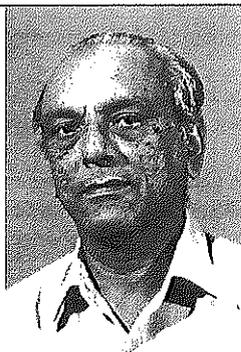**G.G. Chowdhury*, A. Neelameghan** and Sudatta Chowdhury*****
**Addis Ababa University, Ethiopia**

VOCON: Vocabulary Control Online in MicroIsis Databases

Prof.A.Neelameghan(b. 1927),MLib Sci, Post-MA in Special Medical Librarianship. Teaching, research and practice in information science and knowledge organization for over 30yrs. at DRTC, Bangalore and at several places worldwide on Unesco assignments. FIC/CR Chair 1973-80. Author of some 250 articles and 5 monographs, co-authored 25 books.

Chowdhury, G.G.,Neelameghan, A.,Chowdhury, S.:**VOCON: Vocabulary Control Online in MicroISIS Databases**
Knowl.Org. 22(1995)No.1, p. 18-22, 5 refs.
Discusses the need for facilities for online vocabulary control and standardization of terms, codes, etc., so as to secure consistency in naming of subjects, objects, countries, languages, etc., in databases at data entry stage. Most information storage and retrieval packages for microcomputers including MicroIsis provide for online vocabulary control in formulating search expressions for information retrieval, but not at the data entry stage. VOCON.PAS is a Pascal interface program for use with MicroIsis software for (a) online selection of term(s) and/or code(s) from vocabulary control tool, such as, a thesaurus, subject heading list, classification scheme, nomenclature list(s). (Authors)
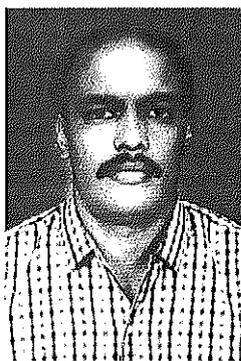
Dr. G.G.Chowdhury, b. 1959, is an Asst. Prof. at the School of Information Studies for Africa, a regional School of Information Studies, established by Unesco and IDRC, at Addis Ababa, Ethiopia. Dr.Chowdhury got his first PhD degree from Jadavpur Univ., India and the 2nd PhD from the Univ. of Sheffield, UK. Publications include three books and more than 30 articles. He is actively engaged in studies and research in different areas of Information Retrieval Systems, particularly Text Retrieval Systems.

Mrs. Sudatta Chowdhury, b. 1965, is a Lecturer in the Department of Library Science at Addis Ababa University, Ethiopia. Mrs.Chowdhury got her BLISc and MLISc degrees from India and MPhi. in Information Science from the UK. She is actively engaged in research in different areas of TextRetrieval Systems and Library Automation with particular reference to the MicroIsis software and the program interface.

1. Introduction

In developing databases, whether bibliographic or factual or other types, need arises to ensure consistency and uniformity in the use of terms to represent subjects, countries, languages, or objects such as, machinery, components of machinery, industrial products and trade items (foods and beverages, chemicals, pharmaceuticals. plants and plant products, animals and animal products, etc.), occupations and job categories, and so on. There may also be need to bring together under a group/category name or code related items, as is done with thesauri, classification schemes, group codes, etc. Securing consistency, compatibility in naming entities becomes crucial if the information in the database(s) is to be exchanged among or used by, different groups, as may be the case in an information network at national or international level. For ensuring such consistency, usually an agreed upon or widely accepted vocabulary control tool or coding scheme is used. Such vocabulary control, standardization, or authority lists include thesauri, subject heading lists, classification schemes, coding schemes (e.g., for diseases, industry and trade names, occupations, and countries), nomenclature lists, terminology lists, etc. There are, for instance, ongoing international programs for developing terminology databases with MicroIsis (1). These tools may be consulted at the time of data entry and appropriate terms and/ or codes selected for assignment to relevant fields in the record. The same procedure may be followed in representing the concepts/elements of a query so as to secure precise match in search and retrieval.

Some information storage and retrieval software packages for use on microcomputers enable online selection of terms from a vocabulary control tool in the formulation of search expressions for information retrieval. Such a facility can be added to, and in fact is available with, the standard MicroIsis (2) as a Pascal interface (THES.PAS). MicroIsis also provides for the linking of different records of a database through the use of REF and REF(L) functions in the display format specification. There is, however, no facility for consulting online vocabulary tool(s). selecting appropriate terms and/or codes, and adding them to user-specified fields in one or more records of a database. Doing this exercise manually is relatively cumbersome and time-consuming. Providing such facilities online is considered to be a useful feature of an integrated thesaurus software (3).

