Textual and Contextual Analysis of the Indus Script:

The Contribution of Corpora and Concordances to the Field of Indus Script Research

CSubramanian

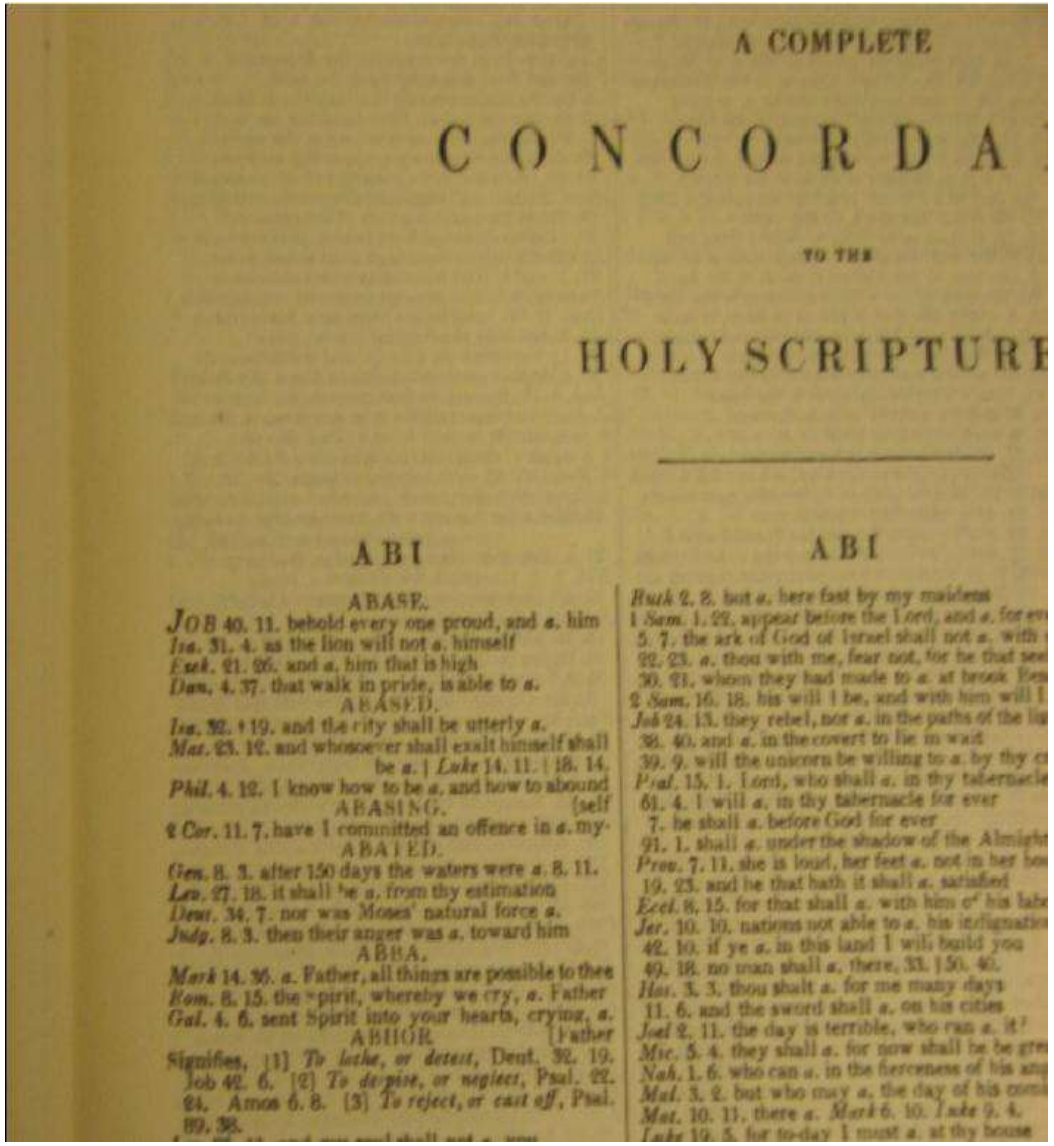
March 14, 2025

Index

- Concordances
- Discovery of Indus Civilization
- Marshall's Report
- Hunter's Concordance
- Mahadevan Concordance
- Parpola Concordances
- Bryan Wells Corpus/Concordance
- Way Forward
- Recent Progress
- Q&A

Concordances

- A concordance is an alphabetical index of the principal words of a book, such as the Bible, with reference to the passage in which the word occurs
- Concordances are also created for texts in languages that are yet to be deciphered



A COMPLETE

CONCORDA

TO THE

HOLY SCRIPTURE

ABI

ABASE.

JOB 40. 11. behold every one proud, and *a.* him
Isa. 31. 4. as the lion will not *a.* himself
Ezek. 21. 26. and *a.* him that is high
Dan. 4. 37. that walk in pride, is able to *a.*

ABASED.

Isa. 32. †19. and the city shall be utterly *a.*
Mat. 23. 12. and whosoever shall exalt himself shall be *a.* | *Luke* 14. 11. † 18. 14.
Phil. 4. 12. I know how to be *a.* and how to abound

ABASING. (self

2 *Cor.* 11. 7. have I committed an offence in *a.* my-

ABATED.

Gen. 8. 3. after 150 days the waters were *a.* 8. 11.
Lev. 27. 18. it shall be *a.* from thy estimation
Deut. 34. 7. nor was Moses' natural force *a.*
Judg. 8. 3. then their anger was *a.* toward him

ABBA.

Mark 14. 36. *a.* Father, all things are possible to thee
Rom. 8. 15. the spirit, whereby we cry, *a.* Father
Gal. 4. 6. sent Spirit into your hearts, crying, *a.*

ABHOR. [Father

Signifies, [1] *To loathe, or detest*, Deut. 32. 19. Job 42. 6. [2] *To despise, or neglect*, Psal. 22. 24. Amos 6. 8. [3] *To reject, or cast off*, Psal. 89. 38.

ABI

Ruth 2. 8. but *a.* here fast by my maidens
1 *Sam.* 1. 22. appear before the Lord, and *a.* for ev
5. 7. the ark of God of Israel shall not *a.* with
22. 23. *a.* thou with me, fear not, for he that see
30. 21. whom they had made to *a.* at brook Be
2 *Sam.* 16. 18. his will I be, and with him will I
Job 24. 13. they rebel, nor *a.* in the paths of the li
38. 40. and *a.* in the covert to lie in wait
39. 9. will the unicorn be willing to *a.* by thy c
Psal. 15. 1. Lord, who shall *a.* in thy tabernacle
61. 4. I will *a.* in thy tabernacle for ever
7. he shall *a.* before God for ever
91. 1. shall *a.* under the shadow of the Almigh
Prov. 7. 11. she is loud, her feet *a.* not in her ho
19. 23. and he that hath it shall *a.* satisfied
Eccl. 8. 15. for that shall *a.* with him of his lab
Jer. 10. 10. nations not able to *a.* his indignatio
42. 10. if ye *a.* in this land I will build you
49. 18. no man shall *a.* there, 33. 150. 40.
Hos. 3. 3. thou shalt *a.* for me many days
11. 6. and the sword shall *a.* on his cities
Joel 2. 11. the day is terrible, who can *a.* it?
Mic. 5. 4. they shall *a.* for now shall he be gre
Nah. 1. 6. who can *a.* in the fierceness of his an
Mal. 3. 2. but who may *a.* the day of his com
Mat. 10. 11. there *a.* *Mark* 6. 10. *Luke* 9. 4.
Luke 19. 5. for to-day I must *a.* at thy house

Cruden's Concordance of the Holy Bible, first published 1737

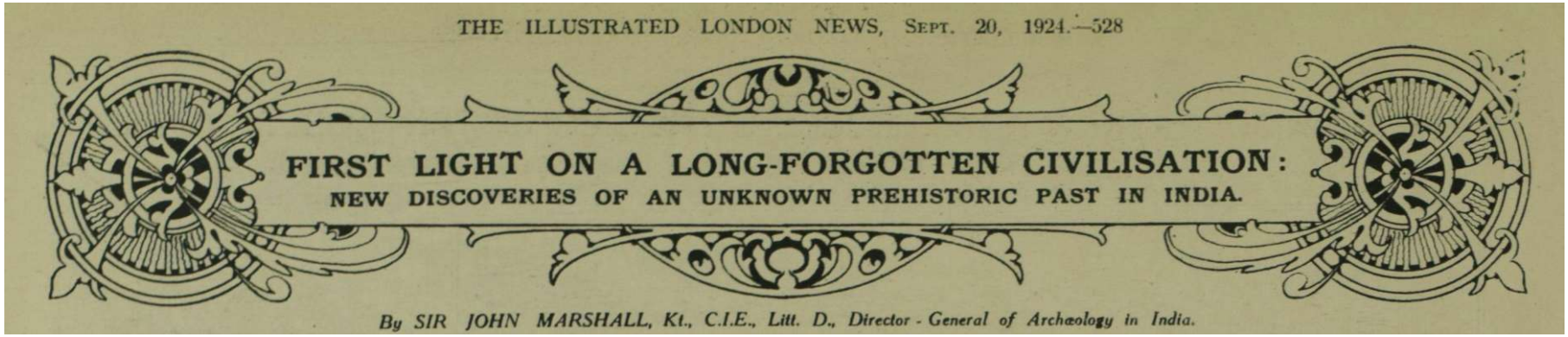
Importance of a Corpus

Houston & Coe (2003) give ***five criteria that must be met before a successful decipherment can take place***:

