On alphabetical ordering:

Some principles and problems

Richard SALOMON
University of Washington, U.S.A.

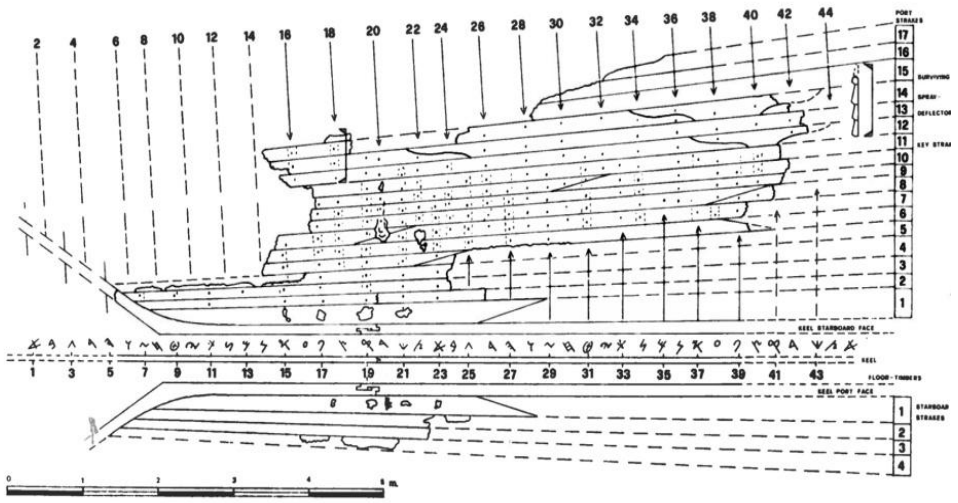
1. Alphabets and archaeology

In researching the topic of my paper, I have found no comprehensive or comparative studies of the principles and patterns of alphabetic ordering.[1] Some particular problems have been addressed in various publications, mainly the problem -- which I will discuss at some length below -- of the underlying rationale of Semitic/European alphabetic order, that is, of the familiar *a b c*... system. Perhaps this neglect of this interesting and complex problem stems from the fact that, for most of the world's population, an alphabetic or syllabic repertoire is among a child's first experience in systematic, intellectual study. This experience is so deeply engrained in our minds and memories that we are hardly inclined even to stop and think about its rationale.

[1] One partial exception is M. O'Connor's brief chapter on "The Alphabet as a Technology" in Daniels and Bright 1996: 787-94, particularly the sub-section on "The order of the alphabet" (pp. 788-90).

For the purposes of this paper, I propose to define "alphabetic order" a conventional and definitive sequencing of the complete set of graphs of a particular writing system and/or language. The primary function of such alphabetic orders is typically pedagogical, as the first step toward literacy. The other main function of alphabetic ordering systems is as a tool for the arrangement of verbal information, for example in lexica and dictionaries, or for material objects, for example in labeling components of a ship or building to ensure their correct positioning.

Image 1: Shipwreck from Isola Lunga (west coast of Sicily) with components labeled in Phoenicio-Punic characters in alphabetic order [< Frost et al. 1981: fig. 113)]



Most if not all languages which use alphabetic or syllabic scripts have some such standardized repertoire. It is also not at all unusual for a given language to have two or more such systems which are used for different functions or cultural contexts. Logographic or mixed logo-syllabic writing

systems such as Chinese or Akkadian tend to develop more complex systems for the arrangement of their much larger character sets, but these fall outside the limited purview of this paper.

In the case of living languages and continuing cultural traditions, for example those of the European languages or of Japanese, the systems of alphabetic ordering are well known -- although their underlying rationales, as I will discuss in detail, often are not. For ancient languages, however, the ordering system often must be deduced from archaeological materials. Most useful among these are abecedaries, that is, inscriptions or other documents on which the character set is written out in full, often as a casual student's exercise (and hence often with irregularities and errors, which cause endless headaches for the modern epigraphist, as they no doubt did for the ancient school teachers):

Image 2: Arapacana (Kharoṣṭhī) abecedary on a potsherd from Termez (Uzbekistan) [<Salomon 2004 (2008)]

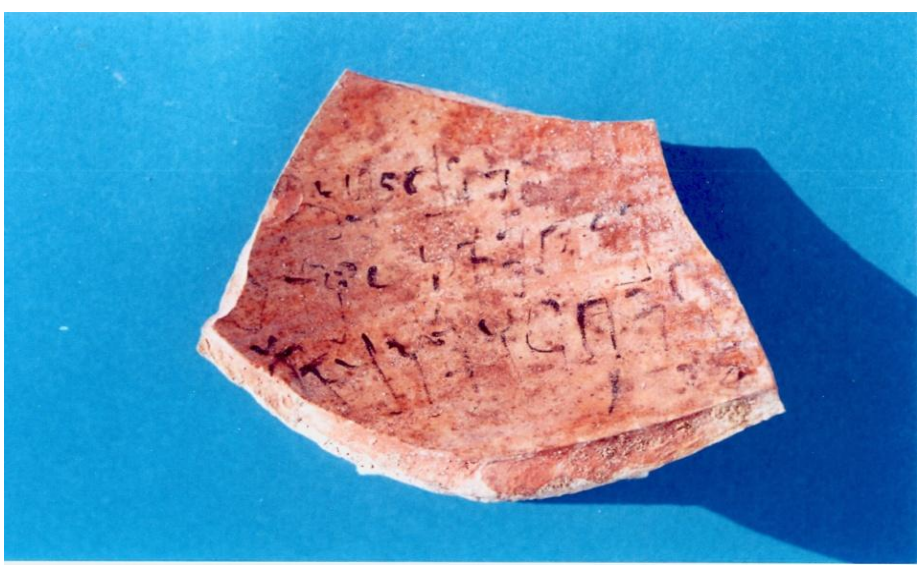
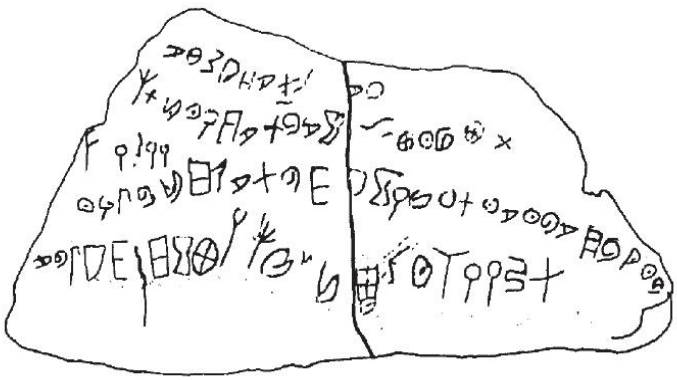


Image 3: Proto-Canaanite abecedary (line 5) from ʻIzbet Ṣarṭah, Israel (ca. 12th century B.C.E.) [< Naveh 1978: 31, fig. 1]



In other cases, an unknown alphabetic ordering system may be deduced from structural components, like the ship parts noted above. In that case, of course, the order of the Phoenician-Punic letters, following the standard Semitic sequence, was already well known, and in such cases the previously known alphabet may help the archaeologists to reconstruct a damaged structure (see, for example, Salomon 2006). But in other situations, an intact lettered archaeological sequence may illuminate an otherwise unknown ancient alphabet, as in the case of a set of stone blocks marked with location letters found at ancient Timna' (modern Hajar Koḥlân, Yeman), capital of the Qatabanian kingdom, dating to about the third century B.C.E. The sequence of letters on these blocks revealed that the order of letters of the modern Ethiopic alphabet was essentially the ancient South Semitic alphabetic order (Honeyman 1952).

