the script was alphabetic (Gnanadesikan, *Revolution*, 85), and at the time of Jensen's publication of *Signs, Symbols, and Scripts* in 1970, scholars had yet to decipher Maya script.

Decipherment of Maya did not progress until Yuri Valentinovich Knorosov correctly guessed based on the number of signs in the script (roughly 700) that Maya should be a logosyllabary and deciphered Landa's notes as such. Around the same time, Tatiana Proskouriakoff identified numbers from dates on standing stone stelae (Macri, Mesoamerican, 175).

These and other scholars deciphered the Maya script by identifying the arrangement of a main logographic block along with smaller syllabic affixes into glyph blocks, squares with rounded corners that appeared in lines, columns, L shapes, T shapes, or a grid (Gnanadesikan, *Revolution*, 91). Each glyph block usually contained a logographic main block along with syllabic affixes (Gnanadesikan, *Revolution*, 92). Complicating the decipherment process, "Maya scribes appear to have taken special delight in using as many different spellings as possible" (Gnanadesikan, *Revolution*, 92). This intricacy also explains the low rate of literacy and the difficulty in decipherment despite the fact that Maya is currently spoken and written in the Roman alphabet (Gnanadesikan, *Revolution*, 94).

3.7 Linear B

On the island of Crete, archaeologists discovered two unknown scripts on tablets and they subsequently named the languages 'Linear A' and 'Linear B' (Chadwick, *Linear B*, 12). While scholars have not yet determined the exact relationship between the two scripts, many believe the two contain parallel structures. However, without a decipherment of Linear A, scholars cannot confirm the hypothesis (Chadwick, *Linear B*, 14). Archaeologists

discovered accounting tablets (Gnanadesikan, *Revolution*, 99) which contained Linear B words, ideograms, and numerals.

The ideograms do not appear to be mixed with the syllabic signs, indicating a visually organizationed function (Gnanadesikan, *Revolution*, 101). Scholars identified roughly 87 core symbols, not counting ideograms and numbers, indicating a somewhat syllabic structure (Gnanadesikan, *Revolution*, 102), and Michael Ventris organized these forms and predicted which symbols might share consonants or vowels (Gnanadesikan, *Revolution*, 105). The syllabary included plain vowels, a unique feature when compared to other typically CV languages (Gnanadesikan, *Revolution*, 105). Eventually, the identification of location names and proper nouns allowed Ventris to begin his decipherment with Greek.

Scholars struggled for many years to decipher Linear B until in 1952, Ventris suggested that Linear B transcribed Greek (Chadwick, *Linear B*, 62). The Greek emerging from transcriptions initially seemed badly transcribed for Greek but contained identifying features such as the Greek dual tense (Gnanadesikan, *Revolution*, 107). It wasn't until an outside excavator used the decipherment on new tablets with success that everyone began to accept Ventris' hypothesis (Chadwick, *Linear B*, 81).

Figure 3.5: From the notes of Michael Ventris

Fig 3.5. Engles, Tom. 66 South.

http://66south.com/Ventris/.

3.8 Sanskrit

The convergence of scripts in in the Indus River Valley resulted from a **linguistic area**[3] rather than stemming from a common source (Gnanadesikan, *Revolution*, 172). Historical Sanskrit manuscripts contained a variety of local scripts (Bright *Devanagari* 384) but in this section, we will focus on Devanagari, the script used to write Sanskrit to this day (Bright, Devanagari, 384).

As described earlier, in section 2.2.5 on page 11, Sanskrit classifies as an alphasyllabary, meaning that elements of both alphabets and syllabaries exist in the script. Early Sanskrit scholars organized the consonants based on five **places of articulation**[4] and five **manners of articulation**[5] (Bright, Devanagari, 384). Each consonant contains an inherent vowel [a] but other vowels attach to the consonant as mandatory diacritics in different positions (Bright, Devanagari, 387).

[3]**Linguistic area:** a region where various languages share characteristics due to mutual borrowing.

[4]**Place of articulation:** The points at which the airstream is modified in the vocal tract to produce sounds (O'Grady, *Linguistics*, 631)

[5]**Manner of articulation:** The various configurations produced by positioning the lips, tongue, velum, and glottis in different ways (O'Grady, *Linguistics*, 625)

3.9 Phoenician

Phoenician script, an old Semitic language coexisted simultaneously with Aramaic and Hebrew and became the basis for the Greek alphabet (Gnanadesikan, *Revolution*, 149). Similar to Egyptian hieroglyphics other Semitic languages, Phoenician did not contain vowels aside from **matres lectionis**[6] (Goerwitz, Scripts, 489). So how did the Greeks result in a fully vocalized script?

3.10 Greek

The Greeks borrowed their alphabet from the Phoenicians evidenced by the fact that the names for letters in Greek (*alpha, beta, gamma*, etc) do not mean anything in Greek but contain meaning in Semitic languages (Threatte, Greek, 271). People also used other scripts to write Greek, such as Linear B (see section 3.7 on Linear B, page 20) (Jensen, *Signs*, 451). The Greek alphabet became the first alphabet to record every vowel and consonant equally (Gnanadesikan, *Revolution*, 215), likely due to misunderstanding and misinterpretation of the Phoenician alphabet (Gnanadesikan, *Revolution*, 214).

The Greeks did not have one uniform script. Instead, different locales had their own regional alphabets (Jensen, *Signs*, 458) until Athens adopted the Ionian alphabet in 403 B.C., solidifying Ionian Greek as the official Greek alphabet (Jensen, *Signs*, 463). The Alexandrians were the first to record pitch-accent in Greek (Gnanadesikan, *Revolution*, 220). Diacritics for breathing and accents (Threatte, Greek, 276) helped separate words until word spacing did not become a consistent feature of the language until the 18th century (Gnanadesikan, *Revolution*, 221).

People often praise Greek for fully representing a language with so few symbols (Gnanadesikan, *Revolution*, 222) but despite the simplicity of the alphabet, normal citizens did not learn to read and they generally regarded writing as a luxury for the upper class (Gnanadesikan, *Revolution*, 217). The spelling of modern Greek illustrates a proclivity for historical spellings versus a current one-to-one matchup of phonemes and letters, something we encounter in many languages throughout the world.

3.11 Etruscan

The Etruscan alphabet derived from Greek, due to an initial fascination with the alphabet, Etruscans retained Greek letters they did not need (Gnanadesikan, *Revolution*, 231) but soon eliminated all the letters they didn't need including letters for [b], [d], [g], and the letter ‹o› (Bonfante, Italy, 297). Latin derived from Etruscan and for our purposes, Etruscan's big historical impact comes from its role as an intermediary between Greek and Latin.

[6]**Matres lectionis:** letters used to spell long vowels, initially used as a coda and later inside words as well (Gnanadesikan, *Revolution*, 151).

3.12 Latin

Latin borrowed its alphabet from the Etruscans, luckily before the Etruscans removed the <b>, <d>, <g>, and <o> from the alphabet (Gnanadesikan, *Revolution*, 229). Due to the fragmented nature of the Roman Empire, the languages and orthographies of the region also varied greatly. Yet the influence of Christianity supplanted other ancient Italian scripts with the Latin alphabet (Jensen, *Signs*, 520). Western Europe mainly spoke colloquial versions of Latin but shared a written language which worked well until the alphabet started to diverge (Gnanadesikan, *Revolution*, 236).

The alphabet did not always contain the twenty-six letters we currently consider our English alphabet. Originally, the alphabet lacked the letters <G>, <J>, <V>, <W>, <X>, <Y>, and <Z>(Gnanadesikan, *Revolution*, 229), and the alphabet only contained capital letters written sometimes left to right, other times left to right, and even boustrophedon[7] (Gnanadesikan, *Revolution*, 229). The bulk of the alphabet however, has remained the same: the modern capital letters parallel Classical Roman forms from 2000 years ago and lowercase letters resemble ninth century Carolingian Minuscule script (Knight, Roman, 33).

Even in the early stages of Latin, the alphabet benefitted greatly from new inventions of paper and movable type in conjunction with the printing press (Gnanadesikan, *Revolution*, 249), and in modern times, electronic communications have once again provided a ripe environment for the proliferation of the Roman alphabet.

3.13 Cherokee

Despite the growing preeminence of the Roman alphabet, some societies still invented new scripts for their languages in more the last two hundred years and some indigenous groups invented new orthographies. Below, we look at two instances from North America, Cherokee and Cree.

Sequoyah, an illiterate Cherokee man, did not speak English (Wilford, Carvings, 1). However, as he witnessed written English, Sequoyah wanted to be able to write his own language and began to develop a script for Cherokee, starting as a picture-script that eventually evolved into the eighty-six syllable script used for many years(Jensen, *Signs*, 241). Some of the symbols resemble English letters although they stand for very different sounds, further evidence that Sequoyah knew about English writing but could not read it. The script read left to right, contained spacing between words, and utilized the same punctuation as English.

In 1824, missionaries tested the system, found it suitable, and began implementing the alphabet in printed materials including the *Cherokee Phoenix*, a bilingual weekly paper with articles in both Cherokee and English, as well as portions of the Bible (Campbell, *Handbook*, 25). Within a couple years, Cherokees created many legal, political, religious,

[7] **Boustrophedon**: *ox-turning*, "the name suggests an ox drawing a plough across a field and turning at the end of each furrow to return in the opposite direction; this is because in boustrophedon every other line of writing reversed the direction of reading. (Campbell, *Historical*, 377)

and informational publications (Scanccurelli, Cherokee, 591), and by 1830 over half of the male population could read (Jensen, *Signs*, 242), greatly benefitting both Cherokee medicine and the Christian church (Scanccurelli, Cherokee, 591).

Similar to some ancient scripts, the Cherokee syllabary did not have spelling rules but represented the sounds enough that a speaker of Cherokee would have no trouble, making the language better suited for L1[8] speakers of Cherokee (Gnanadesikan, *Revolution*, 137). This challenged the continued growth of the syllabary as less people spoke Cherokee as their L1, and the introduction of the Roman alphabet for Cherokee in 1902 ended the use of the syllabary (Jensen, *Signs*, 243).

3.14 Cree

Within ten years, missionary James Evans created the Cree script which accommodates the different regional dialects of Cree with only eleven distinct signs oriented in four different directions (Jensen *Signs* 244). Each geometric symbol stands for one consonant and one of four primary vowels: front vowels invert over the vertical axis and back vowels reflect over the horizontal one (Nichols, Cree, 602). Cree represents codas of the syllables as diacritic marks added to the initial symbol and these finals may also be used as independent consonants (Jensen, *Signs*, 244).

The script gained acceptance quickly despite its initial use by missionaries only to print Christian sacred texts (Nichols, Cree, 599). In the 1970's, aboriginal language use was steadily declining until the Cree Way Project (1973-76) developed Cree classroom and teaching materials for schools (Burnaby, Cree, 197). This process included a standardization of the language, training of teachers and other personnel, and creation of general reading materials (Burnaby, Cree, 202). In 1993 many Cree communities began to establish Cree as a classroom language, revitalizing the script which the Cree still use to this day (Burnaby, Cree, 191).

