On The Brahmi-Indus Connection (A new passage to Indus Script Decipherment)

Sai Venkatesh.B

Centre for Nonlinear Science and Engineering, SASTRA University, Thanjavur, India.

s.venky30@gmail.com, +91-9940784994.

ABSTRACT:

The origins of the Indus Valley Civilization are shrouded in mystery. Two significant questions from this unsolved historical problem are the origins of the Brahmi and Indus scripts. In this light, the present paper performs a omparative feature based study of the Brahmi, Indus and Aramaic scripts, decisively proving that the Brahmi is indeed more closer to the Indus script than to Aramaic. Based on this result, and other established partial decipherments thus far, a hypothesis proposing the two-subsets nature of the Indus script is proposed. Separate phonetic and pictographic sets corresponding to Sanskrit and Tamil are identified, pointing towards a new way of comprehensively deciphering the Indus script.

KEYWORDS:

Brahmi Script, Indus Script, Aramaic Script, Feature Analysis, Aryan and Dravidian Languages.

INTRODUCTION:

It is an undisputable fact that the Indus Valley Civilization of the Indian Subcontinent stands at par with other major Civilizations around the world such as Egypt, Mesopotamia, China and Meso-America [1]. Among various contributions to humanity, the Indus civilization, and hence ancient India will most popularly be remembered for the following:

1. Urban Planning and Flawless Sanitation Design [2].
2. Precise calculations in Astronomy and Cosmology [3].
3. The contribution to Mathematics, including the number system and various other geometric theories [4].
4. Depth in Philosophy, giving rise to the Dharmic religions of Hinduism, Buddhism and Jainism [5].
5. Languages and the Brahmi Abugida (syllabic alphabet system), giving rise to a large number of writing systems used today in South and Southeast Asia [6].

Apart from the contributions listed above, the study of Indus valley Civilization and its associated cultures has gained significant momentum in recent years. This is due to the fact that the Genographic 'Out of Africa' theory places India at a crucial crossroad of human migration [7]. Also, the etymological and linguistic studies linking the ancient Indian languages of Sanskrit and Tamil to other macrolanguage

families such as the Semitic, Japonic, Pama-Nyungan, Mayan, Ural-Altaic, Caucasian and Chukchi, along with new archaeological evidences, each predating the previous, suggest that the Indus Valley Civilization may have had a significant impact on the development of Human Civilization and Culture on a global scale [8-15].

In spite of all the above mentioned significances, it is rather ironic that most of the origins of the Indus Valley Civilization and ancient India remain shrouded in mystery. This is evident from the fact that the dating of the ancient Indian texts, the Vedas have been suggested a wide range spanning millennia from as early as 10,000BC to as late as 1500BC [16-17]. The enigmas surrounding the Indus Valley Civilization are well represented by two mysteries – The Origin and Nature of the Indus Script, and The Origin of the Brahmi Script.

The various theories proposed regarding the origins of Brahmi script can be broadly classified into two categories. The First set of theories claim a West Asian origin of the Brahmi, most probably from the Aramaic alphabet, and such theories rely on the visual similarities between the two scripts [18-19]. The Second set of theories postulate an entirely indigenous origin of the Brahmi, with one sub-sect claiming that Brahmi originated from the Indus Script, the rationale of this claim being that visual similarities exist between Brahmi and Indus scripts and the fact that the dating of edicts in Brahmi are nearly a millennia after the earliest dating of Indus script artifacts [20-21].

With regards to the Indus Valley Script, three categories of theories abound:

1. The investigations carried out by Parpola, Mahadevan and the like suggest a pictorial mapping of Indus Script to an early archaic variant of the Tamil Dravidian Language [22-24].
2. The mapping of certain Brahmi alphabets to the Indus script by frequency of occurrence by Subash Kak and Hunter suggest a phonetic Aryan basis of the Indus Script [25-26].
3. A third category of theories suggest that the Indus Script are nothing more than a collection of fanciful symbols and do not represent any language [27]. However, this theory was disproved by a Markov analysis of the script done by Rajesh Rao [28].

