2.0 Semitic Writing

Writing appears to have developed more or less simultaneously in the Tigris-Euphrates and the Nile valleys. Until recently, the earliest Egyptian texts dated from the end of the pre-dynastic period c. 3100 BCE. A 1998 find at monophthongisation, 400 km south of Cairo, has been dated at 3300 BCE. The earliest Mesopotamian texts are Sumerian tablets from Uruk, dating from the same period. (Sumerian is an apparent language isolate, spoken in a gradually contracting area until about the beginning of the second millennium BCE. It continued as a significant cultural and diplomatic language well into the first millennium BCE.)

2.1 Describing Semitic Writing Systems

Apart from their number and general shape, the form of the glyphs in a writing system is dependent on three not unrelated factors:

i. text genre
ii. writing materials
iii. direction of writing

In general, the more significant a text, the more careful the scribe or other artisan is likely to be about glyph form. A *formal* script style will be used for such texts, when the appearance of the result is more important than the time necessary to achieve that result. A more *cursive* script style develops for less significant texts, like notes or personal correspondence. A *ligatured* script style, in which the writing instrument seldom leaves the writing surface, develops in response to increased writing speed.

In the ancient Near East, stone, clay, pottery, hide, and papyrus were used as writing surfaces. Dressed stone was used for texts whose grandeur and permanence was of particular important, and for which a special formal *monumental* or *lapidary* script style might be developed. Somewhat less formal styles might be used for graffiti on stone. The unique shape of cuneiform glyphs is undoubtedly a function of the use of a reed stylus on damp clay as writing materials. The wedge-and-line form appears initially to have been a cursive development from a more pictographic style.

Potsherds with writing are termed *ostraca*, and were frequently used for short texts or notes. They might be inscribed with a sharp stylus, or used as a smooth surface for ink. Longer texts would be written in ink on dressed hide (*vellum*) or papyrus. The script style would vary from cursive to formal, depending on the genre. The oldest surviving texts tend to be rock graffiti, *stele* (stone dressed as a writing surface), clay tablets, or ostraca, since hide and papyrus deteriorate more rapidly.

Writing direction will influence glyph orientation, and slant and ligature in more cursive styles. All modern Semitic scripts are written right--to-left, with the exception of Ethiopic, which is left-to-right. Cuneiform was in general left-to-right. Direction in other scripts was often variable. For example, some epigraphic South Arabian texts are written left-to-right, other right-to-left, and others *boustrophedon* ('as the ox plows'), in which direction changes for successive lines.

2.2 Cuneiform

Sumerian glyphs were originally rather pictographic in form. With increased use, as writing materials, of damp clay impressed with a wedge-shaped stylus, the glyphs assumed the shapes characteristic of *cuneiform* (< Latin *cuneus* 'wedge') by c. 2500 BCE. It was about this same time that written materials in Akkadian begin to appear. The shapes of individual glyphs changed over time, culminating in the rather simplified style employed by court scribes in the Neo-Assyrian period (early first millennium BCE).

Cuneiform writing was adopted throughout Mesopotamia, the northern Levant (Syria), eastern Anatolia, and the western Persian plateau. Cuneiform texts appear in some dozen languages, including:

Language Isolates	Indo-European Languages	Semitic Languages
Sumerian	Old Persian	Akkadian
Elamite	Hittite	Ugaritic
Hurrian	Palaic	Eblaite
Urartian	Luwian	
Hattic		

Table 1
Languages Using Cuneiform

Cuneiform representations of Amorite (early NW Semitic) proper names and short phrases also appear in cuneiform in Akkadian texts.

All cuneiform glyphs are made up of some combination of five basic wedge shapes:

1 2 3 4 5

given here in modern sign-list order. In Akkadian, cuneiform glyphs had three functions/ interpretations:

i. Cuneiform glyphs could be *logograms*, standing for whole words. By convention, logograms in Akkadian are transliterated (typically in upper case) in their Sumerian readings.

ii. Cuneiform glyphs might be *syllabograms*, standing for V, VC, CV, or (more frequently in the Neo-Assyrian period) CVC prosodic units. A common assumption is that the initial syllabic reading of a cuneiform glyph derived from the phonology of the glyph's Sumerian logographic interpretation. For example, the glyph  represents the *logogram* AN 'sky' (Akk. *šamû*, Sumerian *an*) or the *syllabogram* /an/.

The same *syllabogram* may have multiple readings. This development would perhaps not be unexpected in cases like that of  BI 'its'[1] used for /bi/, /bé/, /pí/, /pé/[2], where the alternate readings are phonetically close. It is less clear how  came to be interpreted as both /aš/ and /rum/, or BE 'if' (Akk. šumma) as /be/, as well as /bad/, /bat/, /baṭ/ and /til/, unless these latter two logograms were also ambiguous in Sumerian.

iii. Cuneiform glyphs might be used as *determinatives*, a sort of orthographic classifier for nouns in Akkadian. Determinatives number only about a couple of dozen in Akkadian. For example, the glyph  (noted above as the *logogram* AN 'sky', Akk. *šamû*) is also the logogram DINGIR 'god' (Akk. *ilum*). It also represents the determinative dingir, used with names of gods. Thus  ^{d}UTU (= Šamaš, the sun god). Determinatives in Akkadian are typically transcribed to the left and in superscript. Thus, the preceding glyph combination would be transcribed ^{DINGIR}UTU.

Cuneiform scripts vanished as the major languages employing them, Sumerian, Hittite, and finally Akkadian, ceased to be used. A further development of cuneiform script, the Ugaritic cuneiform alphabet, will be considered below.

2.3 Egyptian Hieroglyphics

---

1 As a logogram, BI is used to represent the Akkadian 3rd singular possessive suffixes -vu 'his, its' and -va 'her, its' with inanimate (nonhuman) referents.

2 In syllabogram transcriptions, the diacritics are a means of indexing different syllabograms with the same reading; for example, the syllable /be/ has at least two representations in Akkadian: BE 'if' /be/ and BI 'its' (/pé/).

No Semitic language was ever written in hieroglyphics. In fact, the system was not widely exported from Egypt. The only direct derivative of hieroglyphic writing is believed to be the Meriotic alphabet, used in the Sudan between c. 300 BCE and 450 CE. Egyptian hieroglyphics may have had an influence, direct or indirect, on the development of the West Semitic alphabet from which all non-cuneiform Semitic writing systems (and most of the world's other alphabets) descend.

