KAYPRO

THE WORD

T.M.

Plus

The WORD Plus (tm)

(ver 1.2)

Manual
by
Wayne Holder

Oasis Systems
2765 Reynard Way
San Diego, CA 92103

TABLE OF CONTENTS

Using The WORD Plus with Perfect Writer

Exit Perfect Writer and return to CP/M (Get the "A>" prompt on the screen.) Remove the Perfect Writer Edit disk and replace it with The WORD Plus working disk and type "TW" and press the "return" key. The WORD PLus will load. If this is the first time you have used The WORD Plus it will ask you a series of questions. (See page 7 of The WORD Plus manual for details.) Answer these questions as the underlined text shows below:

Save context info? Y/N (Ret=N) :Y

Ignore lines starting with? (Type char, Ret=None) :(press return)

Ignore UPPER case words? Y/N (Ret=N) :(press return)

Ignore text between? (Type char, Ret=None) :(press return)

Mark words that change length? Y/N :Y

Marking character :(press return)

Save new defaults? Y/N :Y

Answering yes to the last question will cause The WORD Plus to save your answers to the previous questions so that you will not have to answer them again next time. Press "return" to proceed when you see the following display:

Checking file : YOURFILE.MSS
Saving context info

These are the current settings
Press return to proceed, or space to change

The WORD Plus will then begin checking your file. When finished The WORD Plus will return to CP/M.

Ignoring Perfect Writer format commands

Perfect Writer allows control of document printing by inserting format commands into the text file. These commands look like this:

                    @INDEX(multiple screens)

Most of these commands consist of perfectly valid words such as "INDEX." However, some use strange words such as "BLANKSPACE" and "NEWPAGE" that The WORD Plus will report as spelling errors.

You have two choices to avoid this problem. If you always type your format commands in all upper case letters you may tell The WORD Plus to ignore all upper case only words. (Be careful that your don't ignore important words.) Or, you may add the strange words to The WORD PLus' update dictionary by using the "U" (update) command in The WORD Plus' REVIEW Program.

INTRODUCING
The WORD Plus (tm)

The WORD Plus is the successor to our popular spelling checking package called The WORD. As you might expect, The WORD Plus has many new features which make it easy to use, flexible and fun. The WORD Plus will not only find spelling errors, but it will correct them too! In addition, The WORD Plus has dozens of other uses. Including: The WORD Plus' automatic correction feature which will help you rewrite your documents automatically. The WORD Plus' dictionary can be used on-line to look up the correct spelling of misspelled words, find rhyming words, solve crossword puzzles and even locate anagrams.

Unlike other programs that you may have purchased, The WORD Plus is modular. Instead of writing one large spelling checking program, we designed The WORD Plus as a collection of separate "tools." Several of the tools work together to check and correct your spelling errors. But, because The WORD Plus is modular you can also use the individual tools for other purposes besides spelling checking. Before describing these other uses in detail we'll give you a quick overview.

Spelling checking and correction are done by four programs (tools) called TW, SPELL, REVIEW and MARKFIX. SPELL finds the spelling errors; REVIEW shows them to you and helps you locate the proper spelling by consulting the dictionary; MARKFIX makes the corrections to your document. TW is the "orchestra leader" program. It asks you questions and then directs the other three programs to do the work.

The WORD Plus also has a tool called HYPHEN which will automatically hyphenate words. HYPHEN will read a text file and automatically insert "soft hyphens" into words longer than a given length. This is useful for people using their text processors to set narrow column text, such as for newsletters or magazines.

FIND, ANAGRAM and LOOKUP are three programs that give you a kind of "Electronic dictionary." LOOKUP is a misspeller's dream come true. If you can't spell a word, LOOKUP will find it for you! Just type it in as best you can and let LOOKUP search the dictionary for words that are similar. Our program FIND can solve crossword puzzles, locate rhyming words, or help you with other word games. ANAGRAM is an anagram finder. Type star, and ANAGRAM will find its anagrams rats and arts.

Lastly, there are three other general purpose tools called WC (Word Count), DICTSORT and WORDFREQ. WC is a quick way of measuring how much you have written. WORDFREQ will compile a list of words from a document and count how often each word was used. This is a valuable check on style which can help eliminate redundancy and overuse of words. DICTSORT will read any text file or word list and alphabetize it. You can use it to compile your own dictionaries.

HOW TO USE THIS MANUAL

Read the next section before reading the rest of the manual. It will tell you what The WORD Plus can and cannot do for you. The section after that tells you how to copy the master disk, what the hardware requirements for The WORD Plus are, and describes which text editors and word processors you can use it with. The section called "Spelling Checking" will show you how to start finding and fixing spelling errors and also gives a quick introduction to the three tools SPELL, REVIEW, MARKFIX. Each tool also has a separate section that describes how it operates and what its limitations are.

* RETURN YOUR REGISTRATION CARD TODAY! *

All registered owners of The WORD Plus will receive a free subscription to our newsletter "The WORD." This newsletter contains up-to-date information on new revisions to The WORD Plus. So return your registration card TODAY or you'll miss out!

THE SPELLING DILEMMA
(a short commentary)

Very few people have perfect recall when it comes to remembering how to spell a word. Everyone makes mistakes, usually because of the numerous inconsistencies built into the English language. The WORD Plus will help you find and fix your spelling errors. But before we can show you how, we should talk about what we mean by a spelling error.

WHAT IS A MISSPELLED WORD?

Proper spelling of a word is often dependent upon how the word is used. In the sentence, "I would like a peace of pie," peace is spelled properly, but it is being misused. The proper word to use is "piece." Words such as peace and piece are called "homonyms." They are words that have the same sound but are spelled differently. Confusion of homonyms is a common "spelling" problem. The WORD Plus cannot automatically detect incorrect homonyms for you. To do so would require that the program "understand" what you are trying to say, which is not a simple thing to do! However, the program MARKFIX can be used to locate and point out potential homonym problems. This will be described in greater detail later.

The spelling checker will list some words as misspelled simply because they do not appear in the dictionary. This is to be expected because the dictionary is not infinite in size. In addition, many words which are unique to a particular part of the country, or which are technical jargon, will not appear in the dictionary. After you get used to using The WORD Plus to check spelling, you will be able to build your own "special" dictionaries. You can then have the spelling checker use these dictionaries in addition to the master dictionary.

The WORD Plus
(a user's guide)

HARDWARE REQUIREMENTS

The WORD Plus can be used with almost any 8080/8085 or Z80 based computer. It runs under the operating system CP/M (tm) by Digital Research. The WORD Plus will also work with most variations of CP/M, such as CDOS by Cromemco (tm) and IMDOS, by the now defunct IMSAI. The WORD Plus will work with either the 1.4 or 2.2 versions of CP/M. We have been careful to avoid using features which are unique to CP/M 2.2 (such as random access) so as not to exclude 1.4 user's. This does not, however, mean that the design of the tools has been compromised in any way.

The WORD Plus can run with almost any amount of memory; however, we think you need a minimum of 32k bytes to do useful work. A 48K system should be more than adequate for most uses. All of the tools will automatically use whatever memory is available, with no need to manually change anything. Likewise, they will print an error message, such as "out of memory," if your system has too little to get the job done.

The only tools which are likely to have a memory problem are SPELL, WORDFREQ, DICTSORT and possibly LOOKUP. This is because they build lists of words within memory. If you have memory size problems, which should be rare, one solution is to reduce the size of the input files. This can be done by breaking your text files up into chapters or sections.

All of the tools will run quite fast on a 2 MHz 8080 system. In fact, The WORD Plus was developed on just such a system. The real limit to operating speed is not really the CPU but rather the floppy disk system. Those of you with 6 mhz Z80's and Winchester disks should have nothing to complain about.

* * * READ THIS BEFORE USING! * * *

The very first thing you should do BEFORE using any of The WORD Plus' tools is make a working copy of the master release disk and then put it away in a SAFE place (not too hot, not too cold, and away from magnets!). Also, before opening the sealed diskette, read the license agreement included with this manual. It describes what your responsibilities are with regard to Copyright restrictions and our proprietary software.

HOW TO COPY THE MASTER DISK

To create a copy of the master disk you should first have a copy of the CP/M program on the disk you want to copy to. You can do this with the SYSGEN program. (Some versions of CP/M may call this program something else.) The WORD Plus is supplied without a copy of CP/M and will therefore not "boot" on your system. Next use the CP/M tool "PIP" to copy The WORD onto your

4

new disk. The CP/M manual explains in detail how to use PIP, but we'll explain briefly here.

1. Place a CP/M disk in drive A that contains a copy of PIP. Put the disk you want to copy to in drive B.

2. Hold down the Control-Shift key and press the letter C.

3. Type PIP and press the return key. The PIP program will load and print a prompting character *.

4. Remove the disk from drive A and insert The WORD Plus master disk.

5. Type:

    B:=A:*.*[V]

    and press the return key. PIP will begin copying the programs and dictionary files and will print each file's name as it does so.