(1) ***A large well-published database***
(2) Correct identification of a known language
(3) One or more bilingual texts
(4) A well understood cultural context
(5) Pictorial references if the script is logographic

Discovery of the Indus Civilization

- John Marshall announced the discovery of the IVC in 1924. Field work by RD Banerji in Mohenjodaro and Daya Ram Sahni in Harappa
- Assyriologists point to similarities with Sumerian and Elamite scripts



THE ILLUSTRATED LONDON NEWS, SEPT. 20, 1924.—528

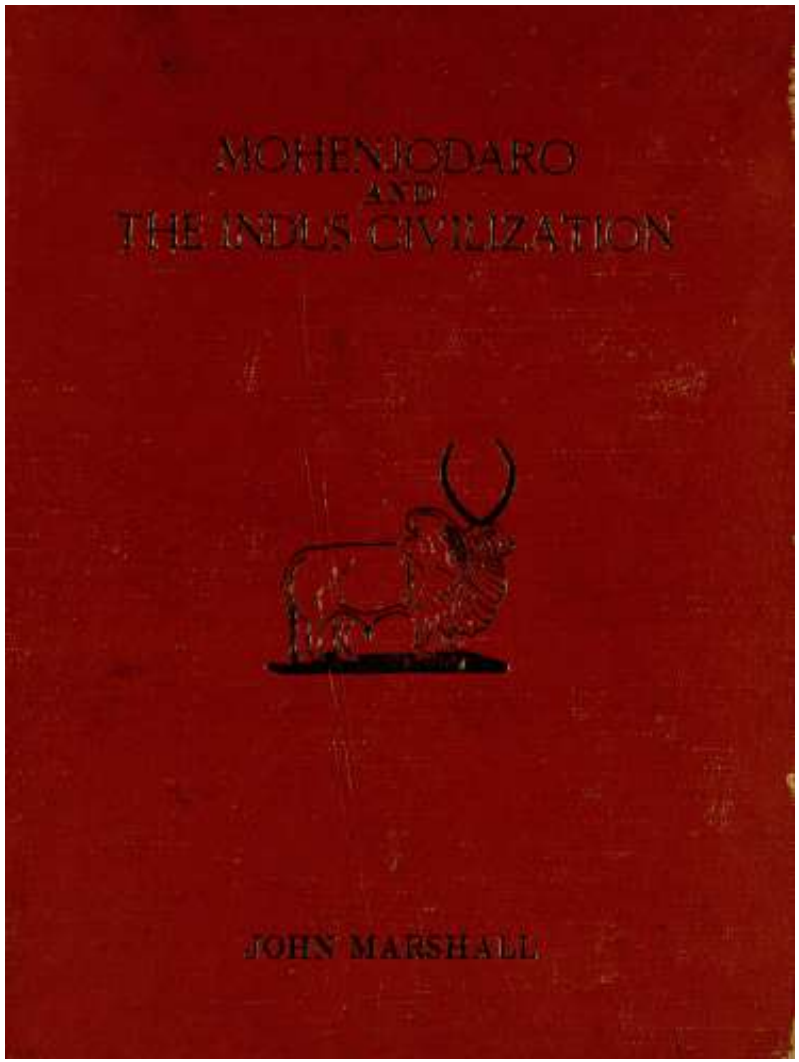
FIRST LIGHT ON A LONG-FORGOTTEN CIVILISATION:
NEW DISCOVERIES OF AN UNKNOWN PREHISTORIC PAST IN INDIA.

By SIR JOHN MARSHALL, Kt., C.I.E., Litt. D., Director-General of Archæology in India.

Sir John Marshall's announcement of the discovery: 1924

Marshall's Report

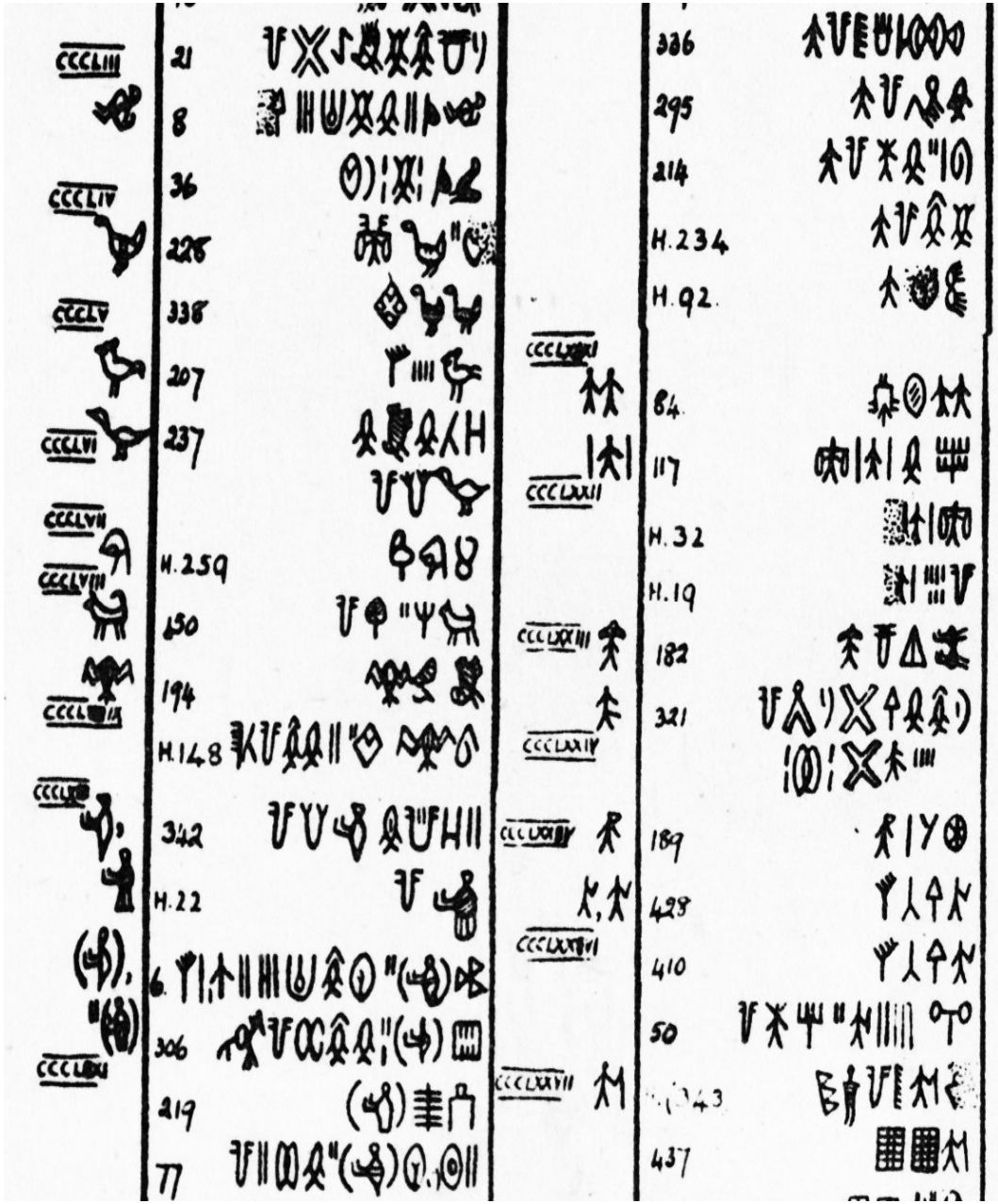
- John Marshall publishes Mohenjodaro and Indus Civilization (MIC) in 1931
- Publication contained "sign manual" by Gadd and Smith
- Also contained Langdon's hypothesis connecting the Indus Script to Brahmi



Mohenjodaro and the Indus Civilization

Gadd and Smith Sign Manual

- First Sign manual of Indus Script
- Contained mix of unique signs and sign variants



Hunter's Concordance

- Originally part of G.R. Hunter's doctoral dissertation at the University of Oxford in 1929. Published as a book in 1934
- Hunter first created a new sign list of 234 distinct signs, eliminating some of the sign variants from the Gadd and Smith sign list
- Based on his sign list, Hunter codified all known (750 or so inscribed objects) Indus texts (both Mohenjodaro and Harappa) into 102 tables. These tables sorted by the unique sign, and aligned on that sign, formed the basis for his analyses

Hunter's Concordance (contd.)

