DICTION 5.0 THE TEXT-ANALYSIS PROGRAM

User's Manual

Windows Version
Software License Agreement

The enclosed software program is licensed by Digitext, Inc. to customers for their non-exclusive use on a single computer system per the terms set forth below. Using the enclosed CD–ROM indicates your acceptance of these terms.

License
You have the non-exclusive right to use the enclosed program on a single computer. You may not electronically transfer the program from one computer to another over a network. You may not distribute copies of the program or documentation to others. You may make one (1) copy of the program solely for backup purposes. You may transfer the software from one computer to another on a permanent basis only, and only when all copies of the software on the original computer are removed on a permanent basis. You may not use, copy, modify, sub-license, rent, lease, convey, translate, or transfer the program or documentation, or any copy, except as expressly provided in this agreement. You may not convert the software to any programming language or format, decompile or disassemble the software or any copy, modification or merged portion, in whole or in part. The software is protected by United States copyright laws and international treaty provisions.

Limited warranty
This program is provided as is without warranty of any kind expressed or implied, including but not limited to the implied warranties of merchantability and fitness for a particular purpose. The entire risk as to the results and performance of the program is assumed by you. Digitext, Inc. does not warrant that functions contained in the program will meet your requirements. The sole and exclusive remedy available to the purchaser or user of this software and accompanying documentation is a refund or replacement of the product at the option of Digitext, Inc.

Customer Remedies
To the original purchaser only, Digitext, Inc. warrants the CD-ROM on which this software product is distributed to be free from defects in materials and faulty workmanship under normal use for a period of ninety days from the date of purchase. If during this ninety (90)-day period the CD-ROM should become detective, it may be returned to Digitext, Inc. with copy of the receipt for a replacement without charge.

Copyright Notices
Program Copyright © 2000 Digitext, Inc. Austin, Texas. All Rights Reserved. Windows is a trademark of Microsoft.

TABLE OF CONTENTS

LIST OF FIGURES

PART 1: INTRODUCTION

Welcome to DICTION

DICTION 5.0 is a Windows-based program that uses a series of dictionaries to search a passage for five semantic features—Activity, Optimism, Certainty, Realism and Commonality—as well as thirty-five sub-features. DICTION conducts its searches via a 10,000-word corpus and the user can create additional (custom) dictionaries for particular research needs. The program writes its results to both alphabetic and numeric files. Output includes raw totals, percentages, and standardized scores and, for small input files, extrapolations to a 500-word norm. DICTION also reports normative data for each of its forty scores based on a 20,000-item sample of contemporary discourse. The user may use these general norms for comparative purposes or select from among thirty-six sub-categories, including speeches, poetry, newspaper editorials, business reports, scientific documents, television scripts, telephone conversations, etc. On a Pentium-based machine with 64 MB of RAM, DICTION processes one hundred and twenty-five passages (62,500 words) in one minute. The program can accept either individual or multiple passages and, at the user's discretion, it provides special counts of orthographic characters and high frequency words.

Installing DICTION

Note: If you have previously installed the demonstration version of DICTION, uninstall it before installing the permanent program.

DICTION will require 4 MB of memory and 2.1 MB of hard disk space. To install do the following:

1. Start Windows and ensure that no other programs are running during the setup.
2. Place the DICTION CD-ROM into the CD-ROM drive. Go to **My Computer** and double-click drive **A**.
3. Double-click on **DICTION set up.exe**. This will begin the setup process.
4. To add a DICTION icon to the desktop after the program has been installed, go to c:\program files\diction and single-click on **Diction.exe**. Then drag the highlighted file to the desktop. Double-clicking o n the icon will start the program. (You can also find DICTIONin the Start/Programs menu.)

PART 2: DICTION 5.0 IMPROVEMENTS

Diction 5.0's look and appearance have been completely redesigned and several new features have been added as well:

- ❖ Full Windows functionalities for selecting and analyzing texts
- ❖ Processes input files of any size (up to 500,000 words)
- ❖ Permits comparison of the user's input file to thirty-six kinds of texts, including speeches, business reports, television scripts, poetry, advertising, etc.
- ❖ Allows easy switching among norms to get multiple views of the same text
- ❖ Input files can carry standard (**.txt**) extensions and can be stored in any directory
- ❖ Project Files let user include any number of DICTIONruns in the same file
- ❖ Numeric file is visible at all times and requires no special reader
- ❖ Program eliminates user errors by never over-writing data fro m the same t ext
- ❖ User is given almost total control over file extensions and file locations
- ❖ Vastly improved on-screen help, with Windows-compatibility and context-sensitive features
- ❖ Program is now highly tolerant of problematic text files
- ❖ Print-preview feature lets user see exact appearance of output before printing (WYSISYG format)

PART 3: DICTION FUNDAMENTALS

Basic Guidelines

1. <u>All passages must be converted into text-only format to be processed</u>. This can be accomplished easily in all word-processing programs. In Microsoft Word, for example, texts can be converted through FILE/SAVE-AS/MSDOS-TEXT. In WordPerfect, it can be done via FILE/SAVE-AS/ASCII-TEXT.
2. <u>A new or existing Project File must be opened (see Figure 1) before a text file can be accessed. (See Figure 2)</u>. The Project File can hold any number of DICTION runs and can be re-formatted and printed at the user's discretion.

Figure 1. Opening a project file
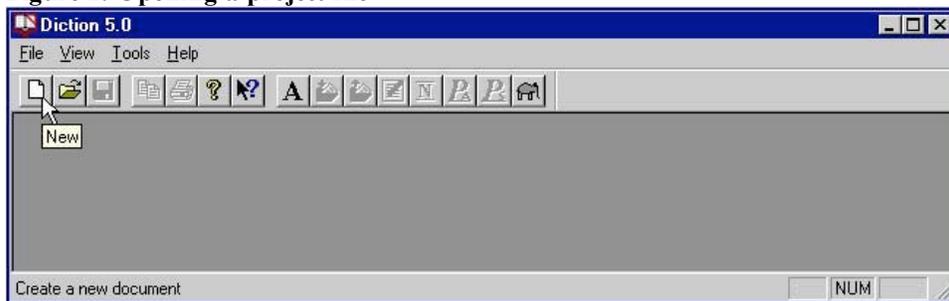
Figure 2. Accessing an input file

3. <u>DICTION will analyze (a) only the first 500 words of a given passage, (b) automatically break up a text into 500-word units, or (c) average the 500-word segments of a passage</u>. This latter opt ion is the program default, thereby allowing passages of any length (up to 500,000 words) to be processed. Short passages can also be processed and the user is given two options for dealing with them. (See Part 7 below).
4. <u>Custom dictionaries must be stored in program-defined directories and carry pro gram-defined (**.dic**) file extensions</u>.

Files and Extensions

The term **Input File** refers to any passage the user wishes to analyze. Its file name is its chief mode of identification and is retained in both the **Project File** (DICTION Sprintable output) and the **Numeric File** (the mathematical results of a search). DICTION is case-insensitive and therefore accepts files in either upper- or lower-case.

When loaded into the Program Files directory of the hard drive, DICTION creates eight directories:

- ❖ DICTION\INPUT: Text passages can be stored in this directory or in any other directory of the user's choosing. Subdirectories can be added to this path as well. Input Files must be in text-only format and can carry any extension (the standard .txt extension is the program's default).
- ❖ DICTION\INPUT\SAMPLES: Includes ten sample passages for familiarizing the user with the program's features.

- ❖ DICTION\OUTPUT: Project Files can be stored in DICTION\OUTPUT or in any directory of the user's choosing. Project Files are assigned a (**.dws**) extension but can be changed at the user's discretion.

- ❖ DICTION\OUTPUT\SAMPLES: Location for storing results from processing the sample documents.

- ❖ DICTION\DATA: The statistical data created by a text-se arch will be added to the bottom of a Numeric File stored within this directory. All Numeric Files must carry a **.num** extension (as in DICTION\DATA\ STUDY#1.NUM).

- ❖ DICTION\EXTRA: All Custom Dictionaries must be stored in this directory. A Custom Dictionary is a word-list (written in text-only or ASCII format) created by the user, a list containing 200 words or less. Ten such dictionaries may be used at any given time. All Custom Dictionaries must carry a .dic extension (as in DICTION\EXTRA\FUNWORDS.DIC).

- ❖ DICTION\STATS: This directory contains a handy utility for converting DICTION's numerical output to a format compatible with SPSS-for-Windows, a popular tool for doing statistical analysis. The conversion file is called SPSS-DIC.SPS. Users uninterested in doing formal statistical analysis can ignore this utility since it has no effect on DICTION's operations.

- ❖ DICTION\INTERNAL: This directory stores files used in Diction's technical operations (specifically for its Elephant files. See Part 12 below).

PART 4: THE MAIN SCREEN

The Main Screen is divided into three distinct areas of (see Figure 3):

1. **File View Window Pane** shows the Input Files in queue for processing by DICTION
2. **Data View Window Pane** contains the verbal report produced when a passage has been analyzed by DICTION
3. **Numeric View Window Pane** shows the mathematical results of a DICTION run (i.e., the Numeric File)

Figure 3. DICTION's main screen
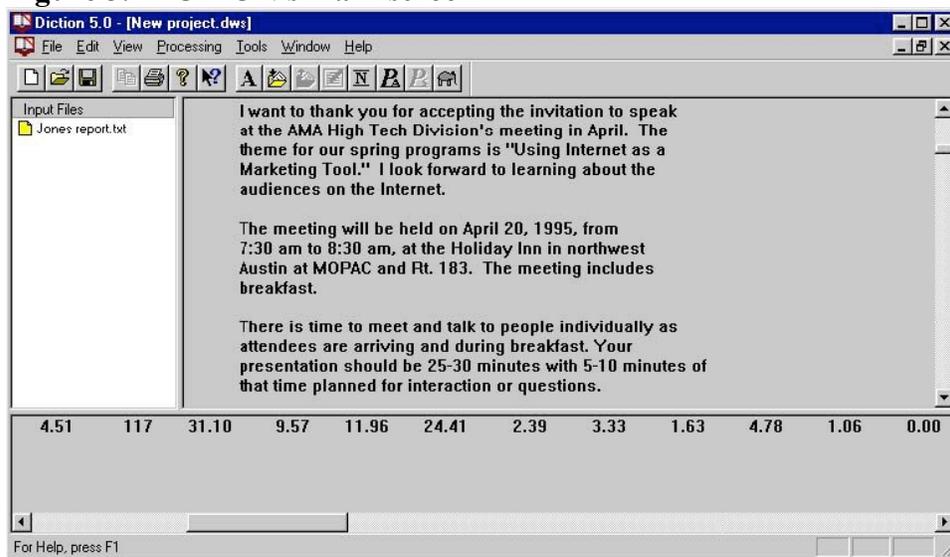
In addition, the Main Screen has the usual Windows functionalities (see Figure 3):

4. **Caption bar** displays the name of the current Project File
5. **Menu bar** the main resources for DICTION processing: opening, closing and printing Project or Input Files, features to include or exclude from the Project report, text processing options, and the usual Tools, Windows and Help features
6. **Tool bar** icons for some of the above operations, as well as buttons for context help, property changes, and font choices
7. **Status bar** reports the DICTION program's operations as they occur
8. **Scroll Bars** permits vertical and horizontal movement within the Main Screen
9. **Resizing border** lets the user adjust the overall size of the Main Screen

PART 5: CREATING PROJECT FILES

Before a text can be accessed by Diction, a new Project File must be created (File/new) or an existing Project File must be called up (File/Open) (See Figures 1 and 2 above.). The Project File holds the results of one or more DICTION runs.