	ə	/ʁ/ > /ʔ/	vulture
	j	/j/ > /ʔ/	flowering reed
	jj y	/j/	two reeds
	o	/ʕ/	forearm
	w	/w/	quail chick
	b	/b/	foot
	p	/p/	stool
	f	/f/	horned viper
	m	/m/	owl
	n	/n/	water
	r	/r/	human mouth
	h	/h/	reed shelter
	x	/ħ/	twisted wick

	e	/x/	placenta
	h̲	/ç/	animal belly
	z	/z/	bolt
	s	/s/	folded cloth
	v	/ʃ/	pool, lake
	q	/q/	hill slope
	k	/k/	basket with handle
	g	/g/	stand for jar
	t	/t/	loaf
	t̲	/ʧ/	tether
	d	/d/	hand
	j	/ʤ/	snake

Table 2
Monoconsonantal Hieroglyphics

from Loprieno 1995:15

The Egyptian hieroglyphic (< Greek 'sacred writing') tradition lasted some 3500 years, from c. 3100 BCE to c. 200 CE. The system remained essentially the same throughout the period, though the number of signs varied (from c.1000 in the Old Kingdom, to 750 at the height of the classical period to perhaps thousands of signs in the Ptolemaic period). The glyphs remained largely pictographic in form throughout, though two cursive varieties did develop; a **Hieratic** (priestly) form documented throughout the history of the system, and a **Demotic** (popular) form appearing first in the 7th century BCE. With the exception of its use in writing the as yet incompletely understood Meroitic language, hieroglyphic

writing remained restricted to Egyptian throughout its history. It was ultimately replaced by a version of the Greek alphabet (termed Coptic), supplemented by some Demotic glyphs. **Coptic** remained in regular use until submerged by Arabic c. 1000 CE. It is still used in the Coptic Church.

The interpretative system for Egyptian hieroglyphics was similar to that of cuneiform. Signs had logographic, determinative, and phonological interpretations. The major difference was in the phonological interpretation; glyphs stood for *consonants* only; some monoconsonantal, others biconsonantal, and still others tri-consonantal. For example,  /b/,  /m-w/,  /m-w/. Phonetic complements were often used to force or reinforce particular readings of logographic or consonantal glyphs.

The existence of monoconsonantal glyphs gave rise to an emergent consonantary, in which all the consonantal phonemes of classical Egyptian (except /l/) had a distinct representation. /l/ was represented as by the glyphs for /n/ or /r/ or those two glyphs in sequence. (Historical /ʔ/ was also unrepresented; signs representing earlier /ʀ/ and /j/ came to represent /ʔ/ through phonetic change.)

2.4 West-Semitic Consonantary

The earliest alphabet is attested in a number of inscriptions from the Middle Bronze to the Early Iron Age (pre-1050 BCE) in Sinai, the Negev. and other sites in Palestine proper, as far north as Shechem. Most date from the 13th or 12 centuries BCE, but some have been dated as early as c. 1700 BCE.

The first Proto-Sinaitic[3] inscriptions were discovered by Flinders-Petrie in 1905, in a temple adjacent to turquoise mines at Serabit al-Kadim in the Sinai, and have been dated to 1700 BCE. The glyphs are mostly pictographic in form, resembling crude hieroglyphics.

The best known Proto-Sinaitic inscription is on a sandstone sphinx, deciphered by Alan Gardiner in 1916. Gardiner hypothesised that the language of the Proto-Sinaitic inscriptions was North-west Semitic, not Egyptian, and that the script was a consonantary like the West Semitic alphabet. He further hypothesised that the basis for the association between pictograph and phoneme was *acrophony*; each glyph represented the initial segment of the North-west Semitic word corresponding to the pictograph. For example,  'serpent' stands for the /n/ of NW Semitic *naḥal* 'serpent'. On this basis, Gardiner interpreted the sandstone sphinx inscription as /lbaʕalat/ 'for the goddess'.

The Proto-Sinaitic inscriptions are fragmentary and not well-preserved , so many glyphs

3

are difficult to identify. It appears that the direction of writing and, hence, the orientation of the glyphs is not constant. Many glyphs have variant forms.

Many investigators doubt that they are alphabetic at well. Daniels (1996:29) stresses that one must first accept all of Gardiner's premises, and notes that some of his reasoning might be flawed. Furthermore, in the case of the pivotal Serabit al-Kadim inscriptions, it is likely that any Canaanites at the turquoise mine site in the Middle Bronze were miners. One wonders how likely it is that a mine labourer in the Middle Bronze Age would have been able to inscribe a devotional inscription. (Daniels has fewer doubts regarding the West Semitic and alphabetic nature of the Proto-Canaanite inscriptions, particularly those from the early Iron Age.)

Albright (1966) assumes that the Proto-Sinaitic (and the Proto-Canaanite) alphabet must have had 27 letters, though he records only 24. The remaining 3, he hypothesises on the basis of:

i) his assignment of phonemic values to the extant glyphs, and
ii) his assumptions regarding the Northwest Semitic consonantal phonemic inventory at the time the alphabet appeared. The direction of writing was not fixed, so many graphemes have mirror-image alternates, as well as other variant forms.

Table 3 gives the Albright's (1966) grapheme/phoneme correspondences for Proto-Sinaitic.

Though there is disagreement regarding its relation to the Proto-Sinaitic inscriptions, there appears to be at least some weak consensus that a Proto-Canaanite consonant alphabet (consonantary) developed during the last half of the 2nd millennium BCE. According to Naveh (1987:10) and O'Connor (1996:88ff), Proto-Canaanite gave rise to two distinct alphabets: a northern form that developed into the Phoenician alphabet by c. 1100 BCE, and a southern form that developed into the scripts of the older north Arabian inscriptions, the South Arabian alphabet, and the Ethiopic scripts.

O'Connor (op. cit.) speculates that the Proto-Canaanite consonantary may have been the inspiration for the Ugaritic cuneiform alphabet. The Ugaritic alphabet was used in texts from the period 1450-1200 BCE. It is also attested in inscriptions found in Syria, Palestine, and on Cyprus. Though alphabetic, it is cuneiform-based. It has thirty graphemes (not including a word separator), but some hypothesise (Naveh 1987:30-31) that its original form had only twenty-seven, as in Albright's Proto-Canaanite reconstruction. Of the thirty with phonological value, three represent glottal stop followed by each of the three vowels /a/, /i/, and /u/.

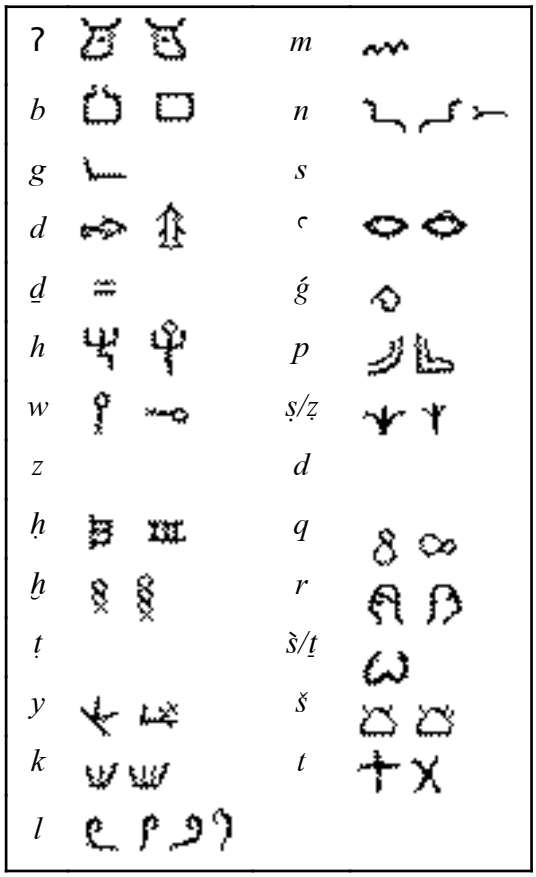


Table 3
Albright's Analysis of Proto-Sinaitic

The Ugarit grapheme order used in the Table 4 reflects an *abecedary* found at Ugarit in 1948. A second *abecedary* found at Beth Shemesh uses South Arabian ordering.