2. Four patterns of alphabetic ordering

My survey somewhat cursory, and far from complete survey of alphabetic ordering systems around the world has revealed four main patterns:

Image 4: The four main patterns of alphabetic ordering

1. Orderings based on phonetic principles

Indic (Brāhmī-derived) sequence: *a ā i ī u ū … ka kha ga gha ṅa…*

2. Orderings based on groupings of graphically similar characters

Modern Arabic alphabet: … *ba' tā' ṯā'…*

3. Orderings based on a poem or other mnemonic device, in which the sequence follows the individual characters or initial characters of the words of the key text.

Japanese *iroha* sequence of *kana* characters: *i ro ha ni ho he to …*

Javanese: *hana caraka, data sawala…*

4. Mixed, unknown, or arbitrary systems with no evident ordering principle, seemingly random or capricious

North Semitic syllabary and its many derivatives (' *b g d … > a b c d …*)

South Semitic/Ethiopian alphabet: *h l ḥ m ś r s …*

Runic futhark: *f u Þ a r k g w …*

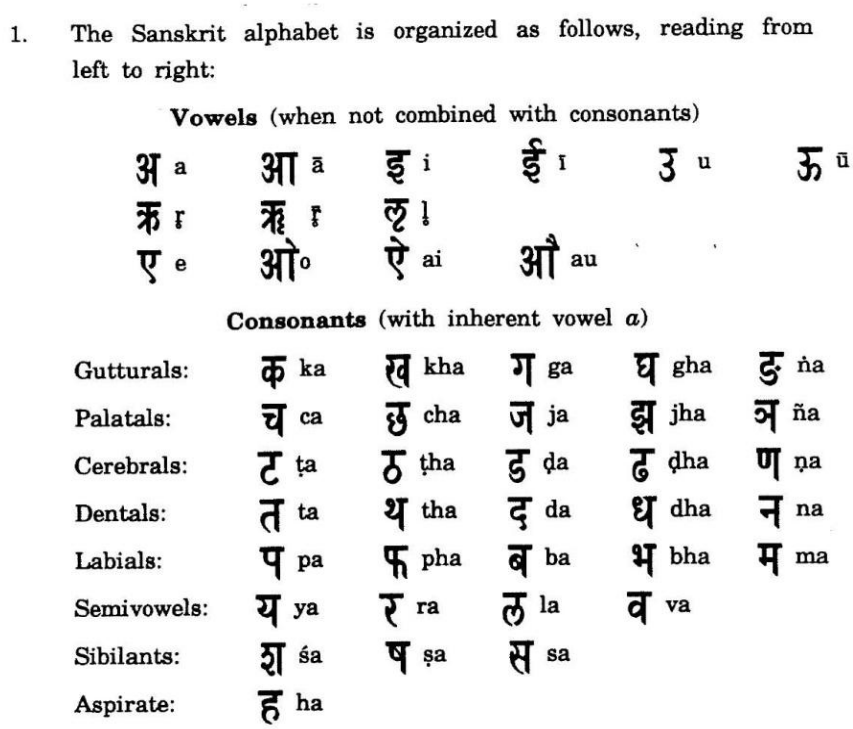
Kharoṣṭhī arapacana: *a ra pa ca na la da ba ḍa ṣa…*

Somewhat surprisingly, the logically transparent systems of the first two groups are relatively rare or geographically restricted, whereas some of the most successful systems and widespread systems -- most notably, the North Semitic alphabet in its many forms and derivatives -- belong to the problematic and puzzling fourth category. Indeed, this paradox constitutes the central mystery of alphabetic order, which I will address in some detail later, after presenting some prominent examples of the various types.

Phonetically structured alphabets

Among alphabetic orders based on straightforward phonetic principles, the super-family of Indian and Indian-derived scripts descended from ancient Brāhmī is the most prominent and widespread example. In this script group, the sets of characters are listed in a sequence, known as *varṇa-mālā* "garland of letters" or *varṇa-samāmnāya* "system of letters," which follows a regular and consistent phonetic pattern: first the vowels, then the consonants, then the semi-vowels, all arranged on the basis of their articulatory postion in the oral apparatus, from back to front.

Image 5: The standard Indic alphabetic order, as represented by Devanāgarī script.



The vowels are set up in pairs of short and long vowels (*a ā i ī u ū*, etc.), while the occlusive and nasals consonants (*sparśa*) are grouped into a grid of five rows representing the points of articulation (velar, palatal, retroflex, dental, labial) and columns for the different manners of articulation (unvoiced unaspirated, unvoiced aspirated, voiced unaspirated, voiced aspirated, nasal).

The earliest direct archaeological attestation of this Indian system of alphabetic arrangement is a terracotta figurine found at Sugh in the north Indian state of Panjab:

Image 6: Terracotta showing a boy writing the vowels of Brāhmī script on a writing tablet. (Line 3 reads: *a, ā, i, ī, u, ū, e, ai* …)

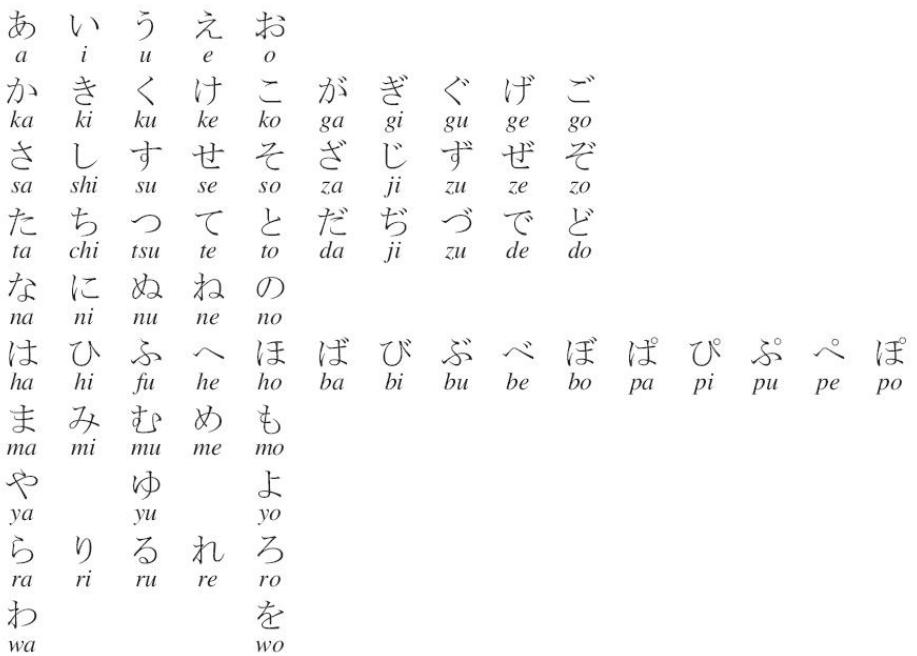
In this piece, a little boy is shown writing out the beginning of the Brāhmī alphabet, that is, the vowels, on a school tablet. The boy is presumably the Buddha as a child, illustrating the legend that on his first day of school he could already write in all of the sixty-four scripts.[2] Although this piece is datable to the early centuries of the Common Era, it is clear from textual evidence that the alphabetic system it represents is much older than that. The system is implied, though not directly attested as such, in the early (Pāṇinian) grammatical tradition, which goes back to at least the fourth century A.D., and it origins probably date back well into the first millennium B.C.E.