Thus the work reported in this paper arose of a practical need to provide an online facility for adding and replacing data to one or more specified fields in one or more records in a MicroIsis database, the data for adding to or replacing existing data in specified field(s) of one or more records of a database being selected from another database or files accessible online to the system.

18

Knowl. Org. 22(1995)No.1
G.G.&S.Chowdhury/A.Neelameghan: VOCON Vocabulary Control Online in MicroIsis Databases

2. VOCON.PAS: A Pascal Interface

An application program, named VOCON.PAS, has been written in CDS/ISIS Pascal. This program has been written in the light of the THES.PAS program, that comes with the MicroIsis software, but has been modified to a large extent to include a number of additional features. In addition to the facilities provided in the THES.PAS program, the VOCON.PAS program can be used:

(a) for online selection of term(s) and/or code(s) from a vocabulary control tool, such as a thesaurus, a subject headings list, a classification scheme, a nomenclature list or a code list that is in the form of a MicroIsis database, and adding the selected term(s) and/or code(s) the specified field(s) in the record(s) of a MicroIsis database, either replacing the existing data in those fields or adding to the existing data. Term(s) or code(s) may be selected from one or more vocabulary lists. The terms and codes may be added to a record specified by its master file number (MFN) or to a contiguous sequence of records specified by a range of MFNs or to the set of records retrieved by a search expression. The code(s) selected may be added to the same field or to one different from the field to which the thesaurus or subject terms are added;

(b) for fast opening of the data entry worksheet of MicroIsis database record(s) retrieved by specifying an MFN or a range of MFNs or through a search expression, for correcting or modifying data in those records globally, if necessary; and

(c) online updating, or modifying one vocabulary control tool with terms and/or codes selected online from other vocabulary tool(s) or code list(s).

3. Implementation of VOCON.PAS

The VOCON.PAS program should be copied to the PROG subdirectory of the ISIS directory or as indicated at parameter 1 of the SYSPAR.PAR file of MicroIsis. The program should be compiled to generate the object code (VOCON.PCD) file by selecting option A of the MicroIsis main menu, then option 'C' (compile) in the sub-menu, typing in, when asked for, the name of the program 'VOCON', and pressing the <ENTER> key.

The program can then be interfaced with and invoked from different points but mainly as a menu exit. For example,

1. Using option 'A' — Advanced Programming Services, in the main menu, then selecting option 'R' (run) in the sub-menu, keying in 'VOCON' for the name of program and pressing the <ENTER> key.
2. Modifying the Data Entry menu EXE1 (using option U — System Utility Services, of the main menu), introducing a new line, say, Y — Vocabulary Control Service (for details see 2), and interfacing VOCON program at option 'Y'. Then pressing 'Y' in the menu EXE1 will invoke the program.
3. While entering data in a record (using either option 'E' for editing of existing record or option 'N' for creating a new record in menu EXE1 — Data Entry Services), pressing <PgDn> and <F9> keys as indicated at the bottom of the worksheet screen, will invoke VOCON. The <PgDn> and <F9> option has been introduced in the sub-menu of data entry worksheet.

If the system hosts two or more vocabulary control tools, then to facilitate choice appropriate tool(s), a text file VOCLIST listing the names and indicating the respective features of the vocabulary tools may be prepared (e.g., using EDLIN) and copied to the DATA subdirectory of the ISIS directory (or as indicated in the parameter 5 of SYSPAR.PAR). The list will be displayed at the appropriate stage to facilitate the selection.

4. Operation of VOCON.PAS

There are two modes in which the program operates. In the first one, the user can run the program from a data entry worksheet by pressing the <PgDn> and <F9> keys as mentioned above. In the other mode, the user can run the program through option 'A' or by choosing option 'Y' in the main menu, as discussed above. Operation of the program in both modes are discussed below.

