## Alphabetic and logo-syllabic scripts: A contradiction between form and content

S. K. Venkatesan

**Abstract**

There were two paths of evolution in writing — one is the modern alphabetic/abugida scripts (abugida being the older consonant-based writing systems used even now in Indic languages, where vowels are indicated through diacritics) and the other is the older logo-syllabic script that is still in wide use in China and Japan. The logo-syllabic script traditionally had strong ideographic tendencies that have been eroded over time as cultural changes produced a semantic diffusion, and the association between signifier (the form) and the signified (the content) started diverging over time.

In this paper we shall study the semiotics of the two writing methodologies from the perspective of economy of size, in terms of the number of glyphs required and the complexity of the glyphs. We illustrate these concepts using the two divergent paths taken by Linear Elamite script and the Indus Valley script.

### 1 Cave painting and art

The old paleolithic factories of the proto-humans produced many wonders of the modern world, including the hands that produced the tools and the brain that developed in relation to the wide availability of food that became possible with the creation of the tools. The limitations of human memory necessitated an external system for memory preservation. The earliest memory devices were cave paintings and art which had the focus on the immediate and now. It conceived hunting scenes as a way of enhancing consciousness, leading to a tantric effect on the planning and focusing on the outcome of hunting. In this direct early form of communication, the story telling was a direct artistic rendering of scene, an enacted drama on the wall. This had an enhancing effect on the consciousness of the brain, improving the outcome, thus causing a reinforcing effect on the rituals.

### 2 Convolutional neural networks and cave paintings

The illustrations and line art on the wall are the abstract perception of a real drama in physical space, as an impressionistic imprint on the wall. The modern analogy is the perception by an artificial deep neural network that captures it using a convolutional neural network.

**Figure 1**: Cave painting from France.

**Figure 2**: Graphics filtered to line-art by convolutional neural network.

Like modern humans, the ancient cave dwellers also had about 200 trillion synaptic neural connections in the cerebral cortex. Assuming that these had the capacity for abstraction and generalization like modern deep neural networks, one can conclude that these illustrations eventually lead to the language of ideograms, quite independent of speech.

These were mostly about the here and now, but were effective in developing consciousness about their actions. This displaced perception in space and time worked like a ritual that would have tantric effects on the outcome of the event, e.g., the hunting of a bull.

### 3 Evolution of speech

The hand that produced the tools also engaged in gestures and sounds leading to the flexibility of the larynx, which further led to a novel mode of communication — speech.

Early understanding of language by the Roman empire was inverted, based incorrectly on the concept of the Tower of the Babel (as elucidated in the Bible), by which it was assumed that all languages were created magically in its full beauty and elegance, in a sudden act of creativity (the modern version being by a sudden act of genetic mutation), and that all languages that exist now are all imperfect diffusion and corruption of the original perfect Platonic form.



doi.org/10.47397/tb/46-2/tb143venkatesan-writing

It was Otto Jespersen [1] who first showed how the early languages like Gothic and all the other primitive languages of Africa were highly complex with an incredible amount of variation. The Gothic word "may", for example, had 30 different variations, which now has been reduced to just one form "may". He showed how these wild variations were tamed over time and simplified to the present modern form.

Quoting Jespersen [1]:

> "People will often use the expressions 'evolution' and 'development' in connexion with language, but most linguists, when taken to task, will maintain that these expressions as applied to languages should be used without the implication which is commonly attached to them when used of other objects, namely, that there is a progressive tendency towards something better or nearer perfection. They will say that 'evolution' means here simply changes going on in languages, without any judgment as to the value of these changes.
>
> But those who do pronounce such a judgment nearly always take the changes as a retrogressive rather than a progressive development: "Tongues, like governments, have a natural tendency to degeneration," said Dr. Samuel Johnson in the Preface to his Dictionary, and the same lament has been often repeated since his time...
>
> ... it is evident *that language ranks highest which goes farthest in the art of accomplishing much with little means, or, in other words, which is able to express the greatest amount of meaning with the simplest mechanism.* (all caps)
>
> ... The only thing I maintain is that the sum total of these changes, when we compare a remote period with the present time, shows a surplus of progressive over retrogressive or indifferent changes, so that the structure of modern languages is nearer perfection than that of ancient languages, if we take them as wholes instead of picking out at random some one or other more or less significant detail."

Most modern languages have been tamed and now have only around 30–40 phonemes. The wilder primitive languages of Africa, like the South African language, Khosa, have many more phonemes, including the famous "click" sounds.[1]

It is not just a mere coincidence that these click sounds are similar to the clipping sound of stone tools. It is an open question on how to represent them in a writing system. It has also been suggested to use click sounds for punctuation like comma and full-stop [2].