This alphabetic sequence has been maintained down to the present day without any fundamental changes in the many Indian scripts used in the subcontinent and in other parts of Asia, all of them derived directly from Brāhmī. In some of scripts used for Indian languages, minor adjustments have been made to adapt to their phonological systems, whole among the Indian scripts which have been adapted for non-Indian languages of Southeast Asia and Central Asia (notably Tibetan), more substantial changes have been introduced. These often take the form of additional or modified characters needed for those languages, typically inserted in the alphabetic sequence after their archetypal letter or at the end of the alphabet. But in every case, the fundamental ordering has not been altered. This illustrates the general principle of the stability of alphabetic order across languages and across millennia, also attested by the history of the Semitic alphabet order which has remained more or less stable for at least three and a half millennia.

Outside of the Indic script group, there are few scripts whose ordering is based on phonetic principles, and at least one of the prominent systems, namely the Japanese *kana* order, is almost certainly based on the Indian archetype.

[2] The same legend in illustrated in figs. 16-17 below.

IMAGE 7: Standard modern order of Japanese *kana* syllabaries (based on Indic alphabetic order) [< Daniels and Bright: 1996 : 211]



The standard *kana* order of the vowels *a i u e o* and the consonants, *ka ki ku ke ko; ga gi gu ge go*, etc. corresponds exactly with the Indic system. Although the basic Indian alphabet cited above does not list all the syllabic combinations, such complete models of consonant-plus-vowel combinations, arranged in exactly this order, are well attested in Indian texts, especially among Sanskrit manuscripts from Central Asia. This script type and ordering system became well known in China and Japan, especially in its association with the Indian script known as Siddham, associated with the East Asia traditions of esoteric Buddhism, so that its influence on Japanese writing is easily explained on historical grounds.[3]

[3] See van Gulik1956.

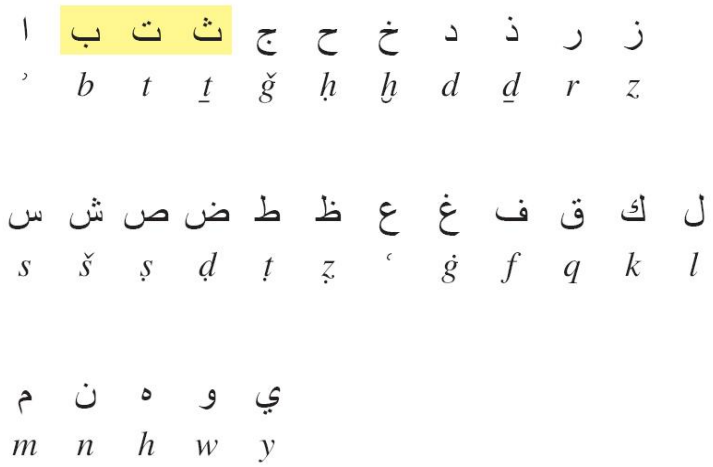
The Korean Hangul script presents another important case of a phonetically organized alphabet. In its original sequence as presented in the *Hunmin chŏng'um* in 1446, the sequence began with the consonants arranged into phonetic sets as *k kh ŋ t th n p ph m c ch s*, etc., followed by the vowels. According to Song (2011: 41), the Hangul order reflects the Chinese model of classification of consonants, and this in turn was presumably influenced by the Indian system (ibid., p. 32). Thus the Hangul system may have been indirectly influenced by the Indian model (ibid., p. 37), but in any case it stands out as another specimen of a phonetically ordered sequence.

A phonetically ordered alphabetic sequence was also introduced for Arabic script by the early Arabic philologist Abu 'Amr al-Khalil ibn Ahmad al-Farāhīdī (718-786 C.E.) in his dictionary *Kitāb al-'Ayn* (The Book of 'Ayn). This dictionary is organized according to letters phonetically sequenced from the back of the mouth (*'ayn*) to the front (*b, h*). Such sequences, using either the last or the first letter of the head word, became standard for Arabic dictionaries, but were restricted in function to reference purposes and did not become the pedagogical order, which arranged the letters on the basis of formal graphic similarities.

Alphabets structured by graphic similarity

This standard Arabic alphabetic ordering is in fact the archetypal example of this class, and in fact is the only clear and indisputable case of a system which was primarily modeled on this system:

Image 8: The Arabic alphabet



Here the grouping by visual similarities between group of graphs is patently obvious, as in the sequences *b t ṯ, ǧ ḥ ḫ, d ḏ, r z*, and so on. It is nevertheless clear that this is a remodeling of the original North Semitic abecedary, traces of which survive in, for example, the retention of *alif* (ʾ) in initial position and the original sequence *l m n*. The case of Arabic thus illustrates two principles of alphabetic ordering: (1) the co-existence of different systems within the same cultural/linguistic tradition and (2) the alteration of pre-existing systems in the course of their application to or development within new languages.

The latter principle seems to apply also to the few other cases -- all of them less thoroughgoing and less certain than that of Arabic -- of alphabetic organization on visual grounds. These principle has been invoked in reference to the South Arabic scripts, and especially their Ethiopic derivatives, which follow an order (*l ḥ m š r g• s*, etc.) that is entirely different from the North Semitic *a b c* sequence. In this system, according to Driver (1976: 271) "the signs seem to be arranged, with occasional exception, according to their forms."

Image 9: G.R. Driver's explanation of South Arabic alphabetic order on grounds of graphic similarity [< Driver 1976: 271]

(Hallo in 'J.B.L.' LXXVII 332–4). Obviously the order of the letters here is not dictated by phonetic principles; rather, the signs seem to be arranged, with occasional exceptions, according to their forms. The first sign is the simplest, being little else than a vertical stroke, and the last is a circle, so that an initial tendency to start from strokes and terminate with box-shaped or circular signs may be detected; then they seem to be arranged in groups according to their shapes (namely, *m-r*, *ġ-k*, *n-ḫ*, *s-'*). Only *ḥ* seems to be misplaced between *l* and *m* and not beside *ḫ*, which closely resembles it in form.