In light of these various theories about the Brahmi and Indus Script, it is intuitively felt that an investigation into the connection between Brahmi and Indus script may clear the air significantly and that there might be a single comprehensive way to correctly decipher the Indus Script. Taking cue from this insight, the present article firstly formulates and performs a feature-based study of the Brahmi, Indus and Aramaic Scripts. The results of such a study reveal that the Brahmi script is indeed closer to the Indus Script, as Hunter had proposed [25-26]. Then, a brief glimpse of the results of Parpola is given [22-24]. Finally, using the results of Hunter and Parpola, a novel hypothesis is formed, whose significant bottomline is that the Indus Script might possibly contain two ‘subscripts’ – one phonetic and one pictographic, used in different contexts for ancient Sanskrit and Tamil respectively. The results and hypothesis proposed in this paper might open the doors for a more comprehensive, more accurate deciphering of one of the biggest unsolved mysteries in human history – The Indus Script.

FEATURE-BASED STUDY OF THE BRAHMI, INDUS AND ARAMAIC SCRIPTS

One of the most significant ways of characterizing various writing scripts and mapping their inter-relationships is to perform a feature analysis of such scripts. Significant literature exists for such analyses, and the 'features' that are typically used therein are strokes, vertical, horizontal and oblique lines, curves, dots and enclosures [29-30]. However, it is understood that due to geographical and evolutionary reasons, scripts may significantly distort over time. An instance of this is the case of extremely 'curvy' natures of the Kannada and Telugu scripts derived from a significantly less-curvy Brahmi script, owing to the writing of the former scripts on palm leaves [31]. Thus, an estimate of strokes, lines and curves may not directly map to the evolution of one script from another.

For the above mentioned reason, the present work defines a different classification of 'features'. Specifically, five classes of features are defined, as follows:

- Class A: Forks, Branches, 'T's, Dead-Ends.
- Class B: Intersections, 'X's, '+'s, Stars.
- Class C: Enclosures, Circles, Boxes.
- Class D: Extensions, Straight Lines, Bends.
- Class E: Isolations, Small Lines, Dots.

An illustration depicting the above mentioned feature classes is shown in Fig. 1.

The feature based comparative analysis of any two alphabet systems X and Y using the five classes can be performed using the following procedure:

1. The number of occurrences of each class of features in the alphabet is noted, and this is represented as a five tuple (A,B,C,D,E). For instance, the alphabet 'P' has an extension and an enclosure, giving the five tuple (0,0,1,1,0) [32].
2. The five tuple values of each alphabet can then be used to map the alphabet uniquely to a point in 5D space.
3. The 'distance' between this alphabet 'i' and its counterpart 'j' in any other writing system can be given by the formula,

$$D_{ij} = \sqrt{(A_i - A_j)^2 + (B_i - B_j)^2 + (C_i - C_j)^2 + (D_i - D_j)^2 + (E_i - E_j)^2}$$

4. D_{ij} gives the distance between the alphabets i and j. The lower the value of D_{ij}, the closer is the match between the features of the alphabets.

The above mentioned feature analysis is done to compare the closeness of Brahmi with the Aramaic and Indus Scripts. Hunter's analysis and identification of the Indus Scripts most closely resembling the Brahmi alphabets are used [25-26]. The distance between Brahmi and Aramic Scripts is denoted by B-A and the distance between Brahmi and Indus script by B-I.

It is noteworthy that while the Brahmi script has more than 40 'alphabets', the Aramaic alphabet has only 25. Thus, the feature analysis for those alphabets phonetically common to both Brahmi and

Aramaic are presented in Fig.2, which also contains the feature analysis for the corresponding Indus Script signs. The sign number, according to the Mahadevan Classification system is also given [22-24].

The average values of B-I and B-A are obtained as 0.4 and 1.46 respectively from Fig. 2.

Thus Fig. 2, decisively confirms the fact that the Brahmi script is indeed closer to the Indus Script than the Aramaic script, true to the predictions made by Hunter.

For the sake of completion, the remaining Brahmi alphabets and their Indus Script Sign matches are shown in Fig. 3.

Also, a frequency analysis of the Indus Signs listed in Fig. 2 and Fig. 3 are performed. The 'frequency' is defined by the number of occurrences of these signs seen in various seals from Harappa, Mohenjodaro, Lothal and other archeological sites [22]. It is found that the grand total of the occurrences of the 40-50 Indus Characters is 1075, amounting to 21.1% of all the Indus character occurrences.