[8] **L1:** mother language, native tongue, first language

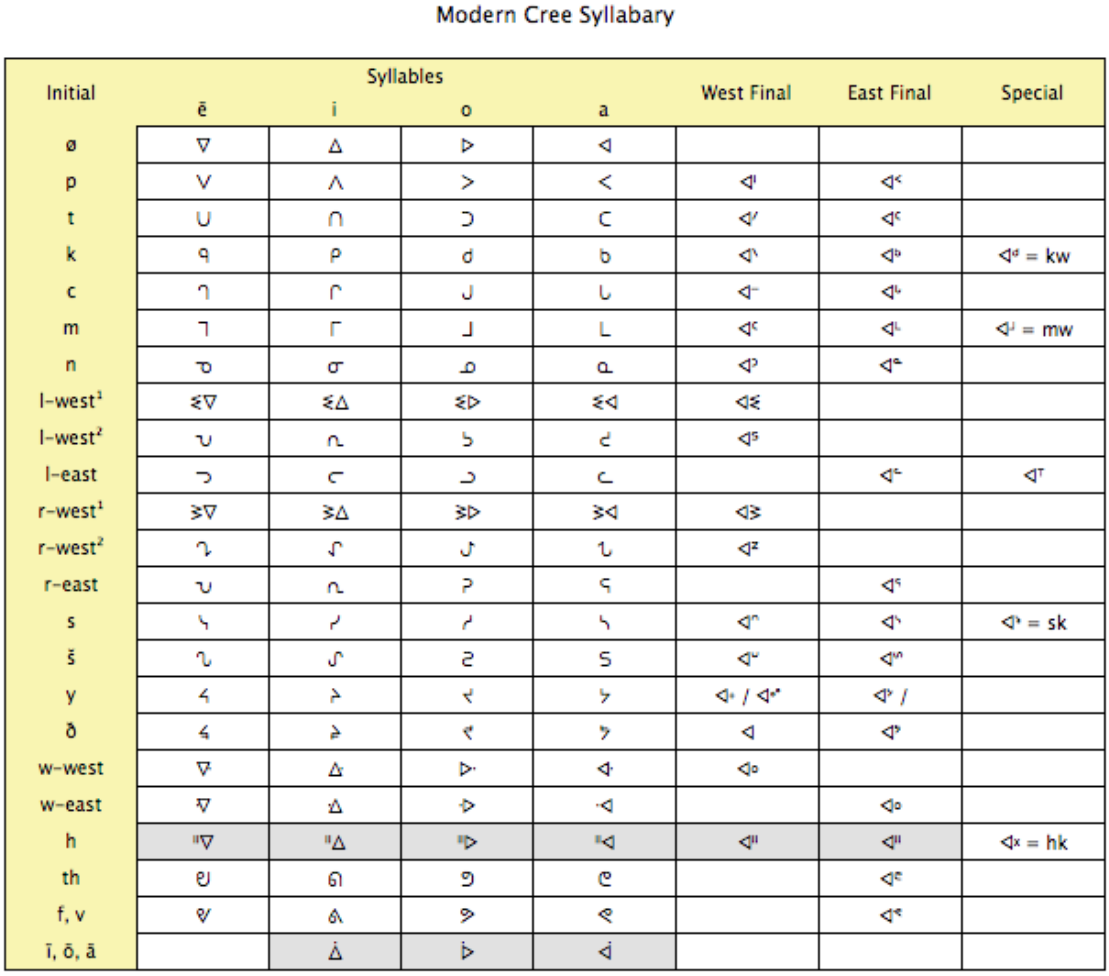


Figure 3.6: Cree syllabics chart

Fig 3.6. 2008. Langauge Geek. http://www.languagegeek.com/algon/cree/cre_syllabarium.html. Accessed 2015-02-20 at 07:51GMT.

Chapter 4

Expansion of writing

4.1 From one to many

Throughout its long history, the Chinese script spread well beyond just the many dialects of Chinese and became the script for other spoken languages such as Korean, Japanese, and Vietnamese. Chinese speakers found the script challenging and developed simplification of logograms and supplementary phonetic systems have developed a simplification of the logograms and supplementary phonetic scripts. Likewise using the Chinese script hindered literacy in the countries which adopted the script, and Korean and Japanese speakers also invented phonological scripts but retained the original Chinese in some contexts. Vietnamese transitioned into the Roman alphabet, and **chũ' qu'ôcngũ'**, the romanized Vietnamese script, has been in use to this day (more in section 4.2.2 Vietnamese on page 31).

4.1.1 Chinese

As we have discussed in section 3.5 on page 19, Chinese remains a logographic writing system. The orthography contains roughly 7,000 characters but only 1,300 spoken syllables, and therefore, a written character represents a specific morpheme but a spoken syllable often represents multiple morphemes (Chung, Morphological, 442). The abundance of characters means learning Chinese takes several years and requires regular practice to retain command of the language (Dreze, Literacy, 2870).

Dissatisfied with the low literacy rates, China's State Language Commission encouraged the use of simplified Chinese characters (Chen, Effects, 430) and implemented reforms to education throughout the country, improving the literacy rate dramatically (Dreze, Literacy, 2871), but only mainland China uses the simplified system while Japan, Korea, Hong Kong and Taiwan all still use traditional characters. Although initially the commission did not announce an official set of simplified character they supplied a list of characters to avoid oversimplification and confusion as people began simplifying characters haphazardly (see figure 4.1)(Wu, China, 1). Due to the simultaneous development of the two movements, we cannot ascribe the literacy growth to either the simplification of Chinese characters or the education reforms, but both together.

Figure 4.2: A set of Chinese characters with different meanings but share a syllabic sound with varying tones

Fig 4.2. How OCR Works. http://www.how-ocr-works.com/languages/chinese-alphabet.html.

Figure 4.1: An example of the same phrase in traditional Chinese and simplified Chinese

Fig 4.1. Hirasuna, Delphine. 2009. At Issue.

http://www.atissuejournal.com/2009/05/28/typography-in-china/.

Some characters contain phonetic portions to aid pronunciation, but these do not always help as the characters used may be complex characters themselves. These supplementary characters also contain individual pronunciations as independent syllables and while it may direct the pronunciation of a new character, it often cannot pinpoint the exact sound (Cheung, Pinyin, 229). Chinese speakers have addressed the difficulty of teaching logographic characters by implementing phonetic systems to go alongside the characters. As described earlier, completely removing Chinese characters and using these phonetic systems would create additional ambiguities, even considering the different tones as different phonemes (see figure 4.2). This resists a complete shift into only phonetic systems, and therefore despite the prevalence of these phonetic systems, Chinese script continues to maintain a stronghold over written Chinese.

Part of the reason Chinese contains so many homonyms relates back to the available sounds in the language. Most syllables are open syllables, meaning they do not typically have final consonants, limiting the number of syllable combinations available. In addition, the monosyllabic Chinese words increase the number of noticeable syllabic homonyms. In English, 'lion' and 'library' share an initial syllable but the following syllables mask the similarities.

Taiwan uses a system of **bopomofo** or **zhuyin fuhao** containing thirty-seven purely phonological characters placed alongside the Chinese characters. Teachers teach early students bopomofo and expect students to use bopomofo to help them comprehend texts with unfamiliar characters. Each symbol stands for a specific phoneme in Chinese, and students learn to blend the sounds together to create Chinese syllables.

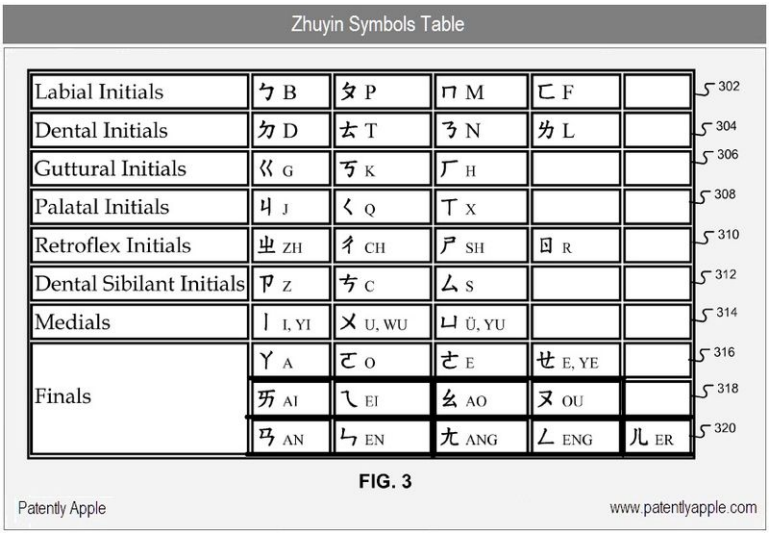


Figure 4.3: Bopomofo or zhuyin fuhao, the phonetic system used in Taiwan to supplement Chinese characters

Fig ??. Purcher, Jack. "Zhuyin symbols table". 2010. Patently Apple.

http://www.patentlyapple.com/patently-apple/2010/09/apple-everything-you-wanted-to-know-about-zhuyin-on-the-iphone.html.

In mainland China, schools also have a phonetic system using the Roman alphabet called **pinyin** meaning *spelled sounds* (Gnanadesikan, *Revolution*, 75), and the system has become conventionalized in other areas of the world learning Chinese as an L2[1] (Chung, Hanyu, 150). Again, students have the ability to decipher unfamiliar characters with the phonetic aid, yet recent research questions the effectiveness of pinyin as an instructional tool. The first obstacle results from the complications of romanization as the letters do not completely match up with other systems. For example, <c> represents [ts^h] in pinyin but [k], [k^h], or [s] in English and therefore, the effectiveness of the system depends on the familiarity of the learner with pinyin (Lee, Effectiveness, 12). The familiarity and ease of pinyin may also hinder students when presented with the pinyin symbols before seeing the Chinese character (Chung, Hanyu, 162).

4.1.2 Korean

Buddhist missionaries and Confucius envoys brought the Chinese script to Korea, and the bureaucracy adopted the language for both administration and culture (Campbell, *Handbook*, 86). However when Koreans attempted to incorporate written Chinese with Korean, the isolating nature of Chinese clashed with the agglutinative nature of Korean (Campbell, *Handbook*, 86). Despite all this, the the educated and literate did not want to learn a new language, and Korea continued to use the Chinese script (Gnanadesikan, *Revolution*, 192).

King Sejong (1397-1450) became dissatisfied with the low literacy rates and commissioned a new alphabet for the Korean language, which would soon become the only indigenous

[1] **L2**: any non-native language (Archibald, Second, 391)

phonetic script of East Asia (Dobrovolsky, Writing, 559). Even now, the script remains successful and has been praised to be "the most efficient and logical writing system in the world" (Gnanadesikan, *Revolution*, 191). The script has only undergone minor modifications since the 1400's, contains twenty eight symbols that fully represent all the phonemes of Korean, and accurately represents speech best compared to any contemporary language (Pae, Korean, 108), making it easy to learn (Gnanadesikan, *Revolution*, 201).

Hangul contains shapes inspired by Chinese script but the forms also relate to phonological theories with the shape of vowels representing the place of articulation of that vowel (Traulsen, Hangul, 105). The consonants fit into classes based on manner of articulation and similar consonants visually have similar form (Gnanadesikan, *Revolution*, 197).

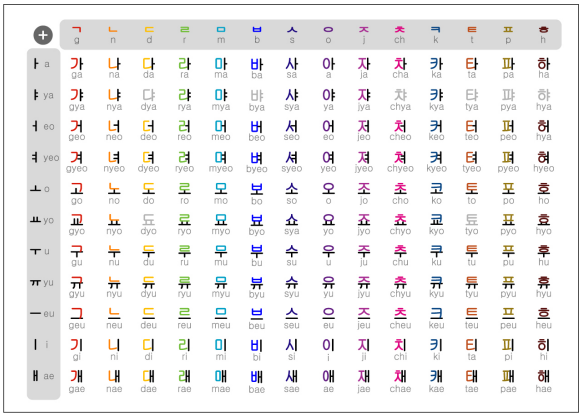


Figure 4.4: Korean script chart of all possible consonant and vowel combinations of open syllables

Fig 4.4. Dorresteijn, Joop. 2008. http://www.joop.in/Archive/learn-to-read-and-pronounce-korean-hangul-in-2-days/.

Despite the practicality of the new script, King Sejong did not wish to impose the script on the stubborn gentry and never mandated a change to the new script (Gnanadesikan, *Revolution*, 205). Due to its simplicity, Hangul nearly fizzled out after King Sejong's death but eventually in 1894 became a nationally sanctioned writing system (Gnanadesikan, *Revolution*, 206).

4.1.3 Japanese

Korean scholars introduced Chinese to Japan through the *Analects of Confucius* (Gnanadesikan, *Revolution*, 113). Japanese, like Korean, does not have the same language typology as Chinese, and the Chinese script could not fully represent the language, especially Japanese names and honorifics (Gnanadesikan, *Revolution*, 114). Initially, Japanese used the Chinese script, retaining Chinese word order and grammatical structure. This allowed people to read Chinese works in either Chinese or Japanese and likewise, individual characters could be pronounced either in Chinese or Japanese. By the 9th century, the Japanese tapped into the advantage of a logographic script and could read Chinese texts in Japanese without learning Chinese (Gnanadesikan, *Revolution*, 119).