Driver likewise explains similar sequences in the Ethiopic script (p. 185), and in this, at least, he is supported by Ryckmans, who thinks that "La plupart des changements apportés par l'éthiopien à l'ordre originel s'expliquent visiblement par des considérations de forme : des lettres de forme analogue sont groupées ensemble, ou inversement, une lettre venu d'ailleurs est inseré, pour mieux les distinguer, entre deux signes très ressemblants." This pattern of locating additional signs to a pre-existing alphabet according to graphic similarities to the original sequence is also invoked by Driver to explain the positions of the eight "extra" characters of the Ugaritic cuneiform alphabet of thirty letters which are not present in the corresponding proto-twenty-two letter Canaanite alphabet (1976: 272).

However, Arabic remains the only totally convincing instance of the principle of alphabetic ordering according to graphic or visual similarities. While some of the other cases mentioned are at least

reasonably persuasive, judging them involves a degree of subjectivity with regard to what features in any given case constitute a significant degree of visual similarity. For example, although Ryckman's agrees with Driver's ideas (though not citing them directly) about the role of graphic similarities in the re-arrangement of the Ethiopic script, he largely rejects Driver's theory that the South Arabic alphabet is based on visual similarities, declaring "Dans la disposition de l'alphabet sud-sémitique, on ne saisit aucun principe d'ensemble de regroupement formel ou phonétique.... En bref, l'ordre de l'alphabet sud-sémitique présente sensiblement la même impression d'arbitraire ou de hasard temperé par l'un ou l'autre rapprochement de formes ... que l'ordre des alphabets nord-sémitiques anciens" (1981: 704-6). In any case, it is certain that this method of organization is of limited frequency among alphabetic and syllabic scripts in general.

Mnemonically-based alphabetic orders

The third ordering principle involves mnemonic systems in which a fixed text determines or organizes the order of the graphic units of a given script. The classic example is the Japanese *iroha* system, which defines a sequence of the forty-seven *hiragana* syllabic characters on the basis of a poem, attributed to the renowned Buddhist monk Kūkai, which uses each syllable once:

Image 10: The Japanese *iroha* syllabary [text and translation by P. Atkins]

In the Indian-derived Javanese script, the twenty basic consonantal characters are similarly ordered on the basis of an abecadaric poem reading *hana caraka, data sawala, padha jayanya, maga bathanga*, "There were (two) emissaries, they began to fight, their valor was equal, they both fell dead" (Joel C. Kuipers and Ray McDermott in Bright and Daniels 1996: 478). Here each syllable consists of the basic consonant together with the neutral or inherent (graphically unmarked) vowel *a*.

Arbitrary or unknown systems

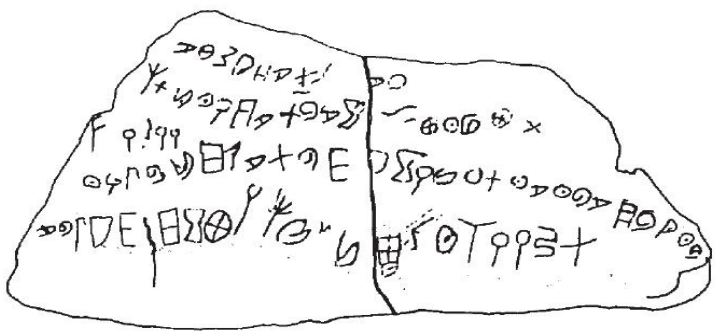
Other traditions, such as the Indic Kharoṣṭhī script (discussed in the following section) are associated with similar mnemonic texts which embody a standard order, but in many of these cases it is difficult to determine whether the text actually determines the ordering, or rather is a secondary creation based on a previously existing conventional sequence. This problem brings us to the fourth type of character ordering, namely those systems -- and there are all too many of them -- whose ordering principles, if any, have no evident rationale, and which therefore seem to be random or capricious.

It has been suggested in some such cases that the ordering was originally based on some poem or other textual basis which has been forgotten over time, and some of these claims will be evaluated or mentioned below. But first, I will introduce the best known case of this problem. This is, of course, the ancient North Semitic ordering *a b g d h w* …, which underlies, with various changes (such as the replacement of *g* by *c* in the Etruscan-Roman node), the alphabetic systems used by all the languages of Europe and the New World. Unlike the much-neglected broader questions of alphabetic ordering which are the main concern of this article, the particular question of the origin of this ordering has been the object of extensive study and argument.

What has become clear as a result of archaeological discoveries in recent decades is that this ordering is very ancient. An inscribed stone from ʻIzbet Ṣarṭah, datable with reasonable certainty and accuracy to the twelfth century B.C.E., contains a crudely written abecedary in the Proto-Canaanite consonant syllabary, the ancestor of the later Semitic scripts such as Phoenician, Old Hebrew and Aramaic. This abecedary agrees, except for a few details, such as the switching of *ḥ* and *z*, with the classic Semitic order:

Image 11: Proto-Canaanite abecedary from ʻIzbet Ṣarṭah, Israel, ca. 12th century B.C.

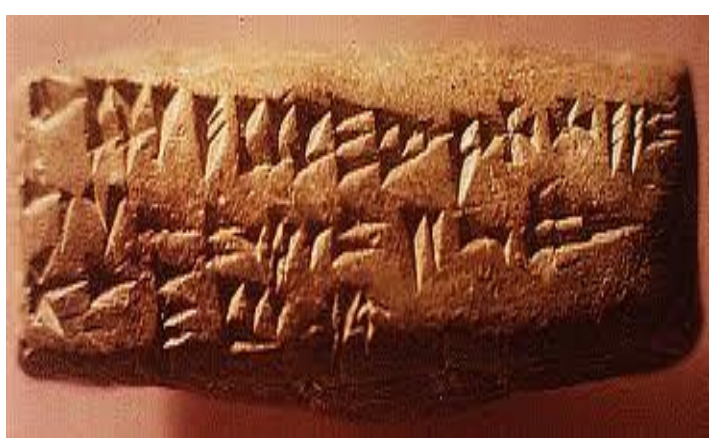
Line 5: *ʾ b g d h w ḥ z ṭ y k l [m] n [s] p ʻ ṣ q r š t*



But we also have even earlier evidence of this system, in the form of an abecedary in the Ugaritic cuneiform syllabary, datable to the fourteenth century B.C.E.:

Image 12: Abecedary in Ugaritic alphabetic cuneiform from Ras Shamra, 14th century B.C.E.

Line 1: *ʾa b g ḫ d h w z ḥ t y k*



Although the Ugaritic syllabary is superficially unrelated to the Proto-Canaanite syllabary, being stylistically modeled on the cuneiform scripts of Mesopotamia, it is systemically virtually identical to them, exceptionally the addition of a few letters peculiar to the Ugaritic language. Thus there can be no question that the same ordering system underlies both scripts, and that this system was already in existence at or at least not long after the time that they were invented.