6. After all the files are copied PIP will print the * prompt again. Remove The WORD Plus master disk and re-insert your CP/M disk. Press the return key to return the CP/M A> prompt. Note: Variations of CP/M including CDOS will have a different procedure. Consult your user documentation for these systems for the proper procedures to copy disks.

    Consult your CP/M users manual or your local CP/M guru if you are still unsure of how to make a copy. Don't use the program on the master release disk. There is a charge to recopy your disk if you erase the programs. You should place the original disk in a safe place away from stray magnets, children, moisture, etc.

5 1/4 DISK USERS

    5 1/4 inch disk versions of The WORD Plus are contained on two disks. Both disks are full. Before using any of the programs we recommend that you copy them onto several disks using the following arrangement:

Copy Disk 1 should contain the files TW.COM, SPELL.COM, REVIEW.COM, MARKFIX.COM, and MAINDICT.CMP. This will be your spelling checker disk. The spelling checker will create several new and temporary files as you use it. This arrangement gives it the maximum amount of room to use.

Copy Disk 2 should contain the files LOOKUP.COM, FIND.COM ANAGRAM.COM, HYPHEN.COM, HYEXCEPT.TXT, and MAINDICT.CMP.

Copy Disk 3 should contain the files MARKFIX.COM, HOMONYMS.TXT, WC.COM, DICTSORT.COM, and WORDFREQ.COM.

BROKEN TOOLS

Every tool included with The WORD Plus has a self-checking feature built into it which will determine if the copy of the program on disk has become damaged. Every time you use a tool the very first thing it does is check to see if all of its program code is intact. If not, it will print the following error message:

* WARNING * This tool is damaged! Use another copy.

This check is included for your protection! A broken tool could wreak havoc with your disk files. We don't want you to lose your important text files because of a bad copy of one of The WORD Plus' tools. If you see this message get the master release disk and PIP (or XFER) over a new copy of the tool in question before proceeding.

COMPATIBILITY WITH EDITORS AND WORD PROCESSORS

The WORD Plus is designed to work in conjunction with almost any text editor or word processor now on the market. The WORD Plus requires only that text files conform to CP/M file conventions. Embedded control characters are ignored as is the setting of the "Most Significant Bit" of characters. A few restrictions are worth noting, however. All of the tools will read text files that contain either a Carriage Return (CR), or a Linefeed (LF), or both, as line delimiters. However, the files which are produced by The WORD Plus will always contain both a CR and LF. The one exception is the file written by MARKFIX, which is an exact copy of the input file except for the inserted marking characters.

Some text editors use certain special characters as "soft hyphens." A soft hyphen is a character inserted into a word to show where it can be divided if it is too long to fit onto a line. The WORD Plus expects to read a Word Star compatible file which uses two special control codes as soft hyphens. Magic Wand can use any character, but usually uses an & character. The WORD Plus needs to know which characters or control codes are used as soft hyphens in order to work properly. If you want to use it with another editor with different soft hyphen codes you should read the section "Special Options" before using any of the programs.

Editors and word processors with which The WORD Plus is fully compatible include: Micro-Pro's Word Master (tm) and Word Star (tm), Digital Research's ED, and Magic Wand (tm) by Peachtree Software. In addition, The WORD Plus will work with CP/M versions of Shrayer Software's Electric Pencil.

6

== SPELLING CHECKING ==
(a guide for first time use)

In this section we'll show you how to find and correct spelling errors in your documents. You'll also learn about the tools SPELL, REVIEW and MARKFIX which are the backbone of the spelling checker. SPELL and MARKFIX are described in much greater detail in later sections. The section "Advanced Options" will show you how to customize these tools in order to use the advanced features of different word processors. Read this section if you are using Magic Wand.

Spelling checking is done in three separate steps. First, SPELL makes a list of all the words from your document that it can't find in its dictionary. Then, REVIEW shows you each word from the list and helps you to locate the correct spelling. If the word was correct, it will help you add it to the dictionary. In the last step, MARKFIX takes the reviewed list of misspelled words, and their corrections, and fixes them in your document.

A master menu program called TW (The WORD) makes it easy to use the three spelling checking tools. TW will ask you for your instructions and then guide the operation of the other tools. For advanced applications, you may want to use the tools directly, but for routine spelling correction TW is much more convenient. In the rest of this document when we talk about "the spelling checker" we mean the combination of the four tools TW, SPELL, REVIEW and MARKFIX.

To check a document, you first need to have written it and saved it on a disk. The spelling checker works by reading and correcting your document file. You start the checking process by running the program TW. If you are using CP/M directly, you must get the A> prompt, type TW and then press return. You can also use the "run" command from within Word Star. TW will not work properly unless it and the three other tools SPELL, REVIEW, and MARKFIX are located on the same drive as TW and that drive is either the "currently logged drive" or an alphabetically lower drive (for example, B is less than C.) TW tries to finds SPELL, REVIEW, and MARKFIX by looking first on the currently logged drive, and then looking on alphabetically lower drives. If TW cannot locate one of the tools it will print a message like:

Can't find tool SPELL.COM

If TW finds everything in order it will start by asking the question:

Name of file to check? :

Answer by typing the name of the document file that you want checked. Include the disk drive letter (A:, B: etc.) in front of the name if the file is on another drive. Examples of file names are B:MYFILE.TXT, B:NEWFILE or DOCUMENT.TXT. Press "return" after entering the name.

Note: If you make a mistake while answering a question, press "Delete," "Back Space" or "Ctrl-U" to edit the answer. If you press any of these keys when there is no answer to edit, you will back up to the previous question.

Special dictionary name? :

TW also needs to know if you want to use an additional dictionary, that is, one besides the built-in main dictionary or the update dictionary. A special dictionary contains words that will not be needed every time you use the spelling checker, such as jargon or technical terms. Normally, you can ignore this question by just pressing return. A later section will describe how to make and use a special dictionary.

SPELL and MARKFIX have built-in options which are turned on and off by "switches." TW will ask the following questions to determine which options you want to use.

Save context info? Y/N (Ret=N) :

This question asks if you want the spelling checker to remember where it found each misspelled word in your document. Normally you want to answer yes by pressing Y. However, using this feature requires more RAM (memory) so there may be times when you will want to answer no. You will find out more about this option in the section describing the C (context) switch for SPELL.

Ignore lines starting with? (Type char, Ret=None) :

Some text processors, such as Word Star, use special instruction codes called formatter commands or "dot" commands to set margins, change indentation, etc. Formatter commands are placed into a text file as lines that begin with a period (or some other special character.) To ignore these lines press the character that your editor uses. If you answer by pressing the return key, any formatter commands will appear as misspelled words to The WORD Plus. Word Star users should enter a period (.) and Magic Wand users should enter a backslash (\) character. If you enter a control code such as control-X it will echo as ^X. (See the section on SPELL which describes the S switch for more details.)

Ignore UPPER case words? Y/N (Ret=N) :

Answering yes will cause any words in your file typed in all capital letters to be ignored. You can use this option to ignore acronyms such as IBM or SYSTAT. However, we recommend that you answer no (press N or return) for normal use. (See SPELL's I switch).

Ignore text between? (Type char, Ret=None) :

This option allows you to ignore any sequence of text that begins and ends with a special character. The Palantir word

processor inserts a special "ruler line" into text files. The
ruler line can contain characters that look like misspelled words.
Since a ruler line always begins and ends with a control-D code
(Ø4 Hex) you can use this option to tell The WORD Plus to ignore
it. Press return if you don't need this option. * WARNING * Be
¬areful if you use this option for any other purpose. If your
file contains an unbalanced number of the start and stop
characters you may cause The WORD Plus to ignore large portions of
your text file. (See SPELL's B switch).

        Mark words that change length? Y/N (Ret=N) :

        Since The WORD Plus can correct misspelled words
automatically it may sometimes change the length of a word. This
will also change the length of the line containing the word.
Also, if the corrected word was hyphenated The WORD Plus will
remove the hyphenation and thereby change the line length. Some
text editors will automatically reform lines. Others, like Word
Star, require you to "reform" paragraphs to fix the changes to the
margins. If you answer yes to this option The WORD Plus will mark
words that change length or are unhyphenated. Word Star's "Find"
command will then quickly locate paragraphs that must be reformed.
(See MARKFIX's L switch.)

        Marking Character :

        During word review you have the option of correcting words
automatically or of marking them in the document for later
correction. Word are marked by inserting a symbol into the
document file just in front of the word. You can then use your
editor's search feature to locate this unique character and fix
the words. Type a character that will not be used elsewhere in
your document.

        You have now told TW everything it needs to know to start
working. As a convenience, however, TW will first ask if you want
to save all of your answers. If saved, TW will remember them
automatically the next time. Your answers are recorded in a file
called TWDFAULT.CMP.

        Save new defaults? Y/N (Ret=N) :

After you answer the question, a display similar to that shown
below will summarize the options TW is going to use.