The Ugarit consonantary was little used outside of the city of Ugarit itself. Texts are mostly in Ugaritic, but some are in Hurrian and some in Akkadian. It was written left-to-right. A shorter version, with only twenty-two glyphs (corresponding to the twenty-two glyphs of the West Semitic consonantary) is found mostly outside Ugarit, and was written right-to-left.

The resemblance between the Ugarit and Proto-Sinaitic forms with similar values is accidental at best. Whether the Ugaritic alphabet was influenced by the Canaanite, or was a completely independent development, is purely conjectural.

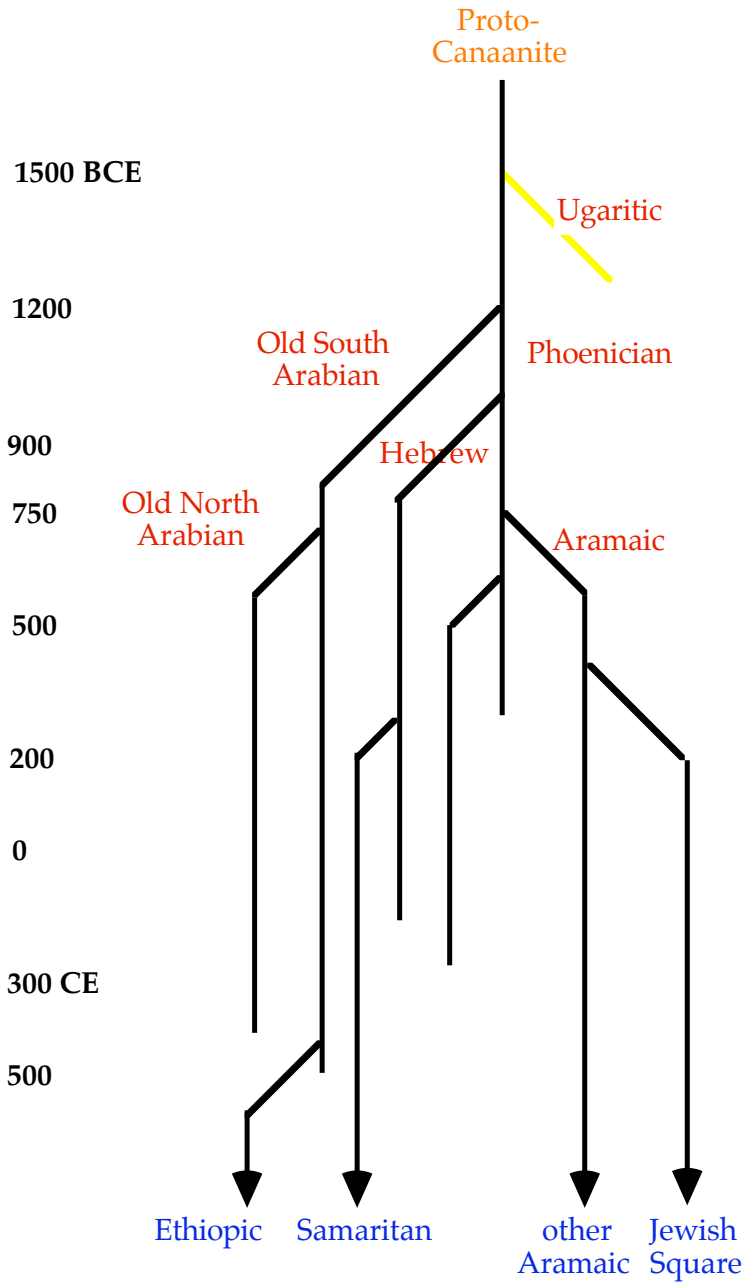


Figure 1
Some Descendants of the Proto-Canaanite Alphabet
after O'Connor (1996:89)

Proto-Sinatic	Ugaritic	TR	IPA
		*a*	ʔa
		b	b
		g	g
		ḫ	x
		d	d
		h	h
		w	w
		z	z
		ḥ	ħ
		ṭ	ṭ
		y	j
		k	k
		š	ʃ
		l	l
		m	m
		ḏ	ð

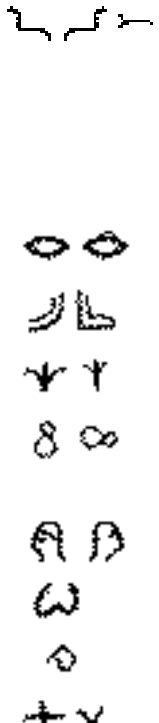
Proto-Sinatic	Ugaritic	TR	IPA
		*n*	n
		ẓ	ẓ
		*s*	s
		ʿ	ʕ
		p	p
		ṣ	ṣ
		q	q
		*r*	r
		ṯ	θ
		*ġ*	ɣ
		t	t
		i	ʔi
		u	ʔu
		*ś*	
		·	

Table 4
Ugaritic Alphabet (with Proto-Sinaitic Comparisons)

2.5 South Arabian Alphabet

The **South Arabian** alphabet has 29 graphemes. Table 5 gives them in South Arabian alphabetical order, left-to-right and top-to-bottom:

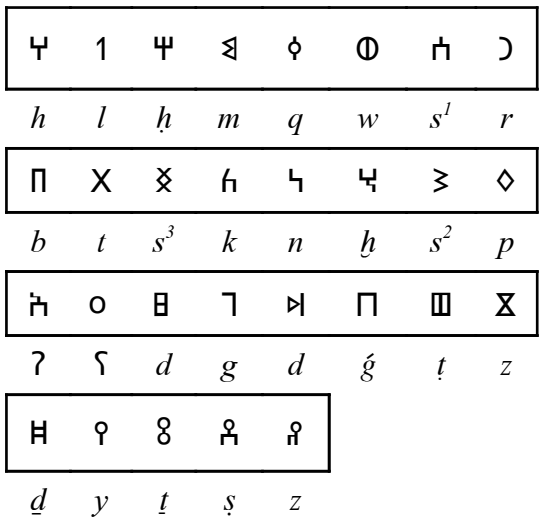


Table 5
Epigraphic South Arabian Alphabet

The transcriptions are those used in the literature on Epigraphic South Arabian. Note that s^1 corresponds to Proto-Semitic *ʃ, s^2 to PSEM *ɬ, and s^3 to PSEM *s.

South Arabian has two symbols more than the 27 postulated by Albright for Proto-Canaanite, a distinct ṭ and a distinct ẓ. In addition, it appears that some symbols have South Arabian values different from the Proto-Canaanite values of their antecedents. Thus, the symbols for *ś* and *š* have apparently swapped values in South Arabian, while the symbol for the glide *w* has taken on the value of the glide *y* (and a new symbol innovated for *w*). Similarly, a variant form of Proto-Canaanite *h* has taken on the value of *ḫ*, replacing the Proto-Canaanite original. Following Albright, these modifications are diagrammed in Table 6 below.