But while the antiquity of this ordering is well established, its rationale remains obscure. The relevant issues and theories, up to 1976, are conveniently summarized in the third edition G. S. Driver's *Semitic Writing from Pictograph to Alphabet*, in his original section on "The Order of the Letters of the Alphabet" (1976: 179-85) and in the supplementary notes appended by S. A. Hopkins (pp. 269-73). Driver dismisses out of hand "fantastic reasons for the order of the letters … based … for example, on astral or lunar theories" (p. 181), and Hopkins (p. 269) considers but dismisses other proposals, including Tur-Sinai's theory of an underlying mnemonic poem, which will be discussed below. Driver then presents his own theory (pp. 182-4), according to which the north Semitic abecedary is determined by a combination of phonetic features, the forms of the signs, the meanings of the names of the letters, and associations between letters due to common features of usage or alternation patterns:

Image 13: G. S. Driver's explanation of the North Semitic alphabetic order

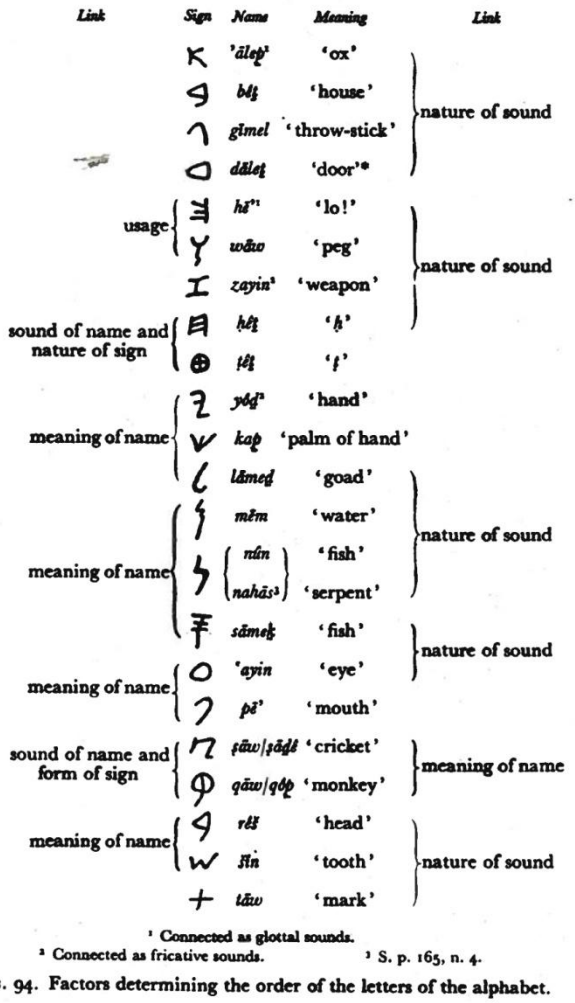


FIG. 94. Factors determining the order of the letters of the alphabet.

He notes, for example, that the first four letters, *' b g d*, are plosives while the following four, *h w z ḥ*, are fricatives. This is a promising start, but to account for the rest of the sequences he has to resort to a grab-bag of miscellaneous factors. Driver himself was evidently quite aware of the

weaknesses in his explanation, remarking that "even if it is fanciful in parts, [it] is not so wholly fantastic as those based on celestial theories," and that "is within the bounds of human possibility" (p. 185) -- terms which hardly inspire a sense of the author's confidence in his own theory.[4]

Among the several other proposed explanations of the North Semitic order, that of H. Tur-Sinai is of particular interest in connection with our comparative approach. He theorized that the sequence reflects an ancient pedagogical poem in which the first character of each word determined the alphabetic order, and even went so far as to reconstruct this poem on the basis of associations of the Hebrew letters with particular key words in later Talmudic traditions: thus, *'lf bynh gml dl h' / vh' zn ḥy ṭb ydh …* "Learn to understand: [God] does good to the poor and feeds all living creatures with bounteous hand…" (Tur-Sinai 1950: 288-9). The principle is in theory valid, since, as we have already seen, other alphabets seem to be organized in this way, but the application can be described at best as highly speculative, so that it is not surprising that Hopkins (in Driver 1976: 269) rejected Tur-Sinai's hypothesis with the curt comment that "the language and style of the supposed poem are enough to condemn the suggestion."

A more linguistically sophisticated attempt to explain the Semitic order has been recently been offered in two articles by W. C. Watt (1987, 1989). He has constructed two phonetically structured grids which he calls the Byblos and Ras Shamra matrices to account for the Proto-Canaanite/North Semitic and Ugaritic orders respectively:

[4] See also the more technical criticism of Driver's theory in Watt 1987: 9-10.

Image 14: W.C. Watt's "Byblos matrix"

	I	II	III	IV	V
	LARYNGEALS	BILABIALS	ALVEOLARS	VELARS	DENTALS
STOPS	ʾ	*b*		*g*	*d*
FRICATIVES	*h*	*w*	*z*	*ḥ*	*ṭ*
ROW 3			*y*	*k*	*l*
ROW 4		*m*	*n*		*tʸ*
ROW 5	ʿ	*p*	*ṣ*	*q*	*r*
ROW 6			*š*		*t*

Watt's matrices present an ingenious argument for a phonetically structured rationale to the ABC order. But they require a good deal of special pleading, including, for example, a "principle of maximum separation" (Watt 1989: 69-76), which is invoked to explain, among other problems, the curious ordering of the columns of the matrix. For these columns represent phonetic categories, specifically places of articulation, in what would on the surface seem a very strange sequence, jumping from the "extreme back" (laryngeals/pharyngeals) to the front (labials), then to the middle (alveolars and palatals, to the back (velars), and finally again to the front (dento-alveolars). This stands in contrast to the "natural" ordering of the phonetic/alphabetic sets in the standard Indic order, directly from the back to the front of the oral cavity, and it seems to presuppose a degree of linguistic sophistication that is hard to imagine in the middle of the second millennium B.C.E.

Even more problematic in Watt's matrices is the logic of the rows. The first row is clearly enough stops, and the second, apparently, fricatives (1987: 2). But the nature of rows 3 through 6 is never clearly specified, though Watt tries to justify them by reference to the pedagogical practice, attested in some ancient documents, of splitting the alphabet into two

halves (1989: 81-4). But this seems to me a rather desperate and convoluted argument. In this and other regards, Watt's matrix theory is in the end more ingenious than convincing. It does bring out, at least, a certain parallelism in place of articulation between the first four letters (*', b, g, d*) and the next five (*h, w, z, ḥ, ṭ*), and it is conceivable that this relationship had some influence on the formulation of the *a b c* ordering, but this is actually little more than a refinement of what had already been pointed out by Driver. Beyond this, Watt's formulation, though ingenious and intriguing, is also largely forced and implausible, and it is no doubt for such reasons that Watt's theory has not, as far as I am aware, won acceptance among specialists.