Syllabic writing suits Japanese well because it contains mostly 'clear syllables', syllables without codas (Mandel, *Japanese*, 30) and the Japanese developed two different phonetic scripts, **hiragana** and **katakana**. *Kana* means 'syllable' and hiragana and katakana mean 'smooth syllable' and 'partial syllable' respectively (Mandel, *Japanese*, 19). The two scripts have similarities often obscured by the varying Chinese sources of the symbols (Gnanadesikan, *Revolution*, 120).

Chinese characters belong to a third script known as **kanji**. Currently, native words without kanji use hiragana while katakana is used for emphasis but theoretically any

Japanese word can be written in either script (Gnanadesikan, *Revolution*, 126). These two syllabaries fully represent the Japanese language, but the Chinese script represented so much heritage and rather than removing it completely, the Japanese now use all three scripts together (Mandel, *Japanese*, 12). In 1946, the Japanese government identified the 1,850 characters necessary for a "good education" to ensure that the schools taught students the necessary characters (Mandel, *Japanese*, 18).

4.2 Growth of the Roman alphabet

4.2.1 Turkish

Sometimes, a script mismatches the spoken language to the extent that gradual shifts cannot adequately accommodate. Turkey has used a number of different scripts throughout its history, but used the Arabic script for over a thousand years. During the early twentieth century, higher class Turkish officials spoke Ottoman Turkish, an administrative and literary language containing a mixture of Turkish, Arabic, and Persian that the large majority of people could not understand (Lewis *Turkish* 8). Arabic and Persian each brought their own grammatical conventions, and Turkish suffixes could be added onto any foreign word (Lewis, *Turkish*, 7). To further complicate things, Arabic script cannot properly represent Turkish phonotactics because Arabic does not typically represent vowels whereas Turkish relies heavily on its vowels (Lewis, *Turkish*, 27). A written word using Arabic script could be read multple ways, for example the written form ولو could stand for *ulu* 'great', *ulu* 'possessors', *ölü* 'dead', *evli* 'married', *avlu* 'courtyard', or *avlı* 'stocked with game' (Lewis, *Turkish*, 26).

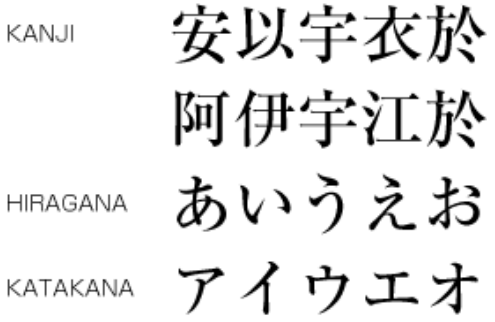


Figure 4.5: Japanese Kanji, Hiragana, and Katakana scripts

Fig 4.5. "Kanji, Hiragana, and Katakana". 2008. W3C Working Draft. http://www.w3.org/TR/2008/WD-jlreq-20080411/.

Lewis quotes Sir Charles Eliot, author of the 1926 *Encyclopedia Britannica*, "the result is that pure Turkish words written in Arabic letters are often hardly intelligible even to Turks and it is usual to employ Arabic synonyms as much as possible because there is no doubt as to how they should be read." Less educated Turks spoke Kaba Türkçe 'crude Turkish' without the Arabic and Persian influence and therefore could not understand Ottoman Turkish.

Throughout the mid-nineteenth century, Turkish scholars began discussing language and writing reform; individuals and private groups made isolated attempts but none succeeded. Some suggestions included using the Arabic script with additional diacritics for vowels, using fully vocalized Arabic letters without connecting them, and adding additional letters

for vowels (Lewis, *Turkish*, 28). The Grand National Assembly voted to accept "international numerals" on May 20, 1928 and a member of the assembly asked about adopting "international letters" as well. This resulted in a committee of nine to consider how Turkish might adopt Latin letters.

In June, the committee split into two groups, one focused on the alphabet and the other on grammar. When the alphabet seemed acceptable to Mustafa Kemal, the president at the time, he introduced it at a gala on August 9, 1928, less than three months after the committee first met. Lessons in the new alphabet for government officials, teachers, and others began two days later and the Grand National Assembly passed the law "on the adoption and application of the new letters" on November 1, 1928. While there were small changes to the alphabet from the first public announcement in August, the government enforced a quick transition. They forbade schoolbooks written in the old script and all books had to be printed in the new alphabet by the end of the new year. Private correspondences had to use the new alphabet by June 1, 1929, and "old Arabic letters" were temporarily allowed in records as shorthand, this was only acceptable until June 1, 1930 (Lewis, *Turkish*, 32-38).

The fears of the opposition that the shift would take Turkey away from its cultural and religious history came true as Turkey became the one of the first Muslim countries to stop using the Arabic script, a noticeable alignment with the West. The younger generation could no longer understand Ottoman literature (Tachau, Reform, 9).

As evidenced through Turkish, challenges of an ill-fitting orthography include low literacy rates, misunderstandings, and dissatisfaction with the system. With phonetic systems especially, it is important that the written forms match up with the spoken system. This does not mean that every alphabet should have a one-to one form to sound ratio, but it does explain why Arabic's lack of written vowels did not represent represent Turkish's vowel heavy language well. In terms of literacy, this writing revolution proved to be an enormous success, improving from 9% in 1924 to 65% in 1975 and 92.3% in 1995 (Lewis, *Turkish*, 37).

4.2.2 Vietnamese

Vietnamese has typological similarities to Chinese such as monosyllabic words and a lack of inflections (Jensen, *Signs*, 183) yet it serves as another example of a language shifting away from Chinese logograms. However, Vietnamese contains six tones as opposed to the four tones of Mandarin Chinese and due to the increased number of syllables, contains less homophones. Arguably, the use of the Roman alphabet for Vietnamese evidences a logistical possibility of romanization for Chinese (Jensen, *Signs*, 184).

Initially, Vietnamese also used Chinese characters with Vietnamese phonetic values which evolved into **chũ' nôm**, a modified Chinese script (Jensen, *Signs*, 184). Catholic missionaries created **chũ' qu'ôcngũ**, their romanization of Vietnamese, with influence from Portuguese, Italian, and French writing (Quinn, *Vietnamese*, viii). The script of twenty-two letters represents each of the six vowels with diacritic marks and each syllable contains two necessary components, the vowel and the tone (Ngo, *Vietnamese*, 17).

Pronunciation of vowels by YourVietnamese.com

Vowels	IPA	English sounds	Examples
a	/a:/	far	xa ("far")
ă	/a/	father (shortened a)*	mắt ("eye")
â	/ə/	but *	đất ("earth")
e	/ɛ/	red	xe ("vehicle")
ê	/e/	may *	hên ("lucky")
i, y	/i/	me	hình ("image"), yêu ("love")
o	/ɔ/	law	lo ("worried")
ô	/o/	spoke *	cô ("she")
ơ	/ə:/	sir	trơn ("slippy")
u	/u/	boo	bún ("noodle")
ư	/i/	uh-uh *	mứt ("jam")

(a) Vietnamese vowel chart

Kính mừng Maria,
đầy ơn phúc. Đức Chúa Trời
ở cùng Bà. Bà có phúc lạ
hơn mọi người nữ, và Giêsu,
Con lòng Bà, gồm phúc lạ.
Thánh Maria Đức Mẹ Chúa Trời,
cầu cho chúng con, là kẻ có tội Khi này
và trong giờ lâm tử. Amen.

(b) A sample Vietnamese text

Figure 4.6: Chữ' qu'ôcngữ, the Romanized Vietnamese script

Fig 4.6b. JWB. "Writing systems worldwide". Wikipedia Commons licensed under CC BY-SA 3.0. http://commons.wikimedia.org/wiki/File:Writing_systems_worldwide.png#mediaviewer/File: Writing_systems_worldwide.png.

Fig 4.6a. Learn Vietnamese Your Way. 2011. https://www.youtube.com/watch?v=_rKYGr-buzI.

With eleven vowels and six tones (Gnanadesikan, *Revolution*, 261), Vietnamese requires many diacritics to represent the sixty-six possibilities with only five symbols and accent marks. However, to those unfamiliar with Vietnamese, the script looks weighed down and the diacritics seem to slow down the reading process. Despite this, Vietnamese proves a capability of the Roman alphabet to functionally incorporate language components such as tone.

4.3 Shift from the Roman alphabet

4.3.1 Shavian

Irish playwright George Bernard Shaw (1856-1950) had a strong interest in English spelling reform, and in his will, he left a sum of money for the development of a new English script based on the pronunciation of King George V "also known as Northern English" (Shaw, *Androcles*, 9). He required the script to contain at least forty symbols, each with a one-to-one correspondence between sounds and forms, and forbade the use of any Latin letters to prevent confusion. Shaw wanted Isaac Pitman, inventor of Pitman shorthand to create the script, but due to legal difficulties, the public trustee assigned to Shaw's will set up a competition and received 450 alphabet designs (Shaw, *Androcles*, 10).

Kingston Read won the competition with an alphabet of 48 letters, each representing one phone in English. Supporters assert the script lends itself to quicker writing compared with the current English writing system, especially with the use of word symbols, 𐑞 [th] for *the*, 𐑓 [v] for *of*, 𐑯 [n] for *and*, and 𐑑 [t] for *to* (Shaw, *Androcles*, 143).

After Read won the competition, the trustee used the remaining money to republish "Androcles and the Lion" with English and Shavian parallels (Shaw, *Androcles*, 9). Soon

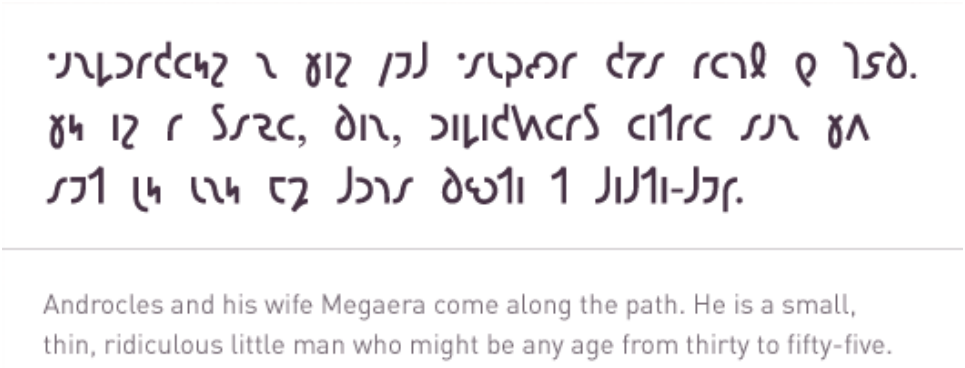


Figure 4.7: A sample Shavian text

Fig 4.7. Burney, Noah. "Sample Text". 2012. The thoughts, works, and discoveries of Noah Burney.

http://oldblog.noahburney.com/language-scripts/the-shavian-alphabet/.

after, the book went out of circulation and the alphabet mostly died out. Since then, the growth of the internet helped Shavian's existence as interested parties started forums and websites dedicated to using Shavian and inspired an advocate to transcribe *Alice in Wonderland* into Shavian in 2013. Today, we can find multiple different scripts with unicode graphs online and an included font on Mac computers, "Apple Symbols" contains many Shavian letters.

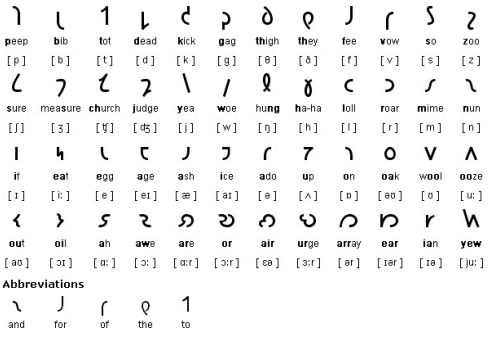


Figure 4.8: The Shavian alphabet

Fig 4.8. Paperblank. "The Shavian Alphabet" 2013. Endpaper.

http://blog.paperblanks.com/2013/11/writing-wednesday-when-does-ghoti-sound-like-fish/the-shavian-alphabet/.

