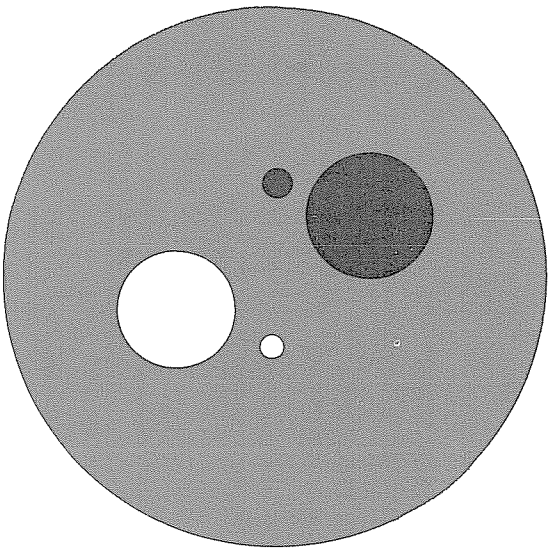
COMPUTER SCIENCES DEPARTMENT

University of Wisconsin - Madison



LEXICO Guide No. 3

# CREATING AND MAINTAINING A COLLECTION

Richard L. Venezky
Nathan Relles
Lynne A. Price

Computer Sciences Technical Report #284

December, 1976

LEXICO Guide No. 3

# CREATING AND MAINTAINING A COLLECTION

Richard L. Venezky
Nathan Relles
Lynne A. Price

Computer Sciences Department
University of Wisconsin
December, 1976

# TABLE OF CONTENTS

## 1. Introduction

This guide describes the maintenance of a collection of texts using the LEXICO system. Most of the tasks involved are the expected ones such as creating and deleting the collection, adding and deleting texts, and obtaining various listings. The one-text collection is briefly mentioned in Section 5.

LEXICO uses a mechanism of default levels to save the user the trouble of repeatedly specifying parameter values. As protection from system crashes and programmer errors, the user may create a copy or backup of his collection. A set of pack commands is available to reduce file charges under certain conditions. These features of LEXICO are discussed in this document.



## 2. Default Levels

Built into LEXICO are values for almost all parameters the user would otherwise have to specify. These values represent the designer's best guesses of the most common values these parameters will assume (e.g., period, question mark, and exclamation point as citation delimiters). When a user creates a collection, he declares only those values that differ from these system default values. Thus, if he plans to use period, question mark, and exclamation point as citation delimters, he need not specify any citation delimiters. If he wishes to use some other delimiter, he must declare them in order to replace the corresponding system default values.

When a collection is created, all of the values specified by the user, and the system default values for all other parameters, become collection defaults--they apply to every text in the collection unless the user overrides them by specific declarations. Since texts in a collection should be uniform in composition, collection defaults minimize (or eliminate) the specification of text parameters. Collection defaults may be changed at any time.

Of those collection defaults that apply to texts (e.g., delimiters), the user may specify that a particular text has different values. The collection defaults together with any overriding specifications become text defaults. These need not be re-specified each time the text is used. Some of these values may be overridden by explicit user declarations for a particular task. Such temporary text values are automatically replaced by text default values for the succeeding operations performed on the text.

The prudent user will try to abide by system defaults wherever possible and thereby eliminate additional specifications each time a task is performed. If this is not possible, he will attempt to keep the texts within a collection as uniform as possible to reduce the number of declarations required each time a task is performed.

A complete list of the parameters used throughout LEXICO, and their system default values, is given in Section 10.



## 3. Creating and Updating a Collection

### 3.1 Selecting a Collection

Before any texts may be entered into a collection, the collection must be defined and given a name. The collection name may be 1-11 alphanumeric (A-Z and 0-9) characters in length. Once created, a collection is referenced by this name. Furthermore, it may be referenced only under the project number used to create it.

At the beginning of each interactive session, LEXICO asks the user to specify a collection with the request

COLLECTION NAME?

After a legal name has been entered, LEXICO will respond with

NEW COLLECTION MAY BE CREATED.

if there is no existing collection with that name. If such a collection does exist, and is not being used by another user, LEXICO replies

COLLECTION ACCESSIBLE; PROCEED.