Thus none of the three attested systems of alphabetic sequencing -- phonetic arrangement, visual resemblance, and mnemonic archetype -- yields a comprehensive and convincing rationale for the *a b c* order. This has led some authorities to conclude, along with Ullman (1932: 20), that "This order seems to be one of chance." This is hardly a satisfying conclusion, since the human mind instinctively seeks for order and pattern, but this does not mean it is wrong.

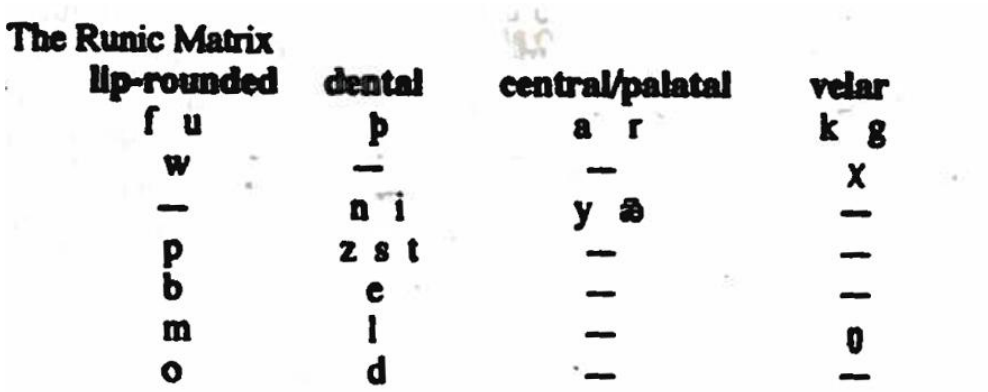
However, before giving up, let us have a look at a few examples of many other script systems whose ordering principles are less than obvious. I have already alluded in the previous section to one important problem in this class, namely the South Semitic order. This is now well attested from partial abecedaries, from the archaeological data from Timna', and from the living traditions of the Ethiopic scripts. Moreover, an astonishing discovery by A. G. Loundine (1987) revealed that a tablet found long ago at Beth Shemesh (Israel) contained an abecedary in the Ugaritic cuneiform alphabet, but following the South Semitic (*h l ḥ m*) order rather than the Northern *' b g d* of the Ras Shamra Ugaritic abecedary. This shows that the South Semitic ordering system, contrary to what had previously been assumed, goes back just as far as the Northern one, that is, to the middle of

the second millennium B.C.E., and also that it was not restricted, as had also been assumed, to the South Semitic language area in the Arabian peninsula. Thus we can now suspect that in the formative period of the Semitic scripts, two ordering systems, apparently unrelated, were in use simultaneously -- a pattern which we already noted in other times and places -- and only later came to be preferred in different regions. This does not, however, bring us any closer to an understanding of the rationale of the South Semitic sequence, for which, as noted above, no cogent explanation has yet been presented, Driver's proposal notwithstanding.

Another such problem, in an entirely different part of the world and period of history, is presented by the Runic alphabet, used in northern Europe and England in the first millennium C.E. The alphabet had twenty-four characters which were arranged in a standard order, well attested from abecedaries, which is referred to by modern scholars as the futhark after its first six letters (*f u Þ a r k*). The origin of the script is in large part obscure, but it is generally agreed that the characters were developed from or at least inspired by the Latin or other southern European scripts. However this may be, the Runic alphabetic sequence has no discernible relationship to that of the Italic scripts, and according to a leading authority "we still have absolutely no idea how this arrangement came about …. the best guess is that it had to do with the manner in which the runes were taught and learned, the result of some mnemonic device which is no longer retrievable" (Antonsen 1989: 140).

However, Miller (1994: 70-6) has recently proposed that the futhark is phonetically arranged in a matrix similar to those which Watt adduced for the North Semitic alphabetic orderings.

Image 15: D. Gary Miller's "Runic Matrix" [< Miller 1994: 71]



Miller points out, for example, that "it is hardly accidental that the order of letters in *fuÞark* begins with a labial (lip-rounded) C-V pair (*f-u*), proceeds to (inter)dental *Þ*, then to 'central' *a-r*, and finally to velar *k*, establishing a phonetic grid." In one regard at least, Miller's futhark grid is more convincing than Watt's Byblos and Ras Shamra grids, in that the four columns (lip-rounded, dental, central/palatal, velar) fall into a natural (front-to-back) sequence, avoiding Watt's dubious "principle of maximum separation."

But here too a considerable amount of manipulation is required to make the grid work, for example by squeezing multiple letters into a single box in the grid, including the first two letters, *f* and *u.* Also, as in Watt's matrices, the rationale of the horizontal rows is never explained. In the end, then, we are left with a similar result: there seems to be some degree of rough phonetic significance in the sequencing of the first several letters of the alphabets concerned, but the system breaks down thereafter into what seems to be an arbitrary or random sequence. Perhaps this parallelism is in itself meaningful, reflecting, for example, primitive and only partially successful attempts at phonetic analysis; but this remains speculation unless and until other similar cases can be adduced. For the time being, in

any case, the futhark must remain in the category of indeterminate alphabetic orderings.

For my penultimate example of obscure orderings, I turn to the Indian Arapacana syllabary. This sequence of forty-two syllables, named, as usual, from its first members (*a ra pa ca na* …), has long been known as a ritual structure for Buddhist texts in Sanskrit, Chinese, and Tibetan. Only recently (Salomon 1990, 2006) has it become clear that this sequence was originally the syllabic ordering of the Kharoṣṭhī script, which was current in the northwestern borderlands of the Indian subcontinent between the third century B.C.E. and the third century C.E. This became clear from sculptural abecedaries such as this one, showing the Buddha demonstrating his ability to write sixty-four scripts on his first day at school [compare image 6]:

Image 16: Gandhāran relief showing the Buddha writing in school

Image 17: Detail of the inscription on the Buddha's writing board: beginning of the Arapacana syllabary. *a ra pa ca na la da [ba]*

It is immediately obvious that the Arapacana syllabary, in terms of its organizing principles, has nothing in common with the transparently phonetic arrangement of its neighboring contemporary script Brāhmī, except that they both begin with *a*. As with the several other problematic alphabets discussed above, no phonetic, visual, or mnemonic pattern is discernible in the Arapacana system, so that its underlying rationale is entirely obscure. Brough (1977: 94) speculated that the Arapacana could have developed from "a list of head-words [which] … might have been in origin a mnemonic device to fix the order of the verses or paragraphs of some important text," and which subsequently was "further reduced to initial syllables," but he was unable to identify any such underlying mnemonic text.

Very recently, however, a manuscript has been discovered which contains a set of verses, beginning with each of the letters of the Arapacana in sequence. But the discoverer of this text prudently avoided

jumping to the conclusion that this was the archetypal text posited by Brough which determined the Arapacana order, stating to the contrary that "More probably, it was the alphabet, thought to be a (complete?) inventory of Kharoṣṭhī signs (not Gāndhārī sounds!), which came first and became subsequently used to arrange the sequence of a certain text or a list of dogmatically important issues" (Strauch 2008: 122). The historical and textual circumstances cited by Strauch strongly support his conclusion, so that now, as when Brough wrote, “the origin of the *Arapacana* is still a mystery.”