Project File Features

Project files can be treated like any other Windows report file. They are assigned a **.dws** extension by default, but the user can change the extension at will. Also by default, Project files are stored in DICTION\OUTPUT but, again, the user can place them wherever desired. To change the default area for Project Files, do the following:

1.   Go to File/Properties or press the File Properties button on the Tool Bar.
2.   On the Output Tab, go to Default Output Directory and choose Browse
3.   Select the file path desired.
4.   Press OK to exit the File Properties screen.

Project File Components:

Once a Project File has been opened and a text processed, DICTION results are read into the Data View Window Pane. The Project File is divided into ten main areas:

1.   **Identification area** contains general language statistics for the Input File as well as two methods of identifying the text passage. (See Figure 4)

Figure 4. Project file: identification/options area
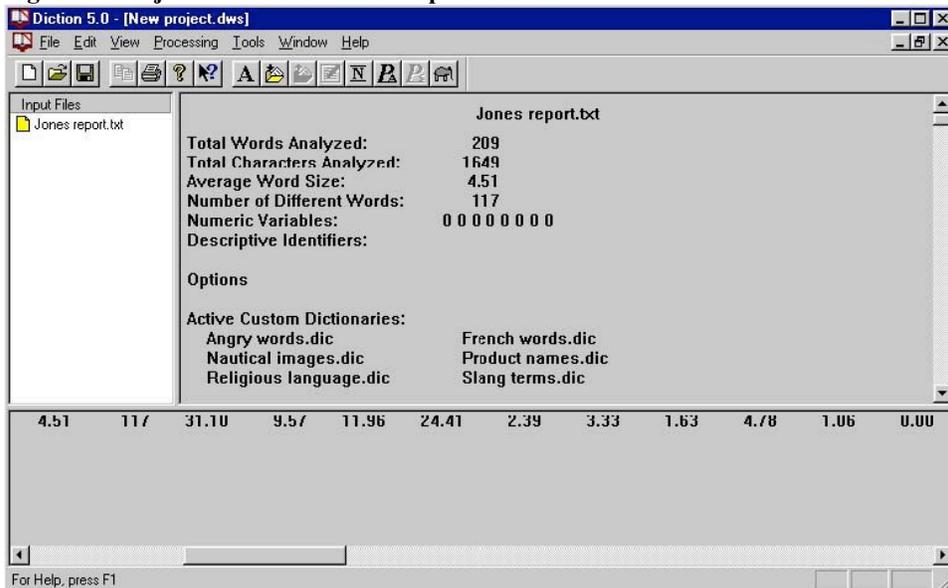
2.	**Options area: Custom dictionaries** indicate which user-created dictionaries, if any, have been used. (See Figure 4)

3.   **Options area: Processing instructions** reports which user-controlled options have been invoked, including file display characteristics, methods of handling large and small files, and print options. (See Figure 5)

Figure 5. Project file: processing/character area
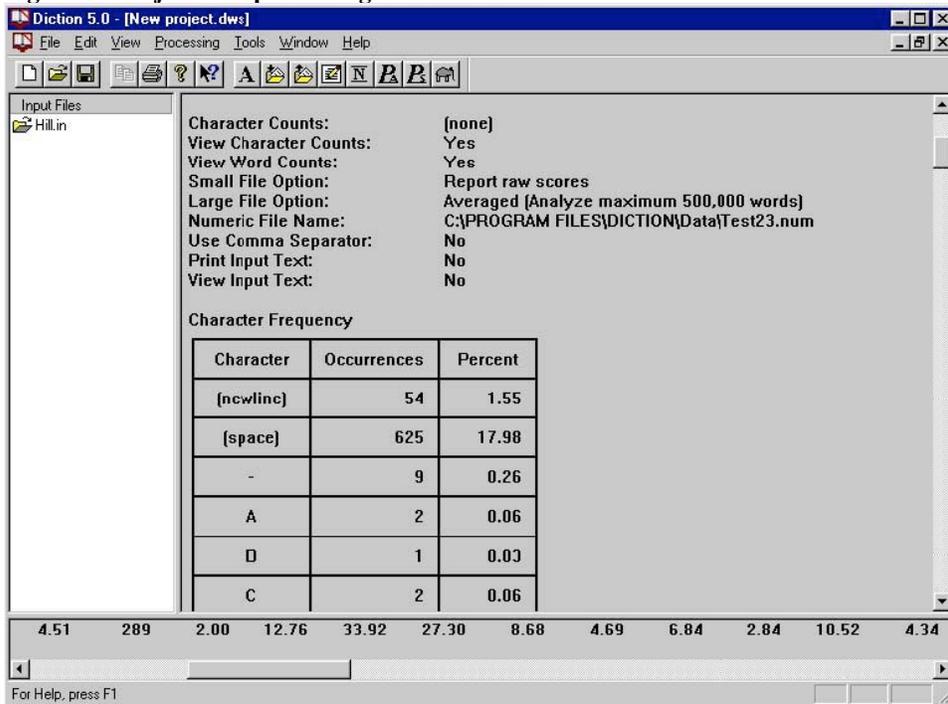
4.   **Character Frequency** lists all alphabetic and orthographic characters (such as brackets, dollar signs, ampersands, etc). These latter markers are often used to pre-treat a text, thereby becoming designators of soft language features such as metaphors, emotional appeals, etc. (See Figure 5.)

5. **High Frequency Words** occur three or more times in a standard (500-word) passage. These words are not employed in any of the program's calculations but are helpful in getting the general flavor of a text. (See Figure 6.)

Figure 6. Project file: high frequency area

6. **Standard Dictionary Totals** includes raw frequencies, standardized scores, and normative standards for DICTION's thirty-one word-lists. Special note is also made of scores lying outside the norm (+1/-1 standard deviation from the mean for the normative grouping chosen). (See Figure 7)

Figure 7. Project file: standard dictionary area

7. **Custom Dictionary Totals** raw frequencies for any special word lists the user has prepared. DICTION permits use of ten such dictionaries, with up to 200 search words in each word list. (See Figure 8).

Figure 8. Project file: custom /insistence area
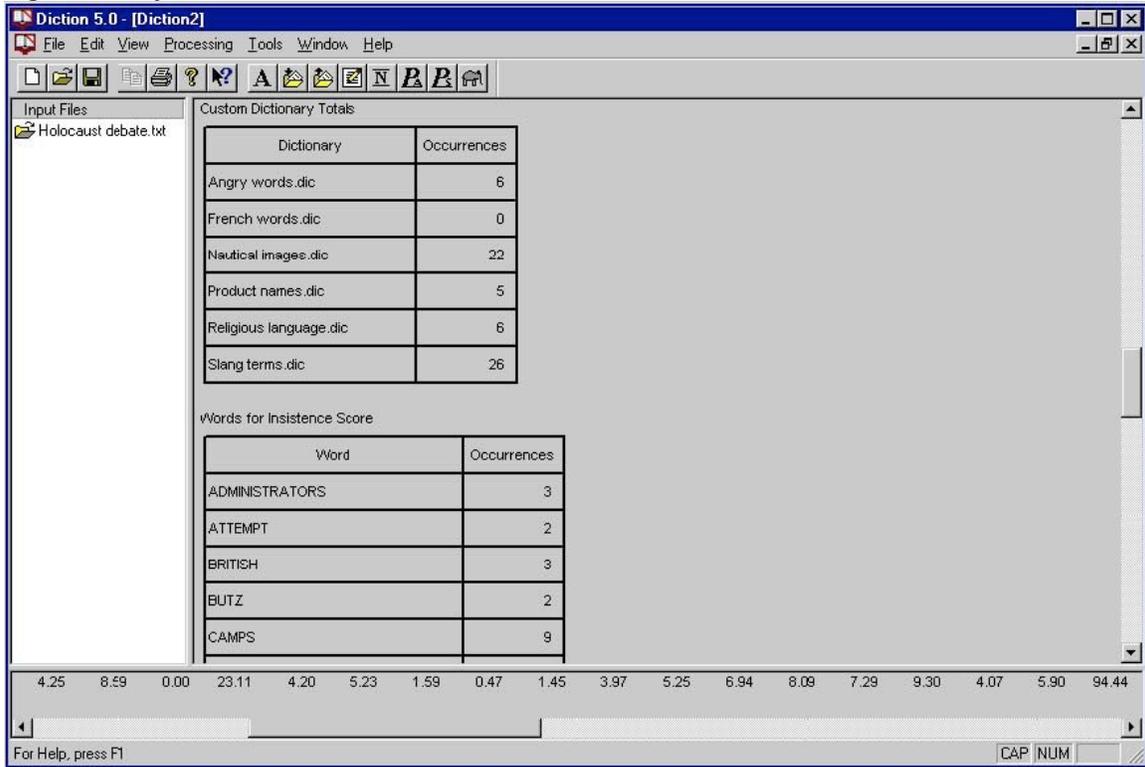
8. **Words for Insistence Score.** The Insistence Score calculates a text's dependence on a limited number of often-repeated words. In calculating this score, DICTION singles out all words used three or more times in a 500-word text. (See Figure 8)

9. **Calculated Variables** four of DICTION's scores result from calculations rather than dictionary matches. These include Insistence (a measure of code-restriction), Embellishment (the ratio of descriptive to functional words), Variety (a measure of linguistic dispersion), and Complexity (word size). (See Figure 9).

Figure 9. Project file: calculated/master variable area

Diction 5.0 - [New project.dws]

File Edit View Processing Tools Window Help

Input Files
Patton.in

Calculated Variables

Variable	Frequency	Normal Range Low	High	Standard Score	Out of Range
Insistence	11.60	9.15	111.15	-0.95	
Embellishment	0.36	0.18	1.10	-0.59	
Variety	0.71	0.45	0.53	5.57	*
Complexity	4.06	4.31	5.01	-1.85	*

Master Variables

Variable	Score	Normal Range Low	High	Out of Range
Activity	58.74	46.74	55.48	*
Optimism	49.53	46.37	52.25	
Certainty	48.53	46.90	51.96	
Realism	52.27	46.10	52.62	
Commonality	48.66	46.86	52.28	

| 4.06 | 257 | 0.00 | 16.58 | 30.00 | 50.56 | 31.50 | 5.50 | 3.10 | 15.50 | 2.69 | 2.00 |

For Help, press F1

10.   **Master Variables** DICTION's five overall measures--Activity, Optimism, Certainty, Realism and Commonality are composed by standardizing all previous scores, combining them via addition and subtraction, and then by adding a constant of 50 (to eliminate negative numbers). When taken together, these five measures provide the most general understanding of a given text. (See Figure 9).