- Tagged each object and referenced them to the source (e.g. museum)
- Provided line drawings of all inscribed objects
- Provided comparison between the Indus Signs and the signs of Sumerian and other scripts
- However, limited by the number of objects included – only about 750

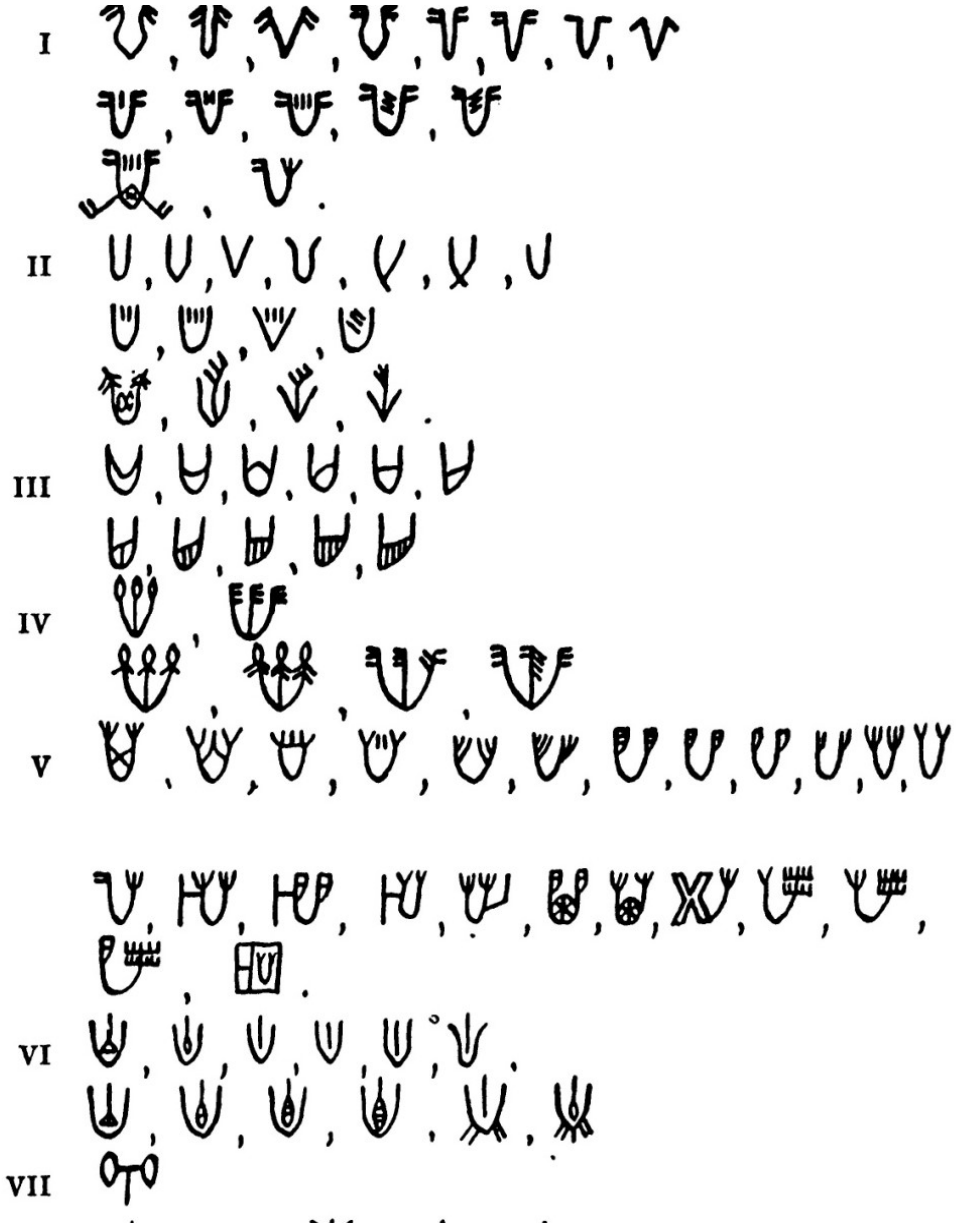
Hunter's Concordance (contd.)

TABLE XIII

No.	Text No.	Text	Sign
241	M. 460		
242	461		
243	M. 427		
244	M. 507		
245	161		
246	231		
247	484		
248	406		
249	363		
250	375		
251	326		

A table from Hunter's Book

Sign Grouping by Hunter (1932)



Hunter's Concordance – Several Firsts

- Hunter's concordance was the first to:
  - Analyze the signs individually, and identify variants
  - Study the structure and formation of Indus signs and analyze how compounding of signs occur
  - Formally tabulate the texts by reference sign so that the text can be aligned and presented.
  - Provide statistical analysis in establishing the direction of writing
  - Study the possible grammatical and textual meaning of signs and sign combinations
  - Develop rules of segmentation of texts
  - Perform contextual analysis Indus text; and
  - Perform comparative linguistic analysis between Indus script and Sumerian, Egyptian, Semitic and Sanskrit scripts

Hunter's Concordance – Contextual Analyses

- Comparison of certain sign combinations at Mohenjodaro vs. Harappa
- Sylistic analysis of the field symbols (animal images that occur on seals).
- Preponderance of copper tablets in Mohenjodaro and their near-absence in Harappa.
- Clay tokens were more common in Harappa.
- While Mohenjodaro and Harappa shared the same set of signs, the sign combinations occurring in Mohenjodaro were very different from those occurring in Harappa

Computer-aided Analysis of the Indus Script

- 1960's saw the deployment of computers in the analysis of the Indus Script
- Soviets and the Finns spearheaded the Indus script analysis
- Very little is known about the Soviet analysis (ie, tools used, corpus details, etc), but the Soviets concluded that the structure of the Indus script is closest to the Dravidian
- The Finnish team announced 1969 the decipherment of the Indus script. Turned out to be rather premature

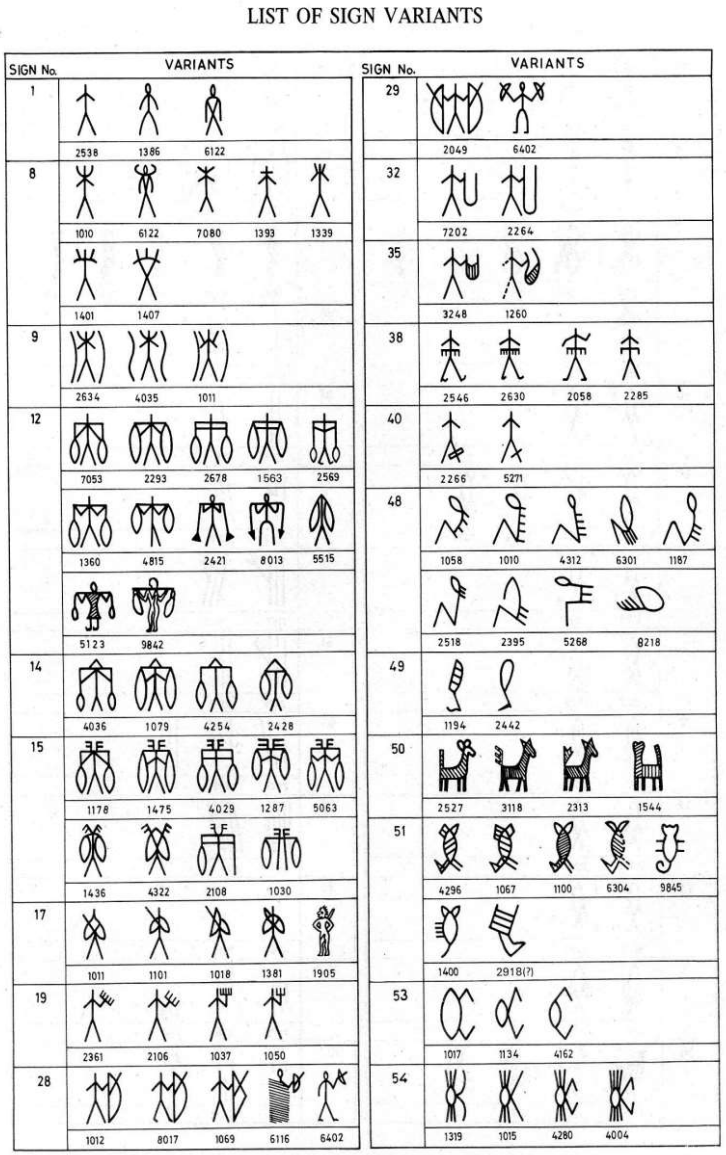
Mahadevan Concordance

- Mahadevan was inspired by Hunter's work
- Broke into the Indus Script Analysis scene in 1970 with a paper drawing attention to the parallels between Dravidian languages and the Indus Script
- Meticulously analyzed all inscribed objects and photographs available in the National Museum, Delhi, and created a photographic card catalogue
- Created an early version of his computer-aided corpus/concordance in 1971

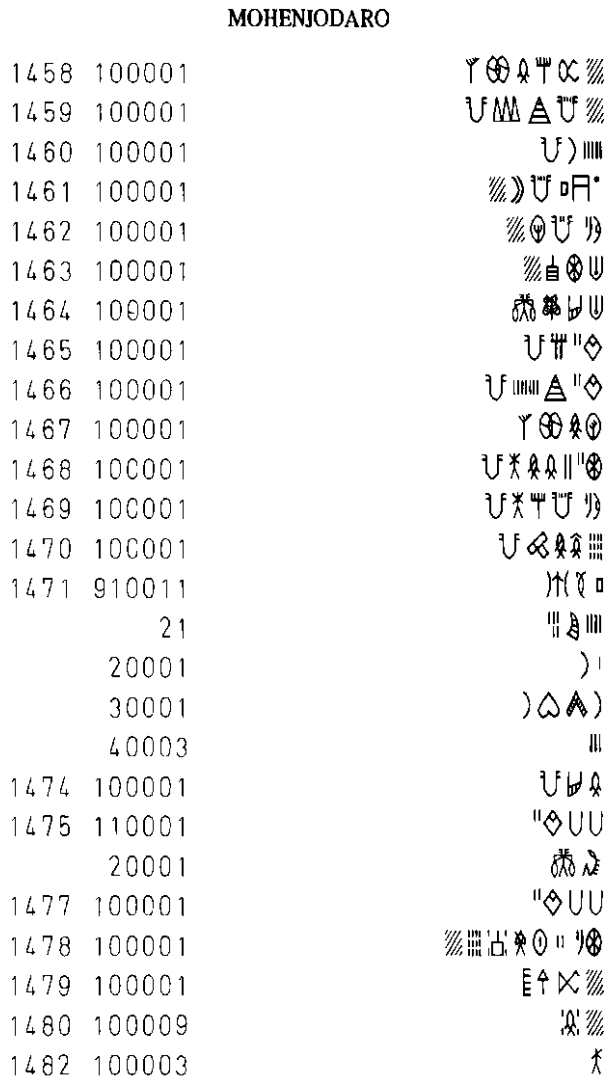
Mahadevan Concordance (contd)