## 4 Ideographic and logo-syllabic writing systems

The first initial attempt at creating a writing system was wedded to the direct semantic-graphemes in the ideographic system, which was a separate graphical system independent of speech. The evolution of speech as a parallel stream of consciousness put pressure on the ideographic system, leading to an association of syllables (molecules of speech) to ideograms. This system then started loosening its grip on the ideographic association, due to rapid cultural changes and diffusion/dissipation, thus producing the stable logo-syllabic system in the Bronze Age.

## 5 Abugida and alphabetic writing systems

A new writing system was initially produced by cutting all its ideographic roots from its logo-syllabic system, producing a phonogram writing system, an intermediate stage before the arrival of alphabetic systems. At the time of the Iron Age, a new revolutionary consonant-based abugida system was conceived, which at the first stage did not produce full explicit representation for the vowels, but indicated them through minor embellishments on the consonants. This consonant-based system is known as an abugida writing system, which is even now widely used in India. At later stages, the vowels were given explicit status as in the Greek and Roman alphabetic systems, thus making vowels first-class citizens of the alphabetic system. The Greek scripts were adopted and evolved by the Byzantine empire, producing the Cyrillic alphabets, while the Roman empire created the Roman scripts, starting initially with the capital letters; modern alphabets involving both the upper- and lowercase letters came later [2].

## 6 Language and grammar

Although like everything else, language and its grammar evolves over time in society, it becomes so dominating that sometimes it is mistaken for the real thing, the underlying "signified", the real objects of life for which the "signifier" is a proxy. The syntactic structure of language in its beauty and perfection is mistaken for the underlying reality that it is a reflection of, the real physical objects of the world. MRW Dixon, who formulated a "Practical Linguistic Theory" [3] for practicing linguists to study foreign languages, has also studied English from a semantic

[1] Click sound: youtube.com/watch?v=WHHGOYu6Fl0


**Figure 3**: Map of Elam.

perspective [4]. Of course, English is not a monolithic construct, as it draws from Norwegian, French, and from its Anglo-Saxon local roots, all pulling in different directions. A study that does not consider this heterogeneity is an exercise in prescription rather than description for the English language. In any case, Dixon's [4] attempt is a novel attempt at extricating its semantic roots, as one should not mistake form for content.

## 7 Linear Elamite script

Here we make a brief study of Linear Elamite as deciphered by François Desset and his team [5]. Although this decipherment has not been widely accepted, we still think there are sufficient elements of reality in this decipherment that makes it useful for our theoretical study here.

Linear Elamite writing is considered as being purely phonographic, produced as an alternative to the logo-syllabic system adopted earlier. The advantage with this method is that it requires one to remember fewer characters, thus increasing literacy. However, the disadvantage is that it requires more effort to write as it requires a larger number of (single-sound) characters to describe the same content, while logo-syllabic systems require fewer (ideographic) characters.

## 8 Indus Valley script

Here we also make a brief study of the Indus Valley script (or simply Indus) as deciphered by the author in [6]. Although this decipherment has not yet been accepted by the academic community, we still think that there are sufficient elements of reality in this decipherment that make it useful for our theoretical study here. Twenty years back, we reported in this journal a similar study [7] of language scripts in relation to artificial languages and proposed a concept of word-embedding using SenseTEX. This area has now blossomed into GenAI and large language models. In that study [7] we also briefly mentioned relationship between the logo-syllabic Indus script and the Chinese Han characters.

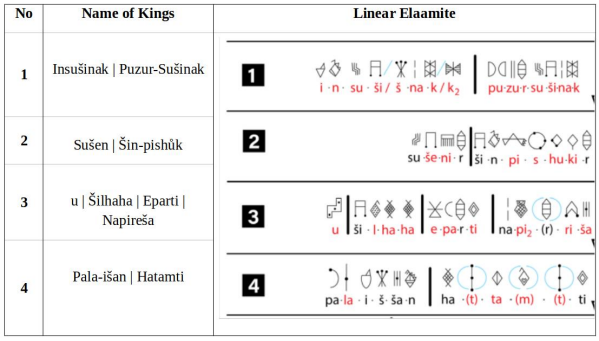


**Figure 4**: A snapshot of the decipherment of Linear Elamite by François Desset and his team.


**Figure 5**: Indus script indicating flooding of houses.

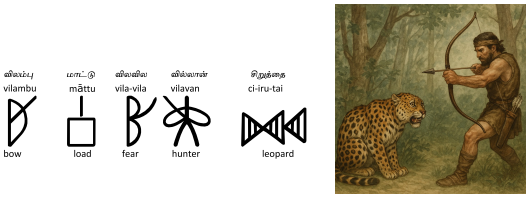


**Figure 6**: Indus script indicating hunting of a leopard.

## 9 Comparison of Linear Elamite and Indus scripts

Andreas Fuls [8] has made a comparative study of both these scripts. He has concluded that Linear Elamite and Indus are very different approaches to writing systems. The Linear Elamite script consists of phonograms progressing slowly towards an alphabetic system, while the Indus script is deepening and developing a logo-syllabic approach. Fuls identifies the Indus script as:

| M77 No | Indus Sign | Dravidian syllable | Elamite sign | Elamite phoneme (Desset) |
|---|---|---|---|---|
| 261 |  | kō |  | ha |
| 210 |  | mēyccal |  | li |
| 233 |  | mumalai |  | hi |
| 397 |  | aniyal, mani |  | la |
| 176 |  | añcal, vanji |  | ni |
| 147 |  | nān-cēri |  | š |
| 321 |  | ñāṇ |  | n |
| 299-287 |  | valai |  | pa |
| 192 |  | paḷḷi |  | po |
| 87 |  | iru |  | zu |
| 245 |  | taṭṭi |  | zo |
| 379 |  | pammu |  | r |
| 135 |  | peru-vaṟi |  | re |

**Figure 7**: Signs common between Linear Elamite and Indus. After [8, fig. 9].

> "Indus culture used a highly mixed logographic-syllabic writing system with many signs representing a logogram, or word root."

Linear Elamite, on the other hand, is a more advanced writing system where its ideographic roots have been discarded (probably deliberately) and represent phonograms that are an intermediate evolutionary stage that lead to alphabets. The fact that phonograms represented vowels as well as consonants is a major advancement from the abugida stage of development. In this regard, it must be mentioned that Ferdinand de Saussure explicitly mentions that the "signifier" must be arbitrary, thus deliberately chopping off its ideographic roots. Of course, it could have also happened through cultural displacement and diffusion/dissipation.

Figure 9 in [8] shows signs that are common between Linear Elamite and Indus scripts. We reconstruct this table (figure 7, next page) with Dravidian decipherment in [6] which uses M77 sign numbers introduced in [9].

This table clearly shows there is little or no correlation between the logo-syllabic values in Linear Elamite and Indus scripts. We can assume that a Saussurean discontinuity and abrupt breaking with ideographic connotations is being applied deliberately here to evolve modern phonograms. This can be considered as an early attempt at evolving a modern alphabet like Greek and Roman alphabets that included both vowels and consonants. The Greek and Roman alphabets evolved from the early Phoenician writing system that used only consonants leaving the vowels as being implicit, the abugida writing system that languages in India use even now.

## 10 Conclusion

Once the ideographic roots of logo-syllabic signs have been cut, a script is then ready for currency in the world of Phoenician merchants for wider usage in trade, like free migrant labor. The Chinese and Japanese scripts, on the other hand, have persevered hard to keep their ideographic roots, marrying them to phonetic attributes, thus keeping the connection between the old and new. This also keeps the connection between the direct cave painting of the early humans and the modern alphabets, the umbilical cord between form and content, a difficult task indeed.

## References


[1] Otto Jespersen (1922), *Language: Its Nature, Development and Origin: Exploring the Evolution and Complexity of Language*, Good Press, 2019

[2] Kumaran Sathasivam, S.K. Venkatesan and Yakov Chandy (2017), Revealing semantics using subtle typography and punctuation, *TUGboat* Vol. 38, No. 2, pp. 165–170, tug.org/TUGboat/tb38-2/tb119sathasivam.pdf

[3] MRW Dixon (2019), *Basic Linguistic Theory. Volume 1. Methodology*, Oxford University Press

[4] MRW Dixon (2005), *A Semantic Approach to English Grammar*, Oxford University Press

[5] François Desset, et al. (2022), The Decipherment of Linear Elamite Writing, *Zeitschrift für Assyriologie*, Vol. 112, No. 1, pp. 11–60

[6] Venkatesan, S.K. (2025), A Dravidian decipherment of Indus script, github.com/Sukii/decipher-ivc

[7] Venkatesan, S.K. (2005), Moving from bytes to words to semantics, *TUGboat*, Vol. 26, No. 2, 165–168, tug.org/TUGboat/tb26-2/venkat.pdf

[8] Andreas Fuls (2024), Comparison of Linear Elamite and Indus Writing Systems, *Iranian Journal of Archaeological Studies*, Vol. 14, No. 1

[9] MASI 77 Indus Script: Texts, Concordances & Tables of Iravatham Mahadevan, indusscript.in



◇ S. K. Venkatesan
TNQ Technologies, Chennai