4.1 Mode 1

While the data entry worksheet of a record is open for entering new data or for editing existing data, the associated vocabulary control tool(s) can be accessed for selecting appropriate term(s), code(s), class number(s), etc., and add them to specified fields (either replacing existing data or simply adding them) of the record. For this purpose, the user has to choose the <PgDn> and <F9> keys as shown at the bottom of the worksheet. Once these two keys are pressed consecutively, the following screen will appear.

```
Use Vocabulary Control — Y[es] N[o]
```

If the response is 'N', the related data entry worksheet will be opened for editing and one can still avail of online vocabulary control by pressing <PgDn> and <F9> keys at the bottom of the worksheet. If the response is 'Y', the following cues will be displayed:

```
A[dd] or R[eplace] term(s)
Key in tag no. for term(s) or <CR> to continue:
Key in tag no. for code or <CR> to continue:
```

and on responding to these cues, a list of vocabulary control tools (VOCLIST) will appear on the screen (Figure 1).

```
Available Vocabulary Tools
_____

AGRIS — AGRIS Subject Codes
AGROV — FAO AGROVOC
COMMON — Common Isolates List
CORPN — Corporate Name Acronyms & Abbreviations
ISOC — ISO Country Codes
ISOL — ISO Language codes
THES — OECD Macrothesaurus
THESB — Information Retrieval Thesaurus
THESM — Artificial Intelligence Thesaurus
                          More — Y/N
```

Figure 1: List of Vocabulary Tools Display

Knowl. Org. 22(1995)No.1
G.G.&S.Chowdhury/A.Neelameghan: VOCON Vocabulary Control Online in MicroIsis Databases

19

Thus the user can view the list of vocabulary control devices available to the system. Responding 'n' to the Cue 'More — Y/N' at the bottom of the box, will invoke the system to ask for the name of the tool to be selected:

```
            Which Vocabulary Tool?
```

Supposing that the THES (OECD Macrothesaurus) has been selected and the name keyed in followed by <CR>, the screen shown in Figure 2 will appear.

```
   ┌─────────────────────────────┐
   │                             │
   └─────────────────────────────┘

   Select Term
```

Figure 2: Initial Screen for Thesaurus Term Selection

If the term selected in the box is 'Enterprises', the thesaurus, as shown in Figure 3, will appear.

```
ENTERPRISES
                          ┌─MFN─┐
                          │1659 │
ENTERPRISES               └─────┘
NT AGRICULTURAL ENTERPRISES
NT2     AGRICULTURAL COOPERATIVES
NT3          FISHERY COOPERATIVES
NT2     FARMS
NT3          EXPERIMENTAL FARMS
NT4             FAMILY FARMS
NT3          SMALL FARMS
NT4             MINIFUNDIA
........

........
 NT FOREIGN ENTERPRISES
 NT2     FOREIGN BANKS

<SpB>Next B[ack] F[irst] P[age] S[elect] T[erm se-
lect] Q[uery] M[ark term(s)]
G[Markcode] ?[Disp query] A[ddrel] D[elete] C[reate]
W[rite in Fld] X[exit]
```

Figure 3: Display of Thesaurus Terms

To select term(s) to be added to the field specified earlier the Space Bar may be used to move the cursor against the desired term(s), say ENTERPRISES, and M (Mark term(s)) pressed. If code or codes are to be chosen, then G (Mark code) is to be pressed. One may select term(s) and/or code(s). To add or write the terms and/or codes onto the record, W (Write in Fld) is to be pressed. The system will then display the following message (in this case, only the term ENTERPRISES has been chosen; no code has been chosen):

```
┌──────────────────────────────────────────┐
│ Keyword chosen : <ENTERPRISES>           │
│ Code Chosen : NONE                       │
│ Do you want to continue — Y or N         │
└──────────────────────────────────────────┘
```

If the response is 'Y', the selected term/codes will be added to the pre-selected fields and the system will return to the worksheet. If the response is 'N', the system will go back to the screen shown in Figure 2, and ask for entering a term.

4.2 Mode 2

Performance of the VOCON.PAS program can be improved by incorporating some of the features of two other application programs written in CDS/ISIS Pascal, viz. MQUICK.PAS (4) and GMOD.PAS (5). The MQUICK.PAS program allows the user to conduct the search in a database, modify (edit) a given record, re-trieved through its MFN, or a set of records retrieved by a given search, and so on. The GMOD.PAS program allows the user to modify data (add, replace, or delete) in one or more fields in the records of a database globally. Thus by virtue of the features of these two application programs enables user to:

– conduct a search in a database to identify one or more records for modification (adding, replacing, or delet-ing) by using a vocabulary control tool;
– select a given record, by its MFN, for modification (adding, replacing, or deleting) by using a vocabulary control tool;
– modify records globally by using a vocabulary con-trol tool, and so on.

Functions of the VOCON.PAS program in this mode are illustrated below.