Shaw never intended for the new script to replace the Roman alphabet but instead desired for a simultaneous use of both scripts. However, the extreme precision of the script challenges the growth of Shavian, as the letters are so precise that many dialectical variations in pronunciation changes how a word should be written. Shavian did not reach a wide audience, in part because "innovation is rare in the history of the written word and the force of literacy prestige is strong" (Gnanadesikan, *Revolution*, 219), meaning that once a script becomes established people are resistant to changing it (Moorhouse, *Alphabet*, 172) and Shavian serves as an example of the difficulty of introducing a new writing system. Along with a resistance to change, the adoption of a new script causes old books and other printed materials to become incomprehensible (Moorhouse, *Alphabet*, 176). In Turkey, old Ottoman Turkic materials became illegible to the Turkish people, especially the younger generations (See 'Turkish' 4.2.1 on page 30).

4.3.2 Unspell

The desire to revise English writing remains and many modern day scholars and educators still create alternatives for English literacy and spelling. Dmitry Orlov, creator of Project Unspell, describes his shapes as an "alternative orthography of the English language" (Orlov, Unspell, 1). He shares a similar vision with George Shaw, a desire to devise a script that represents spoken English. Neither expected their systems to completely replace English writing but supplement the script, as Braille supplements English signs (Vangelova, Spelling, 5). Both scripts also avoid Roman letters to prevent confusion or distraction.

Unlike Shavian, Unspell contains only thirteen symbols, arranged to encompass a variety of dialects and accents (Orlov, Method, 2). Orlov especially desires to use Unspell when teaching new learners of English how to read and even as a tool for accessible literacy training, similar to what the Taiwanese do to teach Chinese characters with the bopomofo system (see section 4.1.1 on page 27) (Vangelova, Spelling, 4).

These movements typically gain little traction as it takes a lot of effort to change a script yet several books including *Treasure Island*, *Tom Sawyer*, *Alice in Wonderland*, and the Bible have Unspell editions. However, as English continues to develop and change, people will continue to discuss spelling reform which many other languages experienced already[2].

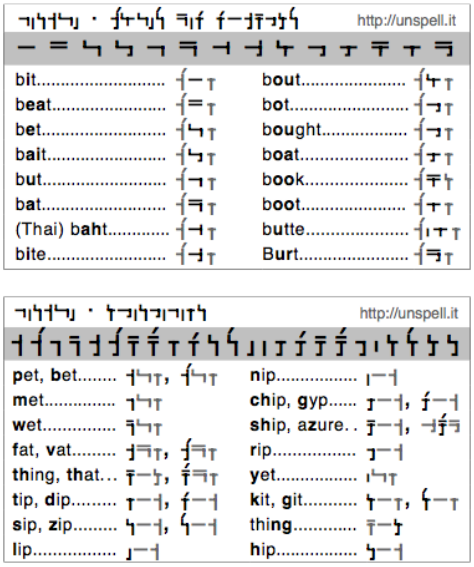


Figure 4.9: The Unspell symbols

Fig 4.9. "Unspell Cheat Sheet". Unspell. http://unspell.blogspot.com/2015/01/the-problem-and-solution.html.

4.4 Arabic expansion

Second only to the Roman alphabet, the Arabic script occupies a large geographical area (Kaye, Arabic, 743). While the language existed prior to Islam, the inscription of the Quran into Arabic letters regularized the script and prepared the way for the script to proliferate beyond the spoken language (Kaye, Arabic 743). Due to the establishment of the Quran, there will never be drastic script reforms (Bauer, Arabic, 563).

The script displaced some native scripts such as Persian and became the first script of many people groups in North Africa (Jensen, *Signs*, 332). Arabic script covers Classical Arabic, Modern Standard Arabic, and a few local dialects of Arabic but also a wide range of non-Semitic languages including some Berber languages, Persian, Pashto, Kurdish, Urdu, Dravidian, and Sulu (Kaye, Arabic 743). Some languages such as Turkic and Malay have

[2]For more information on spelling reform in different countries, research the German orthography reform of 1996 or L'Académie française.

since shifted into using the Roman alphabet (Kaye, Arabic, 743). Interestingly, Maltese, an Arabic language used by Christians in Malta use the Roman alphabet, indicating a religious split between scripts illustrated in section 4.5 on page 36 (Comrie, Languages, 688).

As mentioned earlier in section 2.2.4 on page 10, Arabic classifies as an abjad, meaning it primarily codifies consonants but still has some vowels. While this would cause chaos in English as we would write "tap", "tip", "top", and even "tape" as "tp", this system works in Arabic as vowels occur highly regularly based on part of speech and other grammatical factors (Bauer, Arabic, 563). Arabic does have diacritics for short vowels although typically only the Quran or other important documents will use the highly vocalized script (Bauer, Arabic 562).

Arabic sounds do not perfectly match the sounds of all the other languages which use the Arabic script. Therefore, pronunciation of some letters vary between languages, people will create new letters using the existing diacritics consisting of dots placed either above or below each sign (Kaye, Arabic, 746).

4.5 Cyrillic influence

Two forms of Slavic writing existed: Glagolitic and Cyrillic scripts but we do not have enough context to determine the exact historical origins of the two scripts. Most modern scholars attribute the development of Glagolitic to Constantine (also known as St. Cyril) and not the Cyrillic alphabet. Glagolitic resembles Greek minuscule but encodes Slavonic sounds not present in Greek (Jensen, *Signs*, 492), which suggests a formation of Glagolitic from Greek cursive formalized by Constantine (Cubberly, Slavic, 346). Eventually, Cyrillic replaced Glagolitic (Jensen, *Signs*, 498) everywhere except in Croatia (Cubberly, Slavic, 349).

A simplified version of the Cyrillic alphabet resulted in our modern day Russian alphabet (Jensen, *Signs*, 498) through the removal of redundant letters (Cubberley, Slavic, 350). Despite movements to replace the script with the Roman alphabet, the Russian Cyrillic alphabet still grows (more on this in section 4.5 on page 35) (Jensen, *Signs*, 503).

Similar to the Roman alphabet and Arabic script, the Cyrillic alphabet also spans a large range of languages from different language families. Slavic languages which utilize the alphabet include Russian, Belarusian, Ukrainian, Bulgarian, and Macedonian (Comrie, Adaptations, 700) and non-slavic languages which also use the script include Azerbaijani, Mongolian, Moldovan, Tajik, and Uzbek to name a few (Comrie, Adaptations, 707). Many of these script changes resulted from the 1917 Revolution of the Soviet Union to rid the Islamic people of Central Asia of the Arabic script and when Turkish also switched to the Roman alphabet, the Soviet government forced everyone to adopt the Cyrillic alphabet (Campbell, *Handbook*, 44).

Here, we introduce the idea of scripts as an **isogloss**[3] but instead of separating dialects, we can use this concept as a way of identifying borders between religions. In Russia and the

[3]**isogloss:** lines drawn on a dialect map to represent boundaries between dialects.

surrounding areas, the split between Cyrillic and Roman clearly mirrors a split between the Roman Catholics and Eastern Orthodox Christianity.

4.5.1 Azerbaijani

Azerbaijani's script changes illustrate some of the impact of politics and religion on language and writing. Surrounded by Iran, Russia, and Turkey, Azerbaijan received pressure from each of these countries to use their writing. During the seventh century, Arabs introduced their alphabet alongside Islam (Hatcher, Azerbaijani, 106), but by the nineteenth century, the phonological mismatch similar to that of Turkish (see section 4.2.1 on page 30) urged reforms (Hatcher, Azerbaijani, 107). The Soviet Union also desired to break Azerbaijani cultural ties with the Muslim world and simultaneously pushed for Romanization (Hatcher, Azerbaijani, 108). The change to the Roman alphabet increased literacy from below 10% to "an official literacy rate of 98% within a few decades" (Hatcher, Azerbaijani, 109), illustrating the success of the new script.

However, after Turkey unexpectedly also switched from the Arabic abugida to the Roman alphabet, Joseph Stalin decreed that everyone must use the Cyrillic alphabet instead of the Roman one (Hatcher, Azerbaijani, 109). Not only did this isolate the Turkic states from Turkey, it also encouraged the Russification of these smaller countries. After independence from the Soviets in 1991, the Azerbaijanis faced a choice of who to align with through their choice of script: the Arabic of the Muslims, the Cyrillic of the Russians, or the Roman of the Turks (Hatcher, Azerbaijani, 112). Although Azerbaijan also borders both Georgia and Armenia, two countries with their own writing systems, no one ever pushed for Georgian or Armenian writing as these two countries aligned with Christianity rather than Islam (Hatcher, Azerbaijan, 109). This again underscores the split between religions as an isograph and that a desire to distinguish can be just as powerful an influence over language as the desire to align with a larger political or social force.

4.5.2 Mongolian

Written Mongol, the Mongolian script began in the times of Genghis Khan in 1227. However the script underwent changes and currently the Mongolian script has two alphabetic manifestations, Written Mongol and Cyrillic Khalkha (Jensen, *Signs*, 418) although neither of these orthographies represent any currently spoken dialect but have the ability to encompass multiple dialects of Mongolian (Janhunen, *Mongolian*, 21). Written Mongol traces back to the Mongol Empire of the 12th to 14th centuries while Cyrillic Khalkha reflects the speech of Outer Mongolia in the early 1900's (Janhunen, *Mongolian*, 21).

Written Mongol still contains significant cultural value and has acted as a unifying environment for all Mongolic speakers (Janhunen, *Mongolian*, 6). The orientation of the Written Mongol remains up and down, complicating the adaptation of Mongolian into technology with other scripts. In electronic communications, people use more and more Romanization but without a set pattern of Romanization, many variations exist

(Janhunen, *Mongolian*, 23). Written Mongol still contains significant value and identity and therefore, the script remains in use only peripherally (Janhunen, *Mongolian*, 22).

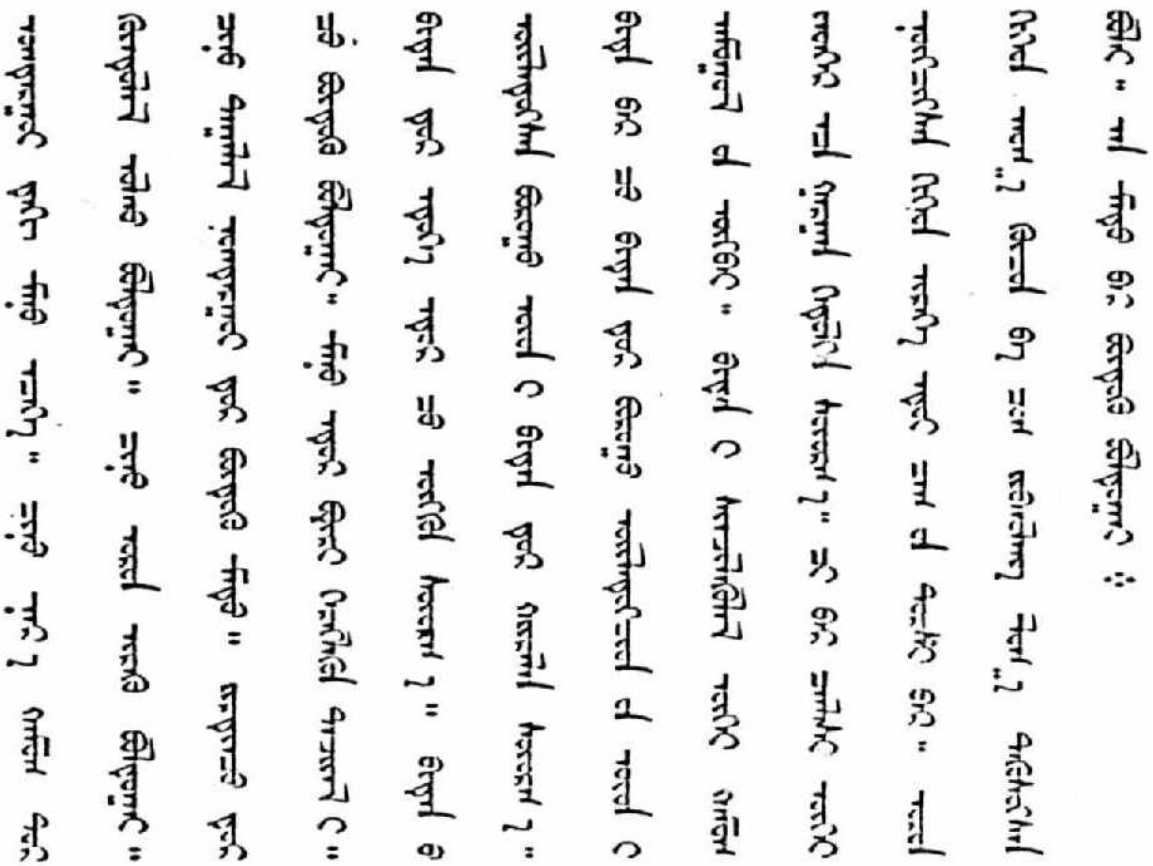


Figure 4.10: The Lord's Prayer in Mongolian

Fig 4.10: "The Lord's Prayer in the languages of Russia" Mark Hürri. 1870. http://www.christusrex.org/www1/pater/JPN-mongol.html