The only surviving descendent of the South Arabian alphabet is the Ethiopic alphabet used in Ge'ez and, with modifications, in modern Ethiopic languages like Amharic. Of the 26 basic letters of the Ethiopic (Ge'ez) alphabet, 24 are derived directly from South Arabian. The letters representing the plain and velarised bilabial stops /p/ and /p̊/, ፐ and ጰ, respectively, appear to be derived from Greek Π and Ethiopic ጸ /ṣ/. Four additional letters, representing the labialised velars /k^w, x^w, g^w, q^w/ are derived from letters representing the corresponding non-labialised velars.

TR	Pr-CA	ESA	Phoe
ʔ			
b			
g			
d			
ḏ			
h			
w			
z			
ḥ			
ḫ			
ṭ			
ẓ			
y			
k			
l		1	

TR	Pr-CA	ESA	Phoe
m			
n			
s			
ʿ			
ǵ			
p			
ṣ			
d			
q			
r			
š			
s			
t			
ṯ			

Table 6
Proto-Sinaitic, ESA, and Phoenician Compared

In Ethiopic order, the basic form of the letters is as follows (where the letters on the lower line are labiovelars):

h l ḥ m s^2 r s^1 q b t x n ʔ k w ʕ z j d g ṭ ṗ ṣ ẓ f p

q^w x^w k^w g^w

In Ge'ez reading pronunciation, the voiceless velar, pharyngeal, and glottal fricatives have fallen together as /h/, the voiced pharyngeal fricative, and glottal stop as /ʔ/, and the glottalised dental fricatives as /s^2/. Of the South Arabian *letters* s^1, s^2, and s^3, only the

former two are reflected in South Arabian. Ge'ez ሠ reflects PSEM *ɬ (Arabic ʃ). In Ge'ez reading, they fall together as /s/.

Table 7 relates corresponding South Arabian and Ethiopic (Ge'ez) basic letters, using South Arabian alphabetical order:

S Arabian	𐩣	𐩡	𐩦	𐩤	𐩧	𐩠	𐩵	𐩪	𐩫	𐩱	𐩬	𐩲	𐩮	𐩯	𐩭
	h	l	ḥ	m	q	w	s^1	r	b	t	s^3	k	n	x	s^2
Ethiopic	ሀ	ለ	ሐ	መ	ቀ	ወ	ሰ	ረ	በ	ተ		ከ	ነ	ኀ	ሠ
	h	l	ḥ	m	q	w	s^1	r	b	t		k	n	x	s^2

S Arabian	𐩨	𐩴	𐩰	𐩶	𐩲	𐩺	𐩸	𐩹	𐩷	𐩳	𐩼	𐩻	𐩽	𐩾	
	p	ʔ	ʕ	ḍ	g	d	ǵ	ṭ	z	d̲	j	ṯ	ṣ	ẓ	
Ethiopic	ፈ	አ	ዐ	ፀ	ገ	ደ		ጠ		ዘ	የ		ጸ		
	f	ʔ	ʕ	ẓ	g	d		ṭ		z	j		ṣ		

Table 7
ESA and Ethiopic Letter Correspondences

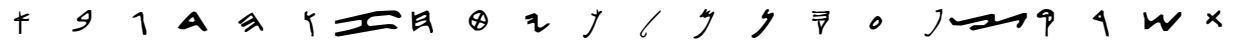
The representation of vowels in Ge'ez orthography, and the form of the seven additional letters used to represent palatals in Amharic, are considered in section 2.7 below.

2.6 West Semitic

The earliest examples of the Phoenician (West Semitic) alphabet are inscriptions from the 11th century BCE, such as the Ahiram sarcophagus inscription from Byblos, c. 1000 BCE. The following alphabet, from the Karatepe inscription (Cilicia, 8th century BCE), is representative:

The oldest Hebrew language inscription is the Gezer calendar (10th century BCE), whose script is still decidedly Phoenician in character. A distinct southern (Hebrew) variant, with left curving glyphs, began to emerge in the 9th century. Its oldest exemplar is the Moabite Mesha stele (c. 850):

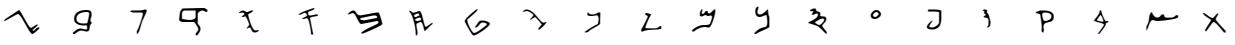
Somewhat later versions of the script are exemplified in the Siloam tunnel inscription (c.700 BCE):



and in the Lachish ostraca (6th century):

A typically Aramaic form of the Phoenician alphabet begins to appear c. 750 BCE, with heavier horizontal and lighter vertical strokes (compared to heavier descending or rightward strokes of Hebrew and Phoenician), and a two-bar *ħet*. The following is a rather late version, from the Elephantine papyri (c. 400 BCE):

Phoenician proper survived in Punic and Neo-punic (and, of course, in Greek and its descendants). The Punic script died out in the early centuries of the Common Era. Hebrew script was largely replaced by Aramaic. It was revived for coins (and other political purposes) in the Hasmonean period and survives in the Samaritan script:



Otherwise, only the Aramaic versions of West Semitic have modern descendants outside of Europe.

2.7 Current Semitic Scripts

The modern Ethiopic script used for languages like Amharic is a direct descendent of the script used for Ge'ez, itself derived from the South Arabian alphabet. The only modification in Amharic is the addition of a number of symbols for palatalised consonants. These are derived from their non-palatalised counterparts by the addition of an upper horizontal line:

Base	ሰ	ተ	ነ	ከ	ዘ	ደ	ጠ	ቀ	በ	ገ
IPA	s	t	n	k	z	ʤ	ʧʾ	q	b	g
Palatal	ሸ	ቸ	ኘ	ኸ	ዠ	ጀ	ጨ	ቐ	ቨ	ጘ
IPA	ʃ	ʧ	ɲ		ʒ	ʤ	ʧʾ		v	

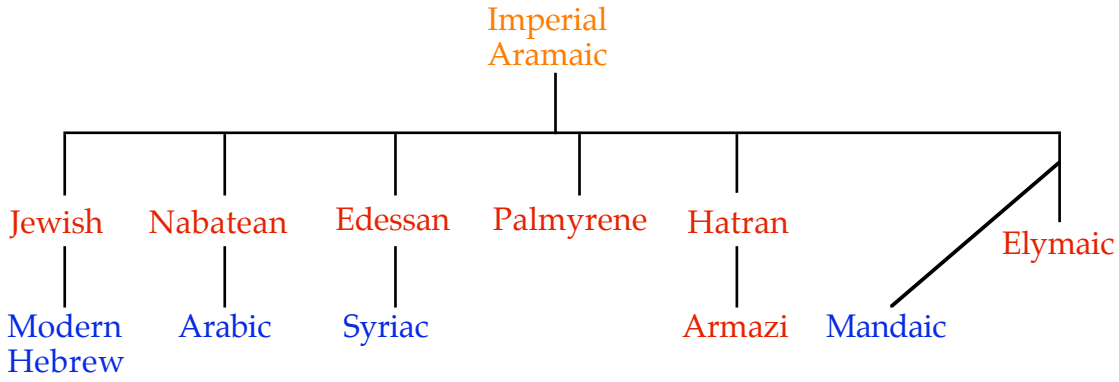
For /ʒ/, the horizontal is broken, while /ʧʾ/ has an esoteric form.

