Tamil Civilization
Vol.4 No.3&4

DATA BASE FOR THE INDUS SCRIPT:A COMPUTERISED CONCORDANCE OF THE INDUS TEXTS

Mythili Rangarao & Iravatham Mahadevan

The Indus Civilization and Script

The Indus or Harappan Civilization[1] came to light with the large-scale excavations of the ancient sites of Harappa and Mohenjodaro (now in Pakistan) from about 1920. The exact chronology of this bronze-age civilization is still a matter of debate; but the inclusive time bracket of 2500-1500 B.C. proposed by Mortimer Wheeler has gained wide acceptance.[2]

The pictographic script of the Harappan inscriptions is now generally known as the Indus Script. The facts known about this script are few and can be quickly summarised. The script is pictographic, each sign being a pictorial representation of some object, though only a few (e.g. the load-bearer, the fish etc.) are recognizable, as the other signs had become too stylised. Only about 3500 short inscriptions, mostly on seals, are known. The number of signs in the script is about 419, too many for an alphabetic system. It is generally accepted that the Indus script is to be read from right to left (with some exceptions). The script exhibits two distinct features, viz., combination of signs by ligature and modification of signs by the addition of strokes or other marks. The Indus script has not yet been deciphered, notwithstanding many claims to the contrary.

A Computer Study of the Indus Script

A project to study the Indus script by using computer techniques was undertaken at the Tata Institute of Fundamental Research (TIFR) during 1972-77.[3] The project essentially consists of compilation of a corpus of the Harappan inscriptions in a numerically coded form and, using this as the input data, the generation of a sign concordance and a set of statistical tables. The present paper briefly describes a part of the work relating to the compilation of a computerised concordance to the Texts in the Indus script. The work was executed on CDC-3600 computer. The programmes for generating the concordance are written in FORTRAN IV language. A special composition software was also developed to print out the results in the Indus script characters by using computer-aided photo-composing techniques on DEC 10 system.[4] Presently the Database for the Indus Script is maintained on the CYBER 170/730 Computer system at the TIFR. A brief description of the elements of the Database is given in Table 1. The authors have also

created a Library of Signs in the Indus Script for graphic reproduction on a CALCOMP Drum Plotter. (See illustrations in Fig. 1).

Sources

The excavation reports on Mohenjodaro, Harappa and Chanhudaro and the publications of the Archaeological Survey of India (ASI) are the principal sources[5] from which the material for the present study was drawn. Additional material reported from excavations and explorations or as stray finds was collected from widely scattered publications. Unpublished photographs and original objects mostly in the collections of the ASI were made use of to add more than 600 unpublished texts to the Corpus (mostly from Harappa, Mohenjodaro, Lothal and Kalibangan).

Coding of Input Data

One of the principal objectives of the study is to correlate the known archaeological data with the patterns of frequency and distribution of signs, sign-combinations and texts to facilitate the interpretation of the texts. The available data on each inscribed object were collected on a cardifle and then coded in a numeric form suitable for computer analysis. The significant data coded for each line of text include (i) the site, (ii) the location of the object within the site, (iii) stratigraphy, (iv) type of inscribed object, (v) number of inscribed sides, (vi) associated 'field symbols' accompanying the inscriptions, (vii) number of lines of text, (viii) direction of writing of each line, (ix) length of each line i.e., the number of signs, and (x) the text. The concepts underlying the coding of the data are briefly explained in the following paragraphs.

Sites

Nineteen Harappan and five West Asian sites are represented in the present collection. The number of inscribed objects included in the corpus from the five major Harappan sites is 2,855 (Mohanjodaro-1540, Harappa 985, Chanhudaro 66, Lothal 165 and Kalibangan 99). The number from other Harappan and West Asian sites is 34 and 17 respectively.

Locus

A 'locus' is a well-defined area designated as such by the excavators within an archaeological site (e.g. DK area in Mohenjodaro, Mound F in Harappa etc.) The available data on locus for most of the published inscribed objects from Mohenjodaro, Harappa and Chanhudaro have been coded. The details of the actual find spots within a locus (e.g. trench, block, house, room, street, etc.) have not been included in the computerised data.

Stratigraphy

The available stratigraphic data in respect of the published inscribed objects from Mohenjodaro, Harappa and Chanhudaro have been coded.