- Second version of the concordance created in TIFR in 1972/73
- Improved version of this concordance was published in 1977
- Standardized sign list of 419 (only 417 were used in the corpus, concordance and analyses) unique signs and all possible sign variants
- Contained 2,906 inscribed objects from 26 sites
- Concordance and statistical analyses
- Textual and Contextual Analyses

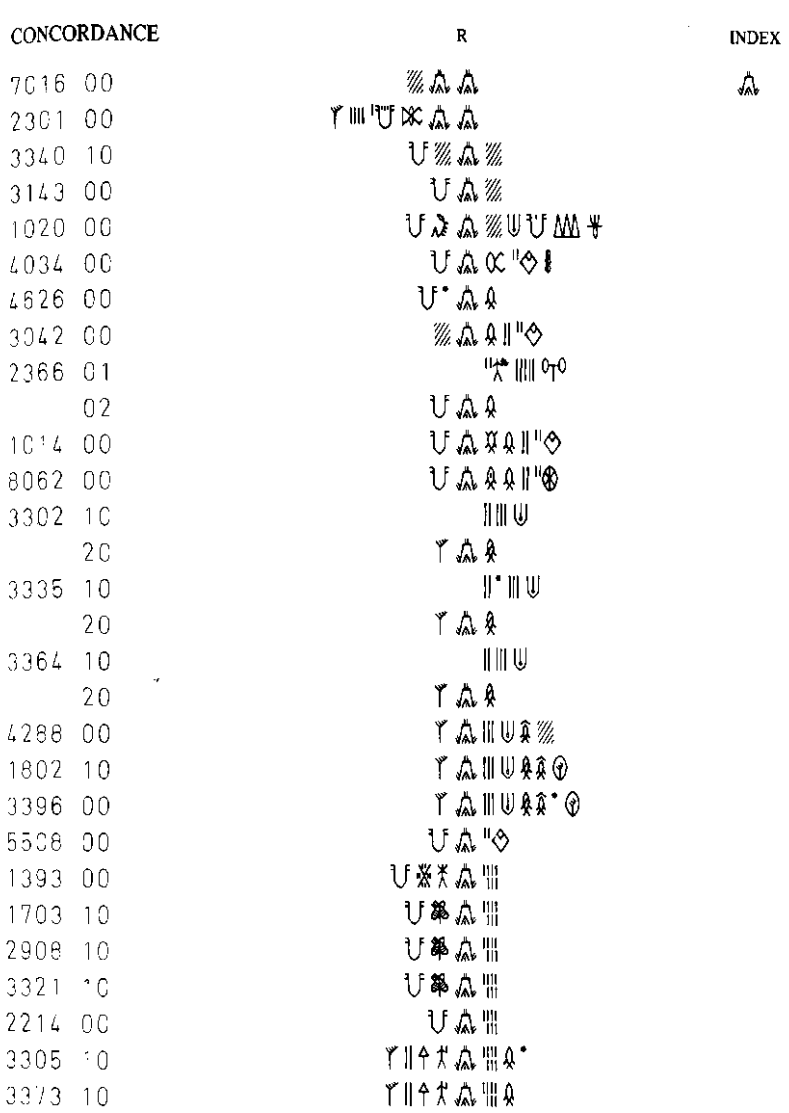
Mahadevan Sign & Sign Variants List



Mahadevan Corpus & Concordance



Mahadevan's Corpus



Mahadevan's Concordance

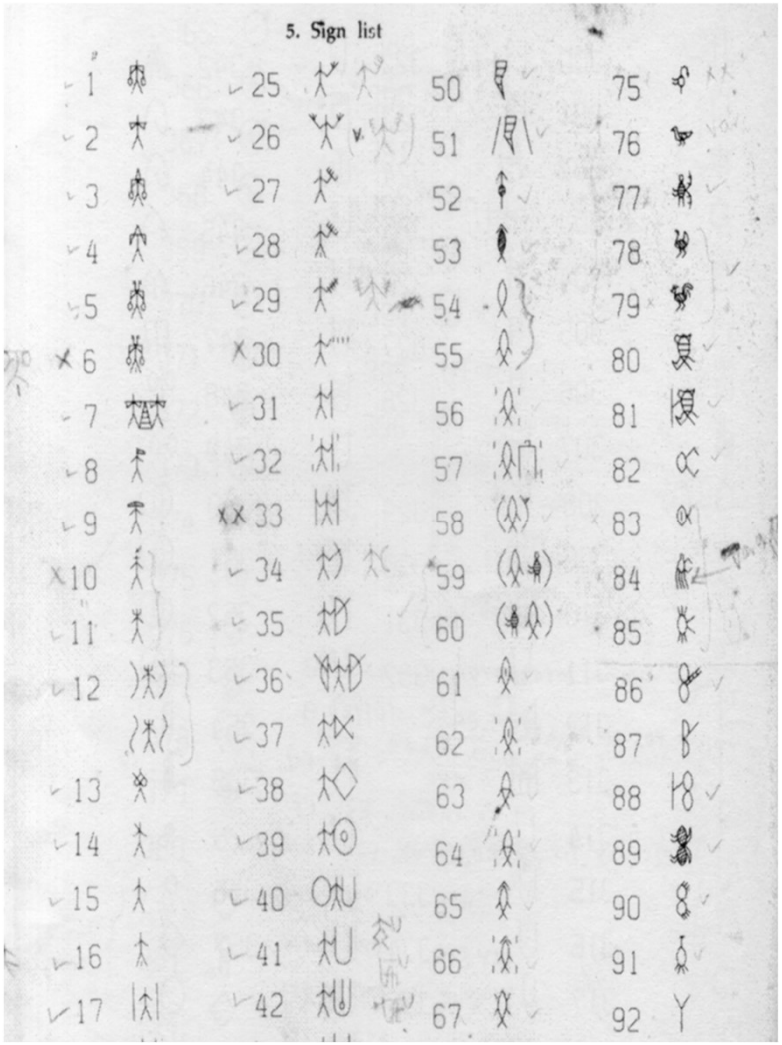
Analyses Provided Mahadevan Concordance

- Textual Analysis - frequency and positional distribution information of signs.
- Pair-wise Frequencies: According to Mahadevan, pair-wise sign combinations represent the simplest form of word formation, and hence this is an important data to analyse.
- Context Analysis: Analyzed the relationships between the inscription and the objects on which they are inscribed; the inscribed objects and the archaeological context of their occurrence; and inscription and the field symbols that occur with them.
- Site of occurrence: Analyzed the relationship between the various Indus sites and (a) signs of the script, (b) direction of writing and (c) field symbols.
- Types of inscribed objects: Analyzed relationship between the object types and (a) signs of the script, (b) direction of writing, and (c) field symbols
- Field Symbols: First to study the field symbols, and their relationship with inscription, object types and sites. This is one of the least studied areas of the Indus script, and only recently research efforts are being made in this area.

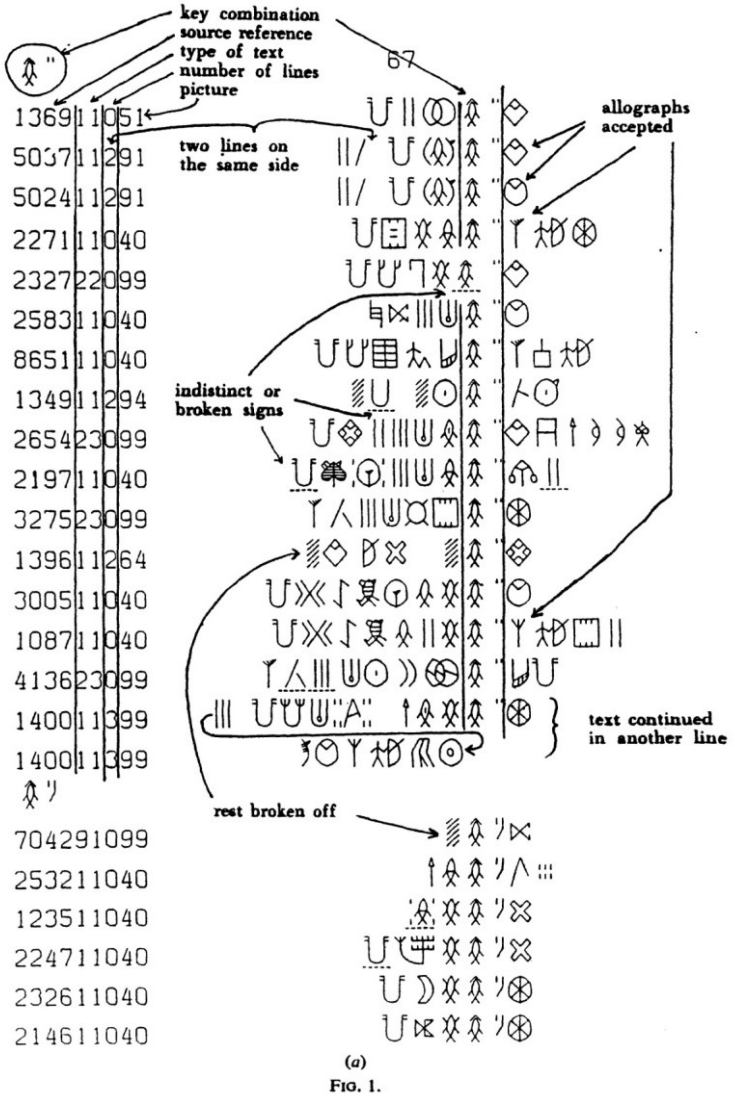
Parpola's Corpora/Concordances

- With Seppo Koskenniemi and Simo Parpola created a "Pairs" concordance in 1973
- With Kimmo Koskenniemi in creating a sign list, corpus and concordance in 1979-82
- Created 5-Volume CISI - the most comprehensive photographic corpus of Indus inscribed objects to-date
  - Vol 1with JP Joshi, 1985
  - Vol 2 with SGM Shah in 1991
  - Vol 3 with P Koskikallio and RH Meadow in 2010 (HARP data)
  - Vol 3.2 with many , including J-F Jarrige and Massimo Vidale
  - Vol 3.3 with Petteri Koskikallio (Iranian Borderlands)