With the exception of Ethiopic and Samaritan[4], all scripts currently used for Semitic languages are descended from the Aramaic script in use in the Persian Empire, itself a

4 Ethiopic and Samaritan scripts are descendents of Proto-Canaanite, if that script did in fact exist.

descendent of the 22 letter NW Semitic (Phoenician) consonant alphabet. The script used in the Elephantine papyri[5] exemplifies this period.

Regional variants of the Aramaic alphabet began to develop after the fall of the Persian Empire (c. 330 BCE). Those used for Aramaic language materials include Jewish, Nabitean, Palmyrene, Hatran, Edessan, and Mandaean.



An early stage of the Jewish (Square Aramaic) script (Table 8, col. 4) now used for Hebrew is exemplified by the Dead Sea Scroll script (Table 8, col. 2). Nabatean script (Table 8, col. 3) developed in the Arabic-speaking trading centre of Petra. It was used for writing in both Aramaic and Arabic. A particularly cursive form of Nabatean developed into the modern Arabic alphabet.

As more cursive variants of Aramaic glyphs developed, the sharp downstrokes of many glyphs began to be curved more to the left. This change led to the development of distinct medial and final forms for some graphemes. As ligatured glyphs forms evolved in scripts like Nabatean/Arabic and Syriac, some distinct non-ligatured word-initial, word-final or isolate glyphs also developed.

Two problems faced those who used Nabatean script to write Arabic. The first problem, shared by Syriac, was that in increasingly cursive forms of Aramaic script, many graphemes began to fall together. Syriac dealt with this problem by introducing diacritics -- see below. The second, uniquely Arabic problem was the need to exploit the 22 graphemes of the Aramaic script to distinguish the 28 consonant phonemes of Arabic. Diacritics were exploited to deal with this issue too.

Nabatean ב /b/ and ת /t/ fell together as Arabic ب /ت; they were then distinguished by

5 The Elephantine papyri are part of the archive of a Judean military colony on Elephantine Island (near modern Aswan) in Upper Egypt. The garrison may have been established as early as the early 6th century BCE, but the documents are from the 5th and 4th centuries BCE.

diacritics. In their ligatured forms (shown here with a following *alif* ا), these are بـا and تـا, respectively. The same glyph was also used for /θ/: ث. Ligatured /n/ and /y/ also joined this set: نـا and يـا, though they kept distinct unligatured forms: ن and ي, respectively. Nabatean 𐢛 /g/ and 𐢕 /ħ/ fell together as Arabic حـ in medial form. This form, disambiguated by diacritics, is used for Arabic /ʤ/, /ħ/, and /ɣ/: in medial form, خـ, حـ, and غـ, respectively. Nabatean 𐢇 /p/ and 𐢑 /q/ also fell together in non-final form as فـ and قـ; in isolated and final forms they remain distinct: ف and ق.

1	2		3	4	
𐡀	𐤀		𐡠	א	
𐡁	𐤁		𐡡	ב	
𐡂	𐤂		𐡢	ג	
𐡃	𐤃		𐡣	ד	
𐡄	𐤄		𐡤	ה	
𐡅	𐤅		𐡥	ו	
𐡆	𐤆		𐡦	ז	
𐡇	𐤇		𐡧	ח	
𐡈	𐤈		𐡨	ט	
𐡉	𐤉		𐡩	י	
𐡊	𐤊	𐤋	𐡪	כ	ך
𐡋	𐤋		𐡫	ל	
𐡌	𐤌	𐤍	𐡬	מ	ם
𐡍	𐤍	𐤎	𐡭	נ	ן
𐡎	𐤎		𐡮	ס	
𐡏	𐤏		𐡯	ע	
𐡐	𐤐	𐤑	𐡰	פ	ף
𐡑	𐤑		𐡱	צ	ץ
𐡒	𐤒		𐡲	ק	
𐡓	𐤓		𐡳	ר	
𐡔	𐤔		𐡴	ש	
𐡕	𐤕		𐡵	ת	

1 **Elephantine Papyri c. 400 BCE**

2 **Dead Sea Scrolls c. 100 BCE**

3 **Nabatean c. 50 BCE**

4 **Modern Hebrew**

Table 8: Imperial Aramaic and Some Early Descendants

The standard alphabetical order in Arabic (reading right-to-left, of course) is as follows:

ي و ه ن م ل ك ق ف غ ع ظ ط ض ص ش س ز ر ذ د ج ح ج ث ت ب ا

j w h n m l k q f ɣ ʕ ʒ ṭ ḍ ṣ ʃ s z r ð d x ħ ʤ θ t b ʔ

In the Arabic of the Maghreb (Morocco, Algeria, Tunisia), another order is sometimes used, one that is closer to the order of the source graphemes in the Aramaic alphabet:

ي و ه ق ش س ف غ ع ض ص ن م ل ك ظ ط ز ر ذ د ج ح ج ث ت ب ا

j w h q ʃ s f ɣ ʕ ḍ ṣ n m l k ʒ ṭ r z ð d x ħ ʤ θ t b ʔ

The Arabic alphabet has two additional glyphs. The ligatured lam-alif /lʔ/ ﻻ, is also found in some Syriac scripts. The *ta marbuta* ة is an /h/ grapheme ه with the diacritic appropriate for /t/. It only occurs word-finally, and is pronounced as /at/ except in Qur'anic pausal forms, when it is /a/.

The Nabatean Aramaic sources of Arabic glyphs are shown in Table 9.

The Syriac language and Syriac script are said to have developed in and around Edessa, at the beginning of the Common Era. (Syrian Christian tradition holds that the Syriac translation of the Bible, the Peshitta, was done in Edessa c. 200 CE.) There are three main versions of the script. The most conservative is the Estrangelo (< Gr. στραγύλη 'round'). A schism in the Syrian Christian church, as well as the political division of the community between Roman and Persian spheres of influence, led to the emergence of two later Syriac scripts. Jacobite or Serto ('simple') script developed in the west in the 8th century CE. Nestorian script developed in the east, maturing in the 12th century CE. Nestorian is currently used for writing the modern Aramaic koiné.

In Syriac scripts we observe two novel features of late Semitic writing systems. Increasing cursiveness has led to some glyph forms becoming identical; in Syriac, this is true of the graphemes for /r/ and /d/, which are distinguished by a diacritic dot above and below the glyph, respectively. Syriac also has a number of special ligatured forms for letter combinations. The combination /tʔ/ has special forms in both Estrangelo and Nestorian: ﭏ and ﭐ respectively. Serto and Nestorian have a special form for /lʔ/, in Nestorian: ﭑ. Nestorian also has a form for /hi/ ﭒ.