To change to another collection, or if the collection name was misspelled, the task command

```
COLLECTION cname ;
```

may be entered, where cname is the desired collection name.

## 3.2 Specifying Collection Defaults

Once the name of a new collection has been specified, the collection is created by a block of the form

```
CREATE ;
  command 1;
  command 2;
     :
END ;
```

where the commands inside the CREATE block specify collection defaults. Since the description of each LEXICO parameter logically belongs with the task to which it applies, the possible collection defaults are not listed here. However, a table in Section 10 shows all system parameters and the guides in which they are explained. These guides also describe the statements used to display the current values of the parameters. Section 11.2 lists the commands which may be entered in the CREATE block.



If for some reason, such as the existence of errors which cannot be corrected immediately, the user does not wish to have the collection created, the command

```
IGNORE ;
```

may be entered instead of

```
END ;
```

Since it is not possible with LEXICO to change the name of a collection, a user who has misspelled the collection name may wish to IGNORE the CREATE block.

Once a collection has been created, collection defaults may be inspected or modified with the block

```
UPDATE ;
  command 1;
  command 2;
     :
END ;
```

However, the UPDATE block may not be IGNOREd.

## 3.3 Inspecting the Collection Directory

The task command

```
SHOW DIRECTORY ;
```

results in a display showing, for each text in the collection, the text name, text code, number of citations and whether the text has

been edited, concorded, respelled, undergone LOOKUP or CLEANUP, or been used to generate slips. The codes S for "scheduled" and C for "completed" are used to present this information. The STATUS OF command, described in Section 4.2 below, may be used to obtain more information about individual texts.

The display also gives some other information, including the daily file charge for the collection. As the day passes, and changes are made to the collection, successive SHOW DIRECTORY commands may give different amounts for this figure. The actual charge will be the maximum that occurs during the day. A sample directory is shown below.

```
COLLECTION POETRY        CREATED BY USER 5687266949   on  2 AUG 1976
COST OF COLLECTION: $1.14  PER DAY.
LAST REFERENCED BY USER 5687266949  ON  4 AUG 1976 AT 22:29
NO BACKUP EXISTS FOR THIS COLLECTION.

TEXT CODE      TEXT NAME    CIT  CON  E  R  L  C  S
    1          POEMS-1       86    C
    2          POEMS-2       88    C
    3          POEMS-3      123    C             C
    4          POEMS-4      123    C             C
    5          POEMS-5        0    S
```

## 3.4 Deleting the Collection

When all processing has been completed, a collection may be removed from the computing system with the command

DELETE COLLECTION;

This command should be used with caution since, under most circumstances, once a collection has been deleted, it cannot be retrieved. However, if the collection existed the day preceding its deletion, the MACC tape librarian can cause it to be reloaded as discussed in Section 5 below.



# 4. Text-Specific Processes

## 4.1 Introduction

Most tasks performed by LEXICO's users refer to specific texts. Each text may be referenced either by text name or text code. Each text-specific task command has one of three general forms:

(1) COMMAND tname;

(2) COMMAND text;

(3) COMMAND textlist;

Here, tname is a text name, text is either a text name or a text code, and textlist is a list of text names, text codes, or ranges of text codes, separated by blanks (e.g., 3 RAVEN 4-9 19-21). A textlist may, of course, refer to a single text.

Many of these commands are described in other guides. The ADD block used to enter texts into a collection is discussed in Guide 4. The CONCORD block is described in Guide 6. The ADDCONCORD block, which is used to concord texts as they are entered, is referred to by both these guides. Guide 5 describes the EDIT block. Guide 7 describes all processes relevant to headword classification and slip generation, including the RESPELL, LOOKUP, CLEANUP, LIST TYPES, and SLIPS task commands. The remaining text-specific tasks are discussed below.