        Checking file : DOCUMENT.TXT
        Saving context info
        Marking character -> *

        These are the current settings
        Press return to proceed, or space to change

        If everything looks correct, press the return key to start
checking. TW will then load and use the SPELL tool to start
looking for misspelled words. As SPELL works it will print a

series of messages to show you what it is doing.

When SPELL has finished, TW will bring in and run the word review program REVIEW. The following display will then appear on your screen:

REVIEW+   Version x.x, Copyright 1981 Oasis Systems

Add word to:	Other options:
U>pdate Dictionary	P>revious word
S>pec. Dict. "SPECIALS.CMP"	N>ext word
M>ark word	R>esume review
D>iscard word	L>ook up word
C>orrect word	V>iew context

    --> ADVIRE -

At the bottom left corner of the screen is the alphabetically first word that SPELL found in your document that it could not locate in the dictionary. Just above is a "menu" of options. Each menu item indicates one of the options for dealing with the word. Pressing the letter key in front of the $\geq$ will select that option.

VIEWING A WORD IN CONTEXT

In the sample screen shown above, REVIEW has located ADVIRE as a misspelled word. The writer probably meant either "advise" or "admire." To see how the word is being used you can press the V key to "view" the context of the word. REVIEW will then show the portion of the sentence where the misspelled word was first used. Review shows control characters as spaces so the end of a line will appear as one or two spaces (Because most word processors insert control characters at the end of a line.) A word divided at the end of a line will show the hyphen followed by one or two spaces. Note: You cannot view a word's context if you answered no to the "Save context info" question.

FINDING THE CORRECT SPELLING

If you are unsure of how to spell the misspelled word, you can get some help by pressing the L key. This will cause REVIEW to scan the dictionary for words similar in spelling to the misspelled word. These words will appear in the upper portion of the screen. To the left of each word will be a single number or a symbol such as # or %. The purpose of this number or symbol will be explained below.

CORRECTING WORDS

Once you have found the correct spelling for your misspelled word, you can tell REVIEW to remember it so that MARKFIX can fix it later. Corrections are made to your document only after you are finished reviewing the entire list of

misspelled words, so you can change your mind at any time. Press
the C key to correct a word and the following prompt message will
appear.

>   ADVIRE Corrected to -->

        If you know the correct spelling, you can type it in and
press the return key. The screen will then change to show you the
next word in the misspelled list. Pressing Delete, Back Space or
Ctrl-U will erase the correction word you typed and allow you to
enter it again. If you change your mind about correcting a word,
delete the correction word and then press the desired option.

        The look up and correction options will work together to
make fixing a word easier. If you had previously used the L key
to look up the correct spelling, you can have the correct spelling
entered automatically. Press the C key to get the "Corrected to
-->" prompt, then press the number or symbol key that appears next
to the correctly spelled word. REVIEW will copy it just as if you
had typed it. If you select the wrong word press the delete key
to remove the word then press the correct selection. Press the
return key to proceed to the next word.

LIMITATIONS ON CORRECTING WORDS

        If the corrected spelling has a different number of
letters than the misspelled word, correcting it will change the
line justification in the original document. Also, if the word
being replaced contained any soft hyphens, these will be lost when
the new word is inserted. If the soft hyphen was "active" (the
word was divided at that point) correcting it will undivide the
word which will also cause the line length to change. If this is
important, or if you are using flush right margins, you should
reform the document to fix the margins. This is usually only
important to people using Word Star or other "what you see is what
you get"-type editors. To help you, the tool MARKFIX will print a
message if it changes the length of a line while correcting a
word. See the section on MARKFIX for more details.

ADDING WORDS TO THE DICTIONARY

        If REVIEW shows you a word that SPELL thought was
incorrect but that you know is right, you have several options.
You can press the U key and have REVIEW add the word to the update
dictionary, press the S key and add it to the SPECIALS.CMP
dictionary, or press the D key and tell REVIEW to ignore the word.
Words placed into the update dictionary (UPDICT.CMP file) will be
recognized by SPELL the next time you check a file. The
SPECIALS.CMP file is a catch-all dictionary into which you can
place words that you want to save, but don't want to put into the
update dictionary.

Note: If you typed the name of a special dictionary when TW asked
you the question "Special dictionary name?" then the name of that
dictionary will appear in the menu in place of the name

11

<u>SPECIALS.CMP</u>. Pressing the S key will then add words to this dictionary instead.

MARKING WORDS FOR LATER CORRECTION

In some instances you may decide that a word is misspelled, but that you don't want to automatically correct it. An example is when you want to examine the word in context using your text editor before deciding which word to replace it with. In this case you can press the M key to have MARKFIX mark the word in the file. As mentioned above, words are marked by inserting a special character in front of the word.

CHANGING A CORRECTION

REVIEW always lets you change your mind. If you had pressed U to add a word to the update dictionary, for example, you could return to the word later and delete it. Press the P key to retreat back up the list. When you get to the old word, a letter, or the prompt "Corrected to -->," will remind you of your previous decision. In this case, you will see a <u>U</u>. To delete the word press <u>D</u>. If you go too far, press the N (Next) key to move back down the list.

You can also change a correction by pressing C to erase the previous correction and get the "Corrected to -->" prompt again. When you're finished, press the R (Resume) key to return to the point in the list where you left off reviewing words.

When you reach the end of the list of words, REVIEW will print the following message:

End of list. Press "return" if done.

If you are finished reviewing the list, press the return key and TW will start correcting the errors using MARKFIX. When MARKFIX is done, control will return to CP/M (or to your editor if you are using a "run" command.)

== SPELL ==

SPELL is The WORD Plus' spelling checker. SPELL is normally used automatically by the TW program. However, you can use SPELL by itself to locate spelling errors. SPELL reads through your document and then writes out a list of all the words that it didn't find in its dictionary. If you are an average speller you will probably find many obviously misspelled words in this list. As noted in a preceding section, the fact that a particular word is not recognized by SPELL does not necessarily mean that the word is misspelled. You should look up all words that you are unsure of in a dictionary. Still, you will find that this is considerably easier than checking your entire text file with a dictionary.

In addition, SPELL will find many more misspelled words than most people could detect in a dozen readings. Studies have shown that for documents longer than about 20 pages, most people fail to find more than 50% of the misspelled words.

The size of the dictionary chosen to be included with The WORD Plus is based upon "common usage." Many words that are used infrequently are excluded. Also, words that might potentially confuse or "mask" misspellings of common words have been excluded. The guidelines used to develop the dictionary are slanted towards practical use.

LIMITATIONS

There are many subtleties about the English language, and its spelling and grammar, which are beyond the abilities of any current spelling checking system. The WORD Plus has a few limitations which are described in the next few sections. In all cases we have tried to reduce the severity of these limitations by compromising on the side of practicality.

HYPHENATED WORDS

SPELL handles hyphenated words by pretending the hyphen is a space. This means that the individual parts of a hyphenated word are checked against the dictionary separately. An important exception to this rule occurs when a hyphen occurs at the end of a line. This almost always means that the word has been "divided" at a syllable to fit onto the line, and that the remainder of the word is at the beginning of the next line. SPELL handles this condition by reassembling the divided word before looking it up.

This procedure of reassembling words before looking them up will cause problems if the word that was divided was a word that is always hyphenated (such as a compound word), and the division occurred at the hyphen. In this case, SPELL will incorrectly assemble the compound word into a single word and will then try to look it up. For example, consider what SPELL would do if it encountered the following text:

This is a sentence which contains a hyphen-
ated word. In this sentence the compound word light-
year is divided between lines.

SPELL would correctly reassemble the word "hyphenated,"
and would then locate it in its dictionary. However, "light-year"
would be incorrectly reassembled as one word "lightyear," without
the hyphen. However, this is easy to spot because REVIEW will
show you the context of a word if you press the View key. Words
divided at the end of a line are shown with the hyphen followed by
one or two spaces (depends upon the word processor you use.)

"SOFT" HYPHENS

Some word processors have the ability to insert what are
called "soft" or "ghost" hyphens into a document text file. These
are optional hyphenation points that the user inserts while
forming paragraphs. They aid the program in deciding where to
break a text line when printing. SPELL will automatically ignore
soft hyphens by always reassembling words divided by them.

The Word Star editor uses two control characters (1E and
1F) as soft hyphens. Spell assumes that you are using it with
Word Star. If you wish to change to a different soft hyphen
character (such as a & for Magic Wand) see the section "Advanced
Options."

APOSTROPHES AND POSSESSIVES

The WORD Plus' dictionary does not include possessive
cases of words. Because of this, SPELL makes a special check to
see whether a word ends in a single apostrophe or an apostrophe
immediately followed by the letter s ('s). Single apostrophes or
an apostrophe s combination are removed from the word before
looking it up. This procedure will cause problems with
contractions such as "it's," but this is a rare occurrence and
should not have a significant effect, because the resulting
shortened word is almost always a correct word.