Nabatean[6]	Arabic			IPA
	Isolate/Final	Medial	Initial/Medial	
𐢀1	ا			ʔ
𐢁2	ب		بـ	b
𐢂22	ت		تـ	t
	ث		ثـ	θ
𐢃10	ي		يـ	j
𐢄14	ن		نـ	n
𐢅3	ج		جـ	dʒ
𐢆8	ح		حـ	ħ
	خ		خـ	x
𐢇4	د			d
	ذ			ð
𐢈5	ه	ـهـ	هـ	h
𐢉6	و			w
𐢊9	ط			ṭ
	ظ			ẓ
𐢋11	ك		كـ	k
𐢌12	ل		لـ	l
𐢍13	م		مـ	m
𐢎15				[s]
𐢏16	ع	ـعـ	عـ	ʕ
	غ	ـغـ	غـ	ɣ
𐢐17	ف	ـفـ	فـ	f [p]
𐢑19	ق	ـقـ	قـ	q
𐢒18	ص	ـصـ	صـ	ṣ
	ض	ـضـ	ضـ	ḍ
𐢓20	ر		ر	r
𐢔7	ز		ز	z
𐢕21	س	ـشـ	شـ	s [ʃ]
	ش	ـسـ	سـ	ʃ

Table 9: Arabic Graphemes Grouped by Form, with Nabatean Sources

6 Subscripts on the Nabatean letters indicate NW Semitic alphabetical order.

Table 10 gives the Estrangelo, Nestorian and Mandaic alphabets:

Estrangelo		Nestorian			Mandaic		IPA
	Final		Final	Isolated Final		Final	
ܐ		ܐ			ࡀ	ࡐ	ʔ
ܒ		ܒ			ࡁ		b
ܓ		ܓ			ࡂ		g
ܕ		ܕ			ࡃ		d
ܗ		ܗ			ࡄ		h
ܘ		ܘ			ࡅ		w
ܙ		ܙ			ࡆ		z
ܚ		ܚ			ࡇ		ħ
ܛ		ܛ			ࡈ		ṭ
ܝ		ܝ			ࡉ		j
ܟ[1]	ܟ	ܟ	ܟ		ࡊ	ࡊ	k
ܠ		ܠ			ࡋ		l
ܡ	ܡ	ܡ	ܡ		ࡌ		m
ܢ	ܢ	ܢ	ܢ	ܲ	ࡍ	ࡍ	n
ܣ	ܤ	ܣ	ܣ		ࡎ		s
ܥ		ܥ			ࡏ		ʕ
ܦ		ܦ			ࡐ	ࡐ	p
ܨ		ܨ			ࡑ	ࡑ	ṣ
ܩ		ܩ			ࡒ		q
ܪ		ܪ			ࡓ		r
ܫ		ܫ			ࡔ		ʃ
ܬ		ܬ			ࡕ		t

[1] non-final forms shown with following *yud* /j/ .

Table 10
Syriac and Mandaic Scripts

Mandaic has a special glyph ࡖ for the high frequency relative clause marker *d-* ([də] or

[ad]), and a set of combined letter pairs. There is also a second glyph ࡀ for *halqa* (as *alef* is called in Mandaic), used where *halqa* functions as *mater lectionis* (see section 2.9.2.0). The Arabic letter ayin ع has been introduced for writing Arabic loans (since the corresponding Mandaic letter is used only as *mater lectionis*). Other Arabic and Farsi sounds not represented in the Mandaic alphabet are marked in loans by adding two horizontal dots under seven Mandaic letters.

2.8 Word Boundaries

Word boundary is marked in the orthography of most Semitic languages. Akkadian cuneiform (inconsistently) and Ugaritic used a vertical wedge 𐎗 as a word separator. South Arabian and Ethiopic employ a colon **:**; Early Hebrew and Aramaic often employ a dot. Later texts employ whitespace.

2.9 Representing Vowels

2.9.1 Vowels in Ethiopic

A notation for vowels is attested in Ge'ez inscriptions from c. 400 CE. The system is grounded in a set of base letters (called *gə'əz*, as in the name of the language), which are modified in fairly consistent ways to identify a following vowel. The vowel system is as follows (with phonetic values as in Amharic):

i u
ə
e ɜ o
a

Figure 3: Ge'ez Vowels

The base letter represents the sequence Cɜ; syllable-final consonants are represented as Cə. The high vowels /i/ and /u/ are represented by horizontal lines attached to the mid and lower right, respectively, of the base character. The mid vowel /e/ and /o/ are marked by circles or upward curves attached to the top and bottom right, respectively. But, with the exception of ሎ /lo/, if the letter 'stands on two or more legs', the second vertical from the right is extended. In practice, the result usually appears like a shortening of the other vertical; for example, ሰ /sɜ/ ሶ /so/. And some Co letters, like ፖ /po/ (cf. ፐ /pɜ/) have unpredictable forms.

The low vowel /a/ is represented by an extended rightmost vertical, curved if the letter has only a single vertical; for example, ላ /la/ ዛ /za/, but ጋ /ga/. ራ /ra/, ና /na/ and the latter's

derivative 𐡍 /ña/ have esoteric forms. Cə is quite variable in form. Often it is a horizontal, sometimes left (𐡅 /wə/) and sometimes right (𐡍 /nə/). Many Cə forms, like 𐡋 /lə/, are unpredictable.

The vowel sequence for /b/ is regular and representative:

𐡁	𐡁𐡅	𐡁𐡉	𐡁𐡀	𐡁𐡄	𐡁𐡄	𐡁𐡅
bɜ	bu	bi	ba	be	bə	bo

2.9.2.0 Vowels in West Semitic

The West Semitic alphabet is a consonantary; its 22 graphemes representing consonants only. As adapted for Aramaic and Hebrew, however, the glides (*yod* and *waw*) and the glottals (*he* and *alef*) were adapted to represent vowels. Final vowels were represented quite regularly by these *matres lectionis* (mothers of reading), but only long word-internal vowels were so represented, and then only inconsistently. Word-internal short vowels were not marked by a mater lectionis.

Yod and *waw* were used for non-low vowels, the former for front, and the latter back. *He* and *alef* were used for low vowels; Hebrew tended to favour the former and Aramaic the latter.

Matres lectionis were never used in Phoenician, though Punic and Neo-Punic did make use of them. Arabic came to use mater lectionis consistently to represent long vowels, both word-internally and word-finally, but did not use them for short vowels. Unique amongst Semitic languages, Mandaic employs *scriptio plena* (that is, matres lectionis) for all vowels except the reduced vowel, which is not indicated. Round vowels are indicated by *ushenna* (corresponding to waw), front non-low vowels by *aksa* (corresponding to yod) or *ayin* (after those left-ligated letters extending below the line), and low vowels by *halqa* (*alef*).

