

Volume 3 : 6 June 2003

LINGUISTICS INFORMATION IN THE INTERNET WITH SPECIAL REFERENCE TO INDIA

B. A. Sharada, Ph.D.

CLICK VIRTUAL LIBRARY IN LINGUISTICS to access the variety of information discussed in this article!

1. INTRODUCTION

Languages are the vehicles of knowledge. Languages can survive only through the continuing processes of creation, re-creation, and wider participation. During the centuries of its existence, a great language like Sanskrit, a treasure house of knowledge, was confined only to books, because some people tried to possess it as their own divine-ordained privilege, by preventing others from using it in their day to day discourse. Now, the 21st century has opened the doors of knowledge to every one through the internet. The digital world has not only shrunk the globe into a miniature web but also expanded the availability of network of knowledge in the open source. This leads not only to knowledge dissemination but also to knowledge expansion and creation. In this virtual world every individual is empowered equally. There is no rich or poor, intelligent or dull. Here, one has to know how to, and where to look for the required information.

So, in order to help the common people and the scholars looking for information on Indian languages, the present study focuses on the availability of electronic information in Linguistics with special reference to India in the internet, and provides a guideline to the users where to locate what, in the enormous digital world.

2. A VIRTUAL LIBRARY

SAVE THE TIME OF THE READER is the first law of library science derived by S. R. Ranganathan that is applicable to library of documents in any form such as books, microforms, digital/electronic, etc. In the broadest sense, a virtual library is a system by which users access information that resides solely in electronic format on computer networks, without respect to physical location of the information.(Pacifici 1997). Many articles are published on both virtual library and electronic information resources and their advantage over print media. Among the electronic formats such as floppy diskettes, CD-ROMs, magnetic tapes, OPAC's of library collection, the information available on the internet, etc., the last one is focused in the present study.

The time has come to information scientists to realize their role to be played in this electronic era. Understanding the users' need in the present day circumstances and

helping them with right information at right time has become absolutely necessary. The information technology has innovated tremendously in achieving these goals. Aspects relevant to our needs could be picked up and applied in appropriate manner. This adds to the responsibilities of the information scientists. Design and development of domain specific Virtual library is one of those. The present study concentrates on the discipline Linguistics.

3. PURPOSE OF THE STUDY

There is no second thought about the abundance of information in the internet. Keeping this in view, getting the right path and going through each and every link file is time consuming for the researcher. Though Meta data is helpful in getting the required document, to get in touch with the domain specific Meta data is also a point to be thought of. One simple and practical example: In the process of collecting the data for the present study, the phrase "linguistics search engine" was given to the following: Google search retrieved 44,500, Altavista 38,992 and Alltheweb 90,430 results. From this, one can ascertain up to what extent we can rely on these search engines and navigational tools. All listed ones are not actual search engines. It includes an individual's contribution, those of organizational, commercial, etc. In fact it is not an easy task to go through all the sites they list out.

Many studies have proved that no single search engine is self sufficient in any information retrieval process. Google is the most popular example of a search facility which is hosted by a third party organization. While discussing search engines, Phil Bradley compares two popular search engines and suggests that Alltheweb has overtaken Alta Vista in terms of general usefulness (Hunter 2001). Alltheweb combines many of the best features of other search engines and has customizable interfaces I have seen in a very long time (Bradley 2001). Each search engines has its own unique features such as,

New! Refine your search with AltaVista Prisma

[Help](#)

Click a term to focus your search or click >> to replace your search.

Applied Linguistics >>
Computational Linguistics >>
Corpus Linguistics >> Historical Linguistics >>
Linguistics Resources >>
Search Engine Marketing >> Search Engine Optimizatio... >>
Search Engine Ranking >>
Search Terms >> Search Tools >>
Internet Resources >>
Bibliography >>

Google lists in the above search "Search Engine of Search Engines". If this site is clicked, Look.Com is the site we get. Following Table displays how many web sites are there in each theme in Linguistics. And most of the sites cover all aspects such as printed and electronic books, journals and dissertations, Institutions, etc.