Parpola's Pairs Concordance

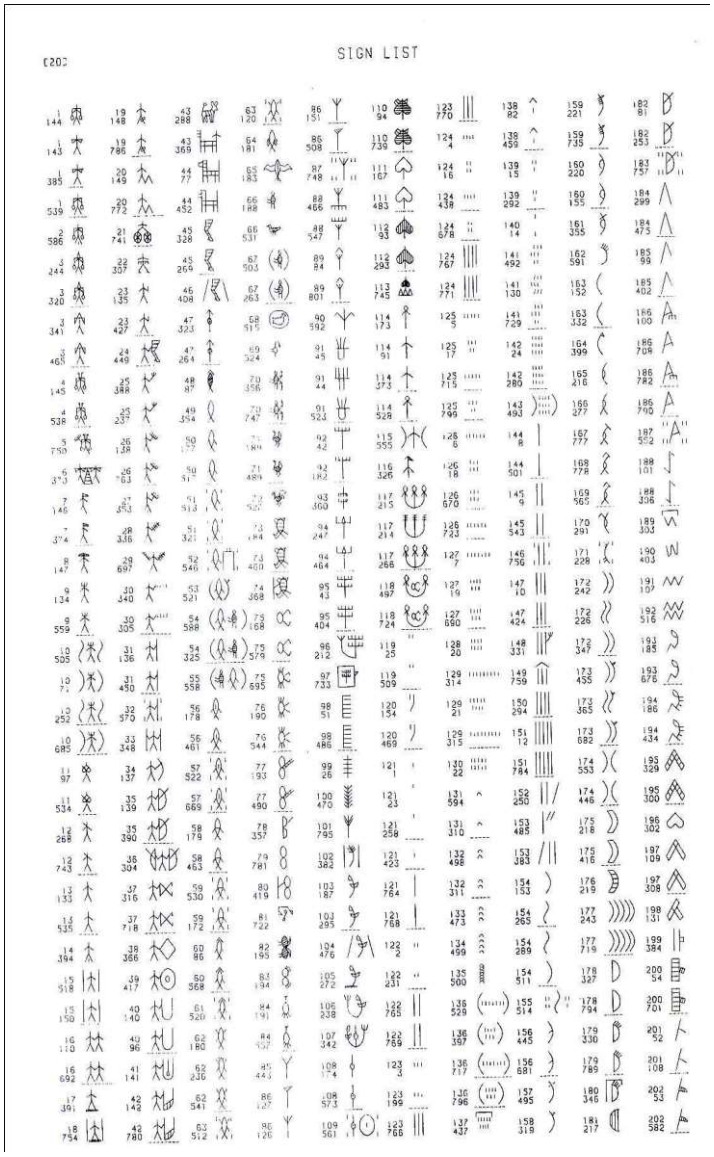


Parpola's Pairs Concordance -Sign List

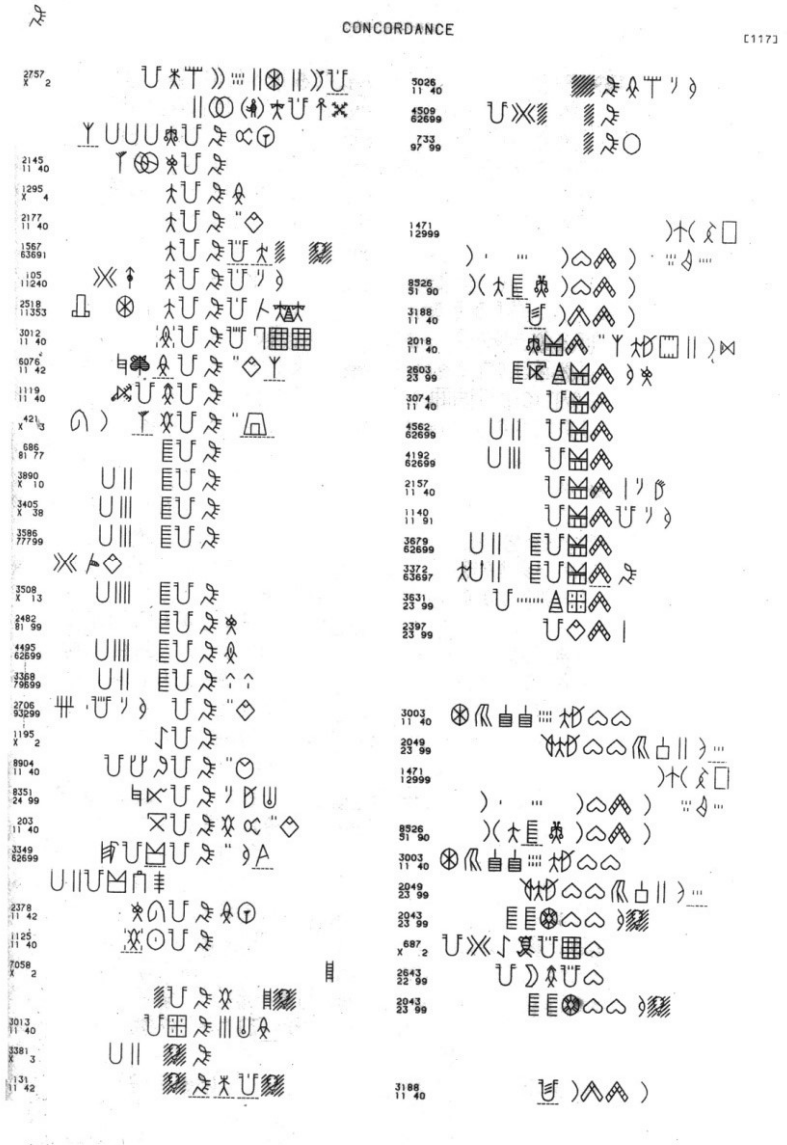


Parpola's Pairs Concordance Layout

Parpola's 1979-82 Concordance

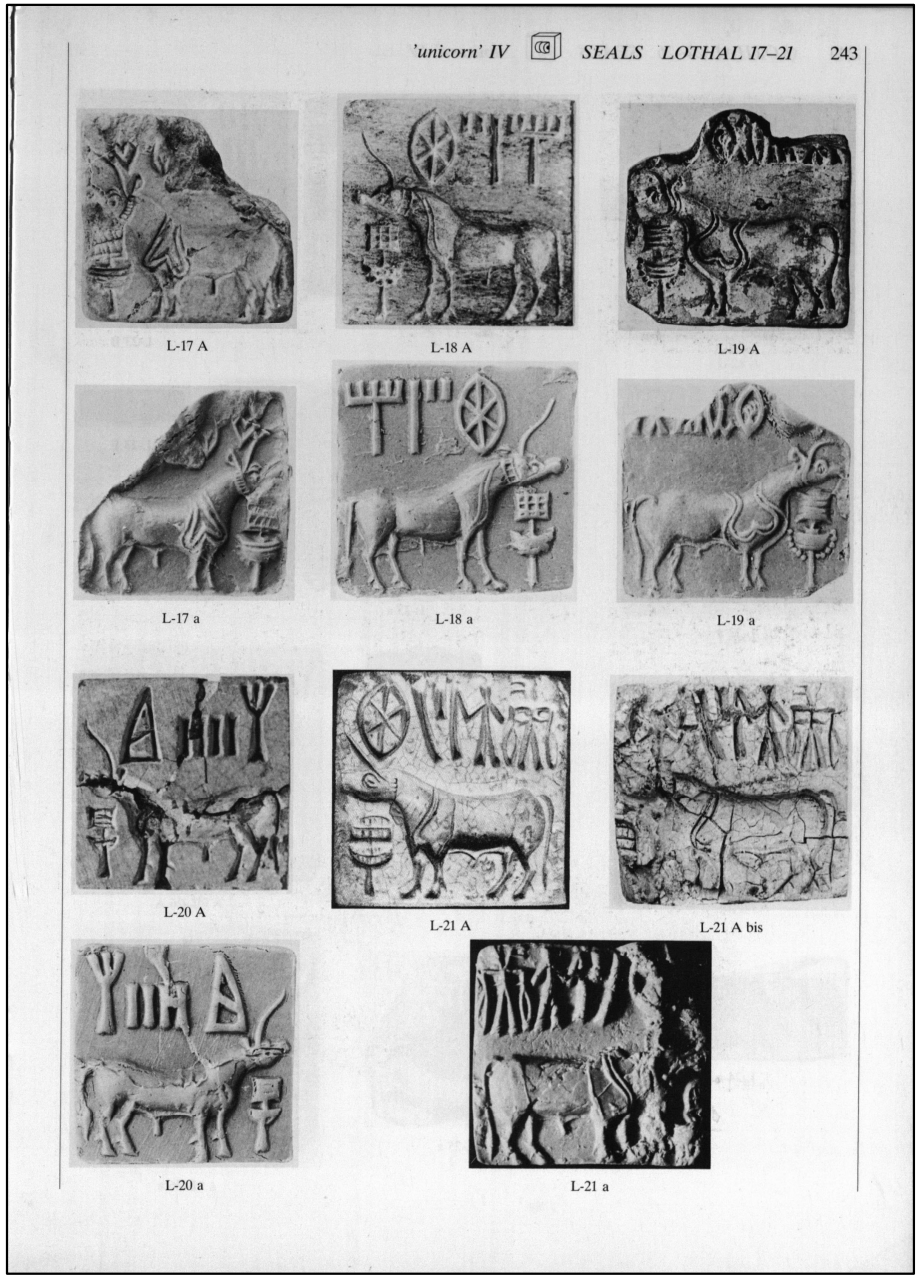


Parpola's Sign List

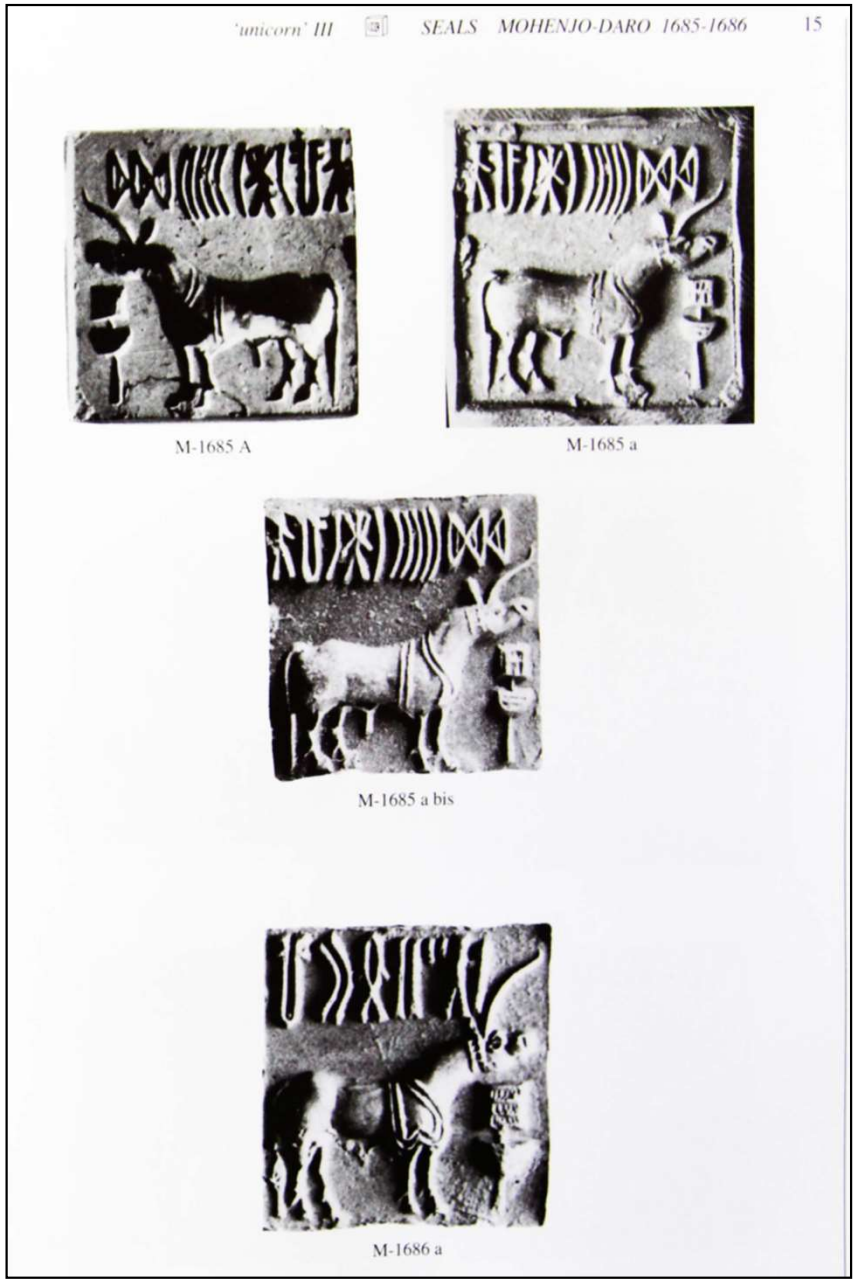


Parpola's Concordance

Parpola's CISI Volumes



Parpola's CISI Vol 1



Parpola's CISI Vol 3.1

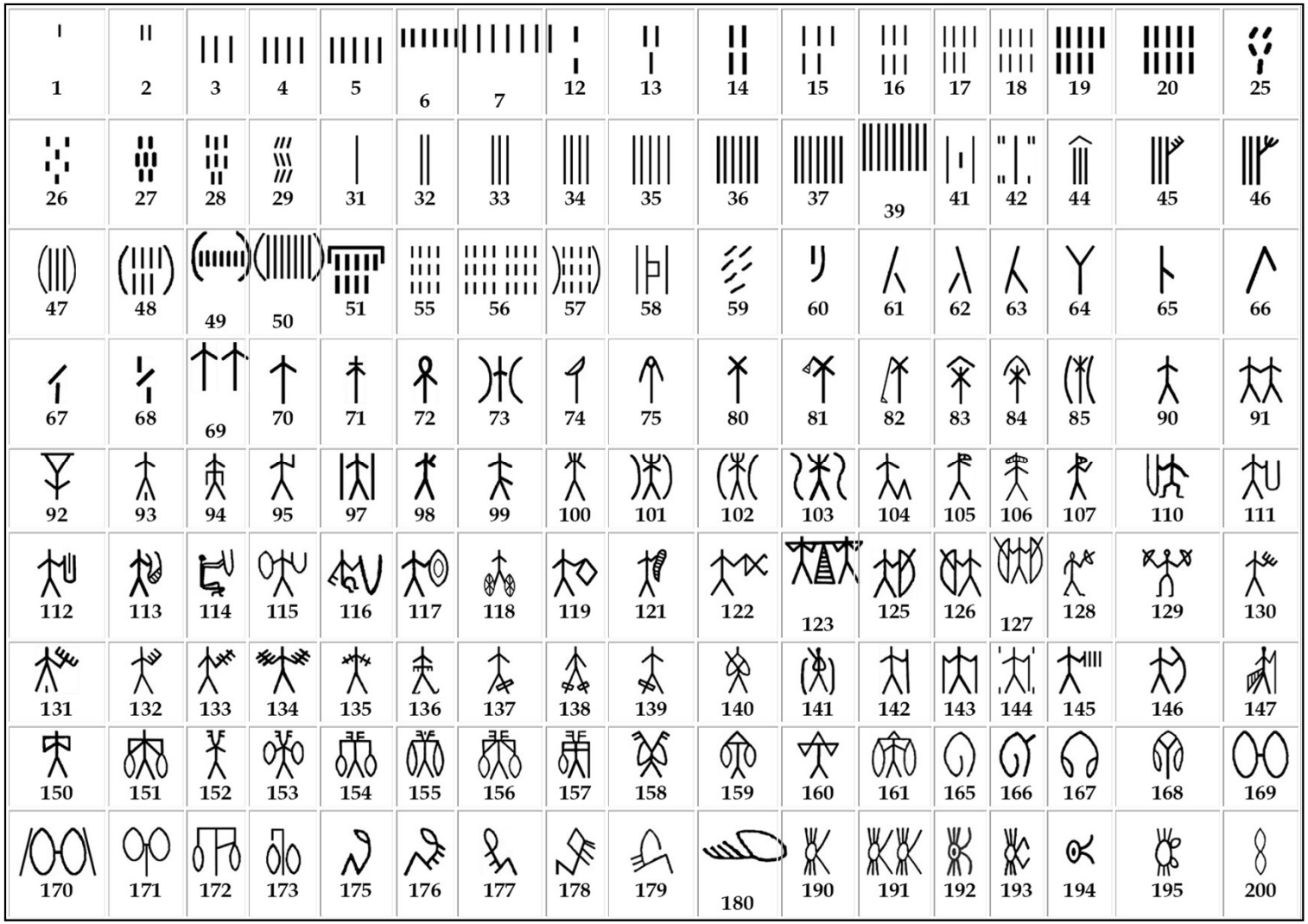
Summary of Parpola's Contribution

- His 1982 Concordance matches Mahadevan Concordance very closely
- Parpola and Mahadevan were in close agreement on the number of signs and sign variants
- Significant contributions to the analysis include segmentation analysis and grid analysis
- CISI is monumental work that resulted in a photographic catalogue of all known inscribed Indus objects. The scope of this multi-volume effort has expanded beyond the core Indus area.