Inscribed Objects

Each inscribed object is given a reference number in four digits, which serves also as the number of the text inscribed on it. An inscribed object can have, by definition, only one text consisting of one or more lines, inscribed on one or more sides of the object. Out of a total

of 3455 Harappan objects in our card file, only 2906 legible ones comprise the Corpus of Texts.

Types of Inscribed Objects

Inscriptions in the Indus Script occur on several types of objects. However these can be grouped under three basic types from the point of view of a study of the writings on them - i.e. (a) Seals (b) Sealings and (c) Other inscribed objects.

Sides of Inscribed objects

Only each distinct surface of an object bearing an inscription or a pictorial motif (termed a 'field symbol') is regarded as a 'side' of the object and is serially numbered. As a matter of convention, sides with inscriptions are numbered before the other sides of the same object without inscriptions but featuring only field symbols. A distinction has been made for purposes of analysis between the only inscribed side of an object and the first side (when there are more than one inscribed side). The following criteria have been adopted while fixing the side order of the lines of text:

(a) Lines of text appearing on different sides of an object may occur as a single line on another object. In such cases the side order is determined with certainty, assuming of course, that the direction of writing is known.

(b) Frequently occurring terminal signs and sign combinations at the beginning or the ending of single lines provide clues to the real order of lines having these signs in similar positions.

(c) In the absence of any evidence to the contrary, the side order given in the source publications is generally followed.

Field Symbols

A pictorial representation or conventional motif (other than the signs of the Indus Script) occurring on an inscribed object is termed a 'field symbol'. Each side of an inscribed object can have, by definition, only one field symbol or none. Different sides of an inscribed object may feature different field symbols with or without accompanying inscriptions.

The Sign List

A sign list of the Indus script was compiled from the photographic card catalogue of the inscriptions. The sign list represents a normalised signary of the script faithfully reproducing the 'neat monumental forms' of the originals seen in the best seals. The most frquently occurring variant of a sign was chosen as its normal form. The sign list records 419 signs in the Indus script. It is however difficult to be precise about the total number of signs in an undeciphered script because it is not easy to distinguish between independent signs and mere graphic variants. The corpus of texts of the present study has recorded 13372 sign occurrences. The following is a summary of the frequency of signs:

FREQUENCY RANGE	NO. OF SIGNS	TOTAL SIGN OCCURRENCES	PERCENT (OF TOTAL OCCURRENCES)
1000 or more	1	1395	10.43
999-500	1	649	4.85
499-100	31	6344	47.44
99-50	34	2381	17.81
49-10	86	1833	13.71
9-2	152	658	4.92
Only once	112	112	0.84
Total	417	13372	100.00

Lines of Text

A line of text is continuously written matter occurring on a side of an inscribed object. The lines are generally written in a straight line, but occassionally radially inscribed on circular sides or even on rectangular sides. The lines are serially numbered within each side of an inscribed object generally from top to bottom (with some exceptions in the case of irregularly inscribed lines). In numbering the lines of text, a distinction has been made for purposes of analysis between the only line of text and the first of more than one line of text appearing on the same side of an inscribed object.

Total number of lines of text is 3573. The majority of the texts are of single lines only. The maximum number of lines is seven in a text and three on one side of an inscribed object. The length of a line of text varies from 1 to 14 signs. The maximum length of a text is 26 signs. Most of the texts are however much shorter. The average length of a text is five signs only.

The unit of textual analysis (distributional statistics) is a line of text. There are two reasons why it is not possible to consider the whole text as a unit for this purpose. Firstly, there is no way of knowing beforehand whether different lines of an inscription appearing on the same object or even on the same side have continuity of sequence or to be regarded as separate texts. Secondly, it is not also possible to ascertain beforehand the real order (if any) of the lines of text appearing on the same object or even on the same side. It is therefore necessary to stress that the definition of a 'text' and serial numbering of the sides and the lines of text are provisional.

Direction of Writing

The following conventions have been adopted in coding the direction of writing of lines of text:

(a) The direction of writing in respect of texts engraved

on seals is assumed to be that on the impressions of the seals.

(b) The direction of writing in all other cases is indicated as in the originals.

(c) The lines of text in the Texts and the Concordance in the numeric code are arranged to be read always from left to right. The actual direction of writing of each line on the original object, or on the impression in the case of seals, is also however indicated separately by a code. (The pictorial versions of the texts and the concordance are arranged in the right-to-left direction).