Applied Linguistics (14)	Morphology and Syntax (25)
Bilingualism (70)	Natural Languages (3,252)
Comparative Linguistics and Typology (6)	Onomastics (13)
Computational Linguistics (85)	Philosophy of Language (154)
Computer Assisted Language Learning (28)	Phonetics and Phonology (71)
Constructed Languages (328)	Pragmatics (20)
Dialectology (21)	Psycholinguistics (25)
Discourse Analysis (11)	Semantics (43)
Graphemics and Orthography (57)	Semiotics (94)
Historical Linguistics (45)	Sociolinguistics (11)
	Translation (151)

Other Categories

To bring the web nearer to the user in a user friendly manner, if research libraries take up the creation of domain specific virtual libraries and design the information media with pointers, it will be of immense help to the whole research community.

4. TOOLS TO ACCESS INFORMATION IN LINGUISTICS

As a preliminary step few Search Engines were chosen for scanning. Following are some examples for search engines. Google, Askjeeves, Looksmart, Lycos, Netscape search, Overture, Teoma, SavvySearch, Alltheweb, Yahoo, Pathfinder, etc. Where and how to search the needed information was the question. Some more search engines are available with special reference to linguistics in the sites: **ZDNet: Search IQ: Directory of Linguistics Search Engines & ... WWW Virtual Library**

Search Engines : Social Science/Linguistics and Human Languages were the two important sites for Linguistics. A click on ZDNet will show the AltaVista site. Here, key in the search terms. For example if we type "e-books in linguistics" many sites are listed. They may not deal with e-books in linguistics only. E-books are treated as one term and linguistics as one term and what ever site that server locates it will list. We have to choose the relevant site such as, <http://www.amazon.com/exec/obidos/ASIN>.

5. KINDS OF AVAILABLE INFORMATION

Information in the internet at the global level is a representative of many kinds of information available in different forms such as

Books, Periodicals, Dissertations, Reports, etc.

The information requirement can also be seen in

- Ø Biographical information of Linguists
- Ø Reference Works

- Ø Institutions/Organizations in Linguistics
- Ø Other information resources on the internet

Also linguists can basically get two kinds of information:

- § Information on 'Publications' both printed and digital that are published
- § Information itself such as papers, articles, etc., that are sometimes unpublished.

The former includes digitized bibliographic database which can be accessed online by registered members and in few cases non-members also. In some cases even the membership is charged. For example EBSCO data base for periodicals should be subscribed so also the ISI (International Science Information) where in citation index journals are published. Some sites are commercial and some are academic. Based on this criterion the sites listed in the present study are chosen. One thing the reader has to keep in mind is that, the web sites change their agenda quite often wherein time is immaterial.

In the conventional card catalogue that existed in the information era, readers had to go to the library card cabinets, find the location of the document and then go to shelf in search of it. The approach would be mainly 'Author', 'Title', 'Series' etc. In this age of information technology, the whole library in the digital form along with the metadata will be available on the readers' desk top. The document can be accessed in number of ways depending upon the defined fields in the database.

Tools to Access the Information

The Home page named as Virtual Library in Linguistics (VLL) is created using the Microsoft front page. Under each heading, its related sites are listed with link to each one of them to the internet. Some files are available offline also when the internet is not on. The image of the Home page of VLL appears as in Picture 1.

Welcome To Virtual Library

About the site	Suggestion from users	E-mail us
--------------------------------	---------------------------------------	---------------------------

CONTENTS

1. [Books](#)
2. [Journals](#)

3. [Linguists](#) 4. [Institutions](#)
 5. [All in One](#) 6. [Other E-Information Resources](#)

Supportingsites-

Picture 1

Under each document types following sites are mentioned as mentioned in picture 2.

My-Simon	Ethnologue	Amazon	D.K.Publishers	Ask.com
----------	------------	--------	----------------	---------

Journals

E-Journals					Print Journals	
Virtual Library	Look Smart	EPSL	Ask.com	EBSCO	Language in India	CIIL
Alta Vista	Excite	Hotbot	The Linguist List			
Alta Vista	Excite	Hotbot	CIIL			
Bibliographies	Census data-India	Tribal Languages-India	Dissertations	Languages of the world	Reference Sources	

In addition to the above Languages of India, Language map of India, and IT News are some more added to Other E- Information.

6. HOW TO USE SPECIFIC TOOLS?

Books in Virtual Library in Linguistics

In the earlier study made on Citation analysis of Indian Linguistics, the usage of books were more in Linguistics that supported other studies in Humanities such as Anthropology, Sociology, etc. (Sharada1983). So, priority in this venture is to help readers to refer to books. First and foremost thing is to create a place for books in linguistics in the VLL Home page. With the result, Home page acts as a base that links to the sites from where one can retrieve information about books. A single mouse click on one of the sites listed under books will open the whole database.