On invoking the program (by running it from the MicroIsis main menu by choosing option A — Advanced Programming Services, or option Y — Vocabulary Con-trol Service, as described earlier in Section 3), the VOCON menu will be displayed (Figure 4). If a database has not already been selected, the system will ask for a database name.

```
┌──────────────────────────────────────────────┐
│ ************** V O C O N **************       │
│                   FOR                         │
│        ONLINE VOCABULARY CONTROL              │
│  ┌─────────────────────────────────────┐      │
│  │ C    Change Data Base               │      │
│  │ W    Change Worksheet               │      │
│  │ O    Reassign Output File           │      │
│  │ Q    Quick                          │      │
│  │ H    Help                           │      │
│  │ X    Exit                           │      │
│  └─────────────────────────────────────┘      │
│                    ?                          │
│ Data Base : Libcat          Worksheet : Libca │
│            Output file : Libcat.txt           │
└──────────────────────────────────────────────┘
```

Figure 4 : Main Menu of VOCON

20

Knowl. Org. 22(1995)No.1
G.G.&S.Chowdhury/A.Neelameghan: VOCON Vocabulary Control Online in MicroIsis Databases

On selecting option 'O' (from Figure 4), the system will ask for another name to replace the output file (LIBCAT.TXT), if desired, for example, LIBCAT.STW (stopword file), to which, by using the QUICK option, selected terms from the index of the database can be assigned to create a stopword file. This is however, optional and need not necessarily be followed in making use of the VOCON program.

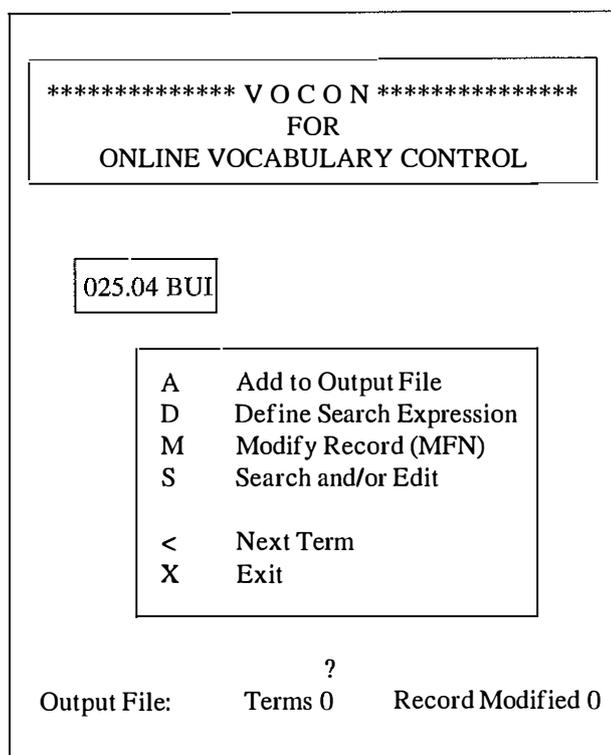On choosing option Q, another menu (Figure 5) will appear.

```
************* V O C O N **************
              FOR
     ONLINE VOCABULARY CONTROL


     025.04 BUI


     A      Add to Output File
     D      Define Search Expression
     M      Modify Record (MFN)
     S      Search and/or Edit

     <      Next Term
     X      Exit


                    ?
     Output File:   Terms 0     Record Modified 0
```

Figure 5: The QUICK Menu

The term appearing in the box (Figure 5) is the first term in the index file for the database. By pressing option 'A', the boxed term can be added to the output file, i.e., LIBCAT.STW (however, this is not absolutely required in the present context, i.e., in using the vocabulary control devices).

To retrieve the record or records to be modified or added to, there are two options, viz.

D Define search expression, and
M Modify Record (MFN)