One of the few well-established and generally accepted facts about the Indus Script is that the direction of the script is generally from right to left. Several investigators have demonstrated this from a study of the external features of the writing and, more importantly, from the evidence of sequences.[6] B.B. Lal has demonstrated from a study of overlapping incisions on pottery graffiti that the inscriptions in question must have been incised from the right.[7]

The computerisation of texts and background data has made it possible to verify the direction of writing by a comparative study of the entire known material and also to quantify the results. The study has confirmed that the general direction of wirting of the Indus script is from right to left, though exceptional cases of writing from left to right and also in the boustrophedon mode are known.

Statistics Relating to Directi

The coding of the directi of writing of the lines of text wa done on the basis of the criter established in the present stud The distribution of lines by directi of writing is shown below:

DIRECTION	NO.OF LINES	PERCENT
Right to left	2974	83.23
Left to right	235	6.57
Top to bottom	7	0.20
Symmetrical seq-ences (which read alike from either end)	12	0.34
Single-sign lines	190	5.32
Doubtful cases (on account of da-maged or ille-gible portions)	155	4.34
Total:	3573	100.00

The Corpus of Texts

The corpus comprises 2906 text in 3573 lines with 13372 sign occurrences, constituting the largest collection of inscriptions in the Indus Script made so far (1977). Texts with more than one line have been arranged with the lines shown separately one below the other. Sides with field symbols but without lines of text are also listed in separate lines. These lines are however not included in the concordance. Signs read doubtfully are indicated by an asterisk (*) prefixed to the sign at left. Lost, damaged or illegible passages are indicated by the code

000 (which may stand for one or more signs as, in the nature of things, it is not possible to count such signs). This has been done so that the non-terminal extant signs at either end of broken lines are not counted as terminal and the signs separated in the inscription by mutilated portions are kept apart. No attempt has been made to restore lost signs conjecturally. Readings have been critically collated from all available sources, both original and secondary. Consistency in readings has been checked with the help of the Concordance, especially in reading doubtful signs. A specimen of two coded lines of text from the corpus is appended with explanatory notes to illustrate the format (Table 2).

Concordance

The processing of the concordance from the corpus is done in five phases. Briefly, the first phase generates as many lines as there are sign occurrences in a text so that every sign occurrence is positioned once in the reference column and written on a file. Only such texts which have more than one line are again written on another file completely, along with a serial number. The second phase sorts lines from the first file in the ascending order of the sign in the reference column. The third phase renumbers the multi-line texts from the second file with respect to the new sorting order of lines on the first file. The fourth phase sorts the multi-line texts from the second file with respect to the new number in the ascending order. The final concordance is generated in the fifth phase by using these two sorted files. The five phases are explained in detail in the succeeding paragraphs.

Generation of lines

The corpus of texts is read line by line and each sign occurrence, including the doubtfully read sign marked with an asterisk, is positioned once in the reference column by shifting the line by one column to the left every time. The generation of such lines consists of two parts. The first part deals with the single line texts and the second part deals with the multi-line texts. The following concepts and definitions are necessary to explain the methods used in two parts.

Let set L = $\{l_j\}$, j = 1, 2, ... 16 be a line of a text in the corpus

A = $\{a_j\}$, j = 1, 2, ... 32 is a generated line or a concordance line

S = $\{s_j\}$, j = 1, 2, ... 30 is an ordered set of sign occurrences in a single line text.

T = $\{s_j\}$, j = 1, 2, ... 60 is an ordered set of sign occurrences in a multi-line text.

C = $\{c_j\}$, j = 1, 2, ... 9 is an ordered set of line numbers and serial number of an multi-line text.

Q = $\{q_j\}$, j = 1, 2, ... 8 in an ordered set of starting positions of lines of a multi-line text in T.

$M = \{L_j\}$, $j = 1, 2, \ldots 7$ be an ordered sequence of lines of multi-line text.

Let Y be a set such that the set Y is defined by concatenating the three sets C, T and Q in the same order.

Let $X = \{A\}$ be a sequence of generated lines and

$Z = \{Y\}$ be a sequence of multi-line texts.

Definitions

Shift

In an operation on a set such that the jth element of the set maps into (j - 1)th element and the first element maps into the last element of the set.

Generate the Sequence

It is a process of appending an existing sequence of sets by a given set thus increasing the number in the sequence by one.

Reference Column

It is defined rank of an element in a set such that every non-blank element in the ordered set is mapped once into this rank while preserving the order.

Blank Sign

It is a sign which has no value including zero.

Assumption

Sequence M, X and Z consist of null set to startwith.