Chapter 5

Predictions

5.1 Maps

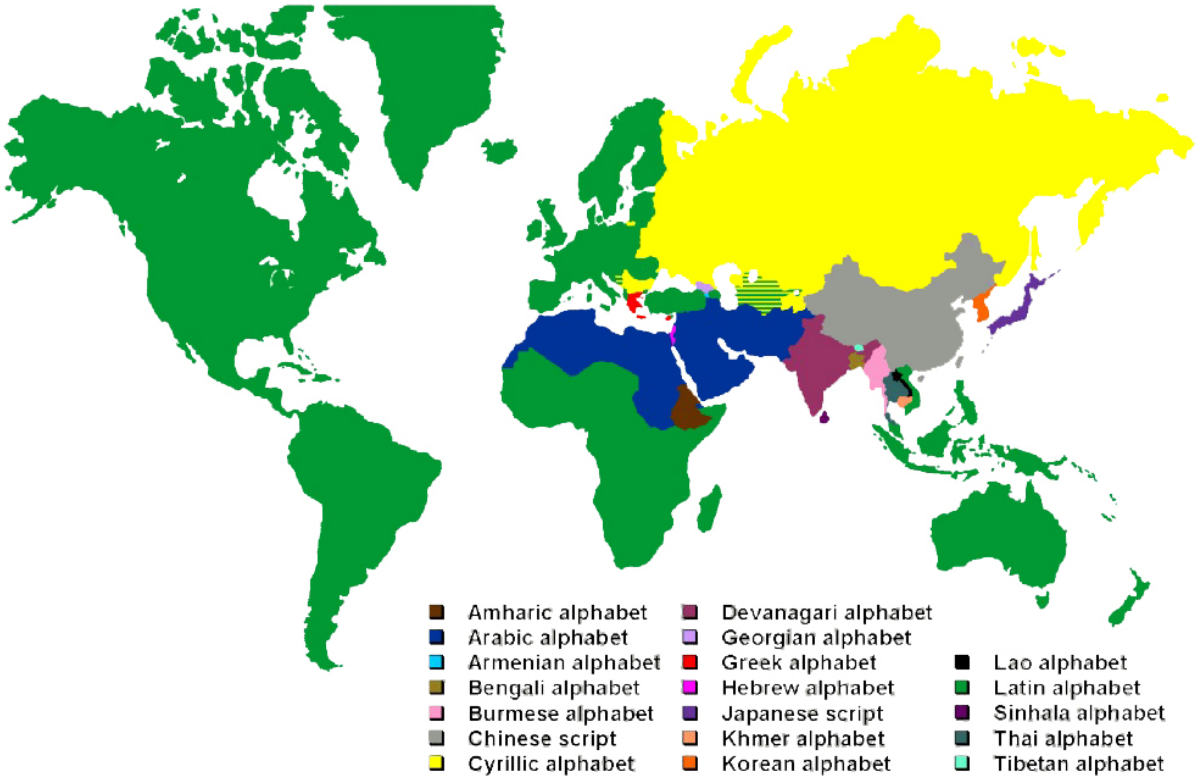


Figure 5.1: Color coded map of script choices throughout the world.

Fig 5.1. Purcher, Jack. "Zhuyin symbols table". 2010. Patently Apple.

http://www.patentlyapple.com/patently-apple/2010/09/apple-everything-you-wanted-to-know-about-zhuyin-on-the-iphone.html.

Maps 5.1 and 5.2 depict the current representation of writing throughout the world. Despite the distorted proportions in the geographic projection, we can still see the overwhelming prevalence of the Roman alphabet across continents (labeled as 'Latin alphabet' on the maps, more on the distinction between the two terms in section 6.1 on

page 43). Both the Arabic and Cyrillic scripts occupy a substantial portion of the world, and the South Asian area contains Devanagari and a variety of related abugidas. As we can see from the modern examples of language shifts, many scripts result from historical conquests or other cultural influences and based on the maps and previous script descriptions, we find a trend of languages using or shifting into one of these popular languages.

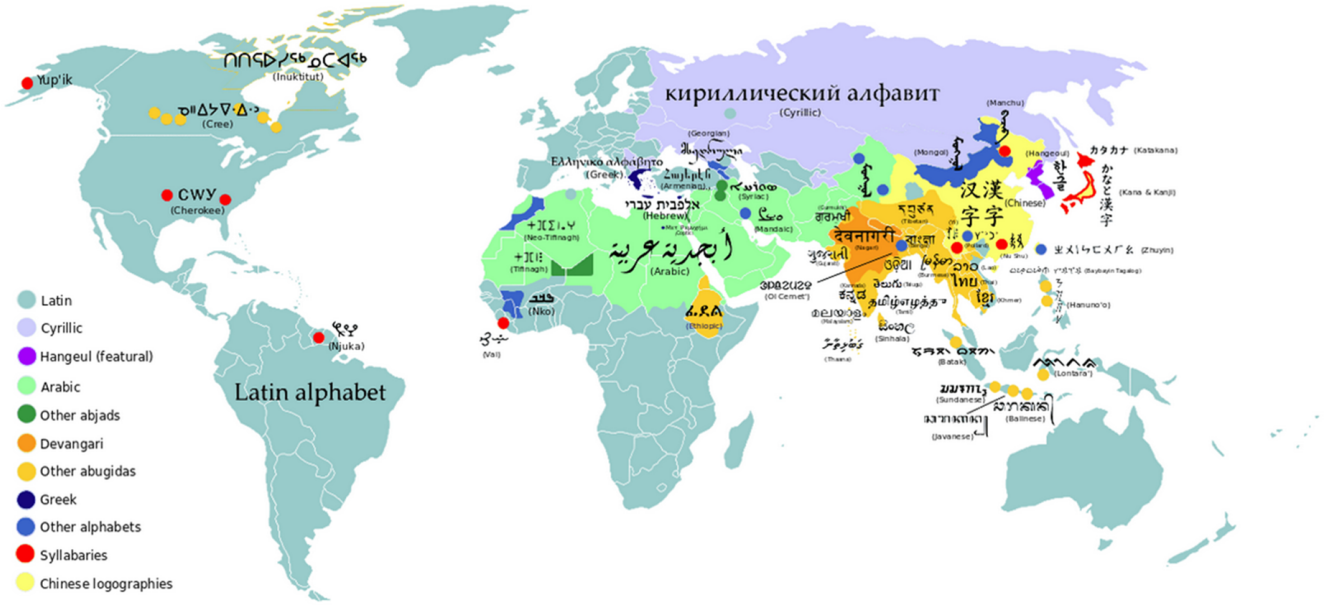


Figure 5.2: Map color coded by dominant script and type of writing system

Fig 5.2. 2014. World standards. http://www.worldstandards.eu/other/alphabets/.

5.2 Principles for a good script

William Smalley, a missionary among the Hmong of Laos created an alphabet for them using Roman letters (Noble, Smalley, 1), and in his chapter "How shall I write this language?" he identifies five key considerations when assigning or creating a script:

1. Motivation and acceptance by the society and those in political authority
2. Representation of speech
3. Ease of learning
4. Transfer value
5. Ease of reproduction (including typing and printing facilities)

These principles provide a good starting point for us to analyze factors that influence preliterate societies. First, the learner must have motivation to learn the script because the script only has purpose if people use it. Overall society and political authorities must also agree with the script choice. As with Roman in Azerbaijan (see section 4.5.1 on page 36), the script fit the language and literacy boomed from less than 10% of the population

reading to almost complete literacy. However, because the government decreed a change, the Azerbaijanis had no choice but to adopt the Cyrillic script.

The script must represent speech. This requirement does not force scripts to represent a language in one specific way or another, but almost all ideographic or logographic languages no longer exist. If not for its historical and cultural tradition, the Chinese script would likely also fall out of use. When teaching everyone to read at the same time, a script that visibly and explicitly represents phonological information helps the learners to use the language. Adequate and complete language representation also does not require over precision, such as exhibited by Shavian (see section 4.3.1 on page 32). If the goal of written language was simply to record all phonemes of the language, everyone could adopt the International Phonetic Alphabet (IPA) or a similar system to record all spoken language.

The script cannot be too complicated. One measure of a script's success stems from its functionality and if it grows too complicated, people's desire to learn the script declines. Cherokee began as a picture-script but this proved too many symbols, and Sequoyah revamped the script to only eighty-six syllabic symbols. Especially for a newly literate group, something too complicated would only frustrate people and hinder the establishment of the script and literacy as a whole. The script must balance ease of learning and representation of speech so that the script still represents the necessary sounds but does not become overburdensome for a new learner.

This idea of transfer value weighs heavily in contemporary language script choices. In general, a preliterate society does not decide to write without context or reason, but oftentimes, someone introduces the idea of writing. If the language of prestige utilizes the same script, members of the society may learn to read and write two languages with one script. However, we must remember that scripts do not often fit multiple languages well and oftentimes in order to accommodate a new language, certain signs must be reallocated or created.

Finally as for ease of reproduction, Smalley wrote these principles in 1963, when typewriters still dominated the production of printed texts. If a new script utilized completely new symbols, the typewriter would struggle to accommodate. With modern technology such as computers, we are no longer limited by the physical capabilities of a typewriter. However, new symbols require coding, installing fonts, and ensuring transferability which may well be more effort than necessary.

Based on the script vignettes seen earlier, all five of Smalley's principles have relevance though not all of them carry equal weight. The idea of transfer relates with the other qualifications. Transfer deeply impacts motivation and acceptance as societies often either adhere closely to neighboring languages or intentionally establish a contrasting script. Transfer impacts ease of learning, depending on the previous literacy of the learners. If everyone already reads and writes English in the Roman script, perhaps using the Roman script would facilitate reading better. On the other hand, if the new script wants to borrow symbols but assign different sounds to the symbols, perhaps a different script would fit the situation better. Ease of reproduction also relates to transfer as scripts with transfer value typically will already have the technology in place for reproduction. Additionally,

languages vary widely and therefore a script that fits one language well may face difficulties in adjusting to a new language (Gnanadesikan, *Revolution* 144).

Political authority also plays a greater role in the process and even if a script matches the language completely and the community or those in power severely oppose the system, it will not be used. Without the Turkish government's enforcement of the shift from Arabic script to Roman alphabet, the change would not have been possible.

How can we tell if a script is a good script? Scripts that fit languages well will continue and scripts that do not fit well will call for changes and revisions. Hangul, created specifically with Korean in mind, has been in use by Koreans for over 600 years. On the other hand, Turkish and Japanese used Arabic and Chinese scripts which did not fit the script and people constantly proposed new ways to write their language.