There is a situation in which SPELL's rules for handling
apostrophes can trip it up. It should occur very rarely, but we
thought you should know about it anyway. It happens in a sentence
involving a quotation embedded within a quotation, such as the
following example:

"He said, 'That's Bob's hat not James''," Tom said.

Many writers will use an apostrophe as a single quotation
mark within a doubly embedded quote. The problem occurs because
the last word of the innermost quote happens to be possessive.
SPELL will lop off the apostrophe being used as a single quote but
will then leave the real apostrophe. SPELL will then try to
locate James' in the dictionary which it won't be able to do.
James' will wind up being listed as misspelled. That's a very
contrived example, but it could happen.

SPECIAL CHARACTERS WITHIN WORDS

SPELL normally recognizes only words containing alphabetic characters or apostrophes. If you want to have SPELL recognize words containing special characters, such as accent marks or tildes, see the section "Advanced Options."

NUMBERS WITHIN TEXT

SPELL performs a special check to ignore numbers followed by letters, such as 32nd or 1st. The rule used is simply:

If a word is immediately preceded by a number without a separating space, then ignore the word. In all other cases simply ignore only the numbers.

== USING SPELL ==

Let's say you have written a document using your favorite text editor and have placed it into a file named <u>DOCUMENT.TXTP</u>. Furthermore, this file is located on a disk in drive <u>B</u>. To use SPELL to list all of the misspelled words to the console, you would type:

        A>SPELL B:DOCUMENT.TXT

        Note: The <u>A></u> is typed by your computer.

As SPELL works, it will print out the following messages:

        SPELL+ ver x.x
        Copyright 1981 - Oasis Systems
        Compiling Word List
        <u>nn</u> unique words
        Checking Main Dictionary
        Listing unmatched words
        <u>nn</u> unmatched words

        The message "nn unique words" refers to how many different words SPELL found in your document. In the last message, "nn" refers to how many of these words SPELL identified as misspelled. If SPELL finds no incorrect words, The message "No incorrect words found" will be printed instead.

OPTION SWITCHES

        If you wished to have the misspelled words listed to the printer instead, you would type:

        A>SPELL B:DOCUMENT.TXT $P

        The characters $ and P are collectively called a "switch." That is, they switch on various options in the SPELL program. The dollar sign simply indicates that the letter <u>P</u> is a switch and not a file name. Several switches can follow the $ sign. For example if you wanted the misspelled words to be printed on both the console and the printer, you would type:

        A>SPELL B:DOCUMENT.TXT $PL

        The L switch means to list the misspelled words to the console. The L switch is normally turned on automatically if you don't specify any other switches when invoking SPELL. If you do use any other switch and you also want the misspelled words listed to the console, you must include the L switch explicitly.

        The next switch is called the file or F switch. It causes all of the misspelled words to be written to a file called <u>ERRWORDS.TXT</u>. This file is used by some of the other tools such as MARKFIX. You can also examine this file with your text editor, as it is sometimes more convenient than reading the words as they

are printed on the console. This is especially useful if you don't have a printer.

The F switch causes the file <u>ERRWORDS.TXT</u> to be written to the currently logged disk. In CP/M jargon, that means the disk that was last selected (usually the same one that contains SPELL). In the above examples the A disk was the currently logged disk (notice the <u>A></u> preceding the command). If you want the ERRWORDS.TXT file to be placed on another disk, then follow the F switch with the letter of the drive containing the disk. For example, to put the file ERRWORDS.TXT on the B drive, type:

    A>SPELL B:DOCUMENT.TXT $FB

CAUTION: If you do not type a drive letter after the F switch, you should make it the last switch on the line. Otherwise SPELL may read your next switch (C or B) and assume you mean to specify a disk drive.

Another switch that may be helpful is the S (Suppress) switch. The S switch causes SPELL to ignore any lines in your text file that start with a period or a colon. This will prevent SPELL from trying to check the spelling of the formatting commands (Sometimes called "dot" commands). Many word processors such as Word Star (tm) use "dot" commands to change line spacing, margins, etc. These programs use a command format that requires a command line to start with a period or a colon. If you try to use SPELL to check a file with these embedded commands, and you don't use the S switch, you'll get a bunch of strange two letter words like "SP," or "LN" included in the list of misspelled words

If you follow the S switch with a special character (such as S\) SPELL will ignore lines that begin with that character instead of a period or colon.

The C switch causes SPELL to remember the location where each word first occurred in you document. When the C switch is included, SPELL writes a file called <u>ERRWORDS.CON</u> which contains index information used by REVIEW to show a word's context. This file is written in binary code, so don't try to examine or edit it. The C switch causes SPELL to use about 30% more memory to compile the word list than would otherwise be needed. It will also write the ERRWORDS.TXT file even if you did not use the F switch. Both files will be written to the same disk selected for the ERRWORDS.TXT file.

The B switch tells SPELL to ignore anything that appears between successive backslash (\) characters. If you follow the B switch with a special character (such as B#) SPELL will ignore text that starts and ends with that character. This switch is useful with some editors that insert special code sequences into text files (such as Palantir's ruler line.) You can also use it to cause SPELL to skip over sections of your document that you do not want checked for spelling. For example, Word Star's .IG command will let you insert a comment into a document that will

not be printed.   Other formatters usually allow a similar option.
If you insert two successive backslash comments as follows:

```
.IG \
This section of text will not be checked for
spelling errors.
.IG \
```

SPELL will totally ignore all words between the two
backslashes and Word Star will not print the two comment lines.
This feature is very useful if you need to insert misspelled words
into a document (as this manual does) or if you have sections of
text (such as computer programs) that would look misspelled to
SPELL. Be careful how you use backslashes. If you forget and
leave out the second backslash, SPELL will not check the rest of
your document.

The last option is the I (Ignore) switch. It causes SPELL
to ignore any words in all capital letters. This option is useful
to ignore acronyms such as "ABC Corp." You can also use the I
switch to ignore programs embedded in your text. The following
passage would produce a lot of spurious misspelled words if it was
checked without using the I switch:

The following program was written to test the new QRP
computer which was installed recently by XYZ Corp.

```
        ; TEST BENCH MARK
        STARTX:
                LXI     H,OMEGAZ
        LOOP:
                DCX     H
                MOV     A,L
                ORA     H
                JNZ     LOOP
        ; END TEST
```

With the I switch, "QRP," "XYZ," "STARTX," etc. will not
be listed as misspelled.

SPECIAL DICTIONARIES

SPELL can also check words in an optional "Special
Dictionary." A special dictionary is useful for people who
routinely use a lot of specific jargon in their documents.   A
later section will describe how to build your own dictionary, but
for now assume that you have already made one called MYDICT.
Typing the following command would cause SPELL to also check
MYDICT before printing out any misspelled words.

A>SPELL B:DOCUMENT.TXT MYDICT $P

Notice that any option switches must be at the end of the
command line.

UPDATE DICTIONARY

If you find the main dictionary is missing some words that you use all of the time, you can put them into the "update dictionary." This is a file called UPDICT.CMP (for Update Dictionary). If this file is present on the currently logged disk, SPELL will automatically check it after checking the main dictionary.

The message:

Checking Update Dictionary.

will be typed to the console by SPELL whenever the file UPDICT.CMP is being checked.

== MARKFIX ==

MARKFIX is a general-purpose document-marking and correction (fixing) tool. MARKFIX is used during spelling checking to make corrections and mark misspelled words in your document files. MARKFIX can also be used to rewrite your documents and make corrections other than spelling corrections. This section will explain how MARKFIX works and how to use it.

MARKFIX needs to know several things before it can start to work. It needs to know the name of the file to be corrected or marked, the name of the file containing the list of corrections and/or words to be marked, and the character to be used for marking. If you run MARKFIX as follows:

A>MARKFIX DOCUMENT.TXT

MARKFIX will read and correct the file DOCUMENT.TXT and will assume that the file ERRWORDS.TXT contains the corrections. Words will be marked with a * if required. If a second file name is typed after the name of the document, MARKFIX will read the correction list from that file. If you do not specify a disk drive, MARKFIX will assume the files are on the currently logged drive.

To mark with an @ character and read corrections from the file CORRECT.TXT type:

A>MARKFIX DOCUMENT.TXT CORRECT.TXT $M@

Whatever character you type following the M switch will be used to mark words instead of a *. Words are marked in a document by placing the marking character in front of the word. The rest of the word is not altered.

Some text editors such as Word Star will require you to reform margins if a line changes length while MARKFIX is correcting words. This would happen, for example, if MARKFIX changed "errorr" into "error." The line containing the word would now be one character shorter and the right margin would off. MARKFIX's L switch will tell it to mark words that change length during correction. To repeat the example given above, except mark words that change length, use MARKFIX as follows:

A>MARKFIX DOCUMENT.TXT CORRECT.TXT $M@L

The correction list is given to MARKFIX as a CP/M text file. Each line in the file specifies a word to be marked or corrected, and if corrected gives a replacement for the word. Lines have the following format:

word/replacement

Using this list, MARKFIX will read a file and change every occurrence of "word" to "replacement." If a word is given in the

list without a replacement, MARKFIX will simply mark all occurrences of that word with the marking character.