Single Line Text

The corpus of texts are read line by line into the set L in which l_1 is the line number of the text and $l_2, l_3, \ldots l_{16}$, are the signs of the text. The generation of the sequence X from the set L is done as follows:

Step 1: Read a line text in L; if the L is null stop generating the sequence X. Map blank sign $\rightarrow$ S_i, $i = 1, 2, \ldots 30$

Step 2: Map 1. L into $a_1 \in A$ and a value O into $a_{32} \in A$

Step 3: Move a value 17 into a constant variable called 'm' where m is the reference column of the set A.

Step 4: Map $l_i \in L \rightarrow S_{m+i-3} \in S$ for $i = 2, 3, \ldots 16$

Step 5: Map $S_j \in S \rightarrow a_{j+1} \in A$ for $j = 1, 2, \ldots 30$

Step 6: If $a_m \in A$ is blank go to Step 1

If $a_m \in A$ is non blank then generate the sequence X by the Set A

Step 7: Shift the set S and go to Step 5.

Multi-line Text

As has already been explained in the "lines of text" section, no

text of the corpus has more than 7 lines and 26 sign occurrences. Hence the sequence M can utmost have 7 elements. The two sequences X and Z are generated as follows:

Step 1: Set a constant variable 'n' to be called a serial number equal to a value O.

Step 2: Set a constant variable 'm' equal to a value 17, a constant variable 'j' equal to 1 and a constant variable 'p' equal to 30. Map a null set to T and M.

Step 3: Generate a sequence M by the set L. Set a constant variable 'k' equal to 1. When the sequence M is null stop generating the sequences X and Z.

Step 4: Consider $L_k \in M$ and if L_k is a null set go to Step 8; otherwise map $l_1(\in l_k) \rightarrow C_k(\in C)$. Set a constant variable 'j' equal 2.

Step 5: If $l_j(\in L_k)$ is blank go to Step 7, else map $l_j(\in L_k) \rightarrow S_{p+j-1}(\in T)$

Step 6: Set 'j' equal to 'j+1' (i.e. increase the constant variable value by 1) and go to Step 5.

Step 7: Map $P \rightarrow q_k(\in Q)$. Set 'k' equal to 'k+1', 'p' equal to 'p+j' and go to Step 4.

Step 8: Map $S_{16+1_3^2}(\in T) \rightarrow a_i(\in A)$ for i = 2, 3, ... 31. If $a_m(\in A)$ is blank go to Step 2.

Step 9: Set the constant 'n' equal to 'n+1' and map $n \rightarrow C_8(\in C)$

Step 10: Map $C_8(\in C) \rightarrow a_{32}(\in A)$ and $C_1(\in C) \rightarrow a_1(\in A)$

Step 11: Generate the sequence X by the set A and define the Set Y.

Step 12: Generate the sequence Z by the set Y.

Step 13: Map Blank Sign $\rightarrow a_{32}(\in T)$. Shift the set T and go to Step 8.

Sorting

The sorting process consists of two stages, viz. internal sorting and the merging of sorted records. The tournament sort, employing the replacement selection techniques, is used for internal sorting. The merging is accomplished either by the balanced or polyphase technique. To sort the sequences X and Z the Polyphase merge is adopted.

Sorting of the generated lines, i.e. the sequences X, is done in the following way. As many as 31 keys are used to sort the sequence X. The first key is the sign in the reference column, that is $a_{17} \in A$ in the sequence X. The next 15 keys are for the preceding signs and the next 14 keys for the succeeding signs with reference to the sign in the reference column. The last key is the line number. The sorted sequence X is generated as follows:

(a) Each of 13372 legible sign occurrences, including the doubtfully read signs, is listed once in the reference column in the ascending order.

(b) When two or more lines have the same sign in the reference

column, as it is bound to, considering that there are only 417 signs, the order of listing the lines is decided by the hierarchy of signs in the columns to the left of the reference column. For this purpose the signs to the left of the reference column are scanned one by one progressively upto the first sign.

(c) When the preceding signs along with reference column sign are same in two or more lines, then the above process is continued on the signs to the right of the reference column sign.

(d) When the sign sequence is the same in two or more lines, then the lines are arranged in the ascending order of the line numbers, which are unique.