## 4.2 STATUS OF

Information about the operations performed on a group of texts may be inspected with the command

STATUS OF textlist;

Shown below is the status of some of the texts in the collection whose directory is shown in Section 3.3, along with the command that caused the display.

```
>status of 4-5;

POEMS-4      (TEXT CODE:   4)
ID TYPE: N
   123 CITATIONS,     812 KEYWORD TYPES,    2290 TOKENS,
  6.10 CHARACTER PER TYPE,   18.60 TOKENS PER CITATION,
  6.60 KEYWORDS PER CITATION.

ADDED TO COLLECTION:  4 AUG 1976
LAST CONCORDED: 4 AUG 1976
LAST CLEANUP:  4 AUG 1976
LAST RUN IDENTIFICATION: XC0334       (SAVERN0804*6124309)

POEMS-5      (TEXT CODE:   5)
SCHEDULED FOR ADDING TO COLLECTION:   4 AUG 1976
SCHEDULED FOR CONCORDING:   4 AUG 1976
```

In any text-specific block (i.e., ADD, CONCORD, ADDCONCORD, UPDATE, TEXT, EDIT, or CLEANUP), the user may enter a memo about a text (see Guide 4, Section 4.13). For example,

```
MEMO 'edited through line 314';
```

Any such memo is also displayed by the STATUS OF command.

## 4.3 Listing Texts

To list a group of texts, showing all changes made during EDIT blocks, enter

```
LIST textlist;
```

## 4.4 Modifying Text Defaults

The following block may be used to update text defaults:

```
UPDATE TEXT text;
  command 1;
  command 2;
     :
     :
END;
```



The commands which may be entered in an UPDATE TEXT block are given in Section 10.3. This block may be IGNORED.

## 4.5 Deleting Texts

After texts have been processed, a user may wish to delete some or all of them from a collection in order to reduce the file charges. This must be done in two steps. The commands discussed here remove the material from the collection, but to reduce the charge, the PACK commands described in Section 7 are needed.

A text is stored in a collection in three parts--the body of the text, its concordance, and the associated word list. The command

```
DELETE TEXTS textlist;
```

is used to remove any of these parts which exist in the collection for the specified texts. This command should be used after these texts have been completely processed, but the collection will not be deleted for some time.

In addition, the command

```
CLEAR TEXTS textlist;
```

causes the removal of the bodies of the texts, but leaves the concordances and word lists. This command is used if it will not be necessary to regenerate the concordances, but headword classification or slip generation will not be completed for several days.

Finally, the command

```
CLEAR CONCORDANCES textlist;
```

causes the removal of the concordances and word lists, but not the body of the text. This is done if the texts need to be reconcorded, perhaps after editing or using different stopwords.

All of these commands may be entered if the concordances for some or all of the texts are stored on tape. No change is made to the tape.

4.6 Permutations of the Word List

A list of the words entered into the concordance of a text, ordered by the frequency with which the words appear in the text, may be printed. The command for creating this listing is

```
LIST TYPES BY FREQUENCY FOR text;
```

The list may also be printed in an order computed by reversing each word (spelling it backwards) and alphabetizing the result. This brings together words with similar endings. The words are printed in normal form, however. The command for generating this listing is

```
LIST TYPES REVERSED FOR text;
```

In both these lists, only the text forms are given; that is, no headword classification that may have been made is shown. Both keywords and stopwords are included. Both lists may also be obtained when a concordance is generated.



## 5. Concording a Single Text

Users who simply wish to concord a single text may easily use LEXICO. In this case, the collection will have only one text. One possibility is to enter all text input (see Guide 4) and concordance (see Guide 6) specifications in the CREATE block. The concordance can then be scheduled with the two-statement block

```
ADDCONCORD text;
END;
```

If, after looking at the output, the user wishes to edit the text, or to change some concordance specifications, it is easy to make changes and reconcord the text with LEXICO. Of course, the user must remember to delete the collection when the concordance has been completed.

## 6. Backups

Execution of LEXICO can be unexpectedly terminated before a task is completed for a variety of reasons. The most common cause is exceeding the dollar limit or time limit. However, a terminal malfunction, trouble with the telephone line, a system crash, or an error within LEXICO are other possibilities. If this occurs while certain tasks are being executed, the collection becomes unuseable, and any attempt to access it will result in an error message.