At this point, it is imperative to review in brief, the primary conclusions made by Hunter and Subash Kak [25-26]:

1. Brahmi is derived from the Indus Script, a syllabary of open and closed syllables.
2. Sabaean, Phoenician and Cypriot writing systems might have been derived from the Indus Script.
3. The basis of the Brahmi and Indus Scripts are phonetic, rather than pictographic.

A NEW HYPOTHESIS

Though the discussion in the previous section validates Hunter's conclusions and confirmed the Indus-Brahmi relation, it asks more questions than it answers. For instance, Parpola and Mahadevan have extensively classified the various Indus Script Signs and have used Dravidian language Tamil combined with pictographic-rebus interpretations to successfully decipher some of the signs [22-24].

The key results from the Parpola-Mahadevan studies are as follows:

1. The Indus script represent a logo-syllabaric system of writing.
2. The language of the Indus valley is Dravidian and might have influenced to other languages of central and west Asia.
3. Rebus techniques were used to represent phonetic equivalents to pictograms.

A glimpse of these key points can be obtained by understanding a small portion of the decipherments made by Parpola, as illustrated in Fig. 4.

In order to reconcile with the interpretations of Parpola-Mahadevan and Hunter-Kak, it is essential to formulate a new hypothesis. Conjectured ***'The Sai Hypothesis'***, the statements are as follows:

1. ***The Indus Script consists of two subsets of scripts, termed for convenience as Subset 'S' and Subset 'N'.***
2. ***Subset 'N' consists of about fifty signs, has a direct correspondence with the Brahmi alphabet, and is phonetic in nature.*** This subset comprises of the signs depicted in Fig.2 and Fig.3 which are the ones analyzed by Hunter and Kak [25-26]. These signs were used by the Indus people when representing phonetically accurate content such as proper nouns, names and sacred chants/mantras. Not surprisingly, most of such content is in the Sanskrit language.
3. ***Subset 'S' consists of more than 500 signs, logographic/pictographic in nature and representing the Dravidian language.*** The signs belonging to this subset correspond to the investigation by Parpola and Mahadevan, and represent a combination of pictographic and logographic signs, sometimes using rebus techniques [22-24]. The Indus people used such signs for normal communication items such as common nouns, verbs, adjectives and adverbs, quantity, gender etc. Thus, this subset is entirely in the Tamil language. This statement makes sense, since the content represented by subset 'S' is not phonetically exacting, and Tamil is significantly easier to attain fluency than Sanskrit, thanks to the simple alphabetic structure.

CONCLUSION

An investigation into the closeness of the Brahmi script with the Indus and Aramaic Scripts is performed using a feature based analysis, and the results reveal that Brahmi is indeed closer to Indus than Aramaic. Further, the key results of Mahadeva-Parpola and Hunter-Kak are briefly reviewed, and these results, along with the feature analysis results are used to formulate a novel hypothesis, the Sai Hypothesis, whose main feature is that the Indus Script consists of two Subsets, 'N' and 'S', representing phonetic Sanskrit and pictographic Tamil characters. This hypothesis, like most other hypotheses will have to withstand the test of time. Nevertheless, the hypothesis does provide new directions to deciphering the Indus Script building on the successes of the Hunter-Kak and Parpola-Mahadevan results, thus solving one of the most unsolved mysteries of human history.

REFERENCES:

[1] J. McIntosh, The Ancient Indus Valley: New Perspectives, ABC-CLIO, 2008.
[2] R. P. Wright, The Ancient Indus: Urbanism, Economy, and Society, Cambridge, 2010.
[3] S. M. Ashfaque, Astronomy in the Indus Valley Civilization: A Survey of the Problems and Possibilities of the Ancient Indian Astronomy and Cosmology in the Light of Indus Script Decipherment by the Finnish Scholars, Centaurus, 21, 1977.
[4] T. A. S. Amma, Geometry in Ancient and Medieval India (Google eBook), Motilal Banarsidass, 2007.
[5] A. Hiltebeitel, The Indus Valley "Proto-Siva", Reexamined through Reflections on the Goddess, the Buffalo, and the Symbolism of vahanas, Anthropos, 73, 1978.
[6] S. P. Gupta, K. S. Ramachandran, The Origin of Brahmi script, D. K. Publications, 1979.