Rearranging

Since the file organization adopted throughout is a sequential type, a file arranged in one sequence needs to be sorted if the same file is required in another sequence. The generated sequence Z is arranged with respect to the serial number, i.e., with respect to $Y_8 \in Y$. But for the desired concordance, the arrangement of sequence Z should be the same as the sorted sequence X. To achieve this, the element $Y_9 \in Y$ is assigned a number according to sequence position in the sorted sequence X. After assigning numbers to every element of the sequence Z, the same sorted in the ascending order of the assigned numbers.

Generation of Concordance

As explained earlier, the sorted sequence X, in file one, comprise both single line and multi-line texts. The second file, which is sequence Z, comprise only multi-line texts. To generate the concordance, the first file is read line by line and if the line represents a single line text, which is indicated by O in $a_{32} \in A$, it is reproduced as such in the concordance. If the line represents a multi-line text, the corresponding line is read from the second file and as many lines as are existing in the text are reproduced in the concordance. Thus the desired concordance is generated.

The generated concordance is a complete index of sign occurrences in the texts. The concordance also sets out textual context of each sign occurrence in full by reproducing the entire text each time along the sign occurrence.

Arrangement of the concordance is as follows:

(a) A six-digit Reference Number for each line of text consisting of the Text, Side and Line Numbers (in that order) is given at the left of the page.

(b) Each of the legible sign occurrences (including the doubtfully read signs marked with an asterisk sign) is listed once in the order of the Sign List in the Reference Column in the middle of the page by shifting suitably the position of the line of text in which the Reference Sign occurs.

(c) Preceding lines (if any) in the same text are given above the Reference line and so arranged as to end immediately before (that is, to the left of) the Reference column. Succeeding lines

(if any) in the same text are given below the Reference Line and so arranged as to begin immediately after (that is, to the right of the Reference Column. Thus the key signs are placed one below the other in a Central Reference column. Each text is read from the top line to the bottom line, and each line from left to right. (In the pictorial version the direction is from right to left).

(d) The listing order of the occurrences of the same sign is positional with reference to the whole text as the unit. The solus occurrences (a single sign constituting the whole text) are listed first, followed by the initial occurrences and thereafter the non-initial occurrences. . The initial occurrences are listed in the order of the succeeding signs. The non-initial occurrences are first listed in the order of the preceding signs. Where the preceding signs are identical, the arrangement is by the succeeding signs. For purposes of sorting, signs occuring in all the lines of a text are taken into account. Identical texts are listed in the order of the Text Numbers. A specimen page from the numerical Concordance is appended (Table 3). The corresponding page from the pictorial version generated by Computer-aided photo-composition method from the numeric version is also appended for comparison (Fig. 2). As mentioned earlier, while the numerical coded version reads from the left the pictorial version reads from the right.

The compilation of the Concordance on these lines has been a complex operation and made possible only by the use of a large computer and elaborate programming. The effort has been worth the while, as a concordance is the basic and indispensable tool for a statistical-positional study of the script. The frequency and positional distribution of each sign and sign combination can be readily ascertained from the Concordance. The juxtaposition of near-identical sequences can be used to determine word division. Frequent sign combinations (especially, pairs and triplets) help us to determine the direction of writing of lines containing these combinations. Doubtful signs can be read with a fair amount of confidence by a comparative study of identical sequences. The order of lines of text as well as the direction of writing of lines can be judged by comparing adjacent terminal signs. Sign variants can be distinguished to a large extent by the use of the Concordance which sets out the identical or similar environments.

Tables

The statistical tables generated from the corpus of texts are in two parts, the first dealing with sign analysis and the second with context analysis. The sign analysis is a detailed study of the frequency and distribution of signs and sign combinations occurring in the corpus of texts. The context analysis seeks to explore the possible relationship between the inscriptions and inscribed objects on the one hand, and between the latter and their archaeological context (i.e. site, locus and stratigraphy). Selected tables have been published (I.Mahadevan 1977).

Acknowledgement

The authors wish to place on record the continued guidance and facilities made available by Prof. R.Narasimhan, Head, Compute Section, TIFR in the preparatio of this paper.

NOTES AND REFERENCES

1. The best general introduction to the subject is provided by *The Indus Valley Civilization*, Sir Mortimer Wheeler, Cambridge University Press, Third Edition, (1968).

2. For a recent discussion on the chronology of the Harappan Civiliation, see *Lothal and the Indus Civilization*, S.R.Rao, (1973), Bombay, esp., the foreword. by Mortimer Wheeler (pp. v-vii) and Chap. XIV (pp. 161-167).