5.3 Creating a new script

What about the creation of a new script? Someone who knows the language well could create the perfect script, similar to what happened with Korean (see section 4.1.2 on page 28). But aside from the Cherokee and Cree scripts, people have not been in the business of inventing new scripts in the past four hundred years.

While the idea of inventing a script may seem intriguing, many practical concerns keep people from doing so. First, one must decide on the letters themselves which brings up the question of shape and number of signs. No one will have significantly more experience with the script and be more trained to teach it to future learners. The new script wouldn't help literacy in any other languages, which may limit people's motivation to learn the script. Additionally, even well-established non-Roman scripts face the pressure to Romanize and a new script would constantly have to defend itself against this pressure. At the end of the day, the effort required to create a new script and maintain its status simply requires too much effort and therefore societies look towards existing scripts.

5.4 Ideal language

Based on the case studies in Chapter 3 and 4, the best choice for new orthographies does not include logographic and ideographic systems; any type of picture writing becomes unwieldy as everyone tries to learn a new system. Additionally, a shallow alphabet or syllabary would fit better because the meaning and representation of the visual stimulus becomes more apparent to the reader.

5.5 Ideographic systems

As mentioned earlier, Chinese still uses an logographic system so evidently, picture-writing can still represent speech. But what makes Chinese unique and why do we not recommend any type of picture-writing for new languages?

(a) American road signs

(b) Standard bathroom sign

(c) I <3 NY

Figure 5.3: American logograms

Fig 5.3a. featurePics. http://www.featurepics.com/online/Road-Signs-Pack-Illustrations220369.aspx.

Fig 5.3b. obob. Bathroom Ideas. http://whningli.com/bathroom-sign.htm/bathroom-sign-4-good-design-bathroom-design.

Fig 5.3c. Glaser, Milton. "I Love New York logo". 2009. http://en.wikipedia.org/wiki/File:I_Love_New_York.svg.

The Chinese script's extensive history provides a secure support for the script, and Chinese speakers' pride in the script secures its place against the trend of Romanization. The existence of literate people in the language provides the biggest contrast between learning the logographic Chinese script and any other language system. If a preliterate society created and began using a pictographic system, everyone has to simultaneously learn hundreds and hundreds of symbols. Chinese on the other hand has literate people who aid new learners.

The phonetic systems of pinyin and bopomofo described in section 4.1.1 on page 26 also assist learners with learning the phonetic sounds of characters until the learners internalize the characters. If a newly literate society chose to go this route, the learners would have to learn two completely new systems without support, lowering motivation to learn the picture-writing orthography.

Once people command a firm understanding of the phonetic script, the number of logograms in a phonologically based script will increase over time (Gnanadesikan, *Revolution*, 25) as people look for ways to save time, energy, or space. Examples of logograms in American culture include road signs, the woman and man signs for bathrooms, and even I<3NY. Even without words, everyone understands what each of these symbols mean.

Chapter 6

Reasons for the Roman alphabet

I predict that almost any preliterate society beginning to codify their language into writing will use the Roman alphabet. The Roman alphabet fits many of the qualities of a suitable script as outlined in the previous chapter, and the structure of the alphabet itself supports its growth. The alphabet contains only twenty six symbols and an expansive arsenal of diacritics to fill in additional sounds. Unlike in Hangul (Korean) or Devanagari (Sanskrit), the forms of Roman letters do not change based on surrounding letters, and the linearity of the script adds to its learnability, especially for new learners (Smalley, Script, 171).

The growth of the Roman alphabet cannot be completely ascribed to its form or flexibility. We highlight some of the other major political and historical circumstances which have contributed to the lasting ubiquity of the Roman alphabet.

6.1 Roman or Latin alphabet?

First, a clarification. World maps 5.1 and 5.2 refer to the **Latin alphabet** while in the rest of this paper, we use the term **Roman alphabet**. We chose the term Roman alphabet because it encompasses a broader spectrum of scripts than Latin alphabet. We find a similar situation when we look geographically northeast and consider the terms **Russian alphabet** or the **Cyrillic alphabet**. Perhaps because people barely use Latin now while Russian remains a living language, the assignment of Russian alphabet seems more limiting than the term Latin alphabet.

Gnanadesikan asserts that "there is no longer one Roman alphabet, but many" (Gnanadesikan, *Revolution*, 260). As we will discuss in chapter 7, the different languages which use the Roman alphabet do not all contain the same phonemes and they often adapt letters to stand for different sounds.

6.2 Religion

Early writing served three functions: bureaucracy, commerce, and religion (Gnanadesikan, *Revolution*, 1), and history displays an especially close relationship between writing and religion. Missionaries often brought education and literacy to people, mostly with their

own native scripts or occasionally the scripts existing in the area (Smalley, Script, 75). Buddhist and Confucius missionaries introduced Chinese to Korea and Korean missionaries of the same religions brought Chinese to Japan. Persian speakers introduced Islam to Turkey, evidenced by the fact that even basic religious terminology such as the words for 'prayer', 'fasting', and 'prophet' source from Persian and not Arabic (Lewis, *Turkish*, 5). Due to the growth and expansion of Islam, the Arabic writing system is the second most widely used writing system (Smalley, Script, 73). Christian missionaries invented the Cyrillic and Cree scripts, and the protestant missionary movement emphasized education and literacy, primarily so that people could read the Bible (Smalley, Writing, 1).

The influence of religion and missionaries cannot be discounted, as many of these societies likely would not have chosen to begin writing without outside influence of missionaries. As we consider different reasons that affect the script decision, Christian missionaries played a large role in the use of Roman script for many contemporary preliterate societies as they began to write. As we observed in Azerbaijani (section 4.5.1 on page 36), script choice often signifies an identification with a certain religion, and a Muslim country such as Azerbaijani would never consider the Christian scripts of Armenia or Georgia.

The spread of the Roman alphabet throughout the world owes a lot to missionary activity, yet in modern times, the Roman alphabet no longer necessiates a Christian connotation and other factors push for the alphabet's continued function.

6.3 Colonization and politics

To some, the Roman script represents the script of colonialism but the script also functions internationally, and many individuals have already been educated in it. The continent of Africa contains three scripts: Ethiopic, Arabic, and Roman. Ethiopia has remained independent throughout its history and therefore been able to retain its written and spoken language, and the Arabic in North Africa stems from Islamic influence. The rest of the continent uses the Roman alphabet, and in many places the colonizer language and script became the language and script of the country.

Morocco, for example, has two official languages: Arabic and Berber. Moroccans often use French in official and bureaucratic settings, adding a third language into the mix. Berber had its own writing system, although it may use the Roman or Arabic scripts depending on context, and French uses the Roman script. Every Arabic speaking country uses two types of Arabic: the formal, bureaucratic, and official Modern Standard Arabic and a local dialect of Arabic. Modern Standard Arabic uses the Arabic script while Moroccans typically write Moroccan Arabic with the Roman alphabet due to residual influence of colonization.

6.4 Technology

The growth of alphabetic writing in the 20th century added to the implementation of the Roman alphabet across the world (de Voogt, *Dispersal*, 7). While people are beginning to use non-Roman scripts online, English and the Roman alphabet have grown to become the

lingua franca[1] of the internet. Despite characteristics of writing on the internet that relate to both speech and writing, the internet is still largely text-based (Crystal, *Internet*, 28).

Prior to the development and use of unicode text encoding, the majority of the internet used the American Standard Code for Information Interchange (ASCII). ASCII, encoded with 8 bytes, limits the characters available to only 256 characters (Crystal, *Internet*, 222). With the development of unicode encoding however, characters each contain 16 bytes, expanding the number of possible characters to 65,000 (Crystal, *Internet*, 222). This enabled the use of more scripts online and currently the majority of the internet uses unicode, enabling the expansion of nonroman scripts online, but new fonts take time to create. Prior to computers, anything that needed to be typed came from a typewriter and creating a whole new alphabet that would be a lot of work when they could use something that already existed?

Figure 6.1: A "dumb phone" with the old typing system

Fig 6.1. O'Brien, Terrance. 2013. Engadget. http://www.engadget.com/2013/12/06/twitter-u2opia-ussd/.

Before computers, everyone used typewriters and this technology strongly imposed the Roman alphabet. With computers, it may require some programming, but people can type non-Roman languages, and the computer even arranges the graphemes, connecting and rearranging as necessary, such as with Hangul or Arabic. Typewriters did not offer such versatility, and the linearity of the Roman alphabet helped its growth. Currently the use of cell phones has grown largely in developing countries. Typically these phones have T9 (more on this in section 6.5 on page 46) which lends itself to the Roman alphabet. Even typing in Chinese from a T9 phone requires spelling things out phonetically.

6.5 Texting and internet slang

Internet slang and text speech have developed with the technology and now features prominently in everyday life. Texting first began with earlier technology, phones that only had nine number buttons which also contained letters. Text messaging could handle 140 bytes or 1,120 bits of information, resulting in 160 characters for Latin letters and punctuation and 70 characters for more complicated Chinese and Japanese symbols (Crystal, *Txtng*, 6).

The process of texting slang already existed in many languages, including English. The use of initialisms or creating acronyms has existed even in Latin times with the abbreviation of *post meridian* as *PM* and *note bene* as *NB* (Crystal, *Txtng*, 41). Nonstandard spellings of words such as *thanx* and *cos* also existed prior to texting (Crystal, *Txtng* 45).

This process shows the use of **rebus principle**, the process of using easily identified visual shapes to stand for abstract homonyms in English (Crystal, *Txtng*, 39). For example,

[1] **Lingua franca:** a medium of communication for people who speak different "first langauges" (Crystal, Language, 452)

Americans pronounce the symbol <4> as [fo:] and they began to use "b4" in place of *before* or simply "4" in place of of *for* (Crystal, *Txtng*, 38).

6.6 Emoticons and emojis

Figure 6.2: Lyrics to the popular song "Call Me Maybe" by Carly Rae Jepsen

Fig 6.2. Funny Emoji Song Lyrics.

http://www.featuredfilm.com/photographheq/funny-emoji-song-lyrics.

In recent times, emoticons[2], have played a large role in internet communications such as texting, email, and Facebook. These emoticons have evolved into small pictures, for example through Apple's emoji keyboard or Google Talk, Google's online chat function which automatically updates ":)" into a yellow animated smiley face. Commonplace through conversations, people often use smiley faces and other expressions either as pictures or spelled out with symbols such as "(:". With 845 emoticons available on iOS8 available to current iPhones, some people choose to communicate primarily with these symbols.

Despite the growing use of emojis, they still face the limitations of pictorial representation. What exactly does each symbol mean? How does the receiver know what the sender intended for the image to stand for? All the different faces clearly express different emotions, but the emoticons represent these emotions ambiguously. Although some of these emojis represent a more iconic image, these may still be difficult to identify exactly. For example, in image 6.2, there is clearly a woman with her right arm up. To someone who knows the lyrics to the song "Call Me Maybe" would assign the meaning "met" to the first emoji of a woman with her arm in the air, interpreting the image as a woman waving. However, the same image could also stand for someone raising their hand or reaching up for a high-five. The meanings of these symbols may shift wildly as they heavily rely on context.

[2]**Emoticon:** "A combination of keyboard characters designed to convey the emotion associated with a particular facial expression. They are typed as a string on a single line. The simplest forms represent basic attitudes: positive, in the case of :) and negative in the case of :(" (Crystal, Txtng, 180).

Chapter 7

Inadequacies and challenges of the Roman alphabet

7.1 Vowels and tone

The Roman alphabet began with Latin and grew to encompass Romance and Germanic languages of Western Europe and the alphabet fits the sounds of these languages well. However, as the individual sounds or range of sounds begin to diverge from these Western European languages, the limitations of the script become increasingly apparent. Vowels in particular exhibit this struggle as most languages contain more than five vowels, and tonal languages augment the problem because the Roman alphabet does not contain a system for expressing tone. As a result, languages must either create new symbols; use diacritic marks; or creating digraphs to represent additional vowels, vowel quality, and tone (Bendor-Samuel, African, 671). People tend to modify the alphabet with diacritics rather than create and expand the basic set of Roman letters (Gnanadesikan, *Revolution*, 259). Although Vietnamese (see section 4.2.2 on page 31) managed to include 66 vowel sounds with the five letters and diacritics, the script looks bulky due to all the additional markings required. Swedish contains nine vowels with long and short distinctions, bringing the total to eighteen (Fant, Swedish, 15) and it adapts by adding the letters ‹å›, and ‹ä›, and ‹ö› to the end of the alphabet for a total of twenty nine letters (Fant, Swedish, 2).