[7] K. TallBear, Narratives of Race and Indigeneity in the Genographic Project, Journal of Law, Medicine and Ethics, 35, 2007.
[8] D. McAlphin, Elamite and Dravidian: Further Evidence of Relationship, Current Anthropology, 16, 1975.
[9] A. Toth, Are all agglutinative languages related to one another?, Mikes International, Holland, 2007.
[10] M. S. Victor, The Babylonian Thamizh, Naller Publications, India, 2007.
[11] S. Ohnu, Nihon-Dravida hikaku gol, nu-hamo, Iwanami, Tokyo, 1984.
[12] A. Butcher, Australian aboriginal languages: consonant-salient phonologies and the place-of-articulation imperative, Psychology Press, USA, 2006.
[13] I. Mahadevan, Tamil-Brahmi Inscriptions, State Department of Archaeology, 1970.
[14] V. Begley, Arikamedu Reconsidered, American Journal of Archaeology, 87, 1983.
[15] F. R. Allchin, D. Erdosy, The Archaeology of Early Historic South Asia: The Emergence of Cities and States, Cambridge, UK, 1995.
[16] K. Chandler, Modern science and Vedic science: An introduction, Consciousness-Based Education and Literature, 491, 1987.
[17] H. Jacobi, On the date of the Rig Veda, Indian antiquary, 23, 1894.
[18] P. T. Daniel, W. Bright, The World's Writing Systems, Oxford, 1996.
[19] G. Buhler, On the Origin of the Indian Braha Alphabet, Indian Studies, Strassburg, 1898.
[20] R. Salomon, Review: On the Origin of the Early Indian Scripts, J. American Oriental Society, 115, 1995.
[21] C. Maloney, The Beginnings of Civilization in South India, J. Asian Studies, 29, 1970.
[22] I. Mahadevan, The Indus Script: Texts, Concordance and Tables, Archaeological Survey of India, 1977.
[23] I. Mahadevan, Dravidian Models of Deciphernent of the Indus Script: a case study, Tamil Civilizations, 4, 1986.
[24] A. Parpola, Deciphering the Indus Script, Cambridge, UK, 1994.
[25] S. C. Kak, A Frequency Analysis of the Indus Script, Vivekananda Kendra Patrika, 40, 2011.
[26] S. C. Kak, Indus and Brahmi - Further Connections, Cryptologia, 14, 1990.
[27] R. P. N. Rao, N. Yadav, M. N. Vahia, H. Joglekar, R. Adhikari, I. Mahadevan, Entropic Evidence for Linguistic Structure in the Indus Script, Science, 324, 2009.
[28] R. P. N. Rao, N. Yadav, M. N. Vahia, H. Joglekar, R. Adhikari, I. Mahadevan, A Markov model of the Indus script, PNAS, 106, 2009.
[29] M. A. Changizi, S. Shimojo, Character complexity and redundancy in writing systems over human history, Proc. R. Soc. B, 272, 2005.
[30] B. Primus, A featural analysis of the Modern Roman Alphabet, Written Language and Literacy, 7, 2004.
[31] D. Anirban, Origin of Sinhalese Script, Mrinmoy Chakraborty: 22.
[32] M. S. Kim, K. Shimada, Geometric Modeling and Processing - GMP 2006, Springer, 2006.

FIGURES:

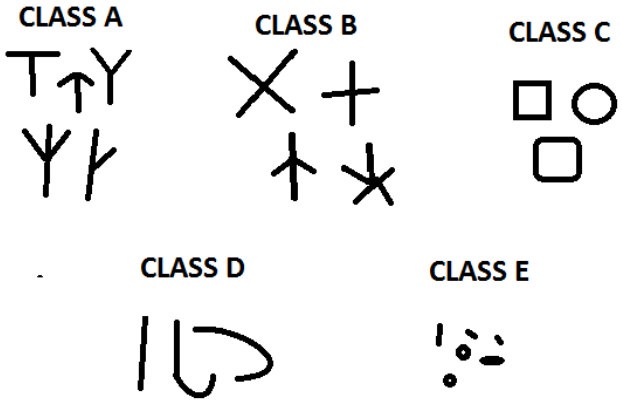

Figure 1 The Five Classes of Features

BRAHMI	A	B	C	D	E	INDUS	A	B	C	D	E	ARAMAIC	A	B	C	D	E	B-I	B-A
A	2	0	0	1	0	584	2	0	0	1	2	ALAP	0	1	0	0	0	2	2.45
V	0	0	1	1	0	469	0	0	1	1	0	WAW	0	0	0	1	0	0	1
E	0	0	1	0	0	309	0	0	1	0	0	`E	0	0	0	1	0	0	1.41
K	0	1	0	0	0	7	0	1	0	0	0	KAP	0	0	0	1	0	0	1.41
KH	0	0	0	2	0	237	0	0	0	2	0	QOP	0	0	0	2	0	0	0
G	0	0	0	1	0	235	0	0	0	1	0	GAML	0	0	0	1	0	0	0
GH	0	0	0	4	0	432	0	0	0	4	0	HETH	0	0	0	3	0	0	1
JH	0	0	0	4	0	334	0	0	0	4	2	ZAIN	0	0	0	1	0	2	3
S	0	0	0	3	0	242	0	0	1	3	0	SAD	1	0	0	1	0	1	2.24
H	0	0	0	3	0	79	0	0	0	3	0	HE	0	0	0	3	0	0	0
T	1	0	0	0	0	234	1	0	0	0	0	TAW	1	0	0	0	0	0	0
TH	0	0	1	0	1	346	0	0	1	0	1	TET	0	0	0	2	0	0	2.45
DH	0	0	1	0	0	561	0	0	1	0	0	DALT	1	0	0	1	0	0	1.73
N	1	0	0	0	0	523	1	0	0	0	0	NUN	0	2	0	0	0	0	2.24
P	0	0	0	1	0	293	0	0	0	1	0	PE	0	0	0	1	0	0	0
B	0	0	1	0	0	522	0	0	1	0	0	BETH	1	0	0	1	0	0	1.73
M	0	1	1	0	0	131	0	1	1	0	0	MEM	0	1	0	0	0	0	1
Y	2	0	0	1	0	308	2	0	0	1	0	YOD	0	0	0	1	0	0	2
L	0	0	0	3	0	307	1	1	0	1	0	LAMD	0	0	0	1	0	2.45	2
S'	1	0	0	0	0	241	1	0	1	0	0	SEMK	0	0	0	3	0	1	3.16
SH	0	1	0	2	0	440	0	1	0	2	0	SHIN	1	0	0	1	0	0	1.73
AVERAGE DISTANCE																		0.4	1.46

Figure 2 Feature Based Comparative Analysis of Indus, Brahmi and Aramaic alphabets

BRAHMI	INDUS	BRAHMI	INDUS	BRAHMI	INDUS	BRAHMI	INDUS
AA	3	O	549	NJ	79	N.	175
I	206	NG	475	T.	572	D	288
EE	218	C	581	TH.	348	PH	73
U	419	CH	520	D.	544	BH	582
AI	77	J	434	DH.	83	R	39

Figure 3 Indus-Brahmi mapping for non-Aramaic characters

SIGN	PICTORIAL-PHONETIC MEANING
	FISH (MEEN)
	ROOF+FISH (MEIMEEN)
	DIVIDE+FISH (PACUMEEN)
	SPACE+FISH (VELLIMEEN)
	FIG+FISH (VADAMEEN)
	FIG+SPACE (VADAVELLI)
	SIX+FISH (CHAARUMEEN)

Figure 4 Pictographic analysis of the Indus Script

Table 1 Frequency Analysis of the Indus Signs considered for the feature analysis

INDUS CHARACTERS WITH ARAMAIC PARALLELS				INDUS CHARACTERS WITHOUT ARAMAIC PARALLELS			
SIGN NO.	OCCURENCE	SIGN NO.	OCCURENCE	SIGN NO.	OCCURENCE	SIGN NO.	OCCURENCE
584	1	346	17	003	47	079	7
469	99	561	8	206	11	572	35
309	4	523	1	218	1	348	16
007	20	293	11	419	6	544	36
237	23	522	1	077	1	83	1
235	4	131	2	549	1	175	4
432	1	308	5	475	27	288	855
334	1	307	12	581	1	73	11
242	43	241	157	520	2	582	1
79	7	440	1	434	1	39	1
234	9						
TOTAL OCCURENCES = 430				TOTAL OCCURENCES = 1075			
GRAND TOTAL = 1505 OCCURENCES = 21.1% OF A TOTAL OF 7131 OCCURENCES FOR ALL CHARACTERS							