2.9.2.1 Pointing

Diacritical points were first introduced into Syriac script to distinguish /ܕ/ d from /ܪ/ r. The concept was extended to distinguish homographs on the word level. Two dots (called syame) were placed over most plural nouns and some feminine plural verbs:

ܡܠܟܐ	/malka/	'king'	ܡܠܟܶܐ	/malke/	'kings'
ܒܝܬܐ	/bayta/	'house'	ܒܝܬܶܐ	/batte/	'houses'
ܟܬܒܘܢ	/kt̲ab̲un/	'they (m) wrote'	ܟܬܒܝܢ	/kt̲ab̲in/	'they (f) wrote'

A single dot subscript marked all perfective verbs (as above), except the 1s, which carried a superscript dot:

3sm	ܟܬܰܒ	/kṯaḇ/	'he wrote'
2sf	ܟܬܰܒܬ	/keṯbaṯ/	'she wrote'
2sm	ܟܬܰܒܬ	/kṯaḇt/	'you (sm) wrote'
2sf	ܟܬܰܒܬܝ	/kṯaḇt/	'you (sf) wrote'
1s	ܟܬܰܒܬ	/keṯbeṯ/	'I wrote'

and as a means of distinguishing masculine singular active participles from 3sm perfective verbs: ܟܬܰܒ /kṯaḇ/ 'he wrote', BtK[/kāṯeḇ/ 'writing (sm)'. The single dot convention seems to have begun as as means of distinguishing isolated common homographs like: ܗܰ /haw/ 'that', ܗܽ /hu/ 'he'; ܡܰ /man/ 'who', ܒܶ /men/ 'from'. A supra- or sublinear line (linea ocultans) is used to indicate some assimilated or lost consonants that are preserved in the orthography: ܐܰܬܬܬܐ /ʔattəta/ 'woman', ܗܘܝ /way/ 'were (pf)'. These conventions predated the East-West schism, since they are employed in both Jacobite and Nestorian scripts.

Two distinct systems, a *Nestorian* (Eastern) and a *Jacobite* (Western), were developed for indicating vowel quality more directly. The Eastern system involves an extension of the diacritics, and was developed between 700-900 CE. The Western system involves superscripted Greek vowel letters, and appeared c. 1000 CE. The are exemplified in Table 11 with the consonant *beṯ*:

Jacobite	value	Nestorian	value
ܒ݂	/bi/	ܒܝ	/bî/
		ܒܹ	/bē/
ܒ݁	/be/	ܒܶ	/be, bi/
ܒ݈	/ba/	ܒܰ	/ba/
ܒ݀	/bɔ/	ܒܳ	/bā/
		ܒ݂	/bo, bō/
ܒ݁	/bu/	ܒ݂	/bu, bū/

Table 11
Vowels in Syriac

These vowel pointings were never part of the regular orthography, but served a purely optional mnemonic function. The Western vowel points could be written either above or below the letters. Only vowel quality, not quantity, was distinguished in this system. (The bF diacritic, indicating a historical long a, is pronounced /ɔ/.) The Eastern tradition did not distinguish length for round vowels (which were written with mater lectionis). Length was

distinguished for non-round vowels, though long /i/ was always marked by mater lectionis.

Fully pointed Syriac texts, both Nestorian and Jacobite, also distinguish stop and fricative allophones of the *begad kefat* consonants (see below) by superscripts for the former and subscripts for the latter. These are not marked as consistently as the corresponding Hebrew *dagesh*, though.

Similar systems of diacritic vowel pointing developed for **Hebrew** roughly simultaneously, and probably influenced by, the Syriac developments. Three different systems are extant in Jewish texts: the Palestinian, the Babylonian, and the Tiberian. The Palestinian and the Babylonian are attested in simple and more complex/fuller versions. Only the Babylonian and the Tiberian systems became widely used, and the Tiberian ultimately emerged as a de facto standard.

The vowel system described by the Babylonian pointing distinguishes six vowel phonemes; that described by the Palestinian and Tiberian pointing distinguish seven (in all cases, ignoring the reduced vowel *shwa*). The phonetics of some of these contrast is a matter of dispute, and will be considered in Chapter 3. The phonetic symbols and positions in the following figures should not be viewed as more than mnemonic at this juncture. Ignoring *shwa* (and other reduced vowels), the Tiberian and Simple Babylonian vowel systems are as in Table 12:

Simple Babylonian		Tiberian		
בִ	בֹ	בִ		בֻ
בֵ	בֺ	בֵ		בֹ
בֶ	בָ	בֶ	בַ	בָ
		ħireq		qibbuṣ
		ṣere		ħolɛm
		sɛgol		pataħ
			qameṣ	
/bi/	/bu/	/bi/		/bu/
/be/	/bo/	/be/		/bo/
/ba/	/bɔ/	/bɛ/	/ba/	/bɔ/

Table 12

Hebrew Vowel Pointing (ignoring reduced vowels)

When the non-low round vowels are marked by mater lectionis *waw* in Tiberian pointing,

the pointing appears on the *waw*: בוּ[7] [bu], בֹּו [bu].

Babylonian pointing, simple and complex, has a single symbol *shwa* used for the reduced vowel בֺ [bə]. Alongside Tiberian shwa: בְּ [bə] are three ultrashort vowels: אֱ [ʔĕ] אֲ [ʔă] אָ [ʔŏ], termed ḥaṭ*ef segol*, ḥaṭ*ef pataḥ*, and ḥaṭ*ef qame*ṣ, respectively, corresponding to the three low vowels of the Tiberian system. These appear with guttural consonants in environments when one would otherwise expect *shwa*.

In Tiberian Hebrew, *shwa* is ambiguous[8]. It marks either the reduced vowel [ə] or the absence of a vowel (typically on the coda of an internal closed syllable). The two functions of *shwa* are distinguished as *mobile* (or vocalic) and *quiescent* shwa (in Hebrew שְׁוָא נָע and שְׁוָא נָח), respectively. Compare: יִגְדַּל [yig.dal] 'he is great' גָּדְלוּ [gā.d̲ə.lu^{w}] 'they were great'. Word-final consonants do not take shwa, except in two cases: word-final kaf (מֶלֶךְ /mɛlɛk̲/ 'king'), and word-final *taw* as the 2sf suffix (perfective) conjugation marker, when itself preceded by quiescent shwa (כָּתַבְתְּ /kāṯab̲t/ 'you (fs) wrote').

The vowel qameṣ בָּ (represented in Table 12 as /כ/) is also ambiguous, though the ambiguity is reflected as a vowel quality contrast only in Sephardi (and standard Israeli) pronunciation. Its usual reading is as identical to the *pataḥ* בַ , where it is assumed to reflect /ā/. In some closed unstressed syllables it is pronounced /o/ (as in כָּל /kol/ 'all'). There it is assumed to have represented a short mid back vowel, perhaps [ɔ] (< PSEM *u). With that value, qameṣ is called *qame*ṣ ḥaṭ*uf*. In Ashkenazic (European) pronunciation, both *qame*ṣ and *qame*ṣ ḥaṭ*uf* are pronounced as short /o/.

The Hebrew pointing systems also employ a number of other diacritics. Tiberian *dagesh* (בּ) has two interpretations. It marks:

i. gemination for those (non-initial) consonants that can occur geminate, or
ii. stop (rather than fricative) articulation for oral occlusives[9] (one of the so called *begad kefat* consonants, oral stops that spirantised post-vocalically in both Hebrew and Aramaic.)