When replacing a word with a correction, MARKFIX will automatically convert the capitalization of the replacement word to match that of the word being replaced. MARKFIX will recognize three kinds of capitalization.

1. All lower case
2. First letter capitalized
3. All letters capitalized

MARKFIX determines which conversion to use by looking at the first two letters of the word to be replaced. Therefore, words such as WINsome will be treated as if they had all letters capitalized.

The correction list should not contain words with possessive endings (i.e. Test's, Tom's and James'). MARKFIX handles possessive endings automatically by removing them from words before trying to find them in the correction list and adding them to replacement words. If your list contained the replacement pair BOB/TOM, MARKFIX would correctly change the sentence "Jim ate Bob's hat." into "Jim ate Tom's hat."

When MARKFIX is finished it will print several messages as follows, to show you what it has done:

    nn words marked
    nn words corrected
    nn words were undivided * reform *
    nn words changed length * reform *

Words that are divided at the end of a line must be undivided when corrected. Also, when corrections are longer or shorter than the word being replaced, the line length will change. With Word Star, you may need to reform your document if this occurs. This will always be true if you want to maintain "flush right" margins.

AUTOMATED REWRITE

MARKFIX is useful for more than just correcting misspelled words. For example, assume you had just written a 300 page novel about six brothers Tom, Bob, Dick, Albert, John and Luke. Each brother also has a nickname and lives in a different town. Your agent calls up and says he has a publisher who wants to buy the story, if only you'll rewrite it to take place in South America. He also tells you to change the characters' names to Raul, Julio, and so on. With MARKFIX you could probably have the story back to him in a few days. Just type in a correction list as follows:

    TOM/RAUL
    BOB/JULIO
    . . .

While the word to be replaced must be a single word, its replacement can be several words. The following example illustrates how you can use this feature to change informal writing into formal writing:

        I'LL/I WILL
        I'VE/I HAVE

Consider the potential for deflating stuffy writing by making a replacement list to change words like "utilize" into "use." The latter word is almost always clearer. Of course, if you want to "pump up" your writing, the reverse is also possible.

== HYPHEN ==

HYPHEN is a tool for automatically hyphenating words. HYPHEN can be used two ways. It can read a document file and insert soft hyphens into words longer than a given length, or it can be used "on-line" to list out hyphenation points for words you type in.

HYPHEN is based on a technique described by Donald Knuth in his book TEX and METAFONT, New Directions in Typesetting (Digital Press, Bedford, MA, 1979). HYPHEN uses an elaborate set of rules that describe how to locate permissible hyphenation points within a word. In addition, HYPHEN uses a disk file of exception words (HYEXCEPT.TXT) that defy its rules. You can add to this list as you discover words in your vocabulary that HYPHEN cannot properly hyphenate.

Knuth's hyphenation algorithm is quite good at finding proper hyphenation points; however, it is not perfect. There are many words that do not obey any "rules" for hyphenation. Also, HYPHEN will not find all of the hyphenation points possible for every word.

You will probably find HYPHEN most useful for setting narrow columns of text, such as for newsletters and magazines. HYPHEN will also help to avoid the "rivers of white space" problem when setting text in "flush right" format.

HYPHEN will not reformat your text. That must be done by your word processor or text formatter. HYPHEN will insert soft hyphens into words so that they will break more cleanly when your text is justified. Unless told otherwise, HYPHEN will assume that you wish to insert soft hyphen characters compatible with Word Star. The section "Advanced Options" will show you how to change to a different soft hyphen character (such as a & for Magic Wand).

You can try out HYPHEN by letting it print out hyphenation points for words you type in. Run it by typing:

A>HYPHEN

HYPHEN will respond by typing a question mark. You can then type in a word and press return. HYPHEN will print back the word with dashes (-) inserted to show where it thinks the word can be divided.

To insert soft hyphens into a file you must invoke HYPHEN with a file name. For example:

A>HYPHEN DOCUMENT.TXT

This will cause HYPHEN to rename DOCUMENT.TXT to the name DOCUMENT.BAK and then write a new file DOCUMENT.TXT with soft hyphens inserted into words longer than 13 letters. Any existing soft hyphens in the file will be preserved. If you want HYPHEN to

only hyphenate words longer than 15 letters type:

    A>HYPHEN DOCUMENT.TXT $15

The number following the $ specifies the maximum length word not to hyphenate. Also, HYPHEN will not disturb any soft hyphen characters which were already in the file.

ADDING WORDS TO THE EXCEPTION LIST

Because of the irregularities of the English language, there will be exceptions to any set of hyphenation rules you care to devise. Such is the case with HYPHEN. Because of this, an exception word list (HYEXCEPT.TXT) has been created to help catch some of the more common exceptions. Even so, you will occasionally run across a word that HYPHEN does not break correctly. When this happens, you have two options: you can manually correct the hyphenation, or you can add the word to HYEXCEPT.TXT.

There are four steps to adding a new word to the exception word list. The first step consists of looking up the correct hyphenation of the word to be added. Using the example of "informant" (which is already in HYEXCEPT.TXT), we find that the correct hyphenation is "in-form-ant," not "in-for-mant," which is how the rules in HYPHEN would break it.

The second step is the trickiest: to determine how many letters are needed to uniquely identify this entry. When HYPHEN is comparing a word against the exception word list, it (usually) only checks the first seven letters. It does this for two reasons. First, that is all it usually needs to identify the exception word. For example, the first seven letters in "mincemeat" readily distinguish it from all other words. Second, this allows words with the same root but different endings to be taken care of by just one entry. For example, putting the word "fore-short-en" into HYEXCEPT.TXT will allow HYPHEN to also handle "fore-short-ened" and "fore-short-ening."

Sometimes, though, seven letters are not enough. For example, the first seven letters of "in-form-ant" will not distinguish it from "in-for-mal" or "in-for-ma-tion." You must somehow tell HYPHEN to look at the eighth letter, "n," in order to avoid mistaking "informal" and "information" for "informant." To do this, you must put an asterisk (*) in front of each letter that HYPHEN must look at beyond the seventh. In the case of "informant," we would write it "in-form-a*nt." Sometimes, several letters will need to be marked. For example, "af-firm-ative" is not hyphenated the same as "af-fir-ma-tion," but the first nine letters of both words are identical. In order to define "affirmative" as an exception, you would have to write it "af-firm-a*t*i*ve."

The third step is relatively simple: just use whatever text editor you normally use to edit HYEXCEPT.TXT and add your new word, complete with hyphenation and, if needed, asterisks. You don't need to put it in any special location--HYPHEN sorts the list as it reads it in--but it would probably be a good idea to put it on a line all by itself.

The fourth step is easy: re-run HYPHEN and see if the exception word is now hyphenated correctly. If it isn't, then check the following:

1. Make sure that HYEXCEPT.TXT is on the same disk as HYPHEN;

2. If so, make sure that it's the version that you just changed;

3. If so, examine it (using TYPE or your editor) and make sure that you correctly spelled and hyphenated the word.

== LOOKUP ==

LOOKUP is a separate program that works the same way as the look up feature in REVIEW. You can use LOOKUP as a convenient electronic "desk dictionary" when you are writing on paper. LOOKUP is able to find the correct spelling for most of the words that people misspell because everyone tends to misspell in four fairly consistent ways. These "Misspellers' Rules" are listed below:

1. Added an extra letter (accute instead of acute.)

2. Dropped a needed letter (absess instead of abscess.)

3. One letter wrong (accur instead of occur.)

4. Two letters reversed (acer instead of acre.)

LOOKUP uses the above rules and tries to "un-misspell" the word and find it in the dictionary. LOOKUP first uses these rules in reverse and builds a list of likely correct spellings in memory. Then LOOKUP searches through the dictionary and prints out any matches.

There are two ways to use LOOKUP depending upon how far off your spelling of a word is. LOOKUP can do a "short" search, meaning that it only checks words in the dictionary that start with the same letter as your misspelling, or it can do a "long" search and check every single dictionary word. Naturally a short search is faster than a long search.

Let's assume that you have misspelled "across" as "accross" (Misspellers' Rule 1). To use LOOKUP with a short search you would type the following:

A>LOOKUP ACCROSS

LOOKUP would respond by printing out a list of likely candidates to your console. If LOOKUP finds a word in the dictionary spelled exactly like the word you typed, it will put a _*_ after the word when it is listed to the console.

A short search is usually sufficient when you are fairly confident of the first letter of your misspelling. But what if you aren't? Well then use a long search by typing:

A>LOOKUP ACCUR $L

Here we have applied Misspellers' Rule 3 to produce "accur" instead of "occur." The $L tells LOOKUP to do a long search (a short search is the normal default). A short search would never find the proper spelling because the misspelling is an incorrect first letter. Always use a long search when you are unsure of the first letter of a misspelling.