On the flip side, languages sometimes have fewer than five distinct vowels. For example, linguists generally accept that Quechua contains three vowels which become five sounds with allophones (Hornberger, Authenticity, 395). Due to Spanish influence in South America, many Quechua speakers want to write with all five vowels to match Spanish (Hornberger, Authenticity, 392), and the tension between the authenticity of general phonemic principles of writing and the acceptance of the community contest with each other.

Romanization of Arabic also illustrates difficulties with vowels, although some of these difficulties source from the nature of an abjad script. In Arabic, patterns of consonants contain the core meaning of words while the vowels fluidly shift depending on the part of

speech. Arabic contains two types of vowels: short vowels and matres lectionis[1]. Short vowels occur in writing only for specialized texts such as the Quran to avoid any possibility of miscommunication, but few other texts utilize full vocalization. The three matres lectionis: ا, ي, و each represent one of three long vowels and a corresponding range of vowels. This complicates Romanization because transcriptions cannot simply replace the matres with a Roman letter, transcribers must know the patterns for short vowels. Even with the three ranges of vowels in Arabic, transcriptions often use all five Roman vowels.

7.2 Consonants

Roman consonants also pose some difficulties. As seen earlier, English <s> signifies a variety of sounds while multiple letters represent the sound [s] (see 2.3 on page 12). This makes the old adage "consonants as in English" quite complicated as these consonants themselves may stand for variable sounds and many sounds in African languages do not exist in English (Bendor-Samuel, African, 670). The letter <c> exhibits this complication across languages: Italian <c> could be [k] or [tʃ] and Turkish <c> is [dʒ] while <ç> is [tʃ].

Hawaiian has one of the fewest phonemes of any language in the world, and the consonants can be represented by as few as seven letters, and therefore multiple distinct consonants represented in the Roman alphabet correspond with one phoneme in Hawaiian (Wise, Hawaiian, 313). Additionally, the glottal stop complicates writing as it functions as a consonant in Hawaiian but Westerners don't always recognize when it shows up and will insert the <'> symbol randomly to make the script look more "Hawaiian" (Wise, Hawaiian, 318). Korean contains a similar issue with writing consonants which vary in voicing such as ㄱ representing both <K>and <G> and ㄷ representing both <T> or <D>.

English uses many digraphs[2]: th as [θ], sh as [ʃ], and ng as [ŋ] to list a few. Digraphs further complicate matters as <th>, <ch>, <ph>, and <ng> use two letters together to stand for [θ], [tʃ], [f], and [ŋ] respectively. Yet when alphabetizing, we don't treat <th> as one unit but instead <t> followed by <h>.

7.3 Letter forms

7.3.1 Capitals and lower case

Only three scripts in the world distinguish between capital (upper case) and minuscule (lower case) letters: Roman, Cyrillic, and Armenian (Gnanadesikan, *Revolution* 222). In the Roman alphabet, most of the letters have similar capital and lowercase forms such as <M> and <m> or <O> and <o>. However, some of the letters have very different lowercase and uppercase forms such as <A> and <a> or <G> and <g>, and lowercase <a> even has multiple forms depending on the font or handwriting. While this does not change the function of the alphabet, it does not meet the boast of only twenty six letters slightly

[1]For definition of **matres lectionis**, see footnote in 3.9 on page 22

[2]**Digraph:** any sequence of two letters pronounced as a single sound (Crystal, *Encyclopedia* 445)

misleading. Should we consider the characters and diacritics separate letters? Or do we only consider the twenty six letters of the English alphabet to encompass the entire Roman alphabet?

7.3.2 Misleading similarities

Earlier, we saw that Cree only uses eleven symbols which invert and reflect across the vertical and horizontal axis to signify different vowels. The Roman alphabet has a set of four letters (<b>, <d>, <p>, and <q>) that exhibit this same inversion and reflection, yet all four letters vary widely in the sound that they represent.

Additionally, letters which have similar forms may not have anything to do with each other or one form could stand for multiple letters. Capital letters <E>and <F>differ by only one short line across the bottom. However, one represents a vowel and the other a consonant, and the two sounds do not contain sound similarities. In contrast, <i>, <I>, <l>, and <1> all look very similar yet they represent capital and lowercase versions of a vowel, a lowercase consonant, and a number.

7.4 Use of diacritics

Different languages use accent marks and diacritic marks differently. Swedish, for example, uses a twenty-nine letter alphabet with <å>, <ä>, and <ö> as separate letters while French and other Romance languages use accent marks on the letter <e>, such as <é> and <è> but the French consider these as variations on the letter <e>.

7.5 Reader assumptions

Perhaps the greatest difficulty of the Roman alphabet stems from people's prior knowledge of the alphabet. When a native English speaker approaches Chinese pinyin or Turkish, they recognize the forms and assume they can read it and pronounce it when in fact the letters and symbols stand for different sounds.

Chapter 8

Conclusion

The history of writing chronicles "a never-ending adaptation of a set of symbols that very much prefer to become fixed and fossilized" (Gnanadesikan, *Revolution*, 263). Depending on the unit of meaning or sound represented by the orthography, writing systems categorize as pictographs, syllabaries, alphabets, abjads, or alphasyllabaries. The relationship between the script and sounds may be shallow or deep, depending on the directness of the representation; pictographic systems envelop a wider range of changes in spoken languages. We begin our list of factors impacting script choice with the process and considerations involving in decipherment.

Full writing systems must include the grammar of the language and cover a range of contexts. We overviewed the origins and contexts of scripts from Sumerian cuneiform; Egyptian hieroglyphics; Chinese logograms; Maya glyphs; Linear B syllabary; Sanskrit akṣara; Phoenician abjad; Greek, Etruscan, Latin, and Slavonic alphabets; to Cherokee and Cree syllabaries. Beyond script creation, Chinese, Korean, Japanese, Turkish, Vietnamese, Shavian, Unspell, Azerbaijani, Mongolian, and Arabic illustrated how scripts adapt and change to challenges, giving insight to environmental elements that impact script choices.

Based on the patterns gathered from the language profiles, a preliterate society most likely will not create a new script or use a logographic system. The community using the language and political authorities must support the script for longevity. The preliterate society will choose the Roman alphabet due to historical and religious transmission of the script, colonization and other political events, and developing technology. The importance of technology grows day by day and the spread of internet communications will continue to augment the use of the Roman alphabet.

Despite the general acceptance of the Roman alphabet, it still faces challenges representing different languages. Only five letters often cannot represent the vowels of the language. Additionally, the claim of only twenty six letters does not include both capital and lower case letters, which sometimes match and sometimes do not. Unlike some other scripts such as Hangul or Arabic, letters which look like each other do not represent similar sounds as with ‹E› [i] and ‹F› [f]. As an increasing number of languages adopt this orthography, many letters and diacritics will no longer represent one set sound or feature but vary representation depending on the language.

With all this in mind, we predict the preliterate society will write their language down

using the Roman alphabet because of its availability. They must remember however, that the Roman alphabet has grown beyond simply twenty six letters A to Z. Rather than blindly using the letters as other languages use them, the society must consider the sounds of their language and adapt the letters to best represent those sounds.

Writing and Preliterate Societies

Cynthia Tung
Advisor: Professor MJ Connolly, Program in Linguistics

THESIS QUESTION

If a preliterate society decided to write their language down, what script will they choose and why?

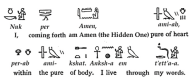
ABSTRACT

This paper explores the above question through an exposition on types of writing systems and a brief history of a few writing systems throughout the world. Societies sometimes invented new scripts, sometimes adapted existing ones, and other times used a combination of both.

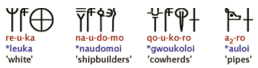
Based on these examples, I identify factors that influence the script decision including neighboring scripts, access to technology, and the circumstances of their introduction to writing. Much of the world uses the Roman alphabet and I present the argument that almost all preliterate societies beginning to write will choose to use a version of the Roman alphabet. However, the alphabet does not fit all languages equally well, and the paper closes out with an investigation into some of these inadequacies and how languages might resolve the issues.

SCRIPTS OF THE WORLD

HIEROGLYPHICS 3200'S BC



LINEAR B 2000'S BC

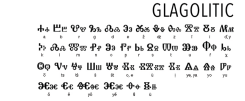


PHOENICIAN 1200 BC

ETRUSCAN 700 BC

MAYA 600 BC

GLAGOLITIC 800'S AD



CHEROKEE 1820'S AD

ᏣᎳᎩ ᎦᏬᏂᎯᏍᏗ
tsalagi gawonihisdi

SHAVIAN 1960 AD

3200'S BC CUNEIFORM

2000'S BC SANSKRIT

संस्कृतम्

1200 BC CHINESE

中文

800 BC GREEK

Ἑλληνική
Hellēnikḗ

700 BC ROMAN

400 AD ARABIC

اللغة العربية

1443 AD KOREAN

1830'S AD CREE

ᓀᐦᐃᔨᐣᐃᐧᐣ
nēhiyawēwin
(plains cree)

REASONS FOR THE ROMAN ALPHABET

- Many languages use the Roman alphabet, which provides easy access to the script
- Ease of learning, only has twenty six letters (not including diacritics)
- Technology facilitated and encourages the use of the Roman alphabet through typewriters, cell phones, and the internet

INADEQUACIES OF THE ROMAN ALPHABET

- Distinctive capital and lower case letters that do not resemble each other
- Limited number of symbols, especially for vowels
- Extensive use of diacritic marks to distinguish similar symbols

MAJOR REFERENCES

- Daniels, Peter T. / William Bright: *The world's writing systems* (1996)
- Gnanadesikan, Amalia E: *The writing revolution: cuneiform to the internet* (2009)
- Jensen, Hans: *Sign, symbol, and script: an account of man's effort to write* (1969)

Works Cited

Archibald, John:
Second language acquisition.
in O'Grady et al.: *Contemporary linguistics*, (2010). chapter 11, 391-428.

Baines, John / John Bennet / Stephen Houston:
The disappearance of writing systems.
(Equinox Publishing) Oakville CT, 2008
978-1-84553-013-6
(ONL P211.D57 2008)

Bauer, Thomas:
Arabic writing.
in Daniels: *The world's writing systems*, (1996). chapter 50, 559-568.

Bendor-Saumuel, John:
African languages.
in Daniels: *The world's writing systems*, (1996). chapter 59, 689-691.

Bennett, Jo Anne / John W. Berry:
Cree literacy in the syllabic script.
Literacy and orality, (1991). 90-104.

Bloomfield, Leonard:
The study of language.
(John Benjamins Publishing Company) Philadelphia, 1983
90-272-1892-7
(ILL P121.B49 1983)

Boltz, William G.:
Early Chinese writing.
in Daniels: *The world's writing systems*, (1996). chapter 14, 191-199.

Bonfante, Larissa:
The scripts of Italy.
in Daniels: *The world's writing systems*, (1996). chapter 23, 297-311.

Bright, William:
The Devanagari script.
in Daniels: *The world's writing systems*, (1996). chapter 31, 384-390.

Burnaby, Barbara / Marguerite Mackenzie:
Cree decision making concerning language: A case study.
Journal of multilingual and multicultural development, (2001). 191-209.

Campbell, George L.:
Handbook of scripts and alphabets.
(Routledge) New York, 1997.
0-415-13715-2
(ONL P211.C25 1997)

Campbell, Lyle:
Historical linguistics. An introduction.
(MIT Pr) Cambridge MA, 2013(3)
978-0-262-51849-9
(SEL LG 24.191)

Chadwick, John:
The decipherment of Linear-B.
(Cambridge UP) New York, 1958.
521-39830-4

Chen, May Jane / Joseph Chak-Kau Yuen:
Effects of pinyin and script type on verbal processing: comparisons of China, Taiwan, and Hong Kong experience.
International Journal of Behavior Development, (1991). 429-448.

