Recent Advances in the Study of the Indus Script*

Iravatham Mahadevan
Joint Secretary
Ministry of Industries, New Delhi.

(Received on 7th March 1979)

1. INTRODUCTION

The Indus Script still remains undeciphered. However, there is a growing understanding of the script, thanks to the steadily increasing number of published texts as well as recent compilations of concordances and statistical tabulations (Parpola et al.: 1973; Mahadevan: 1977). It has now become possible to establish some objective results which, while they still fall far short of an actual decipherment of the script, can nevertheless serve to indicate the direction of further research. This paper deals with some of these results obtained mainly on the basis of my recently published work 'The Indus script: Texts, Concordance and Tables', (Archaeological Survey of India, 1977), referred to hereafter as ISTCT. I shall confine myself in this paper to two fundamental questions relating to the Indus Script, namely, direction of writing and word division, concentrating more on methods rather than on results. Text numbers are cited as in ISTCT.

2. DIRECTION OF THE INDUS SCRIPT

2.2 The direction of the script can be discovered by a study of the external features of the writing like cramping or overflow of end signs, as well as from the internal evidence of sign sequences. It is my view that a conclusive proof of the direction of the script can be obtained only by a combination of both the approaches and by a comparative statistical study of the frequency and distribution of signs.

2.3 In the study of external features of the writing, an early lead was given by Marshall (1931: 40) who pointed out the overflow of one sign at the left end of the text in 1052 (pl. I·A) and the cramping of the left end sign in 1060, both indicating a right to left direction. He also showed that the one and half lines of text in 1247 (pl. I·B) written as a loop opening to the right (thus: ) must be in the boustrophedon mode with the first line starting at the right and the second line continuing from the left. Gadd and Smith (Marshall: 1931: 410) showed that the disposition of the adial line 4254 (pl. I·C) along the top, left and bottom edges of the rectangular side (thus: ) proved that the inscription ran in an anti-clockwise, that is right to left direction. Alekseev (in Knorozov et al.: 1965, 1968; tr. by Zide and Zvelebil: 1976: 18) has collected a number of examples where the signs at the left end of the lines undergo occasional cramping, diminution in size, angular rotation or vertical displacement suggesting that the writing terminated at the left. B. B. Lal (1966: 52)

produced the most important piece of external evidence by demonstrating from a study of overlapping incisions on pottery graffiti (pl. I-D) from Kallibangan that the inscriptions in question must have been incised from the right.

2.4 It is, however, necessary to stress that such external evidence does not by itself constitute conclusive proof of the direction of writing. The limitation of the method lies in the fact that we know, by mere observation, that lines of the script run in either direction, as may be seen from the examples in Fig. 1. Hence all the external characteristics relied upon to prove a direction are, at least in theory, reversible and will, in such cases, lead to the opposite conclusion unless the results are correlated to the internal evidence provided by sign sequences.

2.5 Let me illustrate this point, with special reference to Lal's method of overlapping incisions on pottery graffiti. Fig. 2 shows two pairs of identical inscriptions. In each pair, one inscription is from a seal (impression) and the other is incised on pottery, and the lines in each pair run in opposite directions. If these pottery graffiti had overlapping incisions, they must indicate a direction of writing which is known to be at variance with that occurring on the vastly greater number of sales (impressions). This might lead to an erroneous conclusion, if the normal direction of the script is sought to be deduced merely from the overlapping incisions. In such cases, the results should be corrected by applying the test of sign sequences which will show that these graffiti are incised in the reversed direction.

2.6 The method of overlapping incisions to determine the direction of the script suffers from yet another, and somewhat surprising, limitation. It now appears that it is even possible for the direction of writing to be at variance with the direction of reading. Evidence for this curious situation is furnished by B. B. Lal himself (1967-68: 15), though he does not seem to be aware of the problem. In this paper Lal has published an inscribed potsherd excavated by him at Kallibangan. The inscription* (Pl. I-E) consists of

---

*This is included in ISTCT as no. 8221. Unfortunately the text has been incorrectly copied, as (as unusual variant of ) was confused with (as unusual variant of ) and (as unusual variant of ) with (as unusual variant of ) owing to the indeterminate length of the strokes in the original. As a result, the evidence of sequence was not perceived and the text was copied from the right on the basis of overlapping incisions. The text has to be corrected as indicated in this paper (Fig. 3).

...foursigns written with bold and deep incisions made before firing the pottery. Lal has made an analysis of the sequence of overlapping incisions between the two middle signs as well as of the component parts within each sign and proved that the inscription is written from right to left. However, the test of sign sequences makes it certain that this inscription (as incised in the original) has to be read from left to right. The direction of this text becomes evident when the sign pairs occurring here are compared with those in the Table of pairwise frequencies in ISTCT (Table II: 724). Similarly the sign group formed by the three signs at the left (of the original) can be compared with identical or similar sign groups listed in ISTCT (Concordance: 466-67). The results are shown in Fig. 3. It will be seen that only a left to right reading (of the original) is productive. It may also be mentioned here that the pairs and the triplet arranged in the reversed order (that is, read from the right in the original) are not recorded even once in ISTCT. It is therefore certain that this inscribed sherd from Kallibangan has to be read from the left (in the original) even though it may have been incised from the right as shown by Lal. A plausible explanation of the riddle may be that the potter or the scribe copied the text from a seal (original) which would have been engraved from the left, but wrote from the right according to the normal practice in direct writing. Whatever may be the true explanation, the point to be noted here is that the results of the study of overlapping incisions cannot be applied mechanically to determine the direction of the script without due regard to internal evidence provided by the sign sequences.

2.7 Hunter (1934: 37) was the first scholar to apply systematically the statistics of sign sequences as internal evidence to elucidate the direction of the script. It is instructive to compare his handling of the evidence of the seals 1052 and 1247 (pl. I-A,B) with that of Marshall. In dealing with the overflow sign in 1052, he does not rely merely on the fact that it is found below the left end of the line for want of space. He points out that it is the most common left end sign and that if placed at the right end of the upper line, it does not form a known combination with the adjacent sign. In dealing with the text in 1247, his argument does not rely on the looping of the line at the left end, but on comparison of sign sequences in similar texts which show that the only tenable reading of 1247 is in the boustrophedon mode starting from the right end of the upper line. By adopting the combinatory method he is able to show that most of the second lines in 2-line texts are to be read only from the right and are not boustrophedon
<table>
<thead>
<tr>
<th></th>
<th>Normal Signs</th>
<th>Reversed signs</th>
<th>Normal Pairs</th>
<th>Reversed Pairs</th>
<th>Normal Right end Pairs</th>
<th>Normal Left end Pairs</th>
</tr>
</thead>
<tbody>
<tr>
<td>4508</td>
<td>E</td>
<td>V</td>
<td>V</td>
<td>L</td>
<td>E , V , L</td>
<td>E , V , L</td>
</tr>
<tr>
<td>4516</td>
<td>E</td>
<td>U</td>
<td>V</td>
<td>E</td>
<td>E , V , L</td>
<td>E , V , L</td>
</tr>
<tr>
<td>1025 (Seal)</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U , U</td>
<td>U , U</td>
</tr>
<tr>
<td>9061 (Potsherd)</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U , U</td>
<td>U , U</td>
</tr>
<tr>
<td>1232 (Seal)</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U , U</td>
<td>U , U</td>
</tr>
<tr>
<td>2930 (Potsherd)</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U</td>
<td>U , U</td>
<td>U , U</td>
</tr>
</tbody>
</table>

**Fig. 1**

<table>
<thead>
<tr>
<th></th>
<th>Normal Signs</th>
<th>Reversed signs</th>
<th>Normal Pairs</th>
<th>Reversed Pairs</th>
<th>Normal Right end Pairs</th>
<th>Normal Left end Pairs</th>
</tr>
</thead>
<tbody>
<tr>
<td>6112</td>
<td>E</td>
<td>V</td>
<td>V</td>
<td>L</td>
<td>E , V , L</td>
<td>E , V , L</td>
</tr>
<tr>
<td>2618</td>
<td>E</td>
<td>V</td>
<td>V</td>
<td>L</td>
<td>E , V , L</td>
<td>E , V , L</td>
</tr>
<tr>
<td>3157</td>
<td>E</td>
<td>V</td>
<td>V</td>
<td>L</td>
<td>E , V , L</td>
<td>E , V , L</td>
</tr>
<tr>
<td>5477</td>
<td>E</td>
<td>V</td>
<td>V</td>
<td>L</td>
<td>E , V , L</td>
<td>E , V , L</td>
</tr>
<tr>
<td>1117</td>
<td>E</td>
<td>V</td>
<td>V</td>
<td>L</td>
<td>E , V , L</td>
<td>E , V , L</td>
</tr>
</tbody>
</table>

**Fig. 2**

<table>
<thead>
<tr>
<th></th>
<th>Normal Signs</th>
<th>Reversed signs</th>
<th>Normal Pairs</th>
<th>Reversed Pairs</th>
<th>Normal Right end Pairs</th>
<th>Normal Left end Pairs</th>
</tr>
</thead>
<tbody>
<tr>
<td>4508</td>
<td>E</td>
<td>V</td>
<td>V</td>
<td>L</td>
<td>E , V , L</td>
<td>E , V , L</td>
</tr>
<tr>
<td>4516</td>
<td>E</td>
<td>U</td>
<td>V</td>
<td>E</td>
<td>E , V , L</td>
<td>E , V , L</td>
</tr>
</tbody>
</table>

**Fig. 3**

<table>
<thead>
<tr>
<th></th>
<th>Normal Signs</th>
<th>Reversed signs</th>
<th>Normal Pairs</th>
<th>Reversed Pairs</th>
<th>Normal Right end Pairs</th>
<th>Normal Left end Pairs</th>
</tr>
</thead>
<tbody>
<tr>
<td>6112</td>
<td>E</td>
<td>V</td>
<td>V</td>
<td>L</td>
<td>E , V , L</td>
<td>E , V , L</td>
</tr>
<tr>
<td>2618</td>
<td>E</td>
<td>V</td>
<td>V</td>
<td>L</td>
<td>E , V , L</td>
<td>E , V , L</td>
</tr>
<tr>
<td>3157</td>
<td>E</td>
<td>V</td>
<td>V</td>
<td>L</td>
<td>E , V , L</td>
<td>E , V , L</td>
</tr>
<tr>
<td>5477</td>
<td>E</td>
<td>V</td>
<td>V</td>
<td>L</td>
<td>E , V , L</td>
<td>E , V , L</td>
</tr>
<tr>
<td>1117</td>
<td>E</td>
<td>V</td>
<td>V</td>
<td>L</td>
<td>E , V , L</td>
<td>E , V , L</td>
</tr>
</tbody>
</table>

**Fig. 4**

**Fig. 5**
and that there are many single-line inscriptions, especially from Harappa, which are to be read from left to right. Hunter still remains the best guide in this matter.

2.8 Ross (1939: 554) also applied the test of sign sequences to elucidate the direction of the text in 1247 (PP. I-B). He shows that while a 2-line inscription can be read theoretically in 8 possible ways (by reading each line from left or right and by reading top or bottom line first), there is only one tenable reading of 1247 (first the upper line from the right and then the lower line from the left) when the sign sequences are compared with similar single-line texts. However, Ross’s ‘extremely rigorous treatment’ is marred by his axiom that ‘the direction of all single-line inscriptions is the same as the direction of that line of multilinear inscriptions which is to be read first’. This is just not true as there are many single line inscriptions running demonstrably in either direction. It will be nearer the actual state of affairs to say that the line to be read first in a multi-line inscription will be in the normal direction of the script. At least no exception is known to this modified rule. Ross’s objection to considering the upper line as the first line of a multi-line inscription in an undeciphered script is not also justified, as no exception to this arrangement is known in any script, ancient or modern.

2.9 An independent and complete proof starting from first principles for the direction of the Indus Script has been attempted by me earlier in ISTCT (Introduction: 10). It is based on the premise that the only universal characteristic which determines the direction of a script is the sequence of signs reflecting the fixed order of speech in the language. This order remains the same whether the written text runs from the right or the left or in the boustrophedon mode, and whether we look at a seal or its impression. External evidence is necessary only to identify the fixed order of the more frequent sequences which will then enable us to determine statistically the normal direction of the script. Such external evidence is provided by three sets of ‘direction markers’ viz.,

(i) Orientation of asymmetric signs which can be observed to change along with the direction of the line;
(ii) Terminal signs and sign groups which occur with greater frequency at the left end or right end positions;
(iii) Sequence of the more frequent sign groups;

In each of these categories, the ‘normal’ direction and the ‘reversed’ direction are established purely on a statistical basis. It may be clarified that, the terms ‘normal’ and ‘reversed’ are neutral to actual direction and are merely based on the frequency of occurrences. Examples of these “direction-markers” are given in Fig. 4.

2.10. The statistical data is then applied to “split sequences”, which are identical inscriptions occurring both in single lines and as two lines one below the other on the same side of an inscribed object. If we assume (on the basis of universal experience) that the upper line is written first and the lower line thereafter, we can establish the order of the signs in the corresponding single-line text. It then follows that all other lines containing this sequence have to be read in the same direction. This is a cumulative process and it is possible by means of interlocking evidence in respect of orientation of asymmetric signs, sequence of sign groups and the positional distribution of the more frequent terminal signs and sign groups to establish conclusively the normal direction of the script. The method of ‘split sequences’ is illustrated in Fig. 5. Assuming that in 6112 (Pl. I-F) the upper line is read first, it follows that the single-line text in 2616 (Pl. I-G) has to be read from right to left. The importance of this example lies in the fact that \( \text{end sign in the script. Hence it follows that the normal direction of the script is from right to left.} \)

2.11 The statistical study in ISTCT (p. 14 and Tables VI and VII) shows that about 83 per cent of the lines run from the right and about 7 per cent from the left. (The rest are mostly single-sign or doubtful due to damaged signs.) Writing in the boustrophedon mode is rare, though 18 examples are listed in ISTCT. The evidence seems to suggest that the second line runs in the reversed direction generally when the first line has an incomplete sequence (e.g. 1247, 6402).

2.12 The recent work of Fairservis (1976, 1977) proposing a new model for the decipherment of the Indus Script proceeds on the assumption that the normal direction of the Script is from left to right. This does come as a surprise in view of the near-unanimous and well-established view based on the labours of a long line of scholars for nearly half a century that the normal direction of the script is from the right. It is unfortunate that Fairservis takes hardly any notice of the previous work in the field in this respect. He wrongly cites (Fairservis: 1976: 33) the Soviet scholars in support of his own view that
the script runs from the left. The Soviet Scholars have clearly stated their position (Knorozov: 1968: 99) that the script runs from right to left. Fairservis has perhaps been misled by the Soviet practice of referring to the original seals (which normally run from the left) rather than to the seal impressions (which normally run from the right).

2.13 Fairservis (1977: 6) advances three arguments in support of his view. His arguments may be summarised and briefly countered as follows:

1. Many Graffiti show a slant to the right. Such slants are common to the writing which is written from left to right.

Graffiti on pottery may show a slant depending on the curvature of the surface and the angle in which the potter or the scribe held the vessel while incising the signs. In this respect the few pottery graffiti are not representative of the norm of writing which can be more clearly from hundreds of beautifully carved and well preserved seals.

2. The signs in column 5 of the grid (Fairservis: 1976: Tables 1-75) are positionally so regular that one might presume that they start an inscription rather than end it. It is in fact unlikely that the bulk of the inscriptions would end in the same column.

This is a curious argument considering that the grid itself is so arranged by Fairservis: As the Concordances would show, it is possible to arrange the occurrences of any sign under a regular column. The observed positional regularity is merely a reflection of the set word order found in the Indus texts. It is possible for the bulk of the inscriptions either to begin or to end with the same signs (or columns) if they have stereotyped beginnings or endings, a state of affairs which is not at all unlikely in ancient seal-inscriptions.

3. The tables of Pair-wise frequencies compiled by Fairservis (1976: Tables 57-59) demonstrate left to right reading. Right to left ordering of these signs produced far lower frequencies in all cases.

The absolute frequency of a sign pair must remain constant in whichever direction it is read. What Fairservis means perhaps is that if the pairs are tabulated in a left to right order, the relative frequency of corresponding pairs in the reversed direction would be much less in such a tabulation. This argument would also hold good if the entire arrangement is reversed as in the Finnish Concordance and in ISTC. In these tabulations where the right to left order is followed, the occurrence of corresponding reversed pairs in left to right order is quite rare. This is therefore a circular argument which proves nothing.

2.14 The position taken up by Fairservis in the matter of direction of the script forces him to read the lower lines in the 2-line inscriptions as the first lines (Fairservis: 1977: Appendix IV: 14, 16, 141 and 142). He does not however explain why Harappans chose to write from bottom upwards, an arrangement of lines not known in any other script. Fairservis is also quite inconsistent in this respect. In many other cases, he reads the upper line as the first line (ibid: 41, 66, 99 and 175) without offering any explanation for the variation. He denies (1976: 40) the existence of boustrophedon inscriptions, but does not explain how 1247 would be read by him. In fact if his method of reading is adopted this inscription would not only start from the lower line, but also in the reversed direction (in his system), neither of which is at all likely.

2.15 It is unfortunate that the wrong choice of direction has rendered the model of decipherment set up by Fairservis ab initio invalid. In general it can be stated that in the light of the clear evidence in favour of a right to left direction, no attempted decipherment of the Indus Script based on a left to right direction can be taken seriously.

3. WORD DIVISION IN THE INDUS SCRIPT

3.1 Broadly speaking there are two competing schools of thought about the nature of the signs of the Indus Script, one regarding them as word signs and the other as phonetic signs. I shall briefly summarise the arguments and present the results of my recent work based on the material available in ISTC.

The conclusions will be based purely on the analysis of the script and will take no account of linguistic speculations.

3.2 The earliest analysis of "the mechanical nature" of the script was undertaken by Smith (in Marshall: 1931: 415). He analysed three texts and recorded the following findings:

1120: "A series of intelligible expressions"
1314: "A series of 5 signs, each having a separate meaning"
1435: "The evidence points to the inscription being a succession of separate words". He also analysed the positional distribution of one sign (†) and found that there was a general probability that it possessed "a meaning in and by itself" and that, on the whole, it seemed a separate word, at least in most cases. He made the perceptible observation that this sign could not be a syllable "for, it should then appear in more varied positions". He concluded that most of the signs studied by him had meanings by themselves and that some were probably ideograms in that they conveyed a word as an idea and were therefore not used in syllabic values.

3.3 Hunter (1934: 126) formulated the following criteria in deciding on word division in the script:

(i) That the combination (of signs) is found in a number of cases relatively larger in proportion to the total occurrences of one of its members;
(ii) That the first member of the combination is demonstrably independent of any signs found preceding the combination; and
(iii) That the last member is demonstrably independent of any signs found following the combination.

After carrying out the analysis, Hunter concluded that "no sign of common occurrence (with one exception) is not found as a single word". In view of this finding, Hunter's "words" containing two or more signs are perhaps better described as "phrases". It should also be noted here that these results are independent of Hunter's linguistic speculations viewing the signs as phonetic syllables and hence the Harappan language as monosyllabic, as these deductions do not necessarily follow from his own formal analysis and are in any case unproved.

3.4 The Soviet Scholars (Knorozov et al.: 1965, 1968; tr. Zide and Zvelebil: 1976) analysed the texts by computer techniques involving interval statistics and were able to segment the inscriptions into 'blocks' corresponding to definite grammatical units like words and word combinations. The 'blocks' are regarded as comprising word signs standing for roots, attributes grammatical suffixes.

3.5 The Finnish scholars (Parpola et al.: 1969) arrived at similar results by analysing inscriptions with the same words but different word order, and near-identical inscriptions. Their basic principle in dividing the inscriptions into words is that each of the supposed words should be verifiable from elsewhere either separately (forming the whole of an inscription) or in other contexts so that the neighbouring words can also be ascertained by similar means. In all their subsequent publications, the Finnish scholars have proceeded with their interpretation of the script on the basis of word signs; but a consideration of the linguistic aspect of their work is outside the scope of the present paper.

3.6 I undertook an independent verification of these results utilising the much larger textual material as well as the concordance and the statistical tabulations now available in ISTCT. The procedures followed by me are extremely simple and logical and can easily be verified by other scholars from the published material.

Method (1):

3.7 The first step is to collect single-sign texts from well-preserved and complete inscribed objects. It will be seen from the concordance in ISTCT that 43 signs occur as single-sign texts, which must therefore be regarded as self-contained and independent linguistic units, namely, 'words'. A special category of the single-sign texts which can be readily identified as words, consists of 'numerals' (two to ten short strokes arranged in one or two tiers) incised on the rims of pottery and on the blades of bronze weapons for obvious inventory purposes.

3.8 The second step is to compile two-sign texts in each of which one sign is already identified as an independent word on account of its solitary occurrence elsewhere. In these cases, the remaining sign must therefore be regarded as a separate word. This technique is illustrated in Fig. 6. The search is then extended progressively to longer texts containing three or more signs, of which all but one are already established as word signs, thus proving that the remaining sign must also be an independent word.

Method (2):

3.9 Another method is to list pairs of texts which are identical but for the presence of one additional sign at either end. The additional signs in these cases must therefore be independent 'words' (including grammatical morphs like case endings, suffixes and other particles). For example, ISTCT lists 24 pairs of identical texts but for the presence of one additional sign (E) at the left end in one of the texts in each pair. See the examples in Fig. 7.
Figs. 7—10
3.10 Another extension of this method is to search for pairs of near-identical texts which vary only by the addition of single signs in the medial positions. These are to be considered as independent words. The method can be extended further by comparing pairs of near identical texts which vary only by the substitution of one sign in a text by another in the other text. Such substitutions prove the independent character of both the signs in question.

Method (3):

3.11 Another method is to search for sets of texts which are built up by the progressive addition of one sign at a time at either end. Hunter (1934: 127) has already cited one such interesting set. For another, see Fig. 8. such texts clearly prove the independent word values of each sign occurring in them.

Method (4):

3.12 A few signs occur with very high frequencies and form stable pairs with a large number of other signs. Analysis shows that these few high-frequency signs are integral units having some independent meaning or grammatical function. Examples of such signs are given below:

\[
\begin{align*}
\mathcal{U}, & \quad \mathcal{A}, \quad \mathcal{H}, \quad \mathcal{V}
\end{align*}
\]

It therefore follows that the other members forming stable pairs with the above signs must be separate words. This is a particularly productive method in view of the very high frequency of these "suffixes" and the very large number of stable pairs formed by them.

3.13 The methods described above are overlapping and the results are cumulative. They prove that, almost without exception, the signs in the Indus Script are word signs with specific meanings or grammatical functions and cannot be treated as mere phonetic units with alphabetic or syllabic values as suggested by scholars like S.R. Rao (1973: 127).

3.14 The next step is to segment the texts into 'phrases' consisting of stable combinations of two or more words or compound words. Two methods were followed by me for the purpose:—

Method 1:

A longer text can be shown to consist of two or more shorter texts appearing as complete texts elsewhere, clearly indicating the points of segmentation. See Fig. 9 for examples.

Method 2:

With the help of pairwise frequencies, it is possible to isolate "phrases" (generally containing 2 to 4 signs only) with high frequencies forming integral linguistic units.

It is possible, by following these methods, to segment all but few of the texts included in ISTCT into separate phrases and words. Fig. 10 contains some illustrative "phrases" with high frequencies.

4. CONCLUSION

4.1 In conclusion, I should like to stress two general points about the results summarised in this paper. Firstly, they are totally independent of any theory about the language of the Harappans and will remain valid for any successful decipherment of the script. Secondly, as the foregoing discussions show, my methods are in continuation and extension of those already employed by the earlier investigators in the field and the results are also in general conformity with their findings. The significance of the present work lies in its independent verification and confirmation of the results as well as in the application of the methods to the entire corpus of known texts to test their validity. The results thus from a secure base for further progress towards the eventual decipherment of the Indus Script.

REFERENCES