== FIND ==

FIND is a unique program which can be a lot of fun once you know how to use it. FIND is a general-purpose search utility which searches the dictionary for word patterns. FIND will help you solve a variety of word puzzles, locate rhyming words and much more.

An introductory example might help introduce FIND better than an entire paragraph of explanation.

Let's pretend that you are trying to solve the following crossword puzzle.

```
G R A T I N G
R   D   N   E
A B A N D O N
P   P   O   E
H ? T ? O ? S
I   E   R   I
C O R P S E S
```

You just can't seem to figure out what "H?T?O?S" should be. FIND can easily scan through the dictionary and locate all words that match the pattern H?T?O?S. Just type:

A>FIND H?T?O?S

and FIND will return the answer. Did you get "Hotdogs"?

FIND is very useful to crossword puzzle fans. In fact, FIND is so good it's almost like cheating. You can also use FIND to create your own crossword puzzles. Try and finish the following puzzle.

```
D A R T
I ? ? A
G A N N
S ? ? G
```

(HINT: Work on the last word S??G first.) When you solve that, try and make up your own puzzle from scratch.

The question mark (?) used in the above example is called the "match all" character. You can use a it in any position within a word when using FIND. A question mark can even appear as the first character of a pattern, as for example:

A>FIND ?BLE

FIND will take longer when searching for ?BLE, because it must check the pattern against every word in the dictionary. In contrast, FIND is smart enough to know that it only needs to check the R section of the dictionary when looking for the pattern RE??.

ADVANCED FEATURES

The following command:

A>FIND *HING

is extremely powerful. It tells FIND to locate and list every word which ends in HING. The * character means to match any sequence of one or more characters. The * can also be used as follows:

A>FIND TEST*

This will locate and print all words that start with the letters TEST. You can also combine ? and * characters in the same pattern, as in the example:

A>FIND A??L*

Try this pattern yourself and see what you get. Another pattern you may be tempted to try is this:

A>FIND *

If you try this we hope you have a lot of time, because FIND will list every single word in the dictionary. There is not really much point in trying it, but maybe some evening when you're bored with TV. . . .

The most obvious use for the * in pattern matching feature is in finding rhyming words. For example, if you wanted to find words which might rhyme with "probably." You could try the patterns *ABLY or *LY.

Remember that when using the * character as the first letter of a pattern, FIND must search the entire dictionary. As we mentioned above, this will take more time than searching for a pattern beginning with a letter.

Note: FIND can only accept a single * in a word pattern. Typing:

A>FIND *IOU*

or a similar pattern, is not allowed.

== ANAGRAM ==

    Many newspapers carry a word puzzle called "Jumble" (registered trademark) devised by Henri Arnold and Bob Lee. A jumble puzzle involves unscrambling four words to reveal several letters which, when combined and also unscrambled, will yield the caption to a humorous cartoon. ANAGRAM is a word unscrambling program which will help you solve these puzzles. For example, if you want to find the solution to the jumble LECUN, type:

    A>ANAGRAM LECUN

    After a short delay, ANAGRAM will print "UNCLE."

    ANAGRAM will also let you type ? symbols to represent positions which can be occupied by any letter. This is useful for solving other types of word puzzles. For example, to find all of the words which contain five letters and the vowel A in two positions, type:

    A>ANAGRAM AA???

    We'll leave you to find the answer.

== WC ==
(Word Count)

WC is general-purpose word counting utility. WC will be of interest mostly to writers and professional wordsmiths. It gives a quick way of finding out, in terms meaningful to humans, just how big a document really is. Most professional writers are used to seeing phrases like "submissions 1500--4000 words" in publishers' guidelines. WC provides a quick way to find out whether your latest article or novel is too long or too short. You can use WC just by typing:

A>WC DOC.TXT

Within a few seconds (depending on how long your file is) WC will reply something like: There are 1,325 words in the file.

WC uses a simple technique for counting words adapted from "Software Tools," by Kernigan and Plauger. Any sequence of alphabetic characters (including apostrophes and hyphens) surrounded by nonalphabetic characters is counted as a word. Hyphenated words are thus counted as one word (unless the hyphen (-) occurs at the end of a line, i.e. breaking a word at the hyphenation point).

This technique is not infallible, owing to its simplicity but it is very fast and is certainly much more accurate than counting words by hand.

== HOMONYM HELP ==

As we mentioned in the chapter describing how to use SPELL, there is currently no automatic way to locate incorrectly used homonyms. This is a limitation of most computer spelling checkers. It's really a shame also, because homonyms are exactly the kinds of words that many people have trouble with. Some common examples are:

BAZAAR BIZARRE
CAPITAL CAPITOL
COLONEL KERNEL
COMPLIMENT COMPLEMENT
DINGHY DINGY
NAVAL NAVEL
WAIVER WAVER

Because many of us suffer from homonym confusion we have provided a feature we call Homonym Helper. This feature will not tell you which homonyms are correct in which contexts, because which word you use depends on what you are trying to say. What it will do is give you a way to find words in your document that are potential trouble makers.

Homonym Helper makes use of an option in the program MARKFIX. If you type a second file name after the name of the file to be marked, MARKFIX will read the list of marking words from the second file instead of from the default file ERRWORDS.TXT.

Using this feature MARKFIX can read from a file containing a list of homonyms and mark where they occur in your document. You can then use your text editor's search command to read through the file and verify that all homonyms are used correctly. This is not an ideal solution to the homonym problem, but it does offer help, especially when reviewing a document that may have been written by someone else.

Before using the Homonym Helper feature, you will need to do some initial preparation. On your disk is a file called HOMONYMS.TXT. It contains a list of words many people get confused. Most of these are "homophones" (words that sound alike when pronounced, but that are spelled differently). Some are words that, because of differences in the way people say certain words, can be confused. In order to use this file you should first copy it into another file called HHELP.TXT. Then, using your text editor, look through HHELP.TXT and delete any words which you are sure you know how to spell and use correctly. This will leave HHELP.TXT containing a list of words which you have trouble with. You may also add to this list later as required.

The initial file is organized with two or three alike sounding words per line. This makes it easier to decide which words may give you difficulty. If the words on a line are familiar to you, then delete the line.

Before using HHELP.TXT With MARKFIX you must sort the words into alphabetical order. Do this by using the DICTSORT program as follows:

A>DICTSORT HHELP.TXT

Once you have compiled the list, you can use MARKFIX to mark the homonyms as follows:

A>MARKFIX DOC.TXT HHELP.TXT

You can then use your text editor to look at the marked file.

== THE DICTIONARY ==

The computer dictionary included with The WORD Plus required over a years worth of work to develop. Words included in the dictionary were selected by processing actual text files by computer. This means our dictionary includes the words that people really use.

For accuracy, the dictionary has been checked for errors by several specialized computer programs and also by hand using "The New American Heritage Dictionary of the English Language" 1979, by Houghton Mifflin Company. Common usage of words was checked using "the Pocket Dictionary" also by Houghton Mifflin. Proofreading by hand required over three months and four separate passes through the dictionary.

HOW THE DICTIONARY IS STORED

Some of you may be wondering how we managed to get a 45,000 word dictionary onto a single floppy disk. Well, it isn't as hard as it might seem. The dictionary is actually just a word list in alphabetical order. The original source text file (including carriage returns and line feeds) was over 450k bytes in size. The version included with The WORD Plus has been highly compressed. The compression scheme takes advantage of the fact that an alphabetical list has a lot of redundancy built into it. Successive words in the list normally have a number of letters in common, most notably the first few. Also it is not necessary to use CR and LF characters as delimiters. Instead the last character of each word has its MSB set. In successive words that have common letters, the common letters of the following word are replaced by a one byte count. The next example may clarify this point. Assume the digits 0-9 represent binary values.

```
    normal            compressed

    ABACK             ABACK      (FIRST WORD NOT COMPRESSED)
    ABANDON           3NDON
    ABANDONING        7ING
    ABANDONMENT       7MENT
    ABASEMENT         3SEMENT
    . . .
    chars=42          chars=26
```

The above example illustrates a compression ratio of 42/26. In the actual dictionary the ratio was closer to 3/1. The space gain is also magnified by the fact that CR and LF characters were eliminated.

A side benefit of this compression technique is the compressed dictionary is actually faster to read on a typical floppy disk system. This is because a program using the dictionary can get read more words with fewer disk accesses. The

amount of time taken to uncompress a word is almost negligible
compared to the time it takes to read a sector from a floppy disk
spinning at 6 revolutions a second.

== CUSTOM DICTIONARIES ==

The built-in dictionary (contained in the file
MAINDICT.CMP) provided with The WORD Plus cannot be altered. We
did this to protect you from yourself! It is VERY easy to miss a
few misspelled words when looking down a long list of them. And
it's a lot harder to remove a word from a 45,000 plus dictionary
than to add it.