This is somewhat disturbing in that, if Brough’s mnemonic theory, which *a priori* seems the most probable, is correct, we should be able to identify the underlying text somewhere in the vast and generally well-attested literature of Buddhism. But this has not yet proven possible, and this might mean that the real source of the Arapacana may lie elsewhere, perhaps in some non-Indian tradition of the Iranian or Aramaic-speaking world.

As a final example of a problematic ordering system, I turn to the Tāna alphabet of the Maldive Islands. This relatively modern script, which apparently replaced earlier local scripts in the eighteenth century, consists of twenty-four consonantal characters arranged in the sequence *h š n r b ḷ k ' v m f d t l g ñ s ḍ z ṭ y p j c*. Like so many other scripts, this ordering presents no obvious logic, and it has been suggested that this relates to an "esoteric origin of Tāna, namely a script that was scrambled on purpose in order to keep it secret from other islanders" (Wikipedia, "Maldivian Writing Systems," accessed Sept. 14, 2012). This notion calls to mind the persistent association of the runes with secret lore, but this tradition has been effectively debunked by leading authorites such as Antonsen (1989: 140), and in any case it is not clear to me on what authority, if any, the notion of the Tāna alphabet as secret and intentionally obscure is based. Thus all in all, motives of secrecy do not seem to provide an explanation for opaque systems of alphabetic ordering.

3. Some conclusions

In the end, we are left with a quandary. It is clear that three principles, namely phonetic sequence, graphic resemblance, and textual mnemonics, govern the formation of alphabetic ordering in many script/language systems. But it is equally clear that there is a large residue of systems, including some of the most successful and widespread ones, whose rationale cannot be convincing explained by any of these patterns, and appear to be completely or largely random. Still, it seems counter-intuitive that such fundamental cultural artifacts have no logical underpinnings, and it is mainly for this reason that scholars have assumed that in such cases there must be some lost information which would have provided a logical key to the problem. In such cases, it is often assumed or suspected that a mnemonic text underlies the problematic system, as has been suggested, for example, by Tur-Sinai for the Semitic alphabet and Brough for the Arapacana syllabary. But the latter case, at least, must give us pause, as it shows once again how the relationship between script and mnemonic may be a chicken-and-egg problem. A mnemonic key to the Arapacana syllabary has been recently discovered, but it is almost certainly secondary; that is to say, the poem is based on the pre-existing syllabary, rather than the other way round. It is of course still possible that there was in some remote past time an original poem or text which did determine the syllabic order, but this is, once again, a matter of pure speculation.

So in the end, the matter boils down to a philosophical problem: do we wish to accept the possibility that some alphabetic ordering are completely or largely random or unmotivated, or can we assume that in the many problematic cases there is a lost or hidden rationale which we have failed to find? Most experts have been reluctant to accept Ullman's conception that that the Semitic order "seems to be one of chance," and Song's position that "The order of any writing system would have been

originally accompanied by a rationale, either phonetic or cultural. But, as the rationale was a matter unnecessary to remember for the learners and users of the script, it may be completely forgotten" (2011: 39) seems more convincing, or at least more appealing. Nevertheless, given the many efforts, all more or less unsuccessful, to decipher the rationale of the Semitic alphabet, it is difficult to be optimistic about the chances for a breakthrough, barring some miraculous and unexpected discovery. But it is my hope that a more comparative and comprehensive approach to the study of alphabetic order in general, such as the one for which I have proposed a rough outline here, might eventually lead to some new insights into this and the related problems.

References

Antonsen, Elmer H. 1989. "The Runes: The Earliest Germanic Writing System." In Wayne M. Senner, ed., *The Origins of Writing* (Lincoln NE: University of Nebraska Press), pp. 137-58.

Brough, John. 1977. "The *Arapacana* Syllabary and the Old Lalitavistara." *Bulletin of the School of Oriental and African Studies* 40: 85-95.

Daniels, Peter T. and William Bright. 1996. *The World's Writing Systems.* New York: Oxford University Press.

Driver, Samuel. 1976. *Semitic Writing from Pictograph to Alphabet*. 3rd ed., ed. S. A. Hopkins. London: Oxford University Press for The British Academy.

Frost, Honor et al. 1981. *Libybaeum*. Notizie degli Scavi di Antichità ser. 8, vol. 30, supplement. Rome: Accademia Nazionale dei Lincei.

Hallo, William W. 1958. "Isiaiah 28:9-13 and the Ugaritic Abecedaries." *Journal of Biblical Literature* 77: 324-38.

Honeyman, A. M. 1952. "The Letter-Order of the Semitic Alphabets in Africa and the Near East." *Africa* 22: 136-47.

Loundine, A. G. 1987. "L'abecedaire de Beth Shemesh." *Le Muséon. Revue d'études orientales* 100: 243-50.

Miller, D. Gary. 1994. *Ancient Scripts and Phonological Knowledge.* Amsterdam Studies in the Theory and History of Linguistic Science 116. Amsterdam: John Benjamins.

Naveh, Joseph. 1978. "Some Considerations on the Ostracon from ʻIzbet Ṣarṭah." *Israel Exploration Journal* 28: 31-5.

Ryckmans, Jacques. 1981. "L'ordre des lettres de l'aphabet sud-sémitique. Contribution à la question de l'origine de l'écriture alphabétique." *L'antiquité Classique* 50: 698-706.

Salomon, Richard. 1990. "New Evidence for a Gāndhārī Origin of the Arapacana Syllabary." *Journal of the American Oriental Society* 110: 255-73.

----. 2004 (2008). “An Arapacana Abecedary from Kara Tepe (Termez, Uzbekistan).” *Bulletin of the Asia Institute* 18: 43-51.

----. 2006. “Kharoṣṭhī Syllables Used as Location Markers in Gandhāran *Stūpa* Architecture.” In Pierfrancesco Callieri, ed., *Architetti, Capomastri, Artigiani: L’organizzazione dei cantieri e della produzione artistica nell’asia ellenistica. Studi offerti a Domenico Faccenna nel suo ottantesimo compleanno* (Serie Orientale Rome 100; Rome: Istituto Italiano per l’Africa e l’Oriente), pp. 181-224.

Song, Ki-Joong. 2011. "Ancient Indian and Chinese Models of Sound Classification and their Reflections in the Writing Systems." *Scripta* 3: 25-43.

Strauch, Ingo. 2008. "The Bajaur Collection of Kharoṣṭhī manuscripts – a preliminary survey." *Studien zur Indologie und Iranistik* 25:103-136.

Tur-Sinai [Torczyner], H. 1950. "The Origin of the Alphabet." Jewish Quarterly Review 41: 83-109, 159-79, 277-301.

Ullman, B. L. 1932. *Ancient Writing and its Influence.* New York: Longman, Green and Co.

van Gulik, R. H. 1956 [2001]. *Siddham. An Essay on the History of Sanskrit Studies in China and Japan*. Śatapiṭaka Series 247. Nagpur: International Academy of Indian Culture.