The corresponding Babylonian diacritic *digsha* marks only spirantisation, and is not used consistently. Gemination is marked in complex Babylonian pointing, again inconsistently, by a vertical bar accompanying the vowel point, and indicating that the consonant

---

7 Tiberian ו is termed shureq.

8 The same ambiguity is found in Complex Babylonian pointing, but quiescent *shwa* appears less consistently there than in Tiberian Hebrew.

9 The *begad kefat* consonants {ב ג ד כ פ ת}, /b, g, d, k, p, t/, which spirantise postvocalically in Hebrew and Aramaic.

following the vowel is geminate.

In Tiberian pointing, there is a little used diacritic called *rafe*, the converse of dagesh in the latter's occlusion interpretation. Thus, בּ indicates [b] and בֹּ [v]. The corresponding Babylonian *qipya* is also rare.

In some pointing systems, the difference between consonantal and vocalic interpretations of the consonants {א ה ו י}, commonly used as mater lectionis, is sometimes indicated. In Tiberian pointing, this is commonly done only for word-final consonantal ה, which takes an internal dot (identical to *dagesh*) called *mappiq*. (A *vocalic* value for א and ה is very occasionally marked by *rafe*.)

Babylonian pointing sometimes used *digsha* for consonantal א and final consonantal ה. The latter is also sometimes found with a ה superscript. Consonantal י and ו are sometimes doubled. Vocalic א is occasionally found with *qipya* in Babylonian pointing.

In fully pointed Tiberian texts, and sometimes in Babylonian and Palestinian pointed texts, stress is marked by a system of points called either טְעָמִים *ta'amim* or נְגִינוֹת *neginot*. The former term refers to their function as markers of stress and constituency; the latter for the function of the same points as musical notes used in chanting the texts. For a general discussion of Tiberian stress marking[10], see Aronoff (1985), Joüon (1991:61-69), and Dresher (1994).

Fully pointed Tiberian texts also employ a sublinear perpendicular stroke called *meteg* 'bridle', placed to the left of a vowel point. It usually appears in unstressed open syllables in a variety of functions (see Joüon 1991:59-61). In the pair חָכְמָה 'she is wise' vs. חָכְמָה 'wisdom', it marks the fact that the vowel of the initial syllable in 'wisdom' is *qameṣ ḥaṭuf*.

The *maqqef* is a hyphen joining words that form a single stress group: אֶל כָּל־יִשְׂרָאֵל [ʔɛl kol yiśraʔel] 'to all Israel'.

Classical and Standard Arabic distinguish only three vowel qualities, long and short. The long vowels are marked by mater lectionis: فَا [fa:], فِي [fi:], فُو [fu:]. The symbols above and below the preceding consonant are vowel points marking short vowels: كَ [ka], كِ [ki], كُ [ku]. These are called *fatḥa*, *kasra*, and *ḍamma*, respectively. They are normally not written, except in texts of the Qur'an. In a fully vocalised text, all consonants carry some pointing; the default point, marking a closed syllable (and thus equivalent to Tiberian

10 Complex Babylonian pointing has distinct signs for what some scholars believe are (allophones of) the vowels /a/, /e/, /i/, and /u/ (but not /o/) in unstressed syllables -- see Morag 1962:32-33 for some discussion.

quiescent shwa), is sukūn: مِنْ [min] 'from', كَتَبْتُ [katabtu] 'I wrote'.

Because the dialect of Mecca was one without glottal stop, Arabic | alif does not in fact mark the glottal stop, except initially by default. The sign for glottal stop is *hamza*, a small ḥa ء written above or below an initial or medial consonant (below if the following vowel is /i/: أرض [ʔarḍ-] 'earth' إن [ʔin] 'if'. Non-initial *hamza* may be written with *alif*, *ya*, or *waw*, or independently, depending on environment. (The rules for seating *hamza* are rather complex.)

Consonant gemination is marked by the superscript *shadda*: مُحَمَّد 'Muḥammad'. Written over initial *alif*, it is termed *alif madda* آ and marks initial [ʔā]. Other signs include *wasla*, a *ṣad* superscript ٱ, marking elision of the initial vowel of the article, and doubled vowel points marking final *nunation*.

References (call numbers refer to UWA library holdings)

Albright, William F. 1966. The Proto-Sinaitic Inscriptions and their Development. Cambridge: Harvard University Press.

Aronoff, Mark. 1985. Orthography and Linguistic Theory: The Syntactic Basis of Masoretic Hebrew Punctuation. Language 61, 28-72.

Bergstrasser, Gotthelf 1983 [1923]. Introduction to the Semitic languages. translated with notes and bibliography and an appendix on the scripts by Peter T. Daniels. Winona Lake: Eisenbrauns. HSS 492 1983 INT

Daniels, Peter T. 1983. Semitic Scripts. appendix to Bergstrasser 1983:236-260.
1996. The First Civilizations. in Daniels and Bright, eds. 1996:21-32.
1997. Scripts of Semitic Languages. in Hetzron, Robert ed. 1997:16-45.
and William Bright, eds. 1996. The world's writing systems. Oxford: OUP. HSS 411 1996 WOR

Dresher, Bezalel Elan. 1994. The Prosodic Basis of the Tiberian Hebrew System of Accents. Language 70:1-52.

Hetzron, Robert. ed. 1997. The Semitic Languages. New York: Routledge. HSS 492 1998 SEM

Hooker, J.T. 1990. Reading the Past: Ancient Writing from Cuneiform to the Alphabet. London: British Museum Press. HSS 411.09 1990 REA

Huenergard, John. 1997. A grammar of Akkadian. Atlanta: Scholars. hssmn 492.182421 1998 GRA

Joüon, Paul. 1991. A Grammar of Biblical Hebrew. Part One: Orthography and Phonetics. (tr. and rev. by Takamitsu Muraoka). Rome: Pontifico Instituto Biblico. hssmn 492.482421 1991 GRA v1-2

Loprieno, Antonio. 1995. Ancient Egyptian: A Linguistic Introduction. Cambridge: CUP. HSS 493.1 1995 ANC

Morag, Shelomo. 1962. The vocalisation systems of Arabic, Hebrew, and Aramaic. The Hague: Mouton. HSS 492 1962 VOC

Naveh, Joseph. 1987. Early history of the alphabet. Jerusalem: Magnes Press.

O'Connor, M. 1996. Epigraphic Semitic Scripts. in Daniels and Bright, eds. 1996:88-107.

Sass, Benjamin. 1988. The Genesis of the Alphabet and its Development in the Second Millennium BC, Agypten Und Altes Testament, Band 13, in Kommission bei Otto Harrassowitz, Weisbaden

Sasson, Jack ed. 1995. Civilizations of the Ancient Near East. New York: Scribner HSS Q 939.4 1995 CIV v4

Whitt, William D. 1995. The Story of the Semitic Alphabet. in Sassoon, Jack, ed. 1995:2379-2397