Bryan Wells Corpus

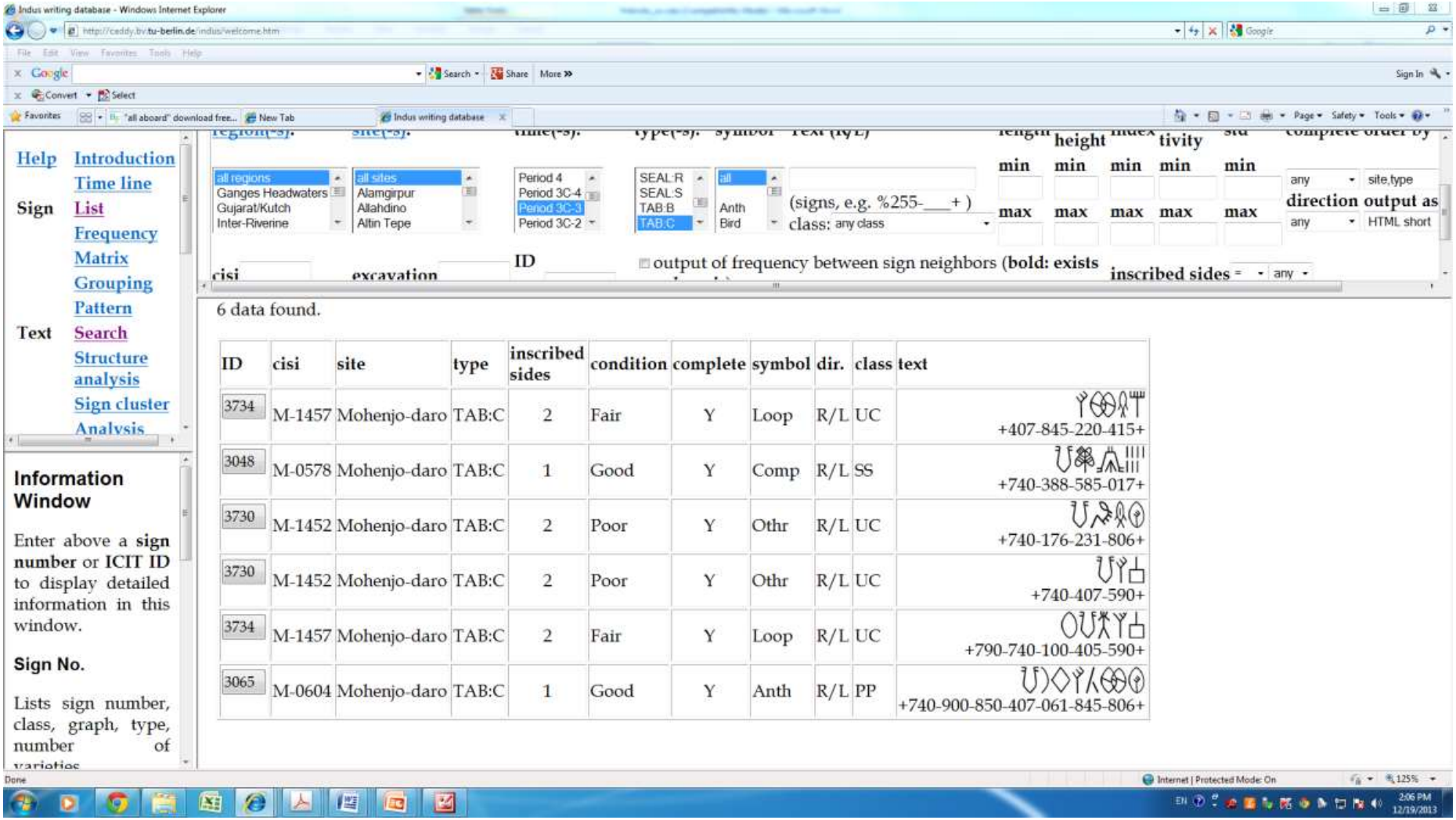
- Initial work in the form of his master's thesis on the Indus Script ; included an expanded sign list
- Created his corpus/concordance as part of PhD dissertation
- Corpus includes Harappa Archaeological Research Project (HARP) data
- Currently available in web-enabled form
- Many dynamic analysis features
- Tabular display of data does not follow the "concordance" format

Wells Sign List



- Significantly more signs (676) than Mahadevan or Parpola. Reduplicated signs, mirror images and sign variants treated as separate signs

Wells Corpus – Web-enabled



- Many dynamic analysis features
- Data presentation not anchored on the key/reference sign

Wells Corpus – Key Contributions

- The most comprehensive Indus script corpus available currently
- Web-enabled and freely available access to scholars
- The first corpus to be created primarily from photographic images of the objects, rather than directly from the objects themselves
- Providing a different approach to handling signs and sign variants, even though has resulted in a large sign list
- Providing analysis of the Indus signs (particularly sign clusters)
- Providing structural, textual and contextual analysis of the Indus script
- Providing hands-on analysis tools
- On-going addition of new features to the on-line corpus
- Needs to be institutionalized

Comparison of the Four Corpora/Concordances

	Criteria	Hunter (1934)	Mahadevan (1977)	Parpola (1979-82)	Wells (2006)
1	Number of Objects	750	2906	Approx. 3000	3835
2	Unit of Analysis - Entire object, each line, etc.	Side	Line	Side	Side
3	Reference Number	Exc. Number	IM Number	Parpola Number	Own internal number; uses CISI Number where available
4	Direction of Writing (General)	Right to Left	Right to Left	Right to Left	Right to Left
5	Sign List	Yes, but not published as a separate list	Yes	Yes	Yes
	Number of Signs	232	419*	394	676†
	Grouping of Signs	Not in the original work, but in his 1932 paper.	Yes	Yes	yes
6	Sign Variance List				
	List of variants	Not in the original work, but in a later, 1932 paper.	Yes	Yes, but identified in the sign list itself	No
7	Sign Structure Analysis	Yes	No	Yes	Yes
8	Textual Analysis				
	Frequency Analysis	Rudimentary	Yes	Yes	Yes
	Positional Analysis	Some	No	Yes	Yes
9	Field Symbols				
	Classification	No	Yes	No	Yes
	Frequency Analysis	No	Yes	No	Yes

Comparison of the Four Corpora/Concordances (contd.)

	Criteria	Hunter (1934)	Mahadevan(1977)	Parpola(1979-82)	Wells(2006)
10	Contextual Analysis	Yes	Yes	No	Yes
	Site information	Yes	Yes	No	Yes
	Locus	No	Yes, but partial	No	Yes, but partial
	Stratigraphy	No	Yes, but partial	No	Yes, but partial
11	Linguistic Analysis	Yes	Yes	Yes	Yes
	Linking Indus to another known language	Yes	Yes	Yes	Yes
	Interpretation of Indus Texts using a known language	No	Yes	Yes	Yes
12	Nature of Language	Phonetic	Logo-Syllabic	Logo-Syllabic	Logo-Syllabic
13	Review by other scholars	Some	Many	Some	Very Few
14	Assumptions about the Harappan language	Yes. Considers it pre-Aryan, possibly Dravidian. Munda is also a possibility	Proto-Dravidian	Proto-Dravidian	Initially Proto-Dravidian, now leans towards Munda or Language X
15	Function of inscribed objects	Discussed	Discussed	Discussed	Discussed
16	Availability	Easily Available	Easily Available	Not easily Available	Easily Available
17	Ease of Use	Not Easy to use, in paper form	Easy to use, but in paper form. Restricted availability of data in electronic form	Easy to use, but in paper form	Easy to use, electronic form
18	Use of Corpus/Concordance by others	None	By Many	By a Few	By Many

Contribution of Corpora and Concordances to the Indus Script Research

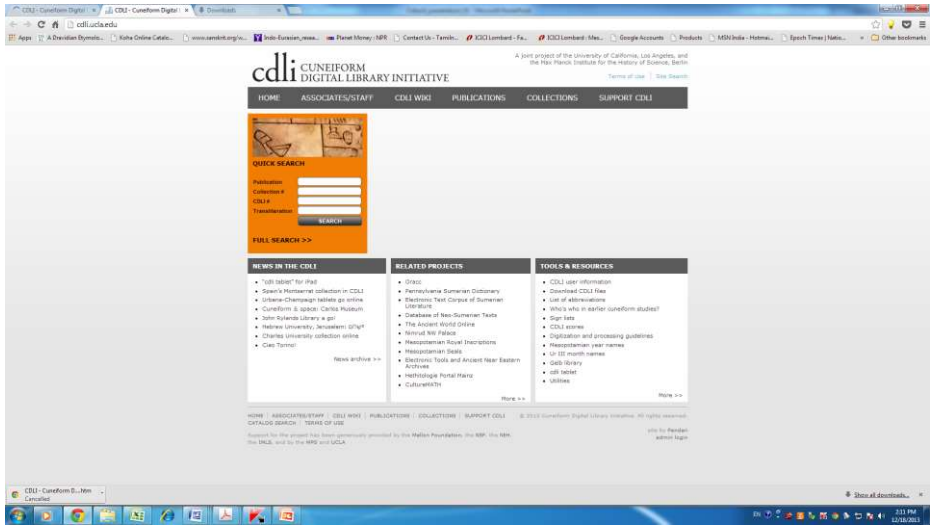
- Provided a comprehensive repository of Indus inscribed objects and a way of organizing them through standard sign and sign variant lists
- Performed statistical analyses of the data and made them available to scholars. This set the stage for further research
- Performed textual and contextual analyses of the Indus script
- Contributed significantly to the understanding of the Indus script
- Identified the potential areas for future studies that will lead to further understanding and potential decipherment of the Indus script.