Project File Options

The user may customize a Project Files in several ways:

❖ The text of the input file may be included or excluded from the report. The default option is to exclude the text.
❖ Character counts can be suppressed or reproduced. The default is to suppress.
❖ High frequency words (those occurring 3 or more times) may be suppressed or reproduced. The default is to suppress.

To make these choices do the following:
1. Go to the Tool Bar and choose View
2. Click (or unclick) Word Counts, Character Counts, and/or View Input Text Other

Project Files options include:

1.   **Font Selection**: Go to Tools/Select Font on the Tool Bar.

2. **Print Preview**: Go to File/Print Preview to see what a Project File will look like when printed.
3. **Printing**: Go to File/Print (or press the printer icon on the Tool Bar) and then follow normal Windows routines.

Saving the Project File

A project file can be saved in any directory of the user's choosing (the default is DICTION\OUTPUT). DICTION can also be prompted to call up the last opened Project whenever the pro gram is start ed, thereby allowing new items to be appended to the same Project across several work sessions. To enact this option do the following:

1. Go to Tools/ System Options
2. Check the box designated Open Last Document on Startup.
3. Click OK to exit.

PART 6: SEARCHING A TEXT

Preparing a Text

As mentioned previously, DICTION requires that all texts be converted to text-only or ASCII format before they can be processed. If the user fails to make the conversion, DICTION will report an error statement (as in Figure 10).

Figure 10. Error report for ineligible file
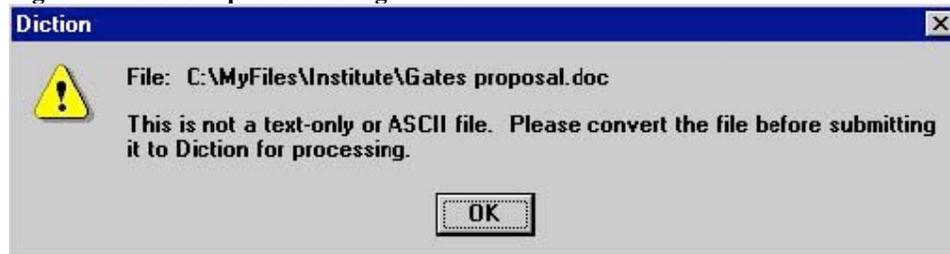
Identifying a Text

To help with later statistical analysis, the u ser may add an **Alpha-Numeric Identifier** at the top of an input file. (See, for example, Figure 11). Once a search is completed, the Alpha-Numeric Identifier will be the first piece of data to appear in the Numeric File. Such an identifier must be structured in the following manner:

1. Do not leave blank lines at the top of the file.
2. A single pound sign must precede the Alpha-Numeric identifier.
3. Up to eight fields of numbers/letters may then be added.
4. Each field of identifiers must be separated by a space o r tab.

DICTION 5.0 Manual

5.   A single pound sign must follow the last field of identifiers.
6.   If no identifier is created, DICTION will report 8 fields of zeroes for that text in the Numeric File.

Figure 11. Sample text file with identification markers
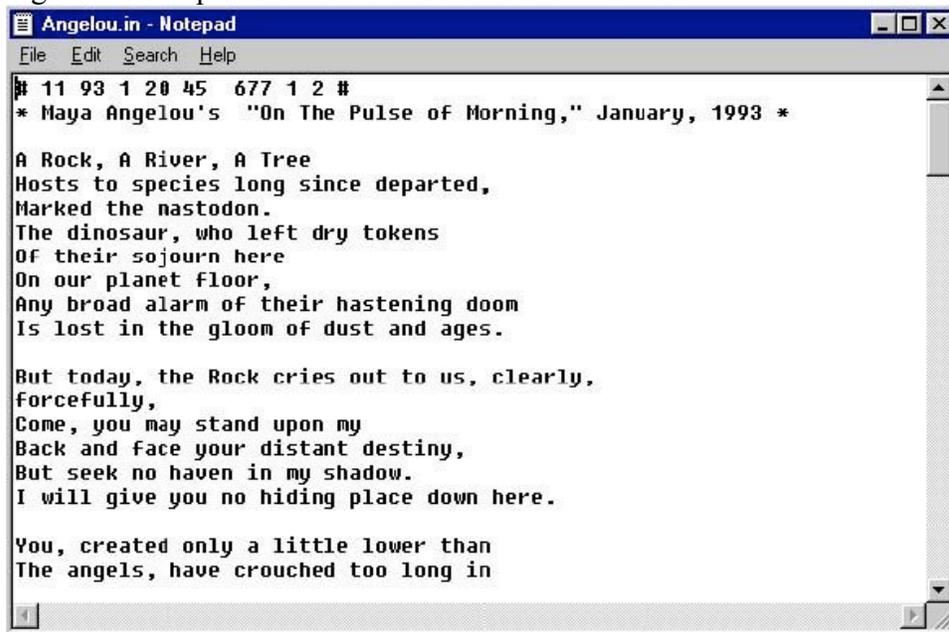
The user may also use a **Descriptive Identifier** to mark a passage (se Figure 11). This identifier will not be included in the words processed by DICTION and will not be appended to the Numeric File. To prepare a Descriptive Identifier follow these steps:

1. The Descriptive Identifier must app ear immediately after the Alpha-Numeric Identifier (if one is used).
2. The Descriptive Identifier must app ear immediately before the passage being searched.
3. The Descriptive Identifier must be preceded and followed by an asterisk (*)

Selecting a Text

To select a text for processing, follow these steps:

1. Make sure you have already opened a Project File.
2. Press the Add File button on the Tool Bar or use Edit/Add File(s).
3. This will produce a standard Windows Open Files box listing the files in the directory you have chosen (Figure 12).

Figure 12 Selecting input files
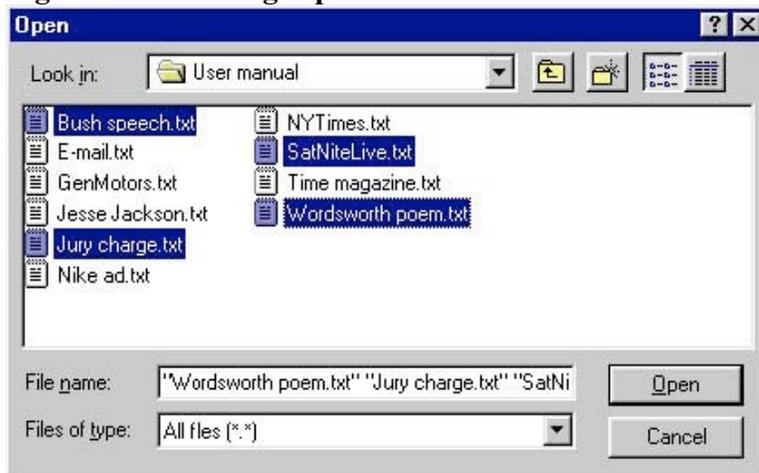
4.   Single-click all or so me of the text passages; then press Open.
5.   The selected files will now appear in the File View Window Pane of the Main Screen.

Processing a Text

1.   To process one file or several alternating files, highlight each and then press the Process Selected button on the Tool Bar.
2.   To process all files at once, press the Process All button on the Tool Bar.

Viewing a Text

Users may find it helpful to view a passage after a search. To bring the text into the Project File, go to View and click Input Text.

Editing a Text

After running a text through DICTION, a user (1) may discover typographical errors in the Input File that somehow comprised the results or (2) may wish to make additions or deletions to the text and then re-process it to see if DICTION's results are affected. (This utility is particularly helpful to those using DICTION to improve writing and editing skills.) To make changes in texts, do the following:

1.   Highlight the desired text file in the File View Window Pane.
2.   Go to Edit/Edit File or press the edit button on the Tool Bar.
3.   Do the editing and then go to File/Close.
4.   This will bring you back t o the main screen where you can run the file again.

Note: If you are running Windows NT, you must re-designate the editor by going to Tools/System Options and changing the Input File Viewer to c:\Winnt\notepad.ese.

PART 7: WORKING WITH INPUT FILES

DICTION has been designed as a general-purpose program suitable for analyzing any sort of English-language text. Because users' projects vary greatly, however, a number of special problems have been anticipated.

Small Input Files

To ensure that DICTION scores mean the same thing to all users at all times regardless of the size of the file being processed, the program has set a standard of 500 words as the textual norm. This norm was also imposed because several of DICTION's forty scores are sensitive to file size.

To handle files smaller than 500 words, the user can choose among two options:

Option One: Report Extrapolations: This opt ion, the default, makes corrective counts of a small file, thereby standardizing it to a 500-word basis and permitting use of the comparative data contained in the DICTION program. To choose this option, do the following:

1. Go to File/Properties or press the File Properties button on the Tool Bar.
2. Click on the Processing tab.
3. Choose Report Extrapolations. (See Figure 13)
4. Click OK and ret urn to the Main Screen.

Figure 13. Changing processing properties in DICTION
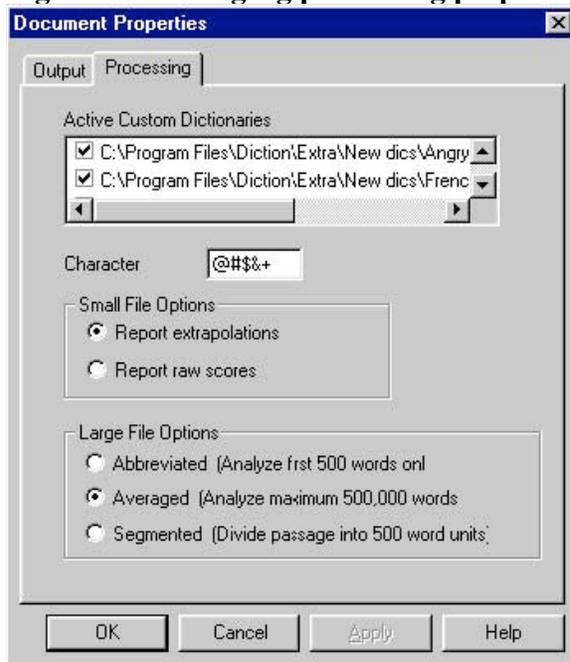
Option Two: Report Raw Scores: This option reports exact numerical counts and ignores the 500-word standard. When selecting this option, however, the user is cautioned that it will make DICTION's norms useless for comparative purposes.

1.   1. Go to File/Properties or press the File Properties button on the Tool Bar.
2.   2. Click o n the Processing tab.
3.   3. Choose Report Raw Scores. (See Figure 13.)
4.   4. Click OK and return to the Main Screen.