Watt, C. W. 1987. “The Byblos Matrix.” *Journal of Near Eastern* Studies 46: 1-14.

----. 1989. "The Ras Shamra Matrix.

Discussion: On Alphabetical Ordering:

Some Principles and Problems

SONG Ki-joong
Seoul National University, KOREA

As Professor Salomon wrote in the opening sentence of this paper, a comprehensive study on the 'alphabetic order' seems yet to appear. However, it has been occasionally mentioned in a number of works on writings. The discussant hopes that Professor Salomon further develop this study and give us in the future a conclusive explanation on, among others, the '*a b c* order'.

Professor Salomon's explanations and discussions are very informative and moderate. However, as the discussant is obliged to comment something out of the contents of the presentation, a few rudimentary opinions and remarks are expressed below.

1.Alphabets and archaeology

A Secondary Function of the "Alphabetic Order" (after Professor Salomon)

As is well known, the 'alphabetic order' of a writing system is commonly used for numbering sequential items or matters. Besides the universally known Arabic numbers (1, 2, 3, ...) and the Roman numbers (I, II, III, ... or i, ii, iii, ...), the Roamn letters in the alphabetic order (A, B, C,... or a, b, c,...) are frequently used for numbering items in Korea and elsewhere in the world and, in Korea, the Korean alphabets (ㄱ, ㄴ, ㄷ,... or 가, 나, 다,...) also.

The traditional, elementary textbook of Chinese characters, *Ch'ŏnjamun* (Chi.*Quianziwen*) 千字文 or *One Thousand Characters*, contains 250 four-character phrases in which a character occurs only once. In the past the characters in the order appearing in the book, as 天地玄黃, 宇宙弘荒, ..., were rarely used in Korea for numbering purpose, especially for numbering pages or folios of voluminous books such as the genealogy of a clan (*chokpo* 族譜).

It is remembered that at all times and countries in the civilized world, in the beginning stage of formal elementary education, the educatees are forced to memorize the writing system in use, the graphic signs and representing sound and, if any, meaning values, as well as the order of the signs. A Korean gentleman in the old times was supposed to remember the order of the 1000 characters found in the *Ch'ŏnjamun*.

2. Four patterns of alphabetic ordering

2a. Phonetically structured alphabets

Professor Salomon's wording "phonetic principles" might be somewhat ambiguous for some readers. The alphabetic order follows "the systematic, (articulatory) phonetic classification of the sounds."

Perhaps, most of Indian writings used in the Indian subcontinent and adjacent regions consist of two series of signs, the first series for the vowels

and the second for the consonants as the case of Devanāgarī. And, the vowels are represented either by independent signs or by diacritical marks.

Among the Indian-derived writings of Southeast Asia and elsewhere, there are those in which independent signs for the vowels lack. The vowels are indicated by diacritics (which are known as ***abugidas***). Thai, Lao and Tibetan writings belong to the latter category. In the ḥPʾags-pa script, which was created in the 13th century on the model of the Tibetan, vowel [a] is inherent in the consonantal signs and the other vowels, [i, u, e, o], are represented by separate signs. In the ʻalphabetic orderʼ of the ḥPʾags-pa script the consonantal signs, in the order of which principle is identical with the Indianʾs, are followed by the vowels.

Therefore, the description of the common feature of the Indian or Indian-derived scripts, "In this script group, the sets of characters are listed in a sequence, ..., which follows a regular and consistent phonetic pattern: first the vowels, then the consonants ..." would be require a modification.

The function of the independent signs for vowels in Indian and/or Japanese writings should not be regarded as identical with that of the vowel signs in the pure alphabetic systems such as Roman and Korean writings. Those Indian and Japanese vowel signs represent ʻa (zero) consonant + [a]ʼ syllables. They are not used to indicate the vowels constituting a part of syllables. Among the writing systems I am familiar with, the ḥPʾags-pa and the Korean systems have independent series of phonemic vowel signs and in the ʻalphabetic orderʼ of both systems the vowels are placed after the consonants.

2b. Alphabets structured by graphic similarity

ʻAn alphabetic order structured by graphic similarityʼ necessarily presupposes that the graphic signs had been created arbitrarily to represent various sounds and that, when arrangement of the signs in a certain order

was needed for a certain purpose, the external forms of the signs were taken as the major criteria.

From the graphic forms of the Arabic letters representing *b t th* (ب ت ث), or from those of *j ḥ kh* (ج ح خ) and so forth, anyone can easily see that the three letters in a group share a common element. One may assume that the letter for *b* (ب) and *t* (ت) arbitrarily share the common stroke since the represented sounds do not reveal a similarity. However, a resemblance between the sounds *t* (ت) and *th* (ث) could be easily recognizable. It would be more plausible that the two letters share the common stroke because their phonetic values are similar, rather than that they came to share the common stroke arbitrarily.

It is also remembered that for the Arabic script different 'alphabetic orders' have been in use for different purposes. (See Wikipedia, under "Arabic alphabet.)

Besides, as Professor Salomon remarks, there are found in the 'alphabetic order' of Arabic, traces of the North Semitic, such as "the retention of alif(') in initial position and the original sequence *l m n.*"

Keeping in mind of what has been written above, one may review Professor Salomon's conclusion regarding the alphabetic order of Arabic, "(1) the co-existence of different systems within the same cultural/linguistic tradition and (2) the alteration of pre-exiting systems in the course of their application to or development within the languages."

2c. Mnemonically-based alphabetic orders

As mentioned above, forcing the learners memorize the graphic forms with phonetic (and semantic) values of writings in use in the beginning of formal education has been a universal tradition of human being's civilization. Most literates in the world would remember how difficult it was to memorize the alphabets in the beginning classes. Historically the methods of easier and effective teaching, or learning, of the alphabets have

been sought, probably, anywhere in the world. Japanese *Iroha* might have been composed for the purpose of easier learning and/or teaching of Japanese writing. (In passing, in a Korean work compiled in the 15th century for the purpose of training official interpreters in Japanese the *Iroha* is found, but the *a-i-u-e-o* listings of the Japanese syllabary.)

In old days Koreans learned, or memorized, the Korean script like a syllabary. system as *ka na ta ra ma pa sa a cha ch'a k'a t'a p'a ha* for the consonants and *ka kya kŏ kyŏ ko kyo ku kyu kŭ ki ka; na nya nŏ nyŏ no nyo nu nyu nŭ ni na*; *ta tya tŏ tyŏ to tyo tu tyu tŭ ti ta* .... for the vowels preceded by consonants.

2d. Arbitrary or unknown systems

It seem to me that Professor Salomon has put a considerable amount of time and efforts for this section, especially on the alphabetic order found in the scripts derived from North Semitic. His presentation would be very informative for those who are interested in this subject. The discussant hopes to see in the near future Professor Salomon's conclusive as well as persuasive explanations on the 'alphabetic order' of the Roman alphabets, currently the most widely used writing in the world.