To protect the user from this situation, LEXICO allows the creation of a duplicate of the collection, called a backup. This is done with the task command

CREATE BACKUP ;

Afterwards, the command

RESTORE COLLECTION ;

causes the collection to revert to its state when the backup was created--any tasks successfully performed in the meantime must be repeated.

A backup may be deleted with the command

DELETE BACKUP ;

Since the file charge for a backup is approximately the same as that of the collection, it is recommended that any backup be deleted if the collection is not to be used for a few days. However, if a backup is deleted and recreated on the same day, file charges will be made for each backup. This is not the case if multiple CREATE BACKUP commands are entered on one day without intervening DELETE BACKUP commands. After creating a backup, LEXICO displays the resulting file charge. If DELETE BACKUP has not been entered, the actual file charge will be the maximum of these. The file charge for any existing backup is also given when the collection directory is displayed.



It is conceivable that a system crash could occur while a backup is being created. Should this happen, the collection is not destroyed. However, since a backup is written over any previous backup the collection may not be restored.

To determine how frequently backups should be created, the cost of doing so must be weighed against the likelihood of having to restore the collection and the cost, in computer time and human time, of repeating tasks. Creating backups is one of the more expensive tasks done on-line by LEXICO. Before scheduling any off-line task that, if interrupted, could result in the destruction of the collection, LEXICO will ask if a backup should be created as the first step in the performance of that task. It is significantly less expensive to have backups created off-line in this manner.

If a collection is destroyed when no backup exists, the user may have to remove it from the system with the DELETE COLLECTION command, and start over. However, each day, at the end of operations, MACC makes a copy of all files which are active at that time. This copy is kept for thirty days. Thus, a collection which existed in valid form within the previous thirty days may be retrieved. To do so(without returning any backup which might have existed at the time) the user must call the MACC tape librarian and ask to have several files reloaded. The librarian will ask for the qualifier and name of each file. The qualifier in each case is the user's project number. The file names are formed by appending a single digit to the collection name. If no concordances have been made, the user should ask to have files cname1 and cname4 reloaded, where cname is the collection name. If any concordances have been generated, files cname2 and cname3 also must be reloaded. Thus, a user who created a

collection called POETRY under project 1234, but did not concord any texts, would ask to have files POETRY1 and POETRY4, both with qualifier 1234, reloaded.

Users should be aware that there is a significant charge for re-loading files.



## 7. Packing

After extensive processing, some of the file space occupied by a collection may not actually be used. If enough of this unused space accumulates, it may contribute significantly to the daily file charge. Therefore, LEXICO provides a set of task commands that may be used to remove unused space from different parts of the collection. When one of these pack commands is entered, LEXICO displays the current file charge for the selected part of the collection along with an estimate of what that charge would be after packing. The system then asks the user if he wishes to schedule an off-line job to perform the packing operations.

The command

PACK TEXTS ;

should be used when any of the following operations have been performed on a significant number of texts in a single collection (or on a small number of the largest texts in the collection):

1. deleting
2. clearing
3. re-adding
4. extensive editing.

Similar guidelines should be followed with

PACK CONCORDANCES ;

after the following operations:

1. re-concording
2. deleting or clearing concorded texts
3. extensive assignment of homographs during LOOKUP or CLEANUP.

Finally, if numerous (i.e., thousands) basetype rules have been added or deleted, the command

PACK BASETYPE RULES ;

should be used.

Packing does not affect the file charge for a collection's backup. If a backup was created when a collection contained unused space, the backup should be deleted, and another created, after the collection has been packed.



## 8. Copying a Collection to Tape

If a collection will not be used for an extended period fo time, it may be economical to store it on tape. It is not feasible to discuss here all the possible variations of copying disk files to tape. However, the runstreams shown below may be punched on cards and submitted to the computing center to copy a collection, without any backup it may have, onto tape, and to read it back in. For other options, refer to the MACC Computing Handbook or ask a MACC consultant.