Cheung, Him:
Pinyin and phonotactics affect the development of phonemic awareness in English-Cantonese bilinguals.
Reading development in Chinese children, (2003). 229-239.

Chung, Wei-Lun / Chieh-Fang Hu:
Morphological awareness and learning to read Chinese.
Reading and writing, (2007). 441-461.

Chung, Kevin K.H.:
Effective use of hanyu pinyin and English translations as extra stimulus prompts on learning of Chinese characters.
Educational psychology: An international journal of experimental educational psychology, (2010). 149-161.

Cleator, P.E.:
Lost languages.
(Mentor Books) New York, 1959.

Cook, Vivian / Jyotsna Vaid / Benedetta Bassetti:
Writing Systems Research: A new journal for a developing field.
Writing Systems Research, (2011) 1-3.

Comrie, Bernard:
Adaptations of the Cyrillic Alphabet.
in Daniels: *The world's writing systems*, (1996). chapter 60, 700-726.

Comrie, Bernard:
Languages of Eastern and Southern Europe
in Daniels: *The world's writing systems*, (1996). chapter 59, 663-689.

Comrie, Bernard:
Writing Systems.
2013.
in: Dryer, Matthew S. & Haspelmath, Martin (eds.): *The World Atlas of Language Structures Online.*
http://wals.info/chapter/141

Cooper, Jerrold S.:
Sumerian and Akkadian.
in Daniels: *The world's writing systems*, (1996). chapter 3, 37-57.

Coulmas, Florian:
The writing systems of the world.
(Basil Blackwell, Inc) Cambridge MA, 1989.
0-631-16513-4
(ONL P211.C27 1989)

Crown, Conrad J:
An alphabet for English-IV.
US Dept of Education, Educational Resources Information Center, (2000). 1-15.
http://files.eric.ed.gov/fulltext/ED438554.pdf

Crystal, David:
The Cambridge encyclopedia of language.
(Cambridge UP) New York, 2010(3).
978-0-521-73650-3

Crystal, David:
Language and the internet.
(Cambridge UP) New York, 2001
0-521-80212-1
(ILL P107.C78 2001)

Crystal, David:
Txtng: the gr8 db8.
(Oxford UP) New York, (2008).
978-0-19-957133-8
(ONL TK5105.73.C79 2008)

Cubberly, Paul:
The Slavic alphabets.
In Daniels: *The world's writing systems*, (1996). chapter 27, 346-355.

Daniels, Peter T. / William Bright:
The world's writing systems.
(Oxford UP) New York, 1996
0-19-507993-0
(ILL P211.W714 1996)

Daniels, Peter T.:
Methods of decipherment.
in Daniels: *The world's writing systems*, (1996). chapter 9, 141-59.

Daniels, Peter T.:
The study of writing systems.
in Daniels: *The world's writing systems*, (1996). chapter 1, 3-20.

de Voogt, Alexander J. / Joachim Friedrich Quack:
The idea of writing: Writing across borders.
(ebrary) 2012.
9004217002
http://site.ebrary.com.proxy.bc.edu/lib/bostoncollege/reader.action?ppg=139&docID=10518837
&tm=1412753736956

Diringer, David:
Writing.
(Praeger) New York, 1962 [3pr, 1967]
(SEL LS 22.23)

Dobrovolsky, Michael / William O'Grady:
Writing and Language.
in O'Grady et al: *Contemporary linguistics*, (2010). 545-567.

Dreze, Jean / Jackie Loh:
Literacy in India and China.
Economic and political weekly, (1995). 2868-2878.

Ehri, Linnea C.:
Learning to read words: theory, finding, and issues.
Scientific studies of reading, (2009). 167-188.

Fant, G.:
Feature analysis of Swedish vowels- a revisit.
STL-QPSR, (1983). 1-19.

Fischer, Steven Roger:
A history of writing.
(Reaktion Books) London, 2001.
1-86189-101-6
(ONL P211.F22 2001)

Gleason, H. A.:
An introduction to descriptive linguistics.
(Holt, Rinehart and Winston) New York, 1955.
0-03-010465-3
(ONL PE1135.G59 1961)

Gnanadesikan, Amalia E.:
The writing revolution: Cuneiform to the internet.
(Wiley-Blackwell) Malden MA, 2009.
978-1-4051-5407-9
(ONL P211.G58 2009)

Goerwitz, Richard L.:
The Jewish scripts.
in Daniels: *The world's writing systems*, (1996). chapter 46, 487-498.

Honeywill, Paul:
Visual Language for the World Wide Web.
(Intellect) Portland OR, 1999
1-871516-96-X
(ILL TK5105.888.H664 1999)

Hatcher, Lynley:
Script change in Azerbaijan: Acts of identity.
International journal of the society of language, (2008). 105-116.

Hockett, Charles Francis:
A course in modern linguistics.
(The Macmillan Company) New York, 1962
(ILL P121.H63)

Hornberger, Nancy H. / Kendall A. King:
Authenticity and unification in Quechua language planning.
Language, Culture, and Curriculum, (2010). 390-410.

Hu, Chieh-Fang / Hugh W. Catts:
The role of phonological processing in early reading ability: What we can learn from Chinese.
Scientific studies of reading, (2009). 55-79.

Janhunen, Juha A.:
Mongolian.
(John Benjamins Publishing Company) Philadelphia, 2012.
978-90-27273055
(eBrary)

Jensen, Hans:
Sign, symbol and script: An account of man's effort to write.
(Putnam's) New York, 1969.
(SEL LS 26.09)

Kaye, Alan S.:
Adaptations of Arabic script.
in Daniels: *The world's writing systems*, (1996). chapter 62, 743-762.

Kim-Renaud, Young-Key et al:
The Korean alphabet: its history and structure.
(U of Hawai'i P) Honolulu, 1997.
0-8248-1989-6
(ONL PL918.K67 1997)

Knight, Stan:
The Roman alphabet.
in Daniels: *The world's writing systems*, (1996). chapter 24, 312-331.

Lee, Chee Ha / Slava Kalyuga:
Effectiveness of on-screen pinyin in learning Chinese: An expertise reversal for multimedia redundancy effect.
Computers in human behavior, (2010). 11-15.

Lewis, Geoffery L:
The Turkish language reform: A catastrophic success.
(Oxford UP) New York, 1999.
978-0-19-925669-3
(ILL PL115.L47 1999)

Longacre, Robert E:
A tone orthography for Trique.
in Smalley: *Orthography Studies*, (1963). chapter 13, 132-137.

Maclean, John:
James Evans: Inventor of the syllabic system of the Cree language.
(Briggs) Toronto, 1890.

Mandel, Gabriele:
Japanese alphabet.
(Abbeville Press) New York, 2007.
978-0-7892-0959-7
(ILL PL528.M34 2008)

Marci, Martha J.:
Maya and other Mesoamerican scripts.
in Daniels: *The world's writing systems*, (1996). chapter 12, 172-182.

McBride-Chang Catherine / Hsuan-Chih Chen:
Reading development in Chinese children.
(Praeger) Westport, 2003.
0-89789-809-5
(ONL LB1573.M157 2003)

McLean, John:
James Evans: Inventor of the syllabic system of the Cree language.
(Methodist Mission Rooms) Toronto, 1890.
(ONL online)

Michalowski, Piotr:
Mesopotamian Cuneiform.
in Daniels: *The world's writing systems*, (1996). chapter 3, 33-36.

Moorhouse, A.C.:
The triumph of the alphabet: A history of writing.
(H. Schuman) New York, 1953.
(ONL P211 .M64)

Nichols, John D.:
The Cree Syllabary.
in Daniels: *The world's writing systems*, (1996). chapter 55, 599-611.

Ngo, Binh Nhu:
Elementary Vietnamese.
(Tuttle Publishing) Boston, 1999.
0-8048-3207-2
(ONL PL4373.N33 1999)

Nguyen Dinh-Hoa:
Vietnamese.
in Daniels: *The world's writing systems*, (1996). chapter 59, 691-695.

Noble, Holcomb B.:
William Smalley, 74, Linguist for the Hmong.
New York Times, (1997-12-26) 1-2.
http://www.nytimes.com/1997/12/26/world/william-smalley-74-linguist-for-the-hmong.html

O'Grady, William / John Archibald / Mark Aronoff / Janie Rees- Miller:
Contemporary linguistics.
(Bedford/St. Martins) Boston, 2010.
978-0-312-55528-3

Olson, David R. / Nancy Torrance:
Literacy and orality.
(Cambridge UP) New York, 1991.
0-521-39850-9
(ILL P211.L725 1991)

Orlov, Dmitry:
The Unspell teaching method.
unspell.it, (2014) 1-5.

Orlov, Dmitry:
Project Unspell.
unspell.it, (2015) 1-53.

Pae, Hye K.:
Is Korean a syllabic alphabet or an alphabetic syllabary?
Writing Systems Research, (2011) 103-115.

Pettersson, John Sören:
Numerical notation.
in Daniels: *The world's writing systems*, (1996). chapter 69, 795-806.

Quinn, Robert M.:
Beginner's Vietnamese.
(Hippocrene Books) New York, 1995.
0-7818-0411-6
(ONL PL4373.Q56x 1995)

Rim-Renaud / Young-Key:
The Korean alphabet: Its history and structure.
(University of Hawaii P) San Francisco, 1997.
0-8248-1989-6
(ONL PL918.K67 1997)

Ritner, Robert K:
Egyptian writing.
in Daniels: *The world's writing systems*, (1996). chapter 4, 73-84.

Scancarelli, Janine:
Cherokee writing.
in Daniels: *The world's writing systems*, (1996). chapter 53, 587-592.

Shaw, George Bernard:
The Shavian alphabet edition of Androcles and the lion.
(Penguin Books) London, 1962
(ILL PR5363.A15 1962)

Smalley, William A. et al:
Orthography studies: Articles on new writing systems.
(United Bible Studies) London, 1963.
(ONL P211.S6x 1963)

Smalley, William A.:
How shall I write this language?
in Smalley: *Orthography Studies*, (1963). chapter 5, 31-52.

Smalley, William A.:
The use of non-Roman script for new langauges.
in Smalley: *Orthography Studies*, (1963). chapter 8, 71-107.

Smalley, William A.:
Writing systems and their characteristics.
in Smalley: *Orthography Studies*, (1963). chapter 2, 1-17.

Smith, Warren S.:
First "Shaw alphabet" book.
The Shaw Review (1963). 34-36.
http://www.jstor.org/stable/40681984

Tachau, Frank:
Language and politics: Turkish language reform.
The Review of Politics (1964). 191-204.
http://www.jstor.org/stable/view/1405748

Taylor, Insup / M. Martin Taylor:
Writing and literacy in Chinese, Korean, and Japanese.
(John Benjamins Publishing Company) Philadelphia, 1995.
90-272-1794-7
(ONL PL1171.T37 1995)

Threatte, Leslie:
The Greek alphabet.
in Daniels: *The world's writing systems*, (1996). chapter 22, 271-280.

Vangelova, Luba:
How spelling keeps kids from learning.
The Atlantic (2015).
http://www.theatlantic.com/education/print/2015/02/how-the-english-language-is-holding-kids-back/385291/

Voorhoeve, J.:
Some problems in writing tone.
in Smalley: *Orthography Studies*, (1963). chapter 12, 127-131.

Wilford, John Noble:
Carvings from Cherokee script's dawn.
New York Times (2009-06-22), 1-3.
www.nytimes.com/2009/06/23/science/23cherokee.html

Wise, C. M. / Wesley Hervey:
The evolution of Hawaiian orthography.
Quarterly Journal of Speech, (2009). 311-325.

Wu, Jing / Guo Likun:
China to regulate use of simplified characters.
ChinaView (2009).
http://news.xinhuanet.com/english/2009-08/12/content_11871748.htm