Large Input Files

Users wishing to process longer files have three options:

1. <u>Analyze first 500 words only</u>. This option automatically cuts-off a passage at 500 words, leaving the remainder of the text unanalyzed. It generates only one report and adds only one line of data to the Numeric File. To choose this option, do the following:

    (a) Go to File/Properties or press the File Properties button on the Tool Bar.
    (b) Click on the Processing tab.
    (c) Choose Abbreviated. (See Figure 13)
    (d) Click OK and ret urn to the Main Screen.

2. <u>Average all 500-word units</u>. This option generates one set of scores for the entire passage (regardless of length) by averaging its 500-word units together. Texts as large as 500,000 words can be processed with this option and therefore it is the pro gram default. To choose this option, do the following:

    (a) Go to File/Properties or press the File Properties button on the Tool Bar.
    (b) Click on the Processing tab.
    (c) Choose Averaged. (See Figure 13)
    (d) Click OK and ret urn to the Main Screen.
    (e) Segment passage into 500-word units.

This option breaks a text into 500-word units and processes each separately, hence exposing a passage s development from beginning to end. When doing so, DICTION creates **Offspring files** from the **Parent file** and then adds the offspring files to DI CTION\INPUT (or their originating directory). The offspring files retain the same alpha-numeric and descriptive identifiers found in the parent file. In addition, they retain the same filename with one exception: the filename is given a second extension (e.g., DICTION\OUTPUT\SAMPLE.txt.1… SAMPLE.txt.2,… etc.). Note, for example, how the long file selected for analysis in Figure 14 is sub-divided into six separate files when introduced into the Main Screen for DICTION processing.

Figure 14. Large file as segmented by DICTION
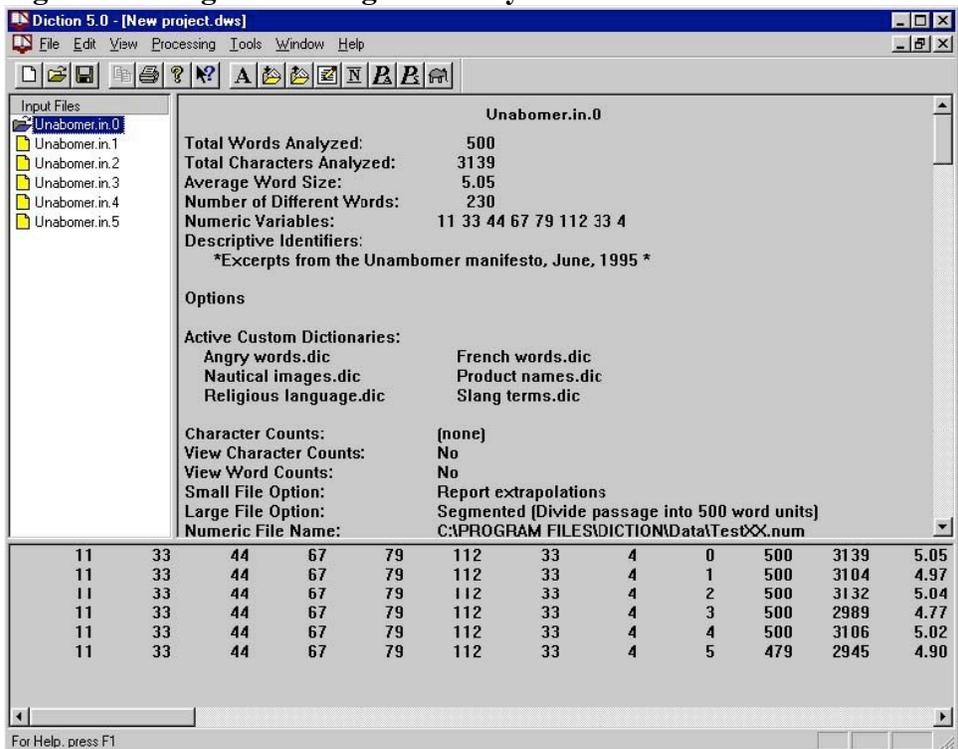
This option introduces a separate report for each segment into the Project File and adds a unique line of data to the Numeric File for each segment. In addition, the ninth field of data (the Segment Number) is incremented by one in the Numeric File. To choose this option, do the following:

1.   Go to File/Properties or press the File Properties button on the Tool Bar.
2.   Click on the Processing tab.
3.   Choose Segmented. (See Figure 13)
4.   Click OK and ret urn to the Main Screen.

System Changes for File Options

The menu of options laid out above is Project-specific. That is, changes mad e in the treatment of large and small files will lapse once the program is exited. One can offset that effect by changing the program s defaults (i.e., DICTION's state when initialized). To do so, follow these steps:

1.   Go to the Processing tab of Tools/System Options.
2.   Make the desired permanent settings for large and small files.
3.   Click OK to exit.

Processing Special Characters

At the user's request, a Project File will include data on all orthographic and numeric characters in a passage. These data are not, however, included in the Numeric File with one exception: The user can request that counts for any five characters be included in the Numeric File in the following manner:

1.   Go to File/Properties or press the File Properties button on the Tool Bar.
2.   Click on the Processing tab. (See Figure 13)
3.   Type any five characters in the box provided. Commas or other delimiters need not be used. *For example*: @#$&+
4.   Click OK and return to the Main Screen.

The user can use such special characters to pre-treat a text and thereby gather quantitative data on contextual phenomena. Metaphors, for example, could be tagged with an ampersand o r colloquial phrases could be demarcated by bracket s. Subsequent reports of these two character counts would therefore have special meaning for the researcher.

System Changes for Special Characters

Changes such as these lapse once the Project is exited. If the user wants the above characters to be searched each time the program is initialized, do the following:

1.   Go to the Processing tab of Tools/System Options.

2. Type any five characters in the box provided.
3. Click OK to exit.

PART 8: CHOOSING NORMATIVE PROFILES

A distinctive feature of DICTION 5.0 is that it comes equipped with a variety of norms the user may employ for getting fine-grained understanding of a given text or body of texts. These normative data have been generated by running more than 20,000 texts through DICTION. These texts range from public speeches to poetry, from newspaper editorials to music lyrics, from business reports and scientific documents to television scripts and informal telephone conversations. All texts were produced in the United States between 1945 and 1998.

Selecting Norms

The default in DICTION is to combine all of these data into a single normative profile so the user can see what a passage looks like in general. Some researchers will want greater specificity. Accordingly, they can select from among thirty-six different sets of norms. To do so, follow these steps:

1. Go to File/Properties/Output or press the File Properties button on the Tool Bar.
2. In Default Normative Values press Change. (see Figure 15, Step 1)

3. To choose the entire assemblage of norms choose All under Class. (See Figure 15, Step 2)
4. To choose a more specific set of norms, make a general selection under Class and then a more specific selection under Type. (see Figure 15, Step 3)

Press OK to leave the Normative Values screen and press OK again to leave the Document Proper ties screen.

Figure 15 Changing norms

Step 1   Step 2   Step 3
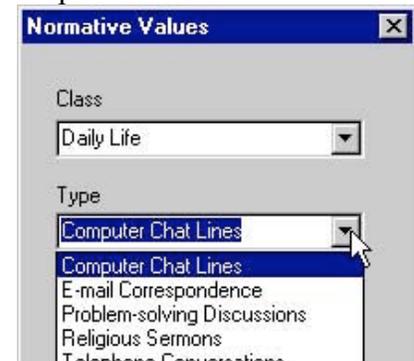
Available Norms

The All Texts option (n = 22,027) consists of all texts processed thus far by the author. These overall norms are divided into six Classes Business, Daily Life, Entertainment, Journalism, Literature, Politics, and Scholarship and the classes are further subdivided into thirty-six distinct Types.

Users can compare a given passage to as many of these sub-categories as desired. So, for example, a political commercial could be viewed from the vantage point of political commentary and, later, from the standpoint of product advertising. To accomplish this, the user would process the passage with one set of norms and then re-process it with a second set. The classes and types of norms built into the DICTION program include the following:

Business

CORPORATE FINANCIAL REPORTS (n = 4 8). A sampling of annual financial reports from a variety of fortune 500 companies, including 3M, Ford, Merk, Dynatech, etc. Reports were collected electronically from such internet sites as Annual Reports Library, Index: Annual Report Gallery, and Barron's Annual Report and Earnings Service.

CORPORATE PUBLIC RELATIONS (n = 1 63). A broad-based collection of official mission statements, public pronouncements, and C.E.O. speeches in behalf of major American corporations from the 1960s through the mid-1990s. Includes manufacturing companies (e.g., Boise-Cascade), mining and construction (e.g., Flour Daniel), transportation and telecommunications (e.g., A.T.&T.), as well as, financial and service-based industries (e.g., Federated Department stores, H&R Block, etc.).

FINANCIAL NEWS (n = 122). A variety of news stories related to financial issues (e.g., tax returns, market predictions, trends in stocks and bonds, tax law, speculation on specific annuities, etc.) obtained from the on-line publication s of Forbes, The San Francisco Chronicle, the Daily News Bulletin, etc.

LEGAL DOCUMENTS (n= 79). An assemblage of argument summaries extracted from county, appellate and Supreme Court briefs obtainable online through Lexis-Nexis. The court cases addressed environmental issues, liability law, child support, tax laws, military tribunals, voter qualifications, and constitutional rights.

MAGAZINE ADVERTISING (n= 9 4). A wide assortment of print advertisements from *Scientific American, Money, Sunset, Discover, Business Week, Vanity Fair, McCall's, MacWorld, Bicycling, Ladies Home Journal*, etc. Products and services included computer hardware, security systems, insurance services, financial investments, beauty aids, travel agencies, automobiles, snack foods, pharmaceuticals, kitchen appliances, stereo systems, and cat food.

TV ADVERTISING (n = 109). A collection of CLIO award-winning commercials from 1984, 1987, 1989, and 1996. The ads feature well-known companies and products (e.g., Nike, Pepsi, Reebok, Levi-Strauss, Apple Computers, Motel 6, Michelob, and McDonalds), sports teams and sporting events (e.g., football, baseball, golf, and hockey), and community and social causes (e.g., World Hunger, anti-tobacco campaigns, anti-drug campaigns, church fund-raisers, Mothers Against Drunk Driving).

Daily Life

COMPUTER CHAT LINES (n = 68). An eclectic group of on-line discussions dealing with such topics as tropical fish, dreams and life choices, parenting advice, family relationships, contemporary politics, the Dallas Cowboys football team, computer hardware and software, and horse racing. The discussions were obtained via the World Wide Web's Chat Archives.

E-MAIL CORRESPONDENCE (n = 75). A selection of miscellaneous messages from over a dozen individuals in eight different work settings, including public relations, elementary education, a state university, county government office, a chiropractor s office, and a car dealership. The messages range from personal notes of encouragement and gossip to formal work deliberations, business memos, and project updates.

PROBLEM-SOLVING DISCUSSION (n = 73). Group decision-making from three settings: a state agency, a school district teacher's caucus, and student peer-mediation. Discussion topics include state contracts, project requirements, financial planning, and salary negotiation. Transcripts were generated during research studies conducted by Dr. Linda Putnam (Texas A&M University), Dr. Craig Scott (University of Texas at Austin), and Dr. Trisha Jones (Temple University).

RELIGIOUS SERMONS (n = 198). Sermons delivered by a wide variety of denominational representatives in the United States between 1935 and 1996. Topics include biblical exegesis, doctrinal disputes, ritualistic remembrances, and general moral and social discussions. Mainline denominations (e.g., Episcopalians, Catholics, and Methodists) as well as an assortment of cults and sects are represented.

TELEPHONE CONVERSATIONS (n= 58). A collection of snippets from everyday telephone conversations conducted in Austin, Texas during the 1990s. Conversations were recorded and transcribed by the language and culture division of the Department of Communication Studies at the University of Texas at Austin under the direction of Professor Robert Hopper.

Entertainment

CELEBRITY NEWS (n = 125). A variety of celebrity profiles (e.g., Sean Penn, Tom Cruise, Madonna, Farrah Fawcett, and Antonio Banderas) taken from on-line versions of Entertainment Weekly, People Magazine, and TV Guide.

ENTERTAINMENT REVIEWS (n = 6 0). An assemblage of book reviews (e.g., *Range of Motion, XY*: *On Masculine Identity, Breaking The Glass Ceilin*g), film reviews (e.g., *Lawrence of Arabia, Nothing to Lose, Crash*), and music reviews (e.g. , *The Archers of Loaf, Soul Coughing, Bruce Cockburn*). Texts were obtained from such web sites as The Music, Film, and Video Review and The Pelican Library.

MUSIC LYRICS (n = 128). A diverse assortment of lyrics from a myriad of musical genres and recording artists, from Barbra Streisand to Jimi Hendrix, from Woody Guthrie to Silverchain, from Bing Crosby to Easy E., from Nat King Cole to the Violent Femmes, from Glen Campbell to Dismemeber). Lyrics were obtained from such web sites as The International Lyrics Server, Grendel's Lyrics Archive and Lost Souls Domain.

SPORTS NEWS (n = 1 39). A collect ion of articles on athletes and athletic events obtained from the on-line versions of Sports Illustrated and the Sporting News. Sports profiles focus on such athletes as Nancy Kerrigan , Arthur Ashe, Joe Montana, John El way, Larry Bird, Tiger Woods, and Venus Williams among others. Athletic events described include basketball, football, baseball and ice hockey.

TV COMEDIES (n = 75). An assortment of prime-time television comedies including *Seinfeld, Friends, Married with Children, Caroline in the City*, and *Cybil*. Characters in these shows comment on urban life, contemporary friendships, the tribulation s of dating, dysfunctional family life, and amusing situations at work. Texts were obtained from a variety of on-line services and from unofficial web-sites.

TV DRAMAS (n = 159). An assortment of prime-ti me television dramas including *My So Called Life, Xena: Warrior Princess, Buffy the Vampire Slayer, Twin Peaks*, and *The Wonder Years*. The shows explore teenage angst, good versus evil, justice and the law, growing up in the 1960s, etc. Texts were obtained from archives included in Drew's Television Transcripts.

Journalism

LETTERS-TO-THE-EDITOR (n = 6126). Letters written on general social and political matters in 12 small-city newspapers between 1948 and the present. Topics include civil rights, inflation, abortion and birth control, taxation, moral and political values, international affairs, etc. Among the newspapers are the *Fall River Herald News*, the *Salinas Californian*, the *Trent on Times*, the *Billing's Gazette*, etc.

NEWSPAPER EDITORIALS (n = 65). An assortment of newspaper editorials from diverse geographical locations (e.g., *Arkansas Democrat-Gazette, Los Angeles Downtown News, Charlotte Sun Herald, South Idaho Press*, etc.). Topics include sexual harassment, infant mortality, casino gambling, drug enforcement and other public policy matters. All editorials were written in the 1990s and obtained from on-line versions of these papers.

POLITICAL REPORTING (n = 73 09). Newspaper coverage of major issues of the day, especially political matters. Stories were extracted from *The New York Times, Washington Post, Christian Science Monitor, Los Angeles Times, Chicago Tribune, Atlanta Constitution*, and AP and UPI wire coverage. All stories were produced between 1948 and 1996.

TECHNOLOGY NEWS (n = 68). A file of news stories about contemporary trends in the computer industry, including hardware and software updates, forthcoming products and services, and other topics related to new technologies. The articles were obtained from online issues of the *Wall Street Journal*, the *Chronicle of Higher Education*, *ZDNe*t, *CNet*, and *Computer News Daily*.

TV NEWS (n = 1219). Transcripts from the nightly news programs of NBC, CBS, ABC, and PBS. News coverage focused on a variety of matters but political/campaign coverage was especially emphasized. All texts were produced during the 1980, 1988, 1992, and 1996 presidential campaigns.

Literature

NOVELS & SHORT STORIES (n = 151). A compilation of fiction from a diverse collection of seminal writers (e.g., Nathaniel Hawthorne, Stephen Crane, J. D. Salinger, James Fenimore Cooper, Lewis Carroll, Kate Chopin, and Francis Bacon). All excerpts were taken from such online sites as the Internet Classics Archive and Book Lover's Resource Page.

POETRY & VERSE (n = 78). An assemblage of English-language poetry including canonical works from the masters (e.g., Tennyson, Whitman, Poe) as well as more contemporary poetry from less well-known artists. Poems were downloaded from sites like American Verse Project, Collection of British Poetry, Digital Library of World Classic Poetry, and *The Atlantic Monthly*.

THEATER SCRIPTS (n = 1 51). A diverse selection of plays (comedy, drama, and mystery) by famous writers from antiquity (e.g., Aristophanes and Aeschylus), from the eighteenth and nineteenth centuries (e.g., Shakespeare, Goethe, and Ibsen), as well as more obscure playwrights from the twentieth century (Agostino, Knag, and Folmer). Texts were obtained electronically from such sites as the Playwrights' Resources Page.

Politics

CAMPAIGN SPEECHES (n = 2357). Speeches delivered by Democratic, Republican, and third-party presidential candidates between 1948 and 1996. Major party candidates were represented equally in the sample and third-party candidates were represented proportional to the vote. Included are stump speeches delivered in local and regional settings as well as nationally broadcast campaign addresses.

POLITICAL ADVERTISING (n = 553). Campaign spots presented on television during the presidential campaigns of 1960 and 1976--1996. Republican, Democratic, and third-party ads are represented in proportion to their prevalence in political campaigns. Promotional, counterattack and negative ads are equally represented in the sample.

POLITICAL DEBATES (n = 652). Includes the entirety of all presidential debates from 1960 through 1996. Debate transcripts were segmented-by-speaker and analyzed separately.

PUBLIC POLICY SPEECHES (n = 615). A broad sampling of speeches delivered by sitting presidents from Harry Truman through Bill Clinton. Included are nationally televised addresses on major policy issues as well as ceremonial addresses, remarks made during awards ceremonies, and radio briefings. Topics include the Korean Conflict, open-housing and job discrimination, the Soviet Union, national economic policy, Vietnam , Watergate, etc.

SOCIAL MOVEMENT SPEE CHES (n = 129). Addresses delivered to marginalized as well as mainstream groups about pressing social matters. Included are speeches by Malcolm X, Andrea Dworkin, Ralph Nader, Paul Ehrlich, etc. Topics include feminism, environmentalism, civil right s, labor grievances, nuclear disarmament, etc.

Scholarship

HUMANITIES SCHOLARSHIP (n = 53). A diverse selection of essays from various periodicals devoted to the humanities. Periodicals include Twentieth Century Literature, British Journal of Aesthetics, American Poetry Review, Monthly Review, Skeptical Inquirer, American Indian Quarterly, and Salmagundi. Texts extracted from electronic sites such as Humanities Research and the Literary Review.

PHILOSOPHICAL ESSAYS (n = 197). A miscellany of philosophical texts, spanning the past two hundred years. Texts include Berkeley's. The Principles of Human Knowledge," Hume s An Inquiry Concerning Human Understanding," Delueze's Postscript of the Societies of Control," Russell s A Man's Free Worship," Hegel s Science of Logic," Dewey s Democracy and Education ," and Baudrillard's Radical Thought. Texts were drawn from Humanities Research, Great Thinkers and Visionaries, and Internet Philosophy Resources.

SCIENCE WRITING (n = 177). A compilation of popular science articles collected from the on-line editions of Discover Magazine and Scientific American. Topics covered include the mechanics of breathing, quantum computers, muscular and metabolic fitness, microorganisms, and gamma-rays.

SOCIAL SCIENCE SCHOLARSHIP (n = 52). A sampling of journal articles in anthropology, sociology, psychology, history, communication, etc. Issues include the social aspects of leisure, fuzzy set theory, American cultural studies, exit-poll experiments, mass media effects on violent behavior, and public policy convergence theory. Texts extracted from web sites like Social Sciences Internet Resources and Arts Humanities, and Social Sciences Resources.

STUDENT ESSAYS (n = 307). An assortment of undergraduate essays from nine college English courses. The essays address racial tensions in school settings, the need for diversity in universities, preferential treatment in college admissions, rhetorical criticism of advertisement s, the homeless in America, technology and mankind, and historical heroes. All texts were made available by the University of Texas at Austin' Program in Rhetoric and Composition.

PART 9: USING NUMERIC FILES

Numeric File Features

All arithmetic results produced by DICTION are stored in Numeric Files. By default, these are stored in DICTION\DATA but the user is free to put them anywhere. The only restriction t o Numeric Files is that they must contain the .num extension. The default option that comes with the program is DICTION\DATA\RESEARCH.NUM.

The Numeric File is reproduced in the Numeric View Window at the bottom of the Main Screen. Scrolling from left t o right reveals its entire contents for a given passage and scrolling up and down shows all data-cases contained in the file.

Changing Numeric Files

The user may wish to separate data collected in one study from data collected in another. To change the Numeric File for a given Project, follow these steps:

1.   Go to the Output tab of File/Properties.
2.   Enter the path name of the new file
3.   Click OK to exit.

System Changes for Numeric Files

Changes like these will lapse once the Project is exited. To change the program s defaults (i.e., to bring up the same Numeric File each time the DICTION program is initialized), do the following:

1.   Go to the Output tab of Tools/System Options.
2.   Designate the path name of the file to be made the default.
3.   Click OK to exit.

Variable Separation

Variables in the Numeric File may be separated in two ways: (1) by spaces or (2) by commas. The default option is to separate by commas. To invoke the spaces option, do the following:
1.   Go to the General tab of Tools/System Options.
2.   Unclick Use Comma as Separator
3.   Click OK to exit.

Variable Sequencing

If an Alpha-Numeric Identifier has not been used in an Input File, DICTION reports a series of eight zeroes. It also reports zeroes if Custom Dictionaries or Character Counts have been left undesignated. It sequences these and the remaining data in the following manner: (1)-(8) user-specified I.D. #1 through #8; (9) Segment number (automatically assigned by DICTION); (10) Words/passage; (11) Characters/passage; (12) Characters/word; (13) Different Words; (14) Numerical Terms; (15) Ambivalence; (16) Self-Reference; (17) Tenacity; (18) Leveling; (19) Collectives; (20) Praise; (21) Satisfaction; (22) Inspiration; (23) Blame; (24) Hardship; (25) Aggression; (26) Accomplishment; (27) Communication; (28) Cognitive Terms; (29) Passivity; (30) Spatial Awareness; (31) Familiarity; (32) Temporal Awareness; (33) Present Concern; (34) Human Interest; (35) Concreteness; (36) Past Concern; (37) Centrality; (38) Rapport; (39) Cooperation; (40) Diversity; (41) Exclusion; (4 2) Liberation; (43) Denial; (44) Motion;
(45) Insistence; (46) Embellishment; (47) Variety; (48) Complexity; (49) Activity; (50) Optimism; (51) Certainty; (52) Realism; (53) Commonality; (54)-(58) Character #1 through #5; (59)-(68) Custom Dictionary #1 through #10; (69) Input File Name.

PART 10: USING CUSTOM DICTIONARIES

Creating Custom Dictionaries

A key feature of DICTION is the user's ability to prepare custom-designed search lists. So, for example, a Custom Dictionary entitled 'funwords' might contain such terms as *party, vacation, laughing,* etc. Dictionaries of this sort must be prepared in the following manner:

1. Type individual words directly into a file, leaving a space between each word.
2. Be sure the file is changed to a text-only (ASCII) format.
3. Name the file and save it under the DICT ION\EXTRA directory.
4. Be sure the file has a .dic extension (e.g. FUNWORDS.DIC).

Limits on Custom Dictionaries

All Custom Dictionaries must be stored in the DICTION\ EXTRA directory. A Custom Dictionary may contain no more than 200 words. No more than 10 Custom Dictionaries may be used at any one time.

Temporarily Activating Custom Dictionaries

To search with a set of Custom Dictionaries for the current Project only do the following:

1. Go to Tools/Custom Dictionaries.
2. A User Dictionaries box will appear (Figure 16, Step 1). Choose Add.
3. This produces a standard Windows File Selection window (See Figure 16, Step 2).
4. Highlight the Custom Dictionaries desired and press Open.

Figure 16. Installing custom dictionaries

Step 1   Step 2   Step 3
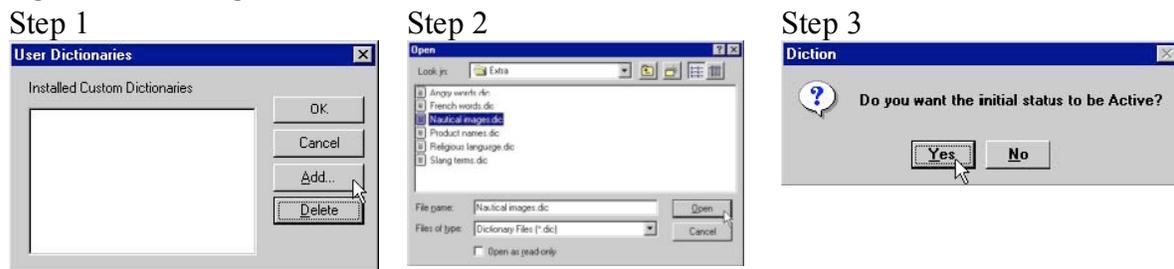
5. A question box will ask if you want the new Custom Dictionary to be activated immediately (See Figure 16, Step 3).
6. If you choose Yes, the dictionary will be employed during the next DICTION search; if you choose No it will merely be added to the menu option in Files/Properties/Processing for later use.
7. In either case, the result is a list of new Custom Dictionaries in the File Selection window (Figure 17) that can be (but do not have to be) accessed in the future.

Figure 17. Installed custom dictionaries
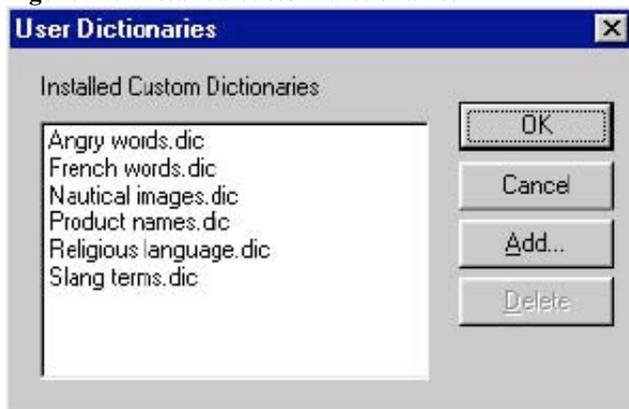
8.   Clicking OK brings you back to the Main Screen.

Temporarily Deactivating Custom Dictionaries

To suspend use of a given Custom Dictionary for the current Project only, do this:

1.   Go to Files/Properties/Processing.
2.   Unclick the dictionary not presently needed.
3.   This particular dictionary will no longer be used in the current Project.
4.   Click OK when finished.

Permanently Selecting Custom Dictionaries

To ensure that certain Custom Dictionaries are invoked each time the DICTION program is initialized, do the following:

1.   Go to the Processing tab of Tools/System Options.
2.   Click or unclick the desired dictionary (see Figure 18).

Figure 18. Permanently selecting custom dictionaries
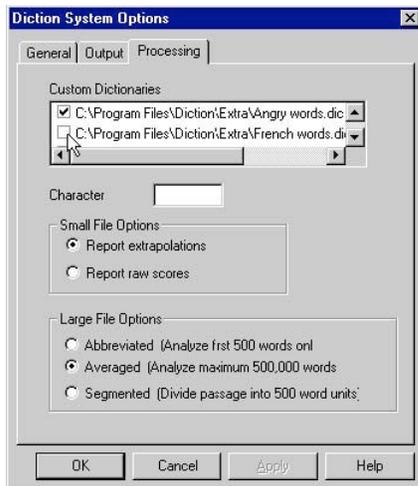
3. This change will remain in effect for all future Projects unless a re-selection is made by the user.
4. Clicking OK brings you back to the Main Screen.

Permanently Removing Custom Dictionaries

To remove a dictionary completely from DICTION, follow these steps:

1. Go to Tools/Custom Dictionaries.
2. Highlight the dictionary no longer needed and choose Delete.
3. Note: This action will not destroy the file; it will just remove it from the DICTION environment
4. Clicking OK brings you back to the Main Screen.

Results of Custom Dictionaries

The results of a Custom Dictionary search are reported in two ways:

The Project File lists raw frequencies and places them immediately after the Standard Dictionary scores.

The Numeric File records Custom Dictionary totals at the end of the file, immediately after Character Counts and immediately before File Name. If no Custom Dictionaries have been selected, DICTION reports a string of ten zeroes.

PART 11: CALCULATING THE INSISTENCE SCORE

Function

The Insistence Score is a measure of code-restriction which calculates a text s dependence on a limited number of often-repeated words. In calculating this score, DICTION singles out all words used three or more times (in a 500-word text). For each word meeting this criterion, DICTION performs the following operations:

- ❖ Consults a list of Exempt Words—function words, verbs, particles, etc.—in the program's Elephant (see below). The Elephant is a device that dynamically learns from the user's actions which words are exempt or eligible and then remembers those actions for future searches.
- ❖ If the high frequency word is already on the exempt list, no decision is made by the user.
- ❖ If the high frequency word is not on the exempt list, DICTION asks the user to determine if the word is a noun or noun-derived adjective (see Figure 19).

Figure 19. Assigning words for insistence score
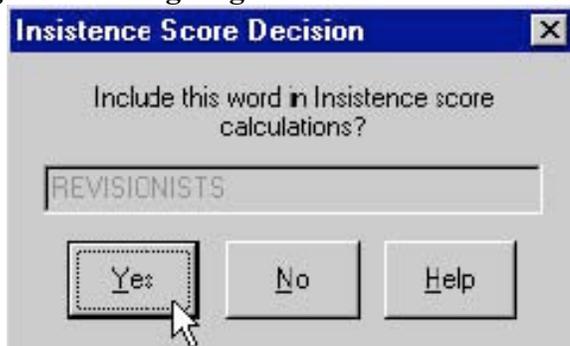
- ❖ If the word is a noun or noun-derived adjective, the Elephant declares the word an Eligible Word, stores that information for future processing, and then adds the new data to the Insistence Score calculation.
- ❖ After a word has been declared Exempt or Eligible, it is no longer presented to the user for judgment.

Eligible Words

A word is defined as an Eligible Word if it meets one of the following criteria:

- ❖ It is a noun (e.g., *prudence*)
- ❖ It is a noun-derived adjective (e.g., *prudent*)
- ❖ It can be used both as a verb and as a noun/noun-derived adjective (e.g., *count assured, perfect, worry, domineering*, etc.)

> **Note**: Treat all abbreviations (e.g., *St., Ave.*), initials (e.g., *George W. Bush*), sums (e.g., *$9.95*), and dates (e.g., *1996*) as whole words.

Exempt Words

An Exempt Word is defined as any word occurring three or more times in a 500-word passage (less than twice in a 350-word or less passage) that is neither a noun nor a noun-derived adjective.

Small Files

DICTION has been developed to process an Input File of 500 words. Because the Insistence Score is particularly sensitive to file size, a special accommodation has been made for this variable when processing smaller files:

- ❖ A passage of 351-500 words is processed normally. This means that any non-exempt word occurring three or more times will be presented to the user to determine eligibility (if it has not already been declared eligible by The Elephant).
- ❖ If a passage is 350 words or less, any non-exempt word occurring two or more times will be presented to the user to determine eligibility (if it has not already been declared eligible by The Elephant).

Optional Signal

Upon request, DICTION will signal the user with a beep when a decision must be made on the Insistence Score. To engage this option, do the following:

1. Go to the General tab of Tools/System Options
2. Click on Enable Sounds (Figure 24).
3. Choose OK to exit.

PART 12: USING THE ELEPHANT

Purpose

The Elephant assists the user by remembering words previously judged Eligible or Exempt. Once such discriminations are made, the Elephant ensures that the user won't have to do so again. DICTION is self-learning to this extent. Approximately 11,000 terms have already been discriminated in these ways and t his knowledge is built into the current version of DICTION. In addition, users will add to the Elephant from time to time during text processing. These supplementary lists can be modified at will by the user.

Adding Words

To add a word to the Exempt Words or Eligible Words lists in the Elephant, do the following:

1. Choose Tools/Modify Elephant on the Menu Bar.
2. Click on the down-arrow to select the appropriate list (Figures 25 and 26).
3. Click on Add. This produces an Edit Elephant Entry box.
4. Type in the appropriate word.
5. Click 'noun or adjective?' box.
6. Choose OK.

Deleting Words

To delete a word from the Exempt Words or Eligible Words list, do the following:

1. Choose Tools/Modify Elephant on the Menu Bar
2. Click on the down-arrow to select the appropriate list (Figures 25 and 26).
3. Click on the word to be removed.
4. Choose Delete.
5. Click OK.

To File Option

This utility is helpful for reviewing the decisions a user has made when building Insistence Scores. To download the user's list of Eligible Words or Exempt Words do the following:

1. Choose Tools/Modify Elephant on the Menu Bar.
2. Choose Modify Elephant.
3. Use the down-arrow to select the desired word list (Figures 25 and 26).
4. Click on To File. This will produce a standard Windows Open File display.
5. Enter the desired file name.
6. Click on OK to exit.

PART 13: DESCRIPTIONS OF SCORES

Overview

DICTION uses thirty-one dictionaries (word-lists) to search a text. In addition, five Master Variables are built by concatenating these dictionary scores. Four Calculated Variables are also used. The dictionaries have the following properties:

❖ They vary considerably in size, ranging from as few as 10 words to as many as 745 words.
❖ The dictionaries contain individual words only (vs. phrases).
❖ No words are duplicated across the thirty-one dictionaries.

DICTION 5.0 Manual   41

❖ Homographs, words that are spelled alike but that have different meanings, are treated via statistical weighting procedures (thereby partially correcting for context). %DICTION's Report Files produce both raw scores and standardized scores for each of the standard dictionaries.

Scanning the Dictionaries

The user may view the standard dictionaries in the following manner:

1. Select Dictionaries from the Main Bar.
2. Click on Scan Dictionaries.
3. Use the upper arrow keys to select a particular dictionary (Figure 27).
4. Use the lower arrow keys to scroll through the corpus of words in that dictionary (Figure 28).
5. To exit the scanning function, click on Done.

Loading the Dictionaries

Dictionaries will be automatically loaded into memory at the beginning of each work session.

THE CERTAINTY SCORE

Definition: Language indicating resoluteness, inflexibility, and completeness and a tendency to speak ex cathedra

Formula: [Tenacity + Leveling + Collectives + Insistence.] - [Numerical Terms + Ambivalence + Self Reference + Variety] TENACITY: All uses of the verb to be (*is, am, will, shall*), three definitive verb forms (*has, must, do*) and their variants, a s well as all associated contraction's (*he'll, they've, ain't*). These verbs connote confidence and totality.

LEVELING: Words used to ignore individual differences and to build a sense of completeness and assurance. Included are totalizing terms (*everybody, anyone, each, fully*), adverbs of permanence (*always, completely, inevitably, consistently*), and resolute adjectives (*unconditional, consummate, absolute, open-and-shut*).

COLLECTIVES: Singular nouns connoting plurality that function to decrease specificity. These words reflect a dependence on categorical modes of thought. Included are social groupings (*crowd, choir, team, humanity*), task groups (*army, congress, legislature, staff*) and geographical entities (*county, world, kingdom, re public).*

INSISTENCE: This is a measure of code-restriction and semantic contentedness. The assumption is that repetition of key terms indicates a preference for a limited, ordered world. In calculating the measure, all words occurring three or more times that function as nouns or noun-derived adjectives are identified (either cybernetically or with the user's assistance) and the following calculation performed: [Number of Eligible Words **x** Sum of their Occurrences] ÷ 10. (For small input files, high frequency terms used two or more times are used in the calculation).

NUMERICAL TERMS: Any sum, date, or product specifying the facts in a given case. This dictionary treats each isolated integer as a single word and each separate group of integers as a single word. In addition, the dictionary contains common numbers in lexical format (*one, tenfold, hundred, zero)* as well as terms indicating numerical operations (*subtract, divide, multiply, percentage*) and quantitative topics (*digitize, tally, mathematics).* The presumption is that Numerical Terms hyper -specify a claim, thus detracting from its universality.

AMBIVALENCE : Words expressing hesitation or uncertainty, imp lying a speaker's inability or unwillingness to commit to the verbalization being made. Included are hedges (*allegedly*, *perhaps, might),* statements of inexactness (*almost, approximate, vague, somewhere*) and confusion (*baffled, puzzling, hesitate).* Also included are words of restrained possibility (*could, would, he'd*) and mystery (*dilemma, guess, suppose, seems).*

SELF-REFERENCE: All first-person references, including *I, I'd, I'll, I'm, I've, me, mine, my, myself.* Self-references are treated as acts of indexing whereby the locus of action appears to reside in the speaker and not in the world at large (thereby implicitly acknowledging the speaker s limited vision).

VARIETY: This measure conforms to Wendell Johnson's (1946) Type-Token Ratio which divides the number of different words in a passage by the passage's total words. A high score indicates a speaker's avoidance of overstatement and a preference for precise, molecular statements.

THE OPTIMISM SCORE
Definition: Language endorsing some person, group, concept or event or highlighting their positive entailments.

Formula: [Praise + Satisfaction + Inspiration] - [Blame + Hardship + Denial]

PRAISE: Affirmations of some person, group, or abstract entity. Included are terms isolating important social qualities (*dear, delightful, witty),* physical qualities (*mighty, handsome, beautiful),* intellectual qualities (*shrewd, bright, vigilant, reasonable),* entrepreneurial qualities (*successful, conscientious, renowned*), and moral qualities (*faithful, good, noble*). All terms in this dictionary are adjectives.

SATISFACTION: Term s associated with positive affective states (*cheerful, passionate, happiness*), with moments of undiminished joy (*thanks, smile, welcome*) and pleasurable diversion (*excited, fun, lucky),* or with moments of triumph (*celebrating, pride, auspicious).* Also included are words of nurturance: *healing, encourage, secure, relieved.*

INSPIRATION: Abstract virtues deserving of universal respect. Most of the terms in this dictionary are nouns isolating desirable moral qualities (*faith, honesty, self-sacrifice, virtue*) as well as attractive personal qualities (*courage, dedication, wisdom, mercy).* Social and political ideals are also included: *patriotism, success, education, justice.*

BLAME: Terms designating social inappropriateness (*mean, naive, sloppy, stupid*) as well as downright evil (*fascist, blood-thirsty, repugnant, malicious*) compose this dictionary. In addition, adjectives describing unfortunate circumstances (*bankrupt, rash, morbid, embarrassing*) or unplanned vicissitudes (*weary, nervous, painful, detrimental)* are included. The dictionary also contains outright denigrations: *cruel, illegitimate, offensive, miserly.*

HARDSHIP: This dictionary contains natural disasters (*earthquake, starvation, tornado, pollution),* hostile actions (*killers, bankruptcy, enemies, vices*) and censurable human behavior (*infidelity, despots, betrayal*). It also includes unsavory political outcomes (*injustice, slavery, exploitation, rebellion*) as well as normal human fears (*grief, unemployment, died, apprehension*) and in capacities (*error, cop-outs, weakness*).

DENIAL: A dictionary consisting of standard negative contractions (*aren't, shouldn't, don't*), negative functions words (*nor, not, nay*), and term s designating null sets (*nothing, nobody, none*).

THE ACTIVITY SCORE

Definition: Language featuring movement, change, the implementation of ideas and the avoidance of inertia.

Formula: [Aggression + Accomplish men t + Communication + Motion] - [Cognitive Terms + Passivity + Embellishment] AGGRESSION: A dictionary embracing human competition and forceful action. Its terms connote physical energy (*blast, crash, explode, collide*), social domination (*conquest, attacking, dictatorships, violation), and* goal-directedness (*crusade, commanded, challenging, overcome*). In addition, words associated with personal triumph (*mastered, rambunctious, pushy*), excess human energy (*prod, poke, pound, shove*), disassembly (*dismantle, demolish, overturn, veto)* and resistance (*prevent, reduce, defend, curbed*) are included.

ACCOMPLISHMENT: Words expressing task-completion (*establish, finish, influence, proceed)* and organized human behavior (*motivated, influence, leader, manage)*. Includes capitalistic terms (*buy, produce, employees, sell),* modes of expansion (*grow, increase, generate, construction*) and general functionality (*handling, strengthen, succeed, outputs)*. Also included is programmatic language: *agenda, enacted, working, leadership*.

COMMUNICATION: Terms referring to social interaction, both face-to-face (*listen, interview, read, speak*) and mediated (*film, videotape, telephone, e-mail*). The dictionary includes both modes of intercourse (*translate, quote , scripts, broadcast*) and moods of intercourse (*chat, declare , flatter, demand*). Other term s refer to social actors (*reporter, spokesperson, advocates, preacher*) and a variety of social purposes (*hint, rebuke, respond, persuade*).

MOTION: Terms connoting human movement (*bustle, job, lurch, leap*), physical processes (*circulate, momentum, revolve, twist),* journeys (*barnstorm, jaunt, wandering, travels*), speed (*lickety-split, nimble, zip, whistle-stop),* and modes of transit (*ride, fly, glide, swim*).

COGNITIVE TERMS: Words referring to cerebral processes, both functional and imaginative. Included are modes of discovery (*learn, deliberate, consider, compare*) and domains of study (*biology, psychology, logic, economics*). The dictionary includes mental challenges *(question, forget, re-examine, paradoxes*), institutional learning practices (*graduation, teaching, classrooms*), as well as three forms of intellection: intuitional (*invent, perceive, speculate, interpret),* rationalistic (*estimate, examine, reasonable, strategies*), and calculative (*diagnose, analyze, software, fact-finding*).

PASSIVITY: Words ranging from neutrality to inactivity. Includes terms of compliance (*allow, tame, appeasement*), docility (*submit, contented, sluggish*), and cessation (*arrested, capitulate, refrain, yielding*). Also contains tokens of inertness (*backward, immobile, silence, inhibit*) and disinterest (*unconcerned, nonchalant, stoic*), as well as tranquility (*quietly, sleepy, vacation)*.

EMBELLISHMENT: A selective ratio of adjectives to verbs based on David Boder's (1940) conception that heavy modification slows down a verbal passage by de-emphasizing human and material action . Embellishment is calculated according to the following formula: [Praise + Blame +1] ÷ [Present Concern + Past Concern +1].

THE REALISM SCORE

Definition : Language describing tangible, immediate, recognizable matters that affect people's everyday lives.

Formula: [Familiarity + Spatial Awareness + Temporal Awareness + Present Concern + Human Interest + Concreteness] - [Past Concern + Complexity]

FAMILIARITY: Consists of a selected number of C.K. Ogden s (1968) operation words which he calculates to be the most common words in the English language. Included are common prepositions (*across, over, through*), demonstrative pronouns (*this, that*) and interrogative pronouns (*who, what*), and a variety of particles, conjunctions and connectives (*a, for, so*).

SPATIAL AWARENESS: Terms referring to geographical entities, physical distances, and modes of measurement. Included are general geographical terms (*abroad, elbow-room, locale, outdoors*) as well as specific ones (*Ceylon, Kuwait, Poland*). Also included are politically defined locations (*county, fatherland, municipality, ward*), points on the compass (*east, southwest*) and the globe (*latitude, coastal, border, snowbelt*), as well as terms of scale (*kilometer, map, spacious*), quality (*vacant, out-of-the-way, disoriented*) and change (*pilgrimage, migrated, frontier.*)

TEMPORAL AWARENESS: Terms that fix a person, idea, or event within a specific time-interval, thereby signaling a concern for concrete and practical matters. The dictionary designates literal time (*century, instant, mid-morning*) as well as metaphorical designations (*lingering, seniority, nowadays*). Also included are calendrical terms (*autumn, year-round, weekend*), elliptical terms (*spontaneously, postpone, transitional*), and judgmental terms (*premature, obsolete, punctual*).

PRESENT CONCERN: A selective list of present-tense verbs extrapolated from C. K. Ogden's list of general and picturable terms, all of which occur with great frequency in standard American English. The dictionary is not topic-specific but points instead to general physical activity (*cough,* taste, sing, take), social operations (canvass, touch, govern, meet), and task-performance (make, cook, print, paint).

HUMAN INTEREST: An adaptation of Rudolf Flesch's notion that concentrating on people and their activities gives discourse a life-like quality. Included are standard personal pronouns (*he, his,* ourselves, them), family members and relations (cousin, wife, grandchild, uncle), and generic terms (friend, baby, human, persons).

CONCRETENESS: A large dictionary possessing no thematic unity other than tangibility and materiality. Included are sociological units (peasants*, African-Americans, Catholics*), occupational groups (*carpenter, manufacturer, policewoman*), and political alignments (*Communists, congressman, Europeans*). Also incorporated are physical structures (*courthouse, temple, store*), forms of diversion (*television, football, CD-ROM*), terms of accountancy (*mortgage, wages, finances*), and modes of transportation (*airplane, ship, bicycle*). In addition, the dictionary includes body parts (*stomach, eyes, lips*), articles of clothing (*slacks, pants, shirt*), household animals (*cat, insects, horse*) and foodstuffs (*wine, grain, sugar*), and general elements of nature (*oil, silk, sand*).

PAST CONCERN: The past-tense forms of the verbs contained in the Present Concern dictionary.

COMPLEXITY: A simple measure of the average number of characters-per-word in a given input file. Borrows Rudolph Flesch's (1951) notion that convoluted phrasings make a text's ideas abstract and its implications unclear.

THE COMMONALITY SCORE

Definition: Language highlighting the agreed -upon values of a group and rejecting idiosyncratic mode s of engagement.

Formula: [Centrality + Cooperation + Rapport] - [Diversity + Exclusion + Liberation]

CENTRALITY: Terms denoting institutional regularities and/or substantive agreement on core values. Included are indigenous terms (*native, basic, innate*) and designations of legitimacy *(orthodox, decorum, constitutional, ratified*), systematicity (*paradigm, bureaucratic, ritualistic*), and typicality (*standardized, matter-of-fact, regularity*). Also included are terms of congruence (conformity, mandate, unanimous), predictability (expected, continuity, reliable), and universality (womankind, perennial, landmarks).

COOPERATION: Terms designating behavioral interactions among people that often result in a group product. Included are designations of formal work relations (*unions, schoolmates, caucus*) and informal association s (*chum, partner, cronies)* to more intimate interactions (*sisterhood, friendship, comrade*). Also included are neutral interactions (*consolidate, mediate, alignment*), job-related tasks (*network, detente, exchange*), personal involvement (*teamwork, sharing, contribut*e), and self-denial (*public-spirited, care-taking, self-sacrifice*).

DICTION 5.0 Manual

47

RAPPORT: This dictionary describes attitudinal similarities among groups of people. Included are terms of affinity (*congenial, camaraderie, companion*), assent (*approve, vouched, warrants*), deference (*tolerant, willing, permission*), and id entity (*equivalent, resemble, consensus*).

DIVERSITY: Words describing individuals or groups of individuals differing from the norm. Such distinctiveness may be comparatively neutral (*inconsistent, contrasting, non-conformist*) but it can also be positive (*exceptional, unique, individualistic*) and negative (*illegitimate, rabble-rouser, extremist*). Functionally, heterogeneity may be an asset (*far-flung, dispersed, diffuse*) or a liability (*factionalism, deviancy, quirky*) as can its characterizations: *rare* vs. *queer*, *variety* vs. *jumble*, *distinctive* vs. *disobedient*.

EXCLUSION: A dictionary describing the sources and effects of social isolation. Such seclusion can be phrased passively (*displaced, sequestered*) as well as positively (*self-contained, self-sufficient*) and negatively (*outlaws, repudiated*). Moreover, it can result from voluntary forces (*secede, privacy*) and involuntary forces (*ostracize, forsake, discriminate*) and from both personality factors (*smallmindedness, loneliness*) and political factors (*right-wingers, nihilism*). Exclusion is often a dialectical concept: *hermit* vs. *derelict*, *refugee* vs. *pariah*, *discard* vs. *spurn*).

LIBERATION: Terms describing the maximizing of individual choice (*autonomous, open-minded, options*) and the rejection of social conventions (*unencumbered, radical, released*). Liberation is motivated by both personality factors (*eccentric, impetuous, flighty*) and political forces (*suffrage, liberty, freedom, emancipation*) and may produce dramatic outcomes (*exodus, riotous, deliverance*) or subdued effects (*loosen, disentangle, outpouring*). Liberatory terms also admit to rival characterizations: *exemption* vs. *loophole*, *elope* vs. *abscond*, *uninhibited* vs. *outlandish*.

PART 14: STATISTICAL ANALYSIS

DICTION data can be analyzed with any standard statistical package. Be cause of the popularity of SPSS -for-Windows, however, DICTION comes packaged with a simple program for converting DICTION's raw data to a form readable by SPSS. This program can be found under the DICTION\STATS subdirectory. The program is called **SPSS-DIC.SPS**. It is depicted in Figure 29.

To use the conversion program do the following:
1. Change the extension of the Numeric File to <**.dat**> (this is SPSS s designation for a data file).
2. Start SPSS and call up the DICTION\STATS\SPSS-DIC.SPS syntax file.
3. Enter these file names:
    (a) The name and location of the existing raw data file
    (b) The name and location of the SPSS data file you wish to create

4. Highlight the entire syntax file.
5. Press the arrow on the SPSS task bar entitled Run Highlighted Command Syntax or use the Cntrl-R command.

These procedures produce a file such as that shown in Figure 30.

PART 15: MISCELLANEOUS

Copyright

DICTION 5.0: THE TEXT ANALYSIS PROGRAM is fully protected under U.S. copyright law. Users making use of its data, whether in published or unpublished form, must have a registered license agreement. © Digitext, Inc., 2000.

Acknowledgments

The developers hereby express their appreciation to the following individuals who have helped in the development of DICTION over the years: George Borden, Rebecca Daugherty, Bob Lemon, Tom Porter, Jon Rutter, Wayne Shamo, David Sinclair, Kris Stanton, Laura Stevens, Mary Triece, Morgan Watkins, Charlie Watson, and Steve Weikal.

Homograph Treatment

DICTION make s a modest, statistical accommodation for homographs, words spelled the same but having different meanings. The word *lead* for example, can denote (1) qualities of command or (2) a metal found in nature. Roughly 10% of DICTION's 10,000 search words are homographs.

Not all homographs are problematic. Benign Homographs imply different ideas but belong to the same domain of meaning. The word *judge*, for example, denotes the act of adjudication as well as a magistrate of the court, but both would be equally relevant to a dictionary entitled Legal Terms. Such meanings are not differentiated by DICTION.

Confounding Homo graphs are terms denoting dissimilar ideas or objects. DICTION deals with such terms by applying different weights to a word s various meanings. So, for example, research shows that the word *saw* serves as the past tense of *see* 67% of the time and as a carpenter s implement 33% of the time. Thus, if *saw* occurred ten times in a text, it would be given a rating of 6.7 in a dictionary called Bodily Processes and a value of 3.3 in a Household Tools dictionary. All such calculations have been extrapolated from statistical counts provided in Helen Easton s edited handbook, Word Frequency Dictionary (New York: Dover, 1940).

On-line Help

DICTION has a full range of Help Screens to guide the user through a text search (see, for example, Figure 31). Clicking on Help/Help Topics/Contents will produce the broadest level of help. In addition, individual terms can be explored by clicking on Help/Help Topics/Index or ... Index/Find and then specifying a term (e.g., Familiarity Score).

Context Help

DICTION also offers limited contextual help within the Project File. To access these help screens do the following:

1. Click on the Help icon on the Tool Bar
2. This will produce a Question Mark icon adjacent to the cursor.
3. Click in any area of the File View Window Pane (Figure 32).
4. The relevant help panel will then appear.
5. Click in the upper right corner of the Help box to exit.

Customer Support

If the User's Manual and Help Screens cannot solve your problem, use Customer Support. Before calling, prepare to explain your problem and have these items available:

DICTION User's Guide, original CD rom, your license number, date of purchase.

Then call 213-821-5415 or send e-mail to help@dictionsoftware.com

References

Boder, D. (1940). The Adjective-Verb Quotient: A Contribution to the Psychology of Language. *Psychology Record*. 3: 310-343.

Easton, H. (1940). *Word Frequency Dictionary*. New York: Dover.

Flesch, R. (1951). *The Art of Clear Thinking*. New York: Harper.

Johnson, W. (1951). *People in Quandaries: The Semantics of Personal Adjustment* (New York: Harper).

Ogden, C. K. (1960). *Basic English Dictionary*. London: Evans Brothers.

Bibliography

Users may find it helpful to examine some of the work done with the current version of DICTION:

R. P. Hart, "Redeveloping DICTION: Theoretical Considerations" in M. West (Ed.), *New Directions in Computer Content Analysis* (New York: Ablex, 2000).

R. P. Hart, *Campaign Talk: Why Elections Are Good for Us* (Princeton: Princeton University Press, 2000)

R. P. Hart and H. Gourgey, "Accepting the Political Mantle: Stylistic Considerations," *Political Communication, 15* (1998).

R. P. Hart and S. E. Jarvis, "Political Debate: Forms, Styles, and Media," *American Behavioral Scientist, 40* (1997), 1095-1122.

R. P. Hart, "Rhetoric, Hope, and American Politics," in J. Trent (Ed.), *At the Helm in Speech Communication* (Boston: Allyn and Bacon, 1997)

Work done with the mainframe version of the program that preceded DICTION 5.0 includes these:

R. P. Hart, "Systematic Analysis of Political Discourse: The Development of DICTION," in K. Sanders, et al. (Eds.), *Political Communication Yearbook: 1984* (Carbondale, IL: Southern Illinois University Press, 1985), pp. 97-134.

R. P. Hart, "The Language of the Modern Presidency," *Presidential Studies Quarterly, 14* (1984), 249-264.

R. P. Hart, *Verbal Style and the Presidency: A Computer-Based Analysis* (New York: Academic Press, 1984).