To copy a collection to tape, use

```
                @RUN,/R   name,proj,1M,10
                @PASS     password
                @ASG,T    TAPE.,T,$num
include these   @COPY,GM  proj*cn1.,TAPE.
cards only if   @COPY,GM  proj*cn2.,TAPE.
some texts have @COPY,GM  proj*cn3.,TAPE.
been concorded  @COPY,GM  proj*cn4.,TAPE.
                @FIN
```

Where name is the user's name, proj his project number, num is the MACC tape number of the tape being used, and cn is the collection name. Thus, for a collection called POETRY created under project 1234, the fourth card would be

```
@COPY,GM  1234*POETRY1.,TAPE.
```

After this job has been run successfully, the collection may be deleted. Then, at a later time, the following job may be used to read it back

```
@RUN,/R   name,proj,1M,10
@PASS     password
@ASG,T    TAPE.,T,num
@CAT,P    proj*cn1.
@SAVE,S   proj*cn1.
@COPY,G   TAPE.,proj*cn1.
```

include these cards only if some texts have been concorded

```
@CAT,P      proj*cn2.
@SAVE,S     proj*cn2.
@COPY,G     TAPE.,proj*cn2.
@CAT,P      proj*cn3.
@SAVE,S     proj*cn3.
@COPY,G     TAPE.,proj*cn3.
@CAT,P      proj*cn4.
@SAVE,S     proj*cn4.
@COPY,G     TAPE.,proj*cn4.
@FIN
```



# 9. Synonyms and Abbreviations

## 9.1 Between Blocks

Members of the following groups of expressions may be used interchangeably in the task commands described in this guide:

```
BEFORE,BY,EXCEPT,FOR,FROM,INTO,OF,ON,OR,OUT,
 THRU,TO,WITH;
BASETYPE RULES, BTR ;
CONCORDANCE, CONCORDANCES;
D,DELETE;
DIRECTORY,DIR;
SHOW,SH,S,DISPLAY;
TEXT,TEXTS;
RULE,RULES;
```

For example

```
SHOW    DIRECTORY;
 S  DIR;
 SH DIRECTORY;
```

are all equivalent.

## 9.2 CREATE and UPDATE Blocks

Members of the following groups of expressions may be used interchangeably within a CREATE or UPDATE block:

```
ABTR,ADD BASETYPE RULES ;
AHR,ADD HOMOGRAPH RULES ;
BEFORE,BY,EXCEPT,FOR,FROM,INTO,OF,ON,OR,OUT,THRU,TO,WITH ;
ASR,ADD SPELLING RULES ;
BASE, BASES ;
BTR,BASETYPE RULES ;
CARD,CARDS ;
CHARACTER,CHARACTERS ;
CITATION,CITATIONS,CIT ;
CONCORDANCE,CONCORDANCES ;
DBTR,DELETE BASETYPE RULES ;
DELETE,D ;
DELIMITER,DELIMITERS,DELIM ;
DHR,DELETE HOMOGRAPH RULES ;
DONOT,NO,NOT ;
DSR,DELETE SPELLING RULES ;
HOMOGRAPH,HOMO ;
HR,HOMOGRAPH RULES ;
LIST,LISTS ;
NO,NOT,DONOT ;
NOTE,NOTES ;
REVERSE,REVERSED ;
RULE,RULES ;
SHOW,S,SH,DISPLAY ;
SR,SPELLING RULES ;
STOPWORD,STOPWORDS ;
RB,REPLACE BASE ;
RT,REPLACE TYPE ;
TYPE,TYPES ;
WORD,WORDS ;
```



## 9.3 UPDATE TEXT Blocks

Members of the following groups of expressions are equivalent in an UPDATE TEXT block:

```
DELIMITERS,DELIMITER,DELIM ;
DONOT,NO,NOT ;
LIST,LISTS ;
REVERSE,REVERSED;
TYPE,TYPES;
```

| Parameter | System Default | Maximum Number | Described in Guide | May be entered in these blocks as a Collection Default | Text Default | Temporary Text Value |
|---|---|---|---|---|---|---|
| (DONOT) LIST TEXT | TEXT OUTPUT ON PRINTER ; | —— | 4 | CREATE, UPDATE | —— | ADD, ADDCONCORD |
| LIMIT CITATIONS TO | 10,000 | —— | 6 | CREATE, UPDATE | —— | ADD, ADDCONCORD |
| PHYSICAL LENGTH | 72 | 1 | 4 | CREATE, UPDATE | —— | ADD, ADDCONCORD |
| CONCORDANCE OUTPUT ON | PRINTER | —— | 6 | CREATE, UPDATE | —— | CONCORD ADDCONCORD |
| COLLATING SEQUENCE | @[]#␣ßABCDEFGHIJKLMNOPQ RSTUVWXYZ)-+<=>&$*(%:?!,\ 0123456789';/." where ß denotes the blank symbol. | —— | 6 | CREATE, UPDATE | —— | CONCORD ADDCONCORD |
| SQUEEZE OUT | DONOT SQUEEZE OUT ; | 64 | 6 | CREATE, UPDATE | ADD,CONCORD, ADDCONCORD, UPDATE TEXT | —— |
| AUTOSTOP | DONOT AUTOSTOP ; | 64 | 6 | CREATE, UPDATE | —— | CONCORD, ADDCONCORD |
| FRONTSTRIP | DONOT FRONTSTRIP ; | 64 | 6 | CREATE, UPDATE | ADD, CONCORD, ADDCONCORD, UPDATE TEXT | —— |
| ENDSTRIP | DONOT ENDSTRIP ; | 64 | 6 | CREATE, UPDATE | ADD, CONCORD, ADDCONCORD, UPDATE TEXT | —— |
| stoppable lengths | none | 1 | 6 | CREATE, UPDATE | —— | CONCORD, ADDCONCORD |
| stopword list | none | varies-- approximately 200 words averaging 6 characters | 6 | CREATE, UPDATE | —— | CONCORD, ADDCONCORD |



## 10. LEXICO Parameters

The paramters which a user may specify to the LEXICO system are listed below.

| Parameter | System Default | Maximum Number | Described in Guide | May be entered in these blocks as a Collection Default | Text Default | Temporary Text Value |
|---|---|---|---|---|---|---|
| CITATION DELIMITERS | . ? ! | 6 | 4 | CREATE, UPDATE | —— | ADD, ADDCONCORD, EDIT |
| WORD DELIMITERS | blank , : ; | 6 | 4 | CREATE, UPDATE | ADD, CONCORD ADDCONCORD, UPDATE TEXT | —— |
| ID DELIMITERS | ( ) | 2 | 4 | CREATE, UPDATE | —— | ADD, ADDCONCORD |
| ID LEVEL DELIMITER | / | 1 | 4 | CREATE, UPDATE | —— | ADD, ADDCONCORD EDIT |
| NOTE DELIMITERS | [ ] | 2 | 4 | CREATE, UPDATE | —— | ADD,CONCORD, ADDCONCORD |
| TEXT DELIMITER | ^^ | 1 | 4 | CREATE, UPDATE | —— | ADD, ADDCONCORD |
| (DONOT) SEQUENCE | SEQUENCE ON CITATION ; | 1 | 4 | CREATE, UPDATE | —— | ADD, ADDCONCORD |
| INPUT ON CARD FILE | —— | 1 | 4 | CREATE, UPDATE | —— | ADD, ADDCONCORD |

## 11. Commands Used in Maintaining a Collection

### 11.1 Task Commands

The following task commands are described in this guide.

```
COLLECTION cname ;
CREATE ;
UPDATE ;
SHOW DIRECTORY ;
DELETE COLLECTION ;

STATUS OF textlist ;
LIST textlist ;
UPDATE TEXT text ;
DELETE TEXTS textlist ;
CLEAR TEXTS textlist ;
CLEAR CONCORDANCES textlist ;
LIST TYPES BY FREQUENCY FOR text ;
LIST TYPES REVERSED FOR text ;

CREATE BACKUP ;
RESTORE COLLECTION ;
DELETE BACKUP ;

PACK TEXTS ;
PACK CONCORDANCES ;
PACK BASETYPE RULES ;
```

### 11.2 The CREATE and UPDATE Blocks

The following commands may be entered in a CREATE or UPDATE block:

```
SPECS ;

WORD DELIMITERS d1 d2...d6 ;
ID DELIMITERS d1 d2 ;
ID LEVEL DELIMITER d1 ;
CITATION DELIMITERS d1 d2...d6 ;
NOTE DELIMITERS d1 d2 ;
TEXT DELIMITER d1 ;

SEQUENCE ON CITATION ;
SEQUENCE ON c1 ;
SEQUENCE ON c1 EXCEPT BEFORE c2 ;
SEQUENCE ON n CHARACTERS;
DONOT SEQUENCE;

INPUT ON CARD FILE file;
PHYSICAL LENGTH n ;
LIMIT CITATIONS TO n ;
```



| Parameter | System Default | Maximum Number | Described in Guide | May be entered in these blocks as a Collection Default | Text Default | Temporary Text Value |
|---|---|---|---|---|---|---|
| LIST TYPES BY FREQUENCY | DONOT LIST TYPES BY FREQUENCY ; | —— | 6 | CREATE, UPDATE | ADD,CONCORD, ADDCONCORD, UPDATE TEXT | —— |
| LIST TYPES REVERSED | DONOT LIST TYPES REVERSED ; | —— | 6 | CREATE, UPDATE | ADD,CONCORD, ADDCONCORD, UPDATE TEXT | —— |
| SPELLING RULES | none | 100 | 7 | CREATE, UPDATE, CLEANUP | —— | —— |
| BASETYPE RULES | none | none | 7 | CREATE, UPDATE, CLEANUP | —— | —— |
| use wordlist for stopwords or keywords | stopwords | —— | 6 | CREATE, UPDATE | —— | CONCORD, ADDCONCORD |

```
LIST TEXT ;
DONOT LIST TEXT ;
TEXT OUTPUT ON PRINTER ;
TEXT OUTPUT ON TAPE ;
TEXT OUTPUT ON file ;

CONCORDANCE OUTPUT ON PRINTER ;
CONCORDANCE OUTPUT ON TAPE ;
CONCORDANCE OUTPUT ON file ;

SQUEEZE OUT characters ;
FRONTSTRIP characters ;
ENDSTRIP characters ;
DONOT SQUEEZE OUT ;
DONOT FRONTSTRIP ;
DONOT ENDSTRIP ;
DELETE ALL STOPWORDS ;
ADD STOPWORDS word1 word2... ;
DELETE STOPWORDS word1 word2...;
STANDARD STOPWORDS ;
SHOW STOPWORDS ;
STOP n ;
STOP n OR LESS ;
AUTOSTOP characters ;
DONOT AUTOSTOP ;
REVERSE ;

COLLATE OE ;
COLLATE STANDARD ;
COLLATE NEW sequence ;

LIST TYPES BY FREQUENCY ;
LIST TYPES REVERSED ;
DONOT LIST TYPES BY FREQUENCY ;
DONOT LIST TYPES REVERSED ;

ADD SPELLING RULES r1:s11 s12...,r2:s21 s22... ;
DELETE SPELLING RULES r1:s11 s12...,r2:s21 s22... ;
DELETE ALL SPELLING RULES ;
SHOW SPELLING RULES ;

ADD BASETYPE RULES b1:t11 t12...,b2:t21 t22... ;
DELETE BASETYPE RULES b1:t11 t12...,b2:t21 t22... ;
REPLACE BASE oldbase WITH newbase ;
REPLACE TYPE oldtype WITH newtype ;
DELETE ALL BASETYPE RULES ;
DELETE ALL BASETYPE RULES FOR letter ;
DELETE BASETYPE RULE FOR type ;
SHOW ALL BASETYPE RULES ;
SHOW ALL BASETYPE RULES FOR letter ;
SHOW BASETYPE RULE FOR type ;
ADD HOMOGRAPH RULES  t1:b11 b12...,t2:b21 b22... ;
DELETE HOMOGRAPH RULES t1:b11 b12...,t2:b21 b22... ;
COPY BASETYPE RULES FROM cname ;

END ;
```



In addition,

```
IGNORE ;
```

may be entered in a CREATE block.

## 11.3 UPDATE TEXT

The following commands may be entered in an UPDATE TEXT block:

```
SPECS ;

WORD DELIMITERS d1 d2...d6 ;

SQUEEZE OUT characters ;
FRONTSTRIP characters ;
ENDSTRIP characters ;

DONOT SQUEEZE OUT ;
DONOT FRONTSTRIP ;
DONOT ENDSTRIP ;

LIST TYPES BY FREQUENCY ;
LIST TYPES REVERSED ;

DONOT LIST TYPES BY FREQUENCY ;
DONOT LIST TYPES REVERSED ;

MEMO "note about the TEXT" ;

END ;
IGNORE ;
```

# 12. Reserved Words

## 12.1 CREATE and UPDATE

The following are reserved words in CREATE and UPDATE blocks:

ABTR
ADD
ADDCONCORD
AHB
AHR
ALL
ASR
AUTOSTOP
BACKUP
BASE
BASES
BASETYPE
BEFORE
BLANK
BTR
BY
CARD
CARDS
CHARACTER
CHARACTERS
CIT
CITATION
CITATIONS
CLEAR
CLEANUP
COLLATE
COLLECTION
COLON
COMMA
CONCORD
CONCORDANCE
CONCORDANCES
COPY
CREATE
D
DBTR
DELETE
DELIM
DELIMITER
DELIMITERS
DHR
DIR
DIRECTORY
DISPLAY
DO
DONOT
DSR
EDIT
END
ENDSTRIP
EXCEPT
FILE
FOR
FREQUENCY
FROM
FRONTSTRIP
HOMO
HOMOGRAPH
HR
ID
IGNORE
INPUT
INTO
LENGTH
LESS
LEVEL
LIMIT
LIST
LISTS
LOOKUP
MEMO
MINUS
NEW
NO
NOT
NOTE
NOTES
OE
OF
ON
OPTION
OR
OUT
OUTPUT
PACK
PHYSICAL
RB
RENAME
RESPELL
RESPELLED
RESTORE
REPLACE
REVERSE
REVERSED
ROUTE
RT
RULE
RULES
S
SEMI
SEQUENCE
SH
SHOW
SLIPS
SPECS
SPELLING
SQUEEZE
STATUS
STANDARD
STOP
STOPWORD
STOPWORDS
TEXT
TEXTS
THRU
TO
TYPE
TYPES
UPDATE
WITH
WORD
WORDS



## 12.2 UPDATE TEXT

The following are reserved words in UPDATE TEXT blocks:

ADD
ADDCONCORD
ALL
AUTOSTOP
BACKUP
BASETYPE
BEFORE
BLANK
BY
CARD
CARDS
CHARACTER
CHARACTERS
CIT
CITATION
CITATIONS
CLEAR
CLEANUP
COLLATE
COLLECTION
COLON
COMMA
CONCORDANCE
CONCORDANCES
CREATE
D
DELETE
DELIM
DELIMITER
DELIMITERS
DIR
DIRECTORY
DISPLAY
DO
DONOT
EDIT
END
ENDSTRIP
EXCEPT
FILE
FOR
FREQUENCY
FROM
FRONTSTRIP
ID
IGNORE
INPUT
INTO
LENGTH
LESS
LEVEL
LIMIT
LIST
LISTS
LOOKUP
MEMO
MINUS
NEW
NO
NOT
NOTE
NOTES
OE
OF
ON
OPTION
OR
OUT
OUTPUT
PACK
PHYSICAL
RENAME
RESPELL
RESPELLED
RESTORE
REVERSE
REVERSED
ROUTE
RULE
RULES
S
SEMI
SEQUENCE
SH
SHOW
SLIPS
SPECS
SPELLING
SQUEEZE
SR
STATUS
STANDARD
STOP
STOPWORD
STOPWORDS
TEXT
TEXTS
THRU
TO
TYPE
TYPES
UPDATE
WITH
WORD
WORDS