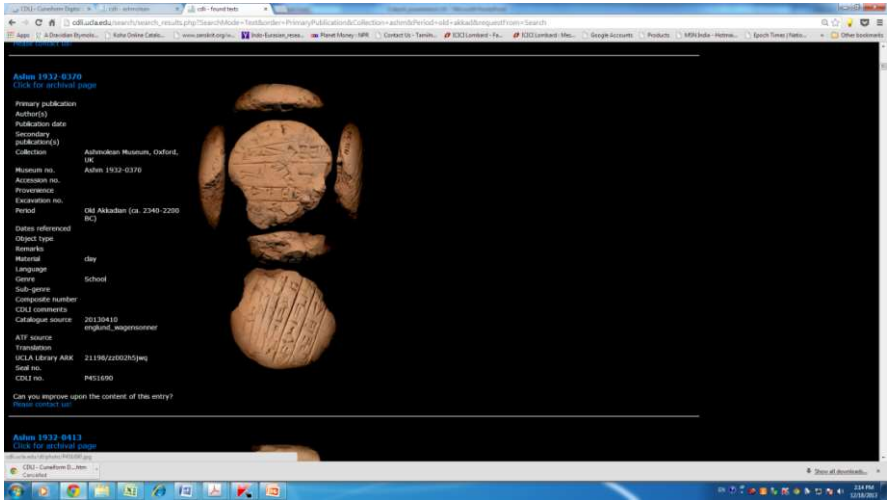
Possible Focus Areas of Future Research

- Contextual Analysis
  - Object type (e.g. seals, terracotta tokens, etc.)
  - Geographical variation
  - Stratigraphy
- Multi-disciplinary Research
- More textual analysis and interpretation of texts
- Development of additional analytical tools

Way Forward

- Two major issues
  - Lack of new talent to continue with research work
  - Availability of up-to-date corpus/concordance of Indus texts
- 3-D rendering of objects will help compensate for lack of physical access
- Cuneiform Digital Library Initiative (CDLI) is worthwhile emulating





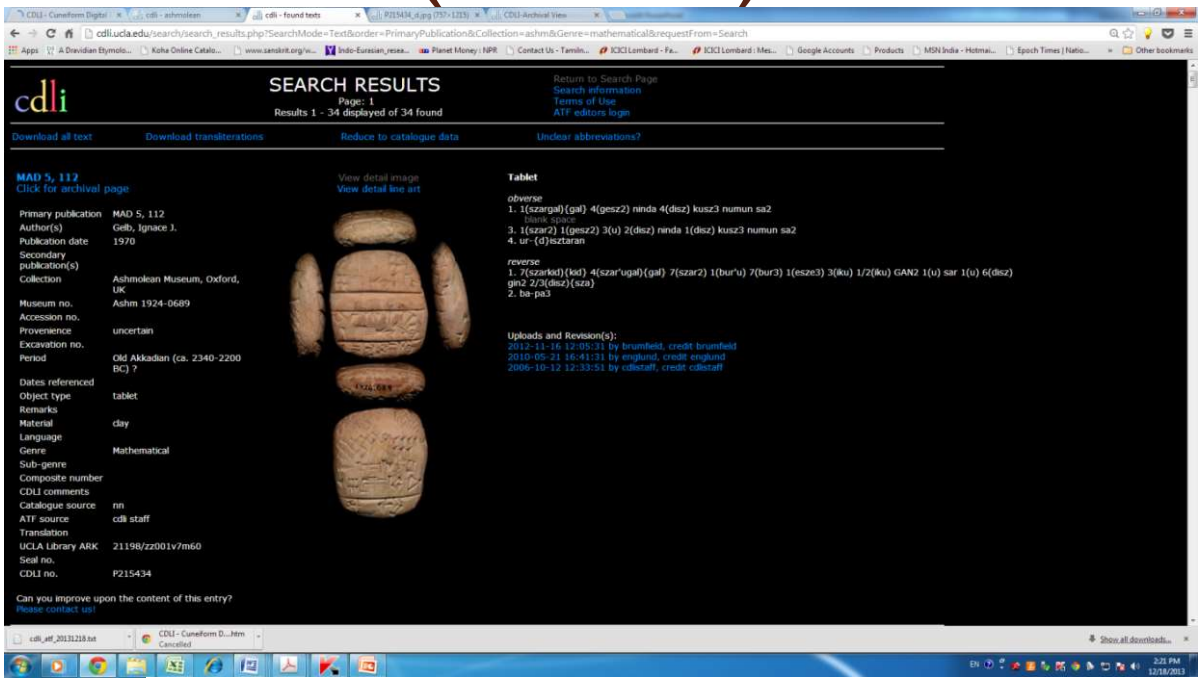
Selected Pages from Cuneiform Digital Library Initiative *(CDLI) Website*

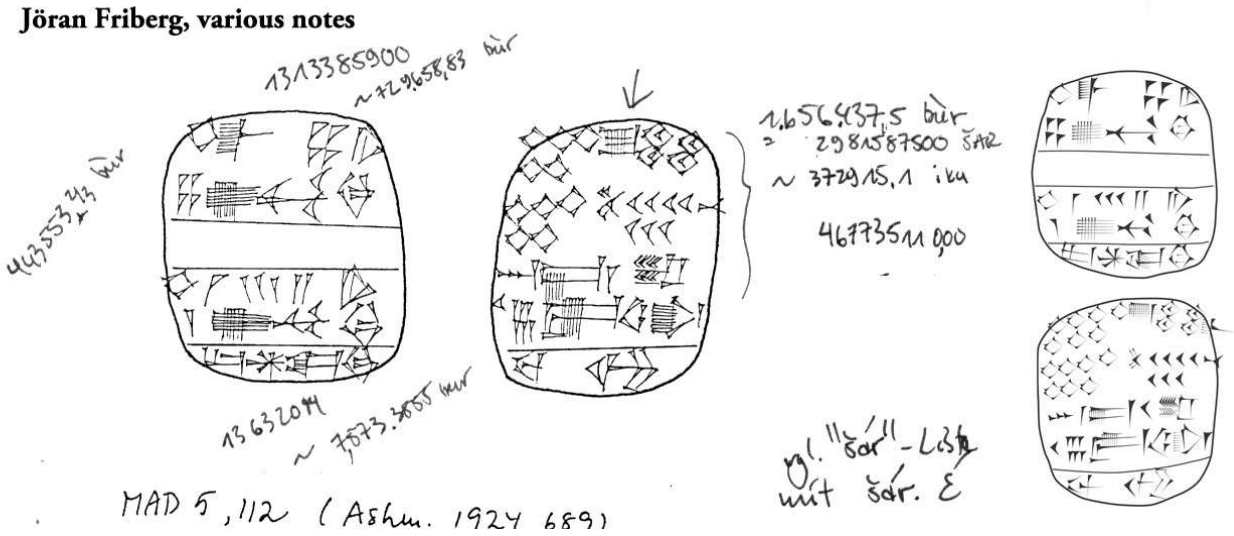
Way Forward (Contd.)

- The Cuneiform Digital Library Initiative (CDLI) represents the efforts of an international group of Assyriologists, museum curators and historians of science to make available through the internet the form and content of cuneiform inscriptions dating from the beginning of writing, ca. 3350 BC.

- A number of these artifacts are currently kept in public and private collections to exceed 500,000 exemplars, of which now more than 360,000 have been catalogued in electronic form by the CDLI.

Way Forward (contd.)





Selected Pages from Cuneiform Digital Library Initiative *(CDLI) Website*

Recent Progress

- **Multi-disciplinary approach in the research of the Indus script**
- **Interactive Corpus of Indus Text (online database of Bryan Wells' work) enhanced to cross reference Mahadevan's data**
- **Omar Khan (Harappa.com) recently mentioned about a concerted effort to raise capital to acquire multiple Indus script data sources currently available and create a comprehensive data source with analysis tools, similar to CDLI**

Q & A

Acknowledgements

I thank the following publications/websites for the material and images used in this presentation:

Marshall, John, First Light on a Long-forgotten Civilisation: New Discoveries of an Unknown Prehistoric Past in India, Illustrated London News, September 20, 1924, pp 528-532.

Marshall, John; Mohenjo-Daro and the Indus Civilisation, Probsthain, London, 1931 (Reprint 1973, by Indological Book House, Delhi).

Hunter, G.R.; The Script of Harappa and Mohenjodaro and its Connection with Other Scripts; Kegan Paul, Trench, Trubner & Co. Ltd., London, 1934.

Mahadevan, Iravatham. The Indus Script. Texts, Concordance and Tables, 1977, Archaeological Survey of India.

Koskenniemi, Kimmo, and Asko Parpola. 1982. A concordance to the texts in the Indus script. Helsinki: Dept. of Asian and African Studies, University of Helsinki.

Wells, Bryan, and Fuls, Andreas: Online Indus Writing Database. Berlin. Last Modified Date: 9-2-2011. http://caddy.bv.tu-berlin.de/indus/welcome.htm. Last Access Date: 19-12-2013.

CDLI - Cuneiform Digital Library Initiative. https://cdli.mpiwg-berlin.mpg.de/

Thank You!