3. This was a follow-up of the earlier work done by one of the authors (Iravatham Mahadevan) under a Jawaharlal Nehru Fellowship during 1970-72.

4. Iravatham Mahadevan, *The Indus Script : Texts, Concordance and Tables*, Memoirs of the Archaeological Survey of India, No.77, New Delhi (1977).

5. The major source publicatio are:
   (i) *Mohenjodaro and the Indu Civilization*, Marshall, J. 3 Vol., London (1931);
   (ii) *Further Excavations at Mohe jodaro* Mackay, E.J.H., Vols., New Delhi (1937-38
   (iii) *Excavations at Harappa*, Vat M.S., 2 Vols., Calcutta (1940
   (iv) *Chanhudaro Excavation* Mackay, E.J.H., Boston (1943

6. Hunter, G.R., *Script of Harapp and Mohenjodaro*, (1934) p.37 Alekseev, *Soviet Studies on Harap pan Script*, Tr. by Pande, H.C. Paper No.6, Field Research Pro jects, (1969), p.1;

7. B.B.Lal, 'the direction of Writing in the Harappan Script', *Antiquity*, XL (1966), p.52; idem, 'A Further Note on the Direction of Writing in the Harappan Script', *Puratattva*, I (1967-68), p.15.

Table 1

DATA DESCRIPTION OF THE FIELDS IN THE DATABASE FOR THE INDUS SCRIPT

Attributes	Description
Reference Number	Each line of text has a unique Ref.No. in 6 digits comprising the Site Number (col.1), the Object Number (cols. 2-4), the Side Number i.e. the number of inscribed faces of the object (col.5) and the Line Number, i.e. the number of lines of text on each inscribed side of the object (col.6).
Locus	Area, Section or Sub-section of the site as determined by the excavator.
Level	The Level in ft. at which the object was found above (+) or below (-) the datum (in Mackay's excavations at Mohenjodaro and Chanhudaro), or below (-) the surface in Marshall's excavation of Mohenjodaro and Vats' excavation of Harappa. (The data on levels on other sites are not available). The levels are rounded off to the nearest foot.
Type	The typology of the inscribed objects. (See Table 4 for list of types).
Field Symbol	The pictorial motif in the field on each side of the inscribed object. (See I.Mahadevan 1977, pp.793-813 for the list of field symbols and illustrations).
Direction of Writing	Mostly from right, occasionally from the left and rarely from top to bottom. (The direction of writing was determined by criteria discussed in I.Mahadevan 1977, pp.10-14).
No. of positions in a Line of Text	This number records the total number of signs and text-breaks (or illegible portions) in a line of text for computational processes.
No.of Signs in a Line of Text	This number indicates the total of extant and legible signs in a line of text.
Line of Text	Each line of text is coded as a series of 3-digit numbers each uniquely defining a sign. (For the Sign List of the Indus Script, see I.Mahadevan 1977, pp.32-35). Doubtful signs are marked by asterisks. Breaks and illegible portions are also indicated by a special symbol.

Note: The Corpus of Texts published by I.Mahadevan (1977) is based on this Input Data; but the format in the book has been slightly re-arranged. Data on Locus, Level, the number of 'positions' and signs have been omitted and the Field Symbol codes have been abridged for want of space.

Table 2
FORMAT OF THE CORPUS OF TEXTS IN THE INDUS SCRIPT

COLS.	1-6	7-8	9-11	12	13-15	16	17-18	19-20	21-80
Codes	100101	17	-03	1	351	1	05	04	000 067 072 008 342
	100102	17	-03	1	351	3	01	01	393

EXPLANATIONS:

COL.NO.	DESCRIPTION	CODE	KEY
1	Site	1	Mohenjodaro (Marshall)
2-4	Object Number	001	Mohenjodaro (Marshall), No.1
5	Side Number	0	Only (inscribed) side
6	Line Number	1,2	Lines 1 and 2
7-8	Locus	17	HR Area B Sec. in Mohenjodaro
9-11	Level	-03	3' below surface
12	Type	1	Seal
13-15	Field symbol	351	Uncertain animal (facing right)
16	Direction of writing	1	Right to left (Line 1)
		3	Only sign (Line 2)
17-18	Number of positions	05	5 positions (including the code '000' standing for lost or illegible signs) in Line 1
		01	1 position in line 2
19-20	Number of signs	04	4 (extant) signs (Line 1)
		01	1 sign (Line 2)
21-80	Coded text	000	Broken portion in Line 1
	other 3-digit numbers	∅ ∅ ∅	each represents one sign in the sign in the sign list

Table 3

SIGN CONCORDANCE OF TEXTS IN THE INDUS SCRIPT

(Numerical version)

TEXT NO.												R.S.						
3312	10									065	336	089	407	169	183	169		
3331	10									065	336	089	407	169	183	169		
3337	00									065	336	089	407	169	183	169		
3354	10									065	*336	*089	407	169	183	169		
3355	10									065	336	089	407	169	183	169		
3381	10									065	336	089	407	169	183	169		
3384	10									065	336	089	407	169	183	169		
4288	00								000	065	336	089	194	169				
4049	00								000	065	336	089	347	342				
4028	00				307	086	099	387	053	065	336	089	211					
3016	00					290	290	343	067	065	336	089	211					
1326	00						267	099	072	065	336	089	211					
2583	00							267	099	065	336	089	222	254				
4081	00								121	065	336	089	389	102	287	342		
1006	00					233	081	099	389	065	336	089	087	155	124	169		
1425	00									066	336	089	124	169				
1147	00								387	066	336	089	211					
4672	10									067	336	089	216	254	176			
	20												177	197	192	341	089	328
6215	00									067	336	089	387	059	000			
6127	00									067	336	089	387	072	012			
5114	00								053	067	336	089	018	211				
1557	00						000	099	053	067	336	089	112	194	342			
6207	00						267	099	053	067	336	089	211					
1008	00					076	402	087	059	067	336	089	211					
2275	00							000	065	067	336	089	211					
2426	00							387	065	067	336	089	048	342				
2123	00							387	072	067	336	089	211					
2918	10									*051								
	20								053	070	336	089	294	008	342			
2654	00	017	293	293	211	202	267	099	065	070	336	089	087	284	342			
1065	00				087	028	257	099	067	070	336	089	211					
3109	00				089	245	245	223	099	070	336	089	178	180				