However, there is a way to add words to the dictionary
without actually altering it. As was mentioned in the section
describing SPELL, you can create a file called UPDICT.CMP which
contains your list of added words. Whenever SPELL is used, it
will check this file as if it were an extension to the main
dictionary.

A word added to the file UPDICT.CMP will be checked EVERY
time you use SPELL on a file. If you have words which you use
only in certain contexts you should add them to a special
dictionary instead (see the section describing how to use SPELL).

Pressing the U key, while using the program REVIEW, will
add a word to the UPDICT.CMP dictionary file. Likewise, when you
press the S key, you are adding the word to the file SPECIALS.CMP.
REVIEW takes care of placing the words into the file in
alphabetical order.

If you are doing a lot of editing and revising on a text
file you can simplify your task by keeping an up-to-date special
dictionary containing the words which are unique to that file.
That is what the file SPECIALS.CMP is intended to be used for.
Then you can have SPELL use that special dictionary the next time
you recheck the file for spelling. This will eliminate having to
see the same words listed as misspelled over and over again.

DICTSORT

Sometimes you may wish to compile your own word list
independent of using REVIEW. For this reason we have provided a
program called DICTSORT which will alphabetize your word lists so
that they can be used by SPELL.

Any file used as a special or update dictionary by SPELL
is simply a list of words in alphabetical order. Each word starts
on a new line (separated by a carriage return and linefeed). The
important part is that the list MUST be in alphabetical order. It
is quite a job to alphabetize a list of 1,000 words using a text
editor. So always use DICTSORT on a word list before using it
with SPELL.

DICTSORT takes a file as input, reads it into memory,
sorts the words into alphabetical order (redundant words are
discarded), and writes the sorted list to a disk file. The sorted
file will be named the same as the input file. Your original word
list will be untouched, but it will be renamed to the old name

plus the extension <u>BAK</u>. For example, let's say you wanted to sort a file called <u>WORDLIST.TXT</u>. You would type:

        A>DICTSORT WORDLIST.TXT

        DICTSORT makes a new file called <u>WORDLIST.TXT</u> and renames the original file to <u>WORDLIST.BAK</u>. We must emphasize that you should ALWAYS DICTSORT a file BEFORE you attempt to use it with SPELL as a user dictionary. It is very difficult to alphabetize a file manually and be sure of getting it right. In addition, SPELL recognizes the special character apostrophe (') as a valid character in dictionaries. This character comes before <u>A</u> when alphabetizing a list of words. Notice that this is different than in printed dictionaries.

        You can also use CP/M's PIP utility, (called XFER on some versions of CP/M) in conjunction with DICTSORT, to add words to an existing special dictionary. Assume that you had a list of words in SPECIALS.CMP which you wanted added to the current update dictionary. The next five steps would do it for you.

1.   A>PIP TEMP.CMP=UPDICT.CMP,SPECIALS.CMP

        The file SPECIALS.CMP is appended to the update dictionary (UPDICT.CMP) and the whole thing is put into a temporary file (TEMP.CMP).

2.   A>DICTSORT TEMP.CMP

        The temporary file is sorted into alphabetical order.

3.   A>ERA UPDICT.CMP

4.   A>ERA TEMP.BAK

5.   A>REN UPDICT.CMP=TEMP.CMP

        The last three commands do a little housekeeping by erasing the old update dictionary and the temporary .BAK file generated by DICTSORT. Then the temporary file TEMP.CMP is renamed to become the new UPDATE dictionary.

OTHER USES FOR DICTSORT

        DICTSORT can also be used to build a vocabulary list from a text file given to it as input. However, DICTSORT will not remove possessive endings from words as SPELL does. You can use DICTSORT to generate a list of all the words in a document and then use the word counting utility WC to see how big your vocabulary is. This manual for example, is 12,023 words in size and contains 1,590 different words.

        Remember that DICTSORT will rename the original file to .BAK, so be careful to make a copy of the original file <u>before</u> using DICTSORT on it. Otherwise you may lose your original file!

BUILDING A CUSTOM MAINDICT.CMP FILE

In some applications you may wish to replace the master dictionary file with a custom dictionary of your own making. This may be useful, for example, to build a restricted vocabulary dictionary for checking a children's book. The procedure for making a new master dictionary is simply to build the dictionary you want with DICTSORT, then delete the old MAINDICT.CMP file and then rename your dictionary to MAINDICT.CMP. SPELL will then read from your dictionary the next time you use it. Be sure and save a copy of the original dictionary so that you can use it again later.

Note: LOOKUP and FIND will <u>not</u> be able to use a custom MAINDICT.CMP file. In addition to being compressed, the master dictionary is also indexed for rapid access. FIND and LOOKUP require this index to operate properly. An attempt to use FIND or LOOKUP with other than the original master dictionary will result in the error message:

Main Dictionary is not on current disk!

== WORDFREQ ==
(lexicographic analysis for beginners)

The last tool described in this manual is WORDFREQ. WORDFREQ works much like DICTSORT does in that it compiles a list of words from a text file. In addition, WORDFREQ also keeps track of how many times each word appears in the document. After the word list is compiled, the program will write a file to disk with the extension .FRQ. This file contains the words and a count of how often they were used. You invoke WORDFREQ as follows:

    A>WORDFREQ DOC.TXT

The word list will appear as follows with the word frequency appearing first followed by the word.

    12   THE
    10   IN
    8   THIS
    . . .

* Note * WORDFREQ does **not** display the frequency list on your terminal screen. You must type, print, or edit the file file.FRQ to see the word frequency information.

In addition, WORDFREQ will print several file statistics to the screen, such as:

    x Total words
    x Unique words
    x Words appearing once

The list will normally be written with the words sorted into decreasing order of occurrence. However, if you invoke WORDFREQ as follows:

    A>WORDFREQ DOC.TXT $A

The A switch will cause WORDFREQ to write the word list in alphabetical order.

WORDFREQ can be very useful to writers who are concerned with how frequently they use certain words. In addition, WORDFREQ can be used on non-text files like computer programs to see how often computer instructions are used. The only rule which must be followed is that WORDFREQ considers a word to be any sequence of alphabetic characters (with embedded apostrophes) surrounded by non-alphabetic characters.

Files read by WORDFREQ should also follow normal conventions for text files with a CP/M End Of File (EOF) character (1AH) appearing at the end. If your editor can read the file then WORDFREQ should also be able to read it.

== ADVANCED OPTIONS ==

When the various tools are reading your text files, they are trying to locate words. A word is defined as any sequence beginning with a letter and containing only other letters or an apostrophe. Any other characters are ignored as long as they don't occur in the middle of a word. Exceptions to this rule are "soft hyphen" or "ghost hyphen" characters. These characters are used by some word processors to indicate where words may be divided during line justification.

As delivered, The WORD Plus is ready to work with Word Star compatible files. Word Star uses two types of soft hyphen codes, Active (ctrl-underline, 1F hex) and Inactive (ctrl-caret, 1E hex). These codes are automatically recognized and removed from words before further processing.

Soft hyphens are invisible characters, unlike the apostrophe which is treated as if it were a valid letter of the alphabet. The WORD Plus has a special way in which you can redefine soft hyphen characters and also the letters that are valid within a word. The latter capability will let you change The WORD Plus to recognize foreign words containing accent marks, tildes, etc.

The built-in definitions for characters that are legal within words, and for soft hyphens can be changed by creating a file with your editor. If the file TWOPTION.CMP is present on the same disk as a particular tool, that tool will read it automatically every time it is used. This file contains new definitions for these characters. Program patching is not required and you can restore normal operation by simply deleting the file.

TWOPTION.CMP is a two line file. The first line should contain all of the nonalphabetic characters which are legal within a word. The second line must contain all legal soft hyphen characters. For example, to customize the tools for Magic Wand (using a & as a soft hyphen) type the following two line file:

```
'

&
```

Be careful not to type any other characters (especially invisible control characters). End each line by pressing return. If you only need to change the definitions for soft hyphens, the first line will still need to contain an apostrophe, otherwise all of the tools will stop recognizing them. Likewise, if you need to redefine the first line you will need to reenter the soft hyphen codes. Each line can contain up to 32 characters.

If you need to insert control characters into the file, type a lower case c in front of the code. For example, to enter a control @ (produces a hex code 0) type c@. The standard definitions for Word Star are produced by the following two lines.

'
c^c_

Note: These definitions are built-in and do not need to be added by making a TWOPTION.CMP file. Also, the inactive soft hyphen character (1E hex) should be defined first. This is the code that HYPHEN will insert into words that it hyphenates.

Note: Be careful when creating or modifying TWOPTION.CMP. A mistake may cause erratic behavior of the programs.