Selection of option 'D' will display the screen, shown in Figure 6, and a search expression for an inverted file-based search or free text search can be entered. If a particular record whose MFN is known is to be modified, then option 'M' is to be selected, and the MFN is to be provided at the prompt, followed by a <CR>.
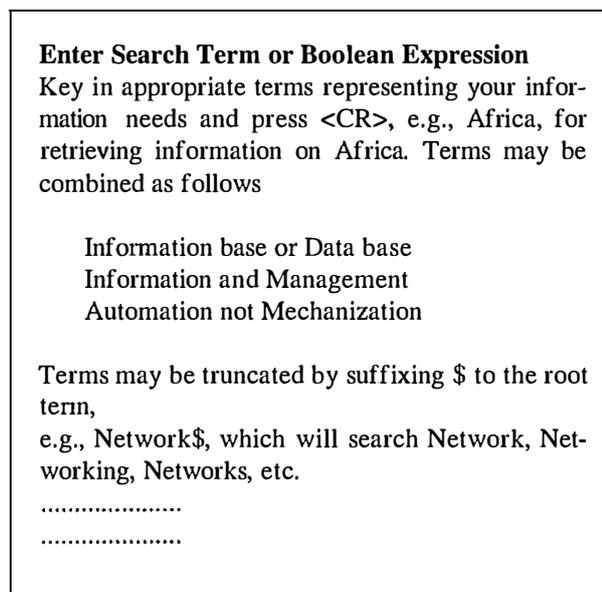
Enter Search Term or Boolean Expression
Key in appropriate terms representing your information needs and press <CR>, e.g., Africa, for retrieving information on Africa. Terms may be combined as follows

> Information base or Data base
> Information and Management
> Automation not Mechanization

Terms may be truncated by suffixing $ to the root term,
e.g., Network$, which will search Network, Networking, Networks, etc.
....
....

Figure 6: Screen for Search Expression

If there are hits matching the search expression, the screen, shown in Figure 7, will appear. Option 'D' (in Figure 7) will display the record for review and edit, option 'E' for editing, option 'S' to skip record(s).
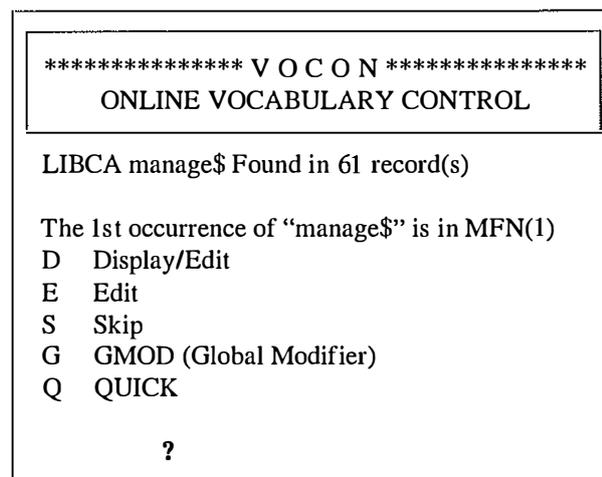
```
************* V O C O N **************
     ONLINE VOCABULARY CONTROL

LIBCA manage$ Found in 61 record(s)

The 1st occurrence of "manage$" is in MFN(1)
D    Display/Edit
E    Edit
S    Skip
G    GMOD (Global Modifier)
Q    QUICK

               ?
```

Figure 7 : Information on Hits

```
The MFN Retrieved is :  9

TITLE          : Management teams: why they
                    succeed or fail
AUTHOR(S)      : Boblin, R. Meredith
PUBLISHER      : London; Heinemann Profes-
                    sional Publishing
YEAR           : 1981
KEYWORDS       : Management; Managers
LOCAT./COPY  : MAIN LIB c.1
CALL NUMBER : 658.401 Bob

    <CR> — Continue;        E — Edit
```

Figure 8 : Display of Hit Record

Knowl. Org. 22(1995)No.1
G.G.&S.Chowdhury/A.Neelameghan: VOCON Vocabulary Control Online in MicroIsis Databases
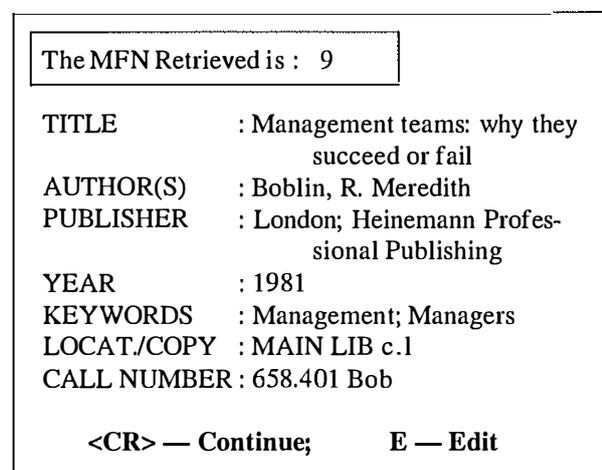
21

Selecting option 'E' (Edit) in the menu in Figure 8 or option 'E' in Figure 7, will open the data entry worksheet for editing the record. If option 'M' is chosen from the menu shown in Figure 5, then also the system will open the worksheet of the appropriate record. The user can then choose one or more term(s) and codes from a selected vocabulary control tool as described in Section 4.1 (**Mode 1** of operation of the VOCON.PAS program).

The VOCON.PAS program, as indicated in the menu, shown in Figure 7, offers the facility to either retrieve records, display and review each and edit selected ones or to edit them globally using the option 'G' (GMOD).

If several records are selected/retrieved for modification/addition of terms, etc., it may be more convenient to do so by selecting the option 'G' (Global Modifier) from the menu shown in Figure 7. If option 'G' is chosen, the system will display the boxes shown in Figure 9.

Once the tag for term(s) and/or code(s) is provided (Figure 9), the system will ask whether it is a case-sensitive search or not. This is because the GMOD program allows one to select a term for searching in the database and allows to edit, replace or delete the term in all the records in the database or in a selected range of MFNs.

Global Change

Database : LIBCAT

Tag for terms and <CR> or <CR> only:

Tag for code and <CR> or <CR> only:

Figure 9 : Messages for Global Modification

case sensitive search (Y/N?):

This feature of the GMOD program can be ignored at this stage by responding negatively (i.e., pressing 'N') at this prompt, because the system has already retrieved a set of records pertaining to the search term 'manage$' (see Figure 7). After this, the system will ask all those questions and the user has to respond to them in the way described in Section 4.1 (**Mode 1** of operation of the VOCON.PAS program). After all the steps have been followed (i.e., one or more term(s) and/or code(s) have been chosen for inclusion in the chosen set of records), the system will produce the following prompt:

Last query (Y/N?):

Here the system wants to know whether the chosen term(s) and/or code(s) are to be added/replaced in all the records those are retrieved in the previous search (Figure

7) or not. If the response is 'Y', at this stage, the system will make necessary modifications in all the records one by one. If the response is 'N', the system will ask for the MFN range, by asking the first and the last MFN, and will make the modifications in all the records in the range.

5. Conclusion

The VOCON.PAS program solves a long-standing problem involved in the development of text databases. As has been described, the system allows users to choose an appropriate vocabulary control device — a thesaurus, a subject heading list, an authority file, a classification scheme, etc., — online at the time of data entry operation, and can select and enter automatically one or more term(s) and/or code(s) in the appropriate field(s) in the record. Thus this program can be used to add/replace/delete one or more term(s) and/or code(s) to one or more records in an existing database by using an appropriate vocabulary control device(s). The program makes use of an improved version (developed by the authors at SISA) of the THES.PAS program which comes with the MicroIsis software. Therefore, the VOCON.PAS program can also be used to update any vocabulary control device. This can be accomplished at the stage when the system asks for choosing and marking term(s) for inclusion in a field (by choosing the appropriate options from the cue in Figure 3). Minor modification of the program will enable users update one database by choosing term(s) and/or code(s) form another database. This will help users update a vocabulary control tool, a thesaurus for example, by picking up terms from another tool, a thesaurus, a subject heading list, a classification scheme and so on. The VOCON.PAS program has been designed such that it will work on any MicroIsis database with very little or no modifications at all.

References
(1) Strategy for MicroIsis-based terminology applications software. Int. Classif. 19(1992)No.1, p.24-25.
(2) Unesco, Division of Software Development and Applications, Office of Information Programmes and Services, Paris: Mini-micro CDS/ISIS Reference Manual (Version 2.3). Paris: Unesco 1989.
(3) Ganzmann, Jochen: Criteria for the Evaluation of Thesaurus Software. Int. Classif. 17(1990)No.3/4, p.147-157.
(4) Srilatha, G., Aparna Sridhar: MQUICK.PAS, Quick Searcher, Editor, and File Maker for MicroIsis. Bangalore: Institute of Information Studies (Under Publication).
(5) Srilatha, G., Apama Sridhar: Manual for GMOD.PAS. Bangalore: Institute of Information Studies (Under Publication).

* Asst. Professor, School of Information Studies for Africa (SISA), Addis Ababa University, P.O. Box 1176, Addis Ababa, Ethiopia e-mail: Chowdhury!SISA@PADIS.gn.apc.org
** Formerly a Professor at SISA. Presently Honorary Prof. DRTC, 8th Mile Mysore Road, Bangalore 560059, India
*** Lecturer, Dept. of Library Science, Addis Ababa University, P.O. Box 1176, Addis Ababa, Ethiopia

22

Knowl. Org. 22(1995)No.1
G.G.&S.Chowdhury/A.Neelameghan: VOCON Vocabulary Control Online in MicroIsis Databases