FIG. 1

SPECIMENS FROM THE LIBRARY OF INDUS SIGNS
PRODUCED ON CALCOMP PLOTTER

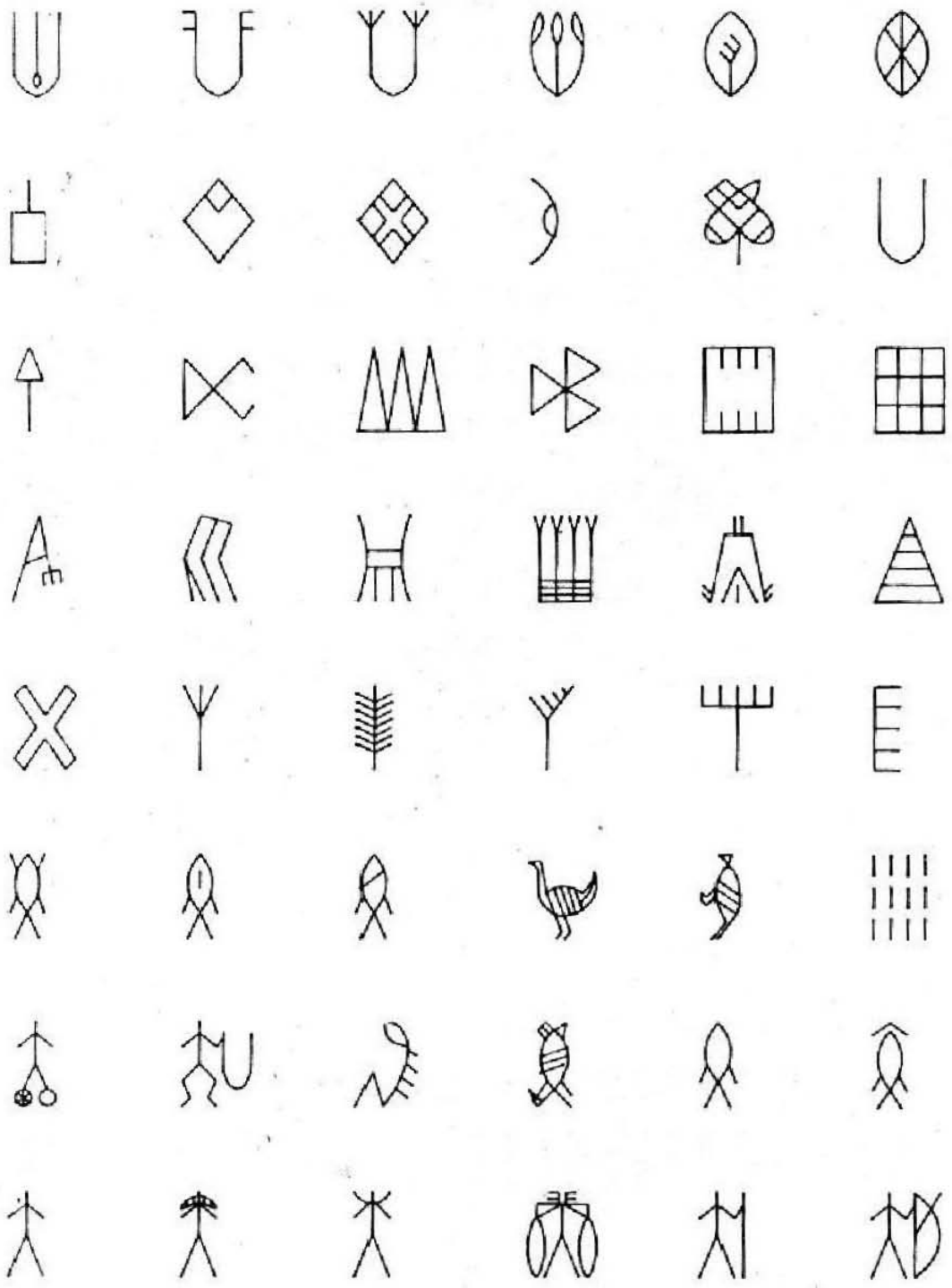
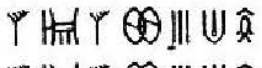
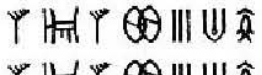
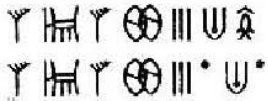
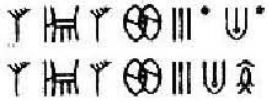
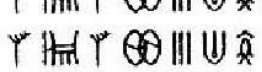
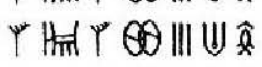
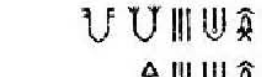
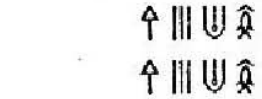
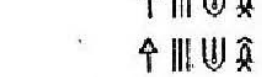
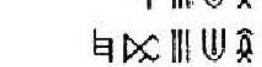
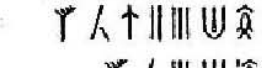
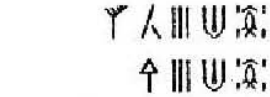
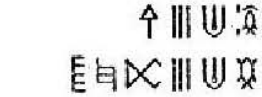
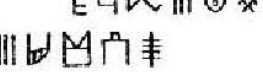
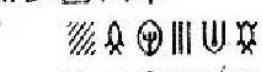
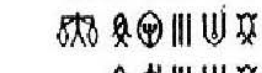
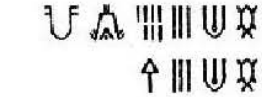
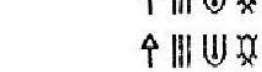
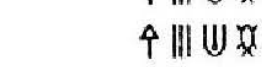
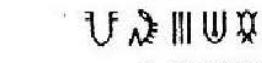
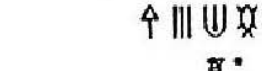
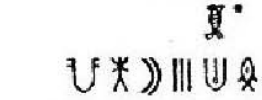
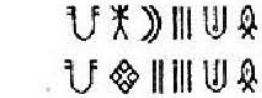
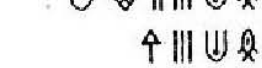
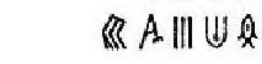


FIG.2.

SIGN CONCORDANCE OF TEXTS IN THE INDUS SCRIPT

(Pictorial Version)

CONCORDANCE	R	INDEX
3312 10		
3331 10		
3337 00		
3354 10		
3355 10		
3381 10		
3384 10		
4288 00		
4049 00		
4028 00		
3016 00		
1326 00		
2583 00		
4081 00		
1006 00		
1425 00		
1147 00		
4672 10		
20		
6215 00		
6127 00		
5114 00		
1557 00		
6207 00		
1008 00		
2275 00		
2426 00		
2123 00		
2918 10		
20		
2654 00		
1065 00		
3109 00		