Let us take an example of a site which is concerned with Books Amazon.com. It is a commercial site where one will get the information on a needed book such as – its availability, number of copies, time taken for shipping, cost of the book, etc. One click on the book will display Cover page, Table of Contents, Excerpts, Full Index, Back cover and an option inside the Book. A click on the above title page will partially describe the physical bibliography part and in addition to it following information is added. It is not only helpful to readers but also helpful in selecting the relevant books for the library.

Look for similar books by subject:

Browse for books in:

- [Subjects](#) > [Reference](#) > [Words & Language](#) > [Linguistics](#)
- [Subjects](#) > [Reference](#) > [General](#)

Search for books by subject:

[Computational linguistics](#)

[Methodology](#)

[Discourse analysis](#)

[Data processing](#)

[English language](#)

[Language Arts / Linguistics / Literacy](#)

[Language Arts & Disciplines](#)

[Linguistics](#)

In the above said options one can check and go to that particular item. In the similar way click on each site such as “My Simon”, “Ethnologue”, “Amazon.com”, “D.K.Publishers” and “Ask.com”. They shall display their content.

Etnologue is available in CD-ROM version also. In addition to these sites mentioned above, few Universities and Academic Institutions where the section on faculty gives publications under each faculty and some professors allow open access and down loading of their papers. For example: The site of Education Policy survey of Arizona State University is useful for information on Language planning.

Journals

Journals are seen in print format and electronic journals. Example : For print journals in linguistics, a click on CIIL journals will list out the journals subscribed by CIIL library in print format. Link file for each journal is under construction.

For electronic journals, a click on virtual library shall lead to many links. For example: Under the title, **Arts - Language Studies and Linguistics NEWCASTLE UNIVERSITY LIBRARY** lists **Electronic Journals**. A click on latest journals in the web page lists those journals. Each journal has link files. The same site further classifies into the following categories and links those items to the document.

- [Working papers in linguistics](#)
- [Monographs](#)
- [Journals](#)
- [Conferences](#)
- [Lectures](#)

In the item Journals mentioned above, list of publishers are given. For example, a click on Oxford University Press will list out few journals with full link files. A click on each journal will open up wide information about the same. In case of EBSCO database since it is ready to use database, no work remains to be done on our part other than just guiding the users to go in proper way retrieve the data. Full text is available either in html or pdf formats for few articles here and abstract available for the most. The EBSCO database monitors journals referred at the global level irrespective of discipline. But one has to subscribe for the database. One important on line journal devoted to Indic articles in Linguistics is **Language in India** available at the site <http://www.languageinindia.com>.

Linguists

[AltaVista](#)

AltaVista has sites dedicated to searching in 21 countries.

The Americas

[Brazil](#)

[Canada](#)

Europe

[Austria](#)

[Belgium](#)

[Denmark](#)

[Finland](#)

[France](#)

[Germany](#)

[Ireland](#)

[Italy](#)

[Netherlands](#)
[Norway](#)
[Portugal](#)
[Spain](#)
[Sweden](#)
[Switzerland](#)
[UK](#)

Asia-Pacific

[Australia](#)
[India](#)
[Korea](#)
[New Zealand](#)

Choose a country and a click on that will display the Linguists of that nation. If the names are not listed, the required information may be provided in a prescribed form in the same site. In the similar way sites such as [Excite](#), [The Linguist List](#) and [Hotbot](#) also provide information on linguists.

Institutions

The <http://www.ciil.org> site gives the Directory of Organizations and Institutions devoted to Language and Literature in India. A provision is made to update the Directory by providing the contact ID in the site. Also a click on [culturopedia](#) shall take to a site listed under Tribal languages. This site covers most of the Organizations and Institutions in India alphabetically. Linguistic Institutions are also included. If this is limited to India, <http://www.altavista> provides information on Institutions in Linguistics that has link to universities at the international level .

All in One refers to the above listed categories, the information of which is available in one site. The site mentioned here <http://linguistlist.org/> covers all aspects of linguistics.

Other E-Information Resources are **Bibliographies, Census data-India, Tribal Languages-India, Dissertations, Languages of the world, Indian languages, Language map of India, IT News and Reference Sources.**

Bibliographies

The sites dealing with linguistic Bibliography that are electronically available are: MLA bibliography, Summer Institute of Linguistics (SIL) bibliography, Koninkljkku Bibliotheek

The famous site [mla.org](http://www.mla.org) is explained for the convenience of the readers.

When you enter the <http://www.mla.org>. Click on publications. On the left side menu click on “ MLA Bibliography”. The *Bibliography* is still available in print from the MLA, as well as on CD-ROM from [SilverPlatter](#) Information. It is accessible on the Internet. The distributors of the electronic format of the bibliography are

[SilverPlatter\[http://www.silverplatter.com/\]](http://www.silverplatter.com/),

[OCLC\[http://www.oclc.org/\]](http://www.oclc.org/),

[Ovid,\[http://www.ovid.com/\]](http://www.ovid.com/),

[Gale Group\[http://www.galegroup.com/about/contact.htm\]](http://www.galegroup.com/about/contact.htm).

One may select any site and get the required information.

7. ELECTRONIC FORMAT OF THE *MLA BIBLIOGRAPHY*

The electronic format of the bibliography contains all the entries in the print issues from 1963 to the present. It is searchable by classification headings, document authors, subject-index terms, language of the book or article, journal name, and other elements. Electronic distribution formats include online, Internet, and CD-ROM; all electronic formats are updated ten times a year.

MLA Directory of Periodicals

The *MLA Directory of Periodicals* contains all information available on the journals and series on the bibliography's Master List of Periodicals. The entries list editorial addresses, phone numbers, fax numbers, e-mail addresses, frequency of publication, descriptions of the periodicals' scopes, circulation figures, subscription prices and addresses, advertising information, and submission guidelines. The directory also provides statistics on how many articles and book reviews the periodicals publish each year, as well as how many are submitted.

Scope

Any regularly published journal available to libraries or universities that publishes articles on language, literature, or folklore with some frequency is eligible for inclusion on the Master List. Any series that includes books on language, literature, or folklore, regardless of the frequency of the series, is also eligible. In order to see the periodicals on the [list, go to the site](http://www.mla.org/www_mla_org/publications/publications_main.asp)

Availability

The directory is available in an electronic format at no charge to subscribers of the *MLA International Bibliography* who subscribe through SilverPlatter, Inc. or OCLC. The *MLA Directory of Periodicals* cannot be purchased separately from the bibliography. It is updated every six months. For information on subscribing to the *MLA International Bibliography*, see <http://www.silverplatter.com/> or <http://www.oclc.org>. The directory will soon be offered by Gale and Ovid.

Census data-India

In order to get information on census data of India <http://www.censusindia.net/> is the site. The language data has been covered in "C - Series Tables"

SOCIO - CULTURAL TABLES

C-7 Distribution of the 18 State/District Part Scheduled		State/District/ Tehsil/Town	
A(i) Languages-1991			
C-7 Part A(ii) Distribution of the Mother Tongues (Having a minimum of 10,000 speakers in India) included under the 18 Scheduled Languages-1991	State/District		State/District/ Tehsil/Town
C-7 Part B(i) Distribution of the 96 Non-Scheduled Languages-1991	State/District		State/District/ Tehsil/Town
C-7 Part B(ii) Distribution of the Mother Tongues (having a minimum of 10,000 speakers in India) included under each of the 96 Non-Scheduled Languages-1991	State/District		State/District/ Tehsil/Town
C-8 Bilingualism and Trilingualism	State		State/District/ Tehsil/Town

Tribal Languages of India

The information on tribal languages of India is available in the site entirely devoted to Indian culture and heritage and well known as treasure house of Indian culture. A click on [culturopedia](http://www.culturopedia.com/Languages/tribal_languages(a).html) in the VLL will lead to the site [http://www.culturopedia.com/Languages/tribal_languages\(a\).html](http://www.culturopedia.com/Languages/tribal_languages(a).html)

The tribal languages are listed alphabetically. Any letter may be selected among the alphabets

Dissertations

The search term “linguistic dissertation India” in the Google search will show 5790 results. Some will be full text and some contain metadata on dissertations.

A site devoted to dissertations in India is [‘www.vidyanidhi.org.in](http://www.vidyanidhi.org.in) where one can find search facility for Meta data and full text dissertations and the dropdown search menu includes Compiler, Title, Subject, Keywords, Guide, etc. There are many institutions devoted for this purpose in Linguistics outside India.

Languages of the world

For information on language data of the world one can refer to Languages Data in the [Ethnologue](#) site which is available off line in the form of CD Title being [Ethnologue : Language of the world](#). About each language following information is given.

- Alternate names
- Number of speakers
- Location
- Dialects
- Linguistic affiliation
- Multilingualism
- Other demographic and sociolinguistic information.

In addition, the CD consists of

[Introduction](#) [Languages Data](#)

[Appendixes: Bibliography Abbreviations Languages of special interest](#)

[Geographic distribution of living languages](#) [Questionnaires](#)

[Links: Language Family Index Language Name Index Maps](#)

If the Ethnologue CD is made available in the system’s CD drive, you can get all the above information.

Indian Languages

Information on Indian Languages is found in <http://www.culturopedia.com/Languages> and also in <http://indic-computing.sourceforge.net/faq/indian-languages.html>.

Language map of India

The language map is illustrated in WebPages of Tamil Electronic Library under the link [Languages and Scripts of India](#). A single click on this will take us to [Languages and Scripts of India](#) and displays the map. Each language and the concerned state have a link file of its own and can get details of each language.

IT News

The IT news in Linguistics and on Indian languages are available in [bharathi bhasha.htm](#) (offline in the Virtual library home page) and <http://VishwaBharat.tdil.gov.in>

Reference Sources

The Oxford English Dictionary, the Grove Dictionary of Art, and the large reference works published by the Gale Group are pioneers in this gradual mobilization of reference resources to the worldwide web. (Hodgkin, 2001). Many Dictionaries and Encyclopedias are available in CDROM form.

For reference sources on Indian languages, one may depend upon <http://www.internetcount.com/~educa>.

For reference sources on Indian languages, one may depend upon <http://www.internetcount.com/~educa>.

8. FURTHER RESEARCH IN THIS FIELD

Researchers who need still fast, complete information dealing from web may go in for software tools such as Web Search Agents like “Copernic Pro 5.02” and “Bulls Eye Pro 3.0”. Copernic costs \$79.95 and the latter \$199.00 and sought as Intelliseek. A comparative study of the performance of these two packages states that Copernic Pro is the preferred product.(Perez,Ernest 2002). Perez in the same article list web-hosted search agents and provides the web addresses of them. In addition there are invisible web sites. As Marieke Napier puts it “There is no central authority to maintain an index and much of the authoritative information accessible over the Internet is invisible to the key search engines”.

They consider 4 levels of invisibility:

- The Opaque Web – files that can be but aren’t included in search engines such as PDFs and word files.
- The Private Web – technically indexable Web pages that have been excluded from search engines by being password protected or through the use of a robots.txt file.
- The Proprietary Web – Web only available to people who have agreed to certain terms in exchange for viewing the pages. Normally some form of registration is needed which may be free or cost money.

- The Truly Invisible Web – The Web, which cannot be seen due to technical reasons, lack of metadata etc. The majority of dynamically created pages created from content-rich databases fall into this category.

Sherman and Price explain that much of this invaluable material is comprised of from universities, libraries, associations, businesses, and government agencies around the world in database format.

The E-world keeps on changing as well as expanding. Best to keep oneself updated in research development in a discipline is to keep in touch with the popular sites in that particular discipline.

[CLICK VIRTUAL LIBRARY IN LINGUISTICS](#) to access the variety of information discussed in this article!

REFERENCES

Bradley, Phil.2001.Goodbye Altavista,hello Alltheweb. *Ariadne* Issue 30

Hunter,Philip 2001. Centering the periphery: A new equity in information access?
Ariadne Issue 30

Marieke Napier 2001. The Invisible Web: Uncovering Information Sources Search Engines Can't see. *Ariadne*, Issue 30 [Review of the book by Chris Sherman and Gary Price, *The Invisible Web: Uncovering Information Sources Search Engines Can't see*. Cyber Age Books, 2001. ISBN 0-910965-51-X

Pacifici(SI)1997.Virtual Libraries:Myth and reality. Posted January. Vol.3; Archived February 14,1997.(<http://www.llrx.com/features/virtual.htm>)

Perez, Ernest 2002. Web Search Agents. *Online*, Mar/Apr,Vol 26(2), p20.

2001. Reference books on the web. *Ariadne*, Issue 30