== QUICK REFERENCE GUIDE ==

- SPELL -

A>SPELL DOC.TXT	List misspelled words to console
A>SPELL DOC.TXT $FB	Write words to file B:ERRWORDS.TXT
A>SPELL DOC.TXT $PL	List words to console and printer
A>SPELL DOC.TXT $S	List to console, suppress '.' lines
A>SPELL DOC.TXT $I	Ignore words in all capital letters
A>SPELL DOC.TXT $B	Ignore Backslash commands
A>SPELL DOC.TXT $C	Save word context information

- REVIEW -

A>REVIEW	Invoke REVIEW program

- LOOKUP -

A>LOOKUP QUIK	Lookup correct spelling for "Quick"
A>LOOKUP INTROPY $L	Long search for "Entropy"

- FIND -

A>FIND T???	Find all 4 letter words starting in T

- ANAGRAM -

A>ANAGRAM LECUN	Unscramble "LECUN" and print the answer
A>ANAGRAM ???EE	Print all five letter words with two e's

- MARKFIX -

A>MARKFIX DOC.TXT	Mark all misspelled words in DOC.TXT
A>MARKFIX DOC.TXT $M@	Mark words with @ instead of *
A>MARKFIX DOC.TXT $L	Mark words that change length when fixed
A>MARKFIX DOC.TXT F.TXT	Read corrections from F.TXT instead of from ERRWORDS.TXT

- HYPHEN -

A>HYPHEN	Invoke HYPHEN for on-line look up
A>HYPHEN DOC.TXT	Insert hyphens into words > 10 letters long in file DOC.TXT
A>HYPHEN DOC.TXT $14	Insert hyphens in words > 14 letters

- WC - Word count

A>WC DOC.TXT	Count how many words are in DOC.TXT

- DICTSORT -

A>DICTSORT DICT.TXT	Alphabetize words in DICT.TXT (Old file becomes DICT.BAK)

- WORDFREQ -

A>WORDFREQ DOC.TXT	Count frequency of words in DOC.TXT and write file DOC.FRQ
A>WORDFREQ DOC.TXT $A	Same as above except DOC.FRQ is kept in alphabetical order

== ERROR MESSAGES ==
(or what to do when your tools break)

Each tool has several error messages which it will print when it is confused, dazed, or damaged. These messages, along with suggested remedies, are listed below. If reading these short "cures" doesn't help, then re-read the chapter describing how to use the particular tool that is giving you problems.

* WARNING * This tool is damaged!
Use another copy.

This message means that the copy of the program that you are using is not loading into memory properly. Each tool checks itself after being loaded from disk to make sure it has all of its marbles intact. If using another copy of the tool in question doesn't help, it could mean that your computer or word processor has a memory problem. Consult your local computer vendor or hardware guru if trouble persists.

* WARNING * Illegal copy!

This is a very serious error message. It means that you are using an illegal or bootleg copy of one of our programs. If you purchased this package through what you believed to be a legitimate vendor, please contact Oasis Systems immediately.

Input file is NOT on disk

This message means that you told one of the tools to read a disk file that doesn't exist. This usually means you made a mistake when typing in the command to invoke the program. Frequently, you may simply forget to add a needed drive specifier, such as <u>B:</u>.

User dictionary is not on disk

This message is a warning message only. It means SPELL tried to read from a special dictionary that does not exist. Aside from the precautions mentioned for the previous error message, this message can be caused by forgetting to put the <u>S</u> in front of switches, thus causing them to look like a special dictionary filename.

Main dictionary is not on current disk!

This message is typed by SPELL, FIND, or LOOKUP when they can't locate the file <u>MAINDICT.CMP</u>. Usually this means you copied SPELL, FIND, or LOOKUP to another disk, but didn't copy the dictionary. Or you accidentally erased the dictionary file. This message can also be caused by using the wrong revision dictionary (see "Incorrect MAINDICT.CMP file!" message.)

Insufficient memory for input file
Program Aborted!

This message is typed by SPELL, LOOKUP, WORDFREQ or DICTSORT. It means that your system doesn't have enough memory to complete the function, or that you are trying to process an extremely large input file. The only solution is to get more memory (64k is the maximum any tool can use), or break your input file up into sections or chapters.

* WARNING * Word too large,
or non-text input file!

This means that the particular tool found a string of alphabetic characters that exceeded its buffer storage of 42 characters for a word. This is only a warning message and means the tool in use has ignored the illegal sequence.

Incorrect MAINDICT.CMP file!

This message means you are using a previous revision dictionary with a newer program.

Illegal switch ignored --> W

This message indicates that the character W was not recognized as a valid option switch for the particular tool you were using. This is only a warning message and means the tool has ignored the invalid switch.

WRITE ERROR

This message means that the tool being used was trying to write a file to the disk and ran out of room. The program has aborted with the file partially written. To proceed, you will either need to ERAse some unused files on the disk, or get a clean disk.

Incorrect Input (word too short, etc.)

This message is typed by LOOKUP when you are trying to look up a word that is less than 4 characters long.

Word too large to look up

This message is typed by the look up feature in REVIEW if you try to look up a word longer than 15 letters. REVIEW has less memory space to build lists than LOOKUP, so it is limited to shorter words.

No similar word

If REVIEW's look up feature cannot locate a similarly spelled word it prompts this message.

Insufficient memory
Program Aborted!

This message should appear rarely. It means that you are attempting to use one of the tools with too little memory (less than 16k).

Incorrect Input Format

FIND types this when you forget to give it a pattern to find.

Marking file is NOT on current disk

This message means that MARKFIX was unable to find the marking file on the currently logged disk (usually ERRWORDS.TXT unless you are using an optional marking file such as HHELP.TXT.) This message usually means that you directed the ERRWORDS.TXT file to the wrong disk by using the F switch in SPELL. MARKFIX expects the file to be on the same disk it is.

Insufficient memory for ERRWORDS file
Program Aborted!

MARKFIX types this when you have too little memory, or if your list of misspelled words is VERY large. You should only see this message if you forget and use a very small CP/M system. If your ERRWORDS.TXT file really is too large, try breaking it in half and making two passes with MARKFIX.

Can't find tool xxxx.COM

TW prints this message if it can't find one of the tools SPELL, REVIEW or MARK. These programs, plus TW, must be on the same disk drive.

Context file ERRWORDS.CON is not present

This message means that you tried to use the look up feature within REVIEW without having had SPELL save a context index file. You must answer yes to the question "Save context info" when you use TW, or use the C switch when you run SPELL directly.

REPORTING PROBLEMS
(what to do when something doesn't work)

PROBLEM--noun, a question or situation that presents uncertainty or difficulty.

Problems with software are frustrating. Perhaps the most frustrating thing about a software problem is that unless you wrote the program yourself you can't do much about it. When it comes to helping you, the fact that you didn't write The WORD Plus' programs is OUR most difficult problem.

The purpose of this section is to help YOU get your problem solved as fast and as painlessly as possible. We'll assume for the moment that you're reading this section because you have found a problem with one of The WORD Plus' tools. What should you do? Call us? That's OK if you need to, but please finish reading this section BEFORE you do.

In order to track down a program bug we need as much information as possible about the nature of the bug. How serious is it? (Did your computer burn up as a result, or did SPELL simply miss a misspelled word); What was going on just before the problem occurred? (Did the room lights blink off); What kind of text file were you using the tool on? (A six hundred page novel, or a one page letter); What error messages, if any, did the program print out?

Before you decide you have a problem with one of our programs we'd really appreciate it if you would re-read the manual and make sure you're doing everything correctly. Many times people attempt things that our programs were never intended to be used for. (Please write and tell us if you find a section of the manual unclear! We'll fix it.)

All of this may seem like nit-picking and you might say to yourself "It's their problem to find out what went wrong, not mine!," but the fact is, without your cooperation, we won't be able to find the problem as fast.

The next section includes a "BUG REPORT FORM." We're not particularly fond of forms, but we have found from experience that it helps to have some guidelines. You don't have to fill out the entire form, just try to put down as much information as possible.

If your problem is relatively minor (be objective) then we would prefer that you mail us your completed form. If the problem is so serious that you just can't wait and you have to call us, please don't call collect! We cannot accept collect calls. If your problem cannot be resolved quickly over the phone then we'll call you back. We really are here to help you!

BUG REPORT FORM

Send to: OASIS SYSTEMS, The WORD Plus bug report, 2765 Reynard Way, San Diego, CA 92103

Your Name, Address, and Telephone Number

Which tool were you using when the problem occurred?

Were there any error messages printed?

What is your system configuration (how much memory, CP/M 1.4, 2.2, CDOS, etc?)

What kind of terminal do you use (Hazeltine 1500 CRT, teletype, etc?)

Were any disk files damaged as a result?

Were you able to duplicate the problem?

Have you made any modifications to The WORD Plus' programs (renamed files, etc.)

What version number tool were you using? (The version number is printed whenever a tool is used.)

What is the serial number of your copy of The WORD? (The serial number is on the master release disk.)

Describe the problem and any related symptoms: