

# B=SONAR=S BOOKENDS InDex Pro

THE COMPREHENSIVE INDEX GENERATOR  
FOR ADOBE® InDesign®

## Professional Version 20.x

Macintosh and Windows  
Combined User's Guide



Copyright © 1999-2023  
Virginia Systems  
All rights reserved

## **COPYRIGHT**

**Copyright © 1999-2023 by Virginia Systems. All rights reserved. No part of this publication may be reproduced, transmitted, transcribed, stored in a retrieval system, or translated into any language or computer language, in any form or by any means, electronic, mechanical, magnetic, optical, chemical, manual, or otherwise, without the prior written permission of Virginia Systems, 5509 West Bay Court, Midlothian, Virginia 23112.**

## **DISCLAIMER**

**Virginia Systems makes no representation or warranties with respect to the contents hereof and specifically disclaims any implied warranties of merchantability or fitness for any particular purpose. Further, Virginia Systems reserves the right to revise this publication and to make changes from time to time in the content hereof without obligation of Virginia Systems to notify any person of such revision or changes.**

## **TRADEMARKS**

**Sonar Bookends InDex and Sonar Bookends InDex Pro are trademarks of Virginia Systems**

**Adobe and InDesign are trademarks of Adobe Incorporated**

**Apple, the Apple logo, Macintosh and the Macintosh logo are trademarks of Apple Computer, Inc.**

**Windows and MS Word are trademarks of the Microsoft® Corporation.**

**All other trademarks acknowledged.**

----- Twentieth Edition: October 2023 -----

# Table of Contents

Section I Introduction .....	1
About Sonar Bookends InDex Pro.....	1
Installing Sonar Bookends InDex Pro.....	2
Activating Sonar Bookends InDex Pro.....	3
Technical Support .....	3
Section II Types of Indices.....	5
Word Frequency List Index.....	5
Word/Phrase List Index.....	5
Single-Level.....	5
Multiple-Level .....	6
Subject List Index.....	7
Proper Noun List Index .....	7
Marked Text Index - Manually Marked.....	8
Marked Text Index - Marked Using Style Sheets.....	8
Index Formats .....	8
Index Format Without Chapter/Section Prefix .....	8
Index Format With Chapter/Section Prefix .....	9
Multiple-Level Index Format .....	10
Table of Contents Format .....	11
Section III Preparing to Make an Index.....	13
Hyphenation.....	13
Setting/Overriding Default Word Order and Proximity.....	13
Document Organization .....	14
File Naming Conventions When Indexing Multiple Files.....	15
Indexing Without Chapter/Section Names .....	15
Indexing With Chapter/Section Names .....	16
Section IV Creating a Word/Phrase List .....	17
Word-Frequency List .....	17
Creating a Word-Frequency List .....	17
Subject List.....	18
Proper Noun List.....	21
Generic Word/Phrase List .....	22
Single- and Multiple-Level Word/Phrase Lists .....	23
Creating a Word/Phrase List Using a Database .....	24
Using Marked Text to Create a Word/Phrase List - Manual .....	25
Using Marked Text to Add Hidden Keywords to a Document.....	26
Using Marked Text to Create a Word/Phrase List - Style Sheets.....	27
Sorting a Word/Phrase List .....	27
Section V Making an Index and Table of Contents .....	31
Handling Multiple Levels .....	32
Specifying Section Prefixes .....	33
Handling Non-Existent Entries .....	34
Previewing and Editing ‘Hits’.....	34

Displaying Frequency Instead of Page Numbers.....	34
Including Group-Letter Titles .....	35
Separating Non-Consecutive Page Numbers .....	35
Abbreviating Consecutive Page Numbers .....	35
Using Database Import Format.....	36
Separating Index Entries and Page Numbers.....	36
Setting Blank Lines Between Entries .....	37
Formatting the Index.....	38
Making a Table of Contents .....	39
Section VI Special Features .....	41
Previewing and Editing ‘Hits’ .....	41
Interrupting a Preview and Edit ‘Hits’ Session.....	44
Rebuilding a Previously Previewed Index.....	44
Wild Cards in Word/Phrase Lists .....	45
Boolean Expressions in Word/Phrase Lists .....	45
Phrases .....	45
Operators .....	45
Modifiers.....	45
Notes on the ‘&’ and ‘^’ Operators.....	46
Sample Boolean Expressions.....	47
Overriding Word Order and Proximity Defaults.....	47
Using the Translation Operator.....	47
Hiding Text .....	47
Adding Cross-References and Blank Lines.....	48
Reversing Names .....	48
Keyword Elimination.....	50
Combining Singular and Plural Instances of a Phrase .....	50
Removing Redundant Pages from Overlapping Phrases .....	52
Converting to a Multiple Level Word/Phrase List.....	53
Redefining the Alphabetic Character Set .....	55
Ignoring Non-Alphabetic Characters.....	56
Indexing Special Characters .....	57
Smart Hyphens.....	58
Filtering Proper Nouns.....	59
Using Style Sheets to Ignore Text.....	59
Controlling Treatment of Phrases That Span Pages.....	60
Section VII Tutorials.....	61
Making a Word Frequency Index.....	61
Making a Multiple Level Proper Noun Index.....	63
Section VIII Appendices .....	67
Appendix A - Glossary.....	67
Appendix B - Error Messages.....	69
Appendix C - Menu Summary .....	71
Appendix D - Trouble Shooting.....	73
Index .....	75

# Section I Introduction

## About Sonar Bookends InDex Pro

Sonar Bookends® InDex Pro™ is a powerful index generator capable of producing an automatic index for Adobe's InDesign™ desktop publishing product. It can also make a table of contents based on style sheets. Several types of indices can be produced, from a basic index to a multiple-level index. The index can include chapter/section names, can be made for any number of files, and can contain cross-reference information.

There are five basic types of indices that can be produced by Sonar Bookends InDex Pro:

1. Word frequency list. A fast, effortless index based on how often words appear in the documents being indexed. Words which are common, and therefore relatively unimportant, can be automatically eliminated.
2. Word/phrase list. Based on a user supplied list of words and phrases stored in a text file. Each word or phrase is indexed. Entries can incorporate advanced indexing techniques such as Boolean operations ("John" or "Jim" within five words of "Smith") and wild cards (all words starting with "jam"). Since the word/phrase list is a text only file, it can be generated in a variety of ways. For example, the list can be created using InDesign or a word processor, downloaded from another computer, exported from a database, etc.
3. Subject list. Using an English-language parsing engine, Sonar Bookends Pro can analyze the contents of the document being indexed and create a list of subjects. Various filtering options generate a more focused list.
4. Proper noun list. Using rules of capitalization and punctuation, Sonar Bookends InDex Pro can find proper nouns (names of people, places, and things) in a document. Unlike a subject list, that is based on interpreting English, a proper noun list is language independent.
5. Marked text list. This list can be generated either by manually highlighting entries to be indexed or by using one or more style sheets to mark entries to be indexed. Marked text can also have hidden keywords that can be indexed.

The five methods (word frequency list, word/phrase list, subject list, proper noun list, and marked text list) can be mixed and matched to produce a final index. For example, a word frequency list can be combined with a subject list, discarding any unwanted words and subjects. The list could also be augmented by typing in additional phrases. The final word/phrase list would then be indexed.

Other important features include:

- **Flexibility** - An index can be created for a single file or a group of related files. The index can be single-level or multiple-level.
- **Enhanced Preview mode** - The context of each occurrence of each item being indexed is displayed, making it easy to determine if any occurrences should not be included in the index. Choices made in preview mode are saved and can be applied automatically to future indices.
- **Singular/plural optimization** - Singular and plural occurrences of a subject can be combined automatically into a single entry.
- **Redundant page number optimization** - Redundant page numbers created when one subject (“bears”) overlaps another subject (“**black** bears”) can be eliminated automatically.
- **Translation** - Translation, designated by a semi-colon (;), allows Boolean expressions to be hidden, blank lines to be inserted, and cross-references to be added to an index.
- **Intelligent name reversal** - First and last names can be switched automatically. For example, “John H. Doe Jr.” would become “Doe, John H. Jr.” Names which are already reversed are not effected. Titles and two-word last names are taken into account.
- **Single-level to multiple-level conversion** - A single-level list of names can be converted to a multiple-level list at the push of a button - an incredible time-saver.
- **Powerful sorting capability** - Single-level and multiple-level lists can be sorted rapidly. Leading articles are ignored. Thus “The United States” would appear under “U” and not “T.” Collating sequences (sorting order) are user-definable.
- **Alphabetic character definition** - The characters that can make up a word are user-definable.
- **Indexing text in tables and footnotes** - Text stored in InDesign tables and footnotes can be indexed in the professional version.

A glossary can be found in Appendix A on page 67.

## Installing Sonar Bookends InDex Pro

To install the Sonar Bookends InDex Pro plug-in, see the ReadMe file in the Sonar Bookends InDex Pro folder.

## Activating Sonar Bookends InDex Pro

When first installed, Sonar Bookends InDex Pro runs in demonstration mode. Demonstration mode has the following restrictions:

- Only words starting with an “s” are in the word frequency list
- Only the first 25 proper nouns or subjects are put in a word/phrase list
- Only the first 10 marked words are put into a marked word list
- Only the first 10 table of contents entries are found
- Only the first two page numbers for each indexed entry are listed

To remove these restrictions, you must purchase Sonar Bookends InDex Pro. After payment is received, you are given a special serial number to enter into a registration box, enabling all features immediately.

To purchase Sonar Bookends Index Pro, contact Virginia Systems at:

Email: [sales@virginiasytems.com](mailto:sales@virginiasytems.com)

## Technical Support

Support contracts are also available to supplement the 30 days of free support that comes with the purchase of the product. Support contracts (Technical Support and Software Update Agreements) include updates free of charge, with email notification of such updates..

Along with any questions, please include your:

- Name
- Serial number
- Telephone number
- Email address

Contact Virginia Systems’ technical support at:

Email: [support@virginiasytems.com](mailto:support@virginiasytems.com)





## Section II      Types of Indices

Sonar Bookends InDex Pro can create several types of indices:

- Word frequency list
- Word/phrase list
  - Single-level
  - Multiple-level
- Subject list
- Proper noun list
- Marked text list (manually or using style sheets)

This section gives a brief description of each type of index.

### Word Frequency List Index

A word frequency list is an alphabetical list of all words which occur no more than a given number of times in a document. A word frequency list is easy to generate, but usually requires a lot of cleaning up. Setting a reasonable word frequency helps to eliminate words which occur too frequently to have much value. A frequency of 10, for instance, would tell Sonar Bookends InDex Pro to index all words in the document that occur no more than 10 times. Therefore, a word such as “the” would occur far more than 10 times in a document and would not be indexed. Once the list is cleaned up, the list is used to generate the index.

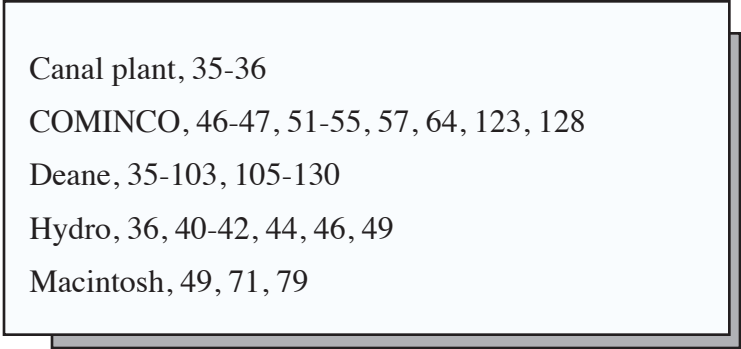
### Word/Phrase List Index

An index is built from a list of words and phrases. A word/phrase list is a text file containing both individual words and combinations of words (phrases) which are to be indexed. Since a word/phrase list is a text file, it can be created in many ways including:

- Using InDesign
- Exporting the list from a database
- Generating a list of words based on their frequency
- Generating a list of all proper nouns
- Generating a list of words and phrases marked manually or marked using style sheets
- Generating a list of subjects
- Any combination of the above

### Single-Level

A single level index is a type of word/phrase list index. A single level index consists of one word or phrase per line. The list can be sorted alphabetically, taking into account non-English characters such as “ñ” and leading articles such as “The.” Part of a single level index is shown in figure 2.1. A non-sorted single-level index may be preferred if the list of words and phrases has already been sorted or is in a special non-alphabetic order.

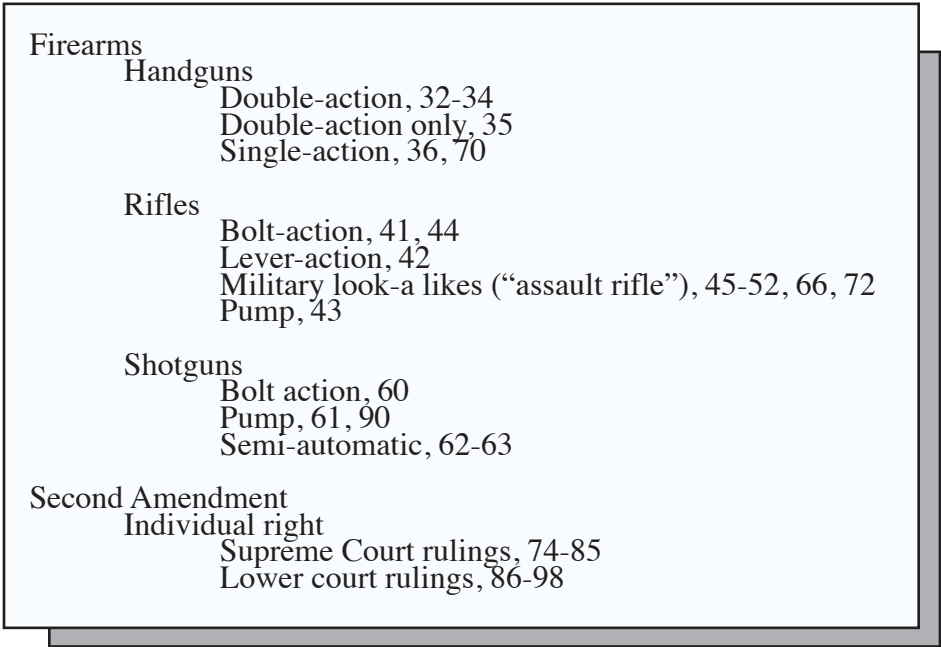


Canal plant, 35-36  
COMINCO, 46-47, 51-55, 57, 64, 123, 128  
Deane, 35-103, 105-130  
Hydro, 36, 40-42, 44, 46, 49  
Macintosh, 49, 71, 79

*Fig. 2.1*

## Multiple-Level

A multiple-level index is a type of word/phrase list index. Part of a multiple-level index is shown in figure 2.2. Each entry is on a line by itself and subentries are indented with tabs. A multiple-level index can have an unlimited number of levels and can be sorted by Sonar Bookends InDex Pro. Only the most indented (lowest level) entries at any given level receive page numbers. The list of entries in a multiple-level index can be created with InDesign, any word processor, or any database.



Firearms  
    Handguns  
        Double-action, 32-34  
        Double-action only, 35  
        Single-action, 36, 70  
  
    Rifles  
        Bolt-action, 41, 44  
        Lever-action, 42  
        Military look-a likes (“assault rifle”), 45-52, 66, 72  
        Pump, 43  
  
    Shotguns  
        Bolt action, 60  
        Pump, 61, 90  
        Semi-automatic, 62-63  
  
Second Amendment  
    Individual right  
        Supreme Court rulings, 74-85  
        Lower court rulings, 86-98

*Fig. 2.2*

**Note:** By default a multiple-level index does not take into account any higher levels for a given entry. For example in figure 2.2, the page numbers for “Lever-action” were found by simply looking for the phrase “Lever-action.” “Firearms” and “Rifles” did not figure into the search. This method works fine if the only places where the phrase “Lever-action” appears is in reference to rifles and firearms.

There are cases where simply searching for a phrase will not work. For example, if “Lever-action” appeared in the book not only as a rifle, but also as a shotgun, then some of the page numbers would be wrong.

Sonar Bookends InDex Pro has two ways to deal with items that need to be qualified:

- The Translation operator can be used to specify a hidden search command. For example, the hidden command could direct Sonar Bookends InDex Pro to find the page numbers for “Lever-action” by looking for “Lever-action” on the same page as “Rifles” and “Firearms.” Translation is covered on page 47.
- Using the “smart” option when indexing. The “smart” option tells Sonar Bookends InDex Pro to require that all higher levels appear on the same page as the phrase being indexed. Thus, for the page number of a page having the phrase “Lever-action” to appear in the index, that page would also have to contain the words “Rifle” and “Firearms.” If necessary, the Translation operator (on page 47) can be used to specify an alternate hidden search phrase for any level. For example, say that the word “exports” did not actually appear in the book per se, but “exported item” does appear. Translation would allow “exported item” to be silently substituted for “exports” when doing a smart multiple-level index. The “smart” option is discussed further on page 32.

## Subject List Index

Sonar Bookends InDex Pro can create a list of subjects found in a document and then index that list. A subject is a complete phrase such as “John Q. Smith,” “United States of America,” or “automatic index generator.” The index is made by using the list of subjects as a word/phrase list. The list can be edited using InDesign.

## Proper Noun List Index

Sonar Bookends InDex Pro can create a list of people, places, and things found in a document and then index that list. A proper noun list is created automatically using rules of capitalization and punctuation. The index is made by using the list of proper nouns as a word/phrase list. InDesign can be used to edit the proper noun list before making an index.

## Marked Text Index - Manually Marked

Words and phrases in a document can be manually marked using Sonar Bookends InDex Pro. Once marked, the words and phrases can be indexed in one of two ways:

1. All occurrences of each word/phrase are found and indexed, even if some those occurrences were not marked. With this mode, words and phrases only need to be marked once in the document being indexed.
2. Only the marked occurrences of each word/phrase that match exactly are found and indexed. With this mode, each occurrence to be indexed must be marked.

Once all of the entries have been marked, a word/phrase list can be created in seconds with a single command. The word/phrase list is then used to make the index.

## Marked Text Index - Marked Using Style Sheets

Words and phrases in a document can be marked using either character or paragraph style sheets. Style sheets that are being used to mark words and phrases for indexing must have names that start with the characters “WPL0” (like “WPL0 Part number”). Once marked, the words and phrases can be indexed in one of two ways:

1. All occurrences of each word/phrase are found and indexed, even if some those occurrences were not marked. With this mode, words and phrases only need to be marked once in the document being indexed.
2. Only the marked occurrences of each word/phrase are found and indexed. With this mode, each occurrence to be indexed must be marked.

Once all of the entries have been marked, a word/phrase list can be created in seconds with a single command. The word/phrase list is then used to make the index.

## Index Formats

### Index Format Without Chapter/Section Prefix

An index without chapter/section prefixes consists of alphabetically sorted words and phrases. By default, each word or phrase is followed by a comma and a list of page numbers (“Memory card, 24, 32, 56”). If two or more page numbers are consecutive, then the sequence of page numbers is abbreviated to be the first and last page number separated, by default, with a hyphen (“23, 24, 25, 26” would be shortened to “23-26”). Both the comma following each word or phrase and the hyphen separating consecutive page numbers can be changed. A portion of a typical index for a document is shown in figure 2.3. The string of asterisks after “Elizabeth Andrews” indicate that the phrase was not found during indexing.



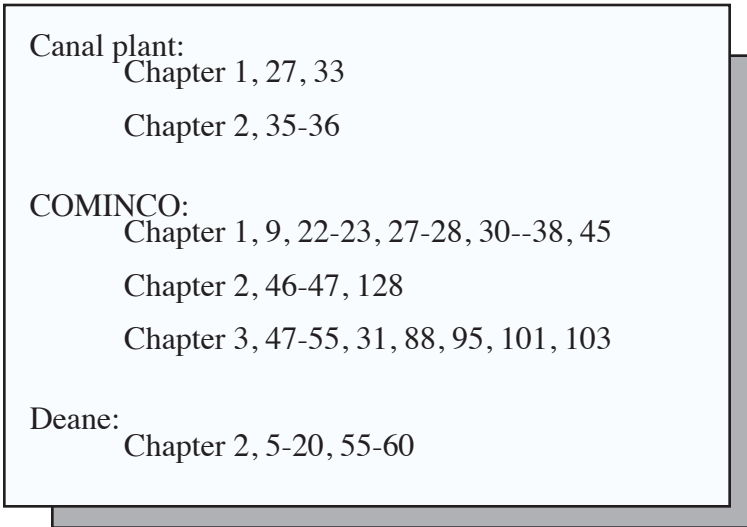
Canal plant, 35-36  
COMINCO, 46-47, 128  
Deane, 35-103, 105-130  
Elizabeth Andrews\*\*\*\*\*  
Hydro, 36, 40-42, 44, 46, 49  
Johnson, 45, 56, 90  
Macintosh, 49, 71, 79

*Fig. 2.3*

## **Index Format With Chapter/Section Prefix**

There are two options for an index containing chapter/section prefixes:

1. Each word or phrase is placed on a separate line and is followed by a colon. The names of all chapters which contain the word or phrase appear next. The chapter names are in alphabetical order, each on a separate line, preceded by a tab, and followed by a list of page numbers. Figure 2.4 has a portion of a typical index including chapter names.



Canal plant:  
    Chapter 1, 27, 33  
    Chapter 2, 35-36  
  
COMINCO:  
    Chapter 1, 9, 22-23, 27-28, 30--38, 45  
    Chapter 2, 46-47, 128  
    Chapter 3, 47-55, 31, 88, 95, 101, 103  
  
Deane:  
    Chapter 2, 5-20, 55-60

*Fig. 2.4*

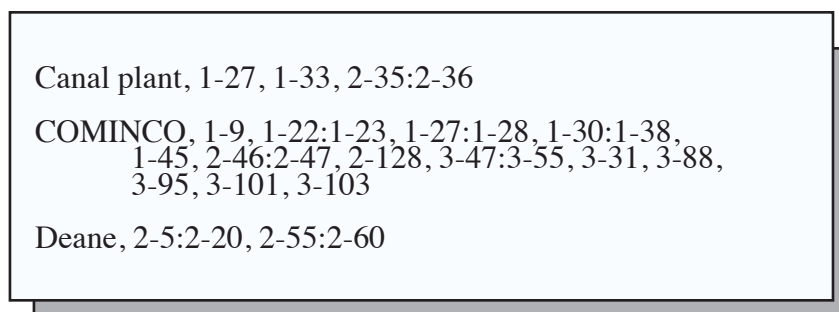
2. Each word or phrase is followed by a comma and a list of page numbers. Each page number is prefixed with the corresponding chapter name. For example:

**Memory card, 2-24, 2-32, 2-56, 4-12, 4-22**

In the example, “Memory card” is found on pages 24, 32, and 56 of chapter 2 and pages 12 and 22 of chapter 4. If two or more page numbers are consecutive, then the sequence of page numbers is abbreviated to be the first and last page number separated, by default, with a hyphen. To avoid confusion when the chapter name contains a hyphen (as in the above example), a tilde “~” or a colon “:” are commonly substituted for the hyphen separator:

**Administrative offices, 2-6, 2-34:2-56**

Figure 2.5 shows a portion of an index using this option.



*Fig. 2.5*

## Multiple-Level Index Format

Sonar Bookends InDex Pro can produce a multiple-level index with an unlimited number of levels. A multiple-level index is created using a list of words and phrases which are indented using tab characters. See figure 2.6. A multiple-level word/phrase list can be manually generated or it can be automatically generated from a single-level word/phrase list.

Accidents	
Environmental damage,	108
Nuclear,	122-128, 150
Countries	
Canada,	7-8, 10, 14-16, 32, 36, 51
France,	56-70
United States,	36-37, 46, 50-51
Fuel	
Gas,	34-35, 53, 57-58
Hydroelectric,	22, 48, 52
Nuclear,	205

*Fig. 2.6*

## Table of Contents Format

A table of contents entry consists of the text for that entry followed by a tab character and the page number. The page number can be a simple page number or it can be a page number prefixed with a chapter/section prefix. Tab leaders, if desired, can be put in using InDesign after the table of contents has been generated.





## Section III     Preparing to Make an Index

Before creating an index with Sonar Bookends InDex Pro it is important to:

- Understand how hyphenated words are treated
- Set the word order and proximity defaults
- Set the pagination information
- Organize the document according to indexing preferences (when the document consists of multiple files)

### Hyphenation

If *Smart Hyphens* in the *Preferences* menu is checked, hyphens are removed by Sonar Bookends InDex Pro if the hyphen is followed by a space, tab or carriage return. Thus, a hyphenated word (like “hy-phen”) is properly indexed as a single word (“hyphen” as opposed to “hy” and “phen”).

There are instances where removal of hyphens is not desirable. An example would be a part number like “123-ABC-789.” If the hyphens were removed, there would be no way of distinguishing between “123-ABC-789” and “1-23A-BC78-9.” For hyphens to be treated as part of a word or number, the hyphen must be touching an alphabetic or numeric character on both sides. Therefore, characters which make up a part number should remain together, and not be separated at the end of a line like this:

123-ABC- 789
-----------------

The above example would produce the part number “123-ABC789,” instead of “123-ABC-789,” since the hyphen following “ABC” was not touching a character or number on its right side.

**Note:** This behavior can be changed to keep all hyphens by unchecking *Smart Hyphens* in the *Special* menu. **Also, discretionary hyphens are always removed.**

### Setting/Overriding Default Word Order and Proximity

Sonar Bookends InDex Pro indexes a phrase as a series of words and not just a string of characters. Therefore, spacing and punctuation are not significant and are ignored. By default, a phrase is found if the words in the phrase appear in the document being indexed in the same order and immediately next to each other. For example, the phrase “John Doe” would be found even if it appeared as “John, Doe.” because periods, commas, extra spaces, and carriage returns are ignored.

By default, the phrases in figure 3.1 would not match “John Doe:”

"Doe, John"	Words are in the wrong order
"John E. Doe"	Words are not next to each other ("E" is between them)
"Doe, Edward John"	Words are not next to each other and are in the wrong order

Fig. 3.1

The default word order and proximity can be set globally for every entry, and overridden for individual entries. To set the defaults for all entries, select *Default word order and proximity...* in the *Preferences* menu, as shown in figure 3.2. The dialog box shown in figure 3.3 will then appear.

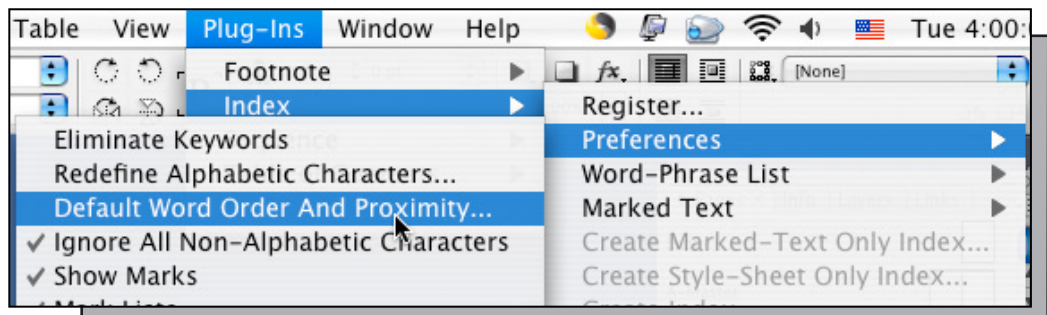


Fig. 3.2

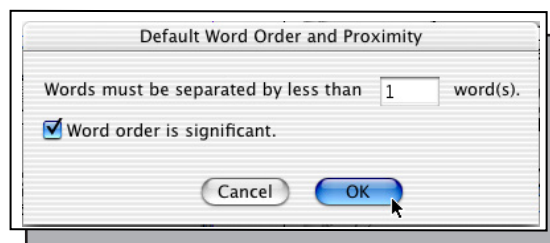


Fig. 3.3

By changing the word separation to 2 (to allow for "John Edward Doe") and unchecking word order significance, all of the phrases in figure 3.1 would match "John Doe."

Word order and proximity values can also be overridden on individual entries in a word/phrase list. See Section VI, "Special Features" on page 41.

## Document Organization

An index or a table of contents can be made for a document whether the document consists of either one file or multiple files. All of the files making up the document must be open in InDesign. You can either open the documents individually or open a book. *Since all open files (with the exceptions listed in the note below) are included in the index or table of contents, be sure that the only files which are open belong to the document to be indexed.*

**Note:** Files which are named “Word Phrase List” or “Index” or any file which has been selected as a word/phrase list using *Make Current File a List* in the *Word-Phrase List* menu are exceptions to the rule and are not indexed.

## File Naming Conventions When Indexing Multiple Files

If the document consists of multiple files, the naming of those files is critical. The names of the files being indexed should be in alphabetical order according to the order that each file appears in the document. For example, if a file named “B” contains pages 1-25, then the file which contains pages 26 through 45 could be named “C,” but should not be named “A.” If the second file was named “A,” then the page numbers would be correct, but would appear out of sequence in the index. For example, page 28 might appear before page 6.

File names that start with numbers are processed in numeric order. There is no need to put leading zeros on shorter numerical sequences. The following file names are shown in the order they would be processed:

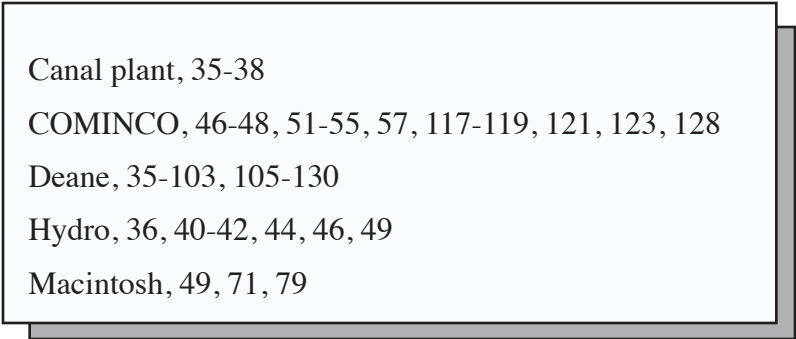
8 Moving forward  
9 A new chapter in life  
10 Where we go from here

## Indexing Without Chapter/Section Names

If the page numbering scheme for a document is simply a series of consecutive page numbers (like the page numbering for this manual), then the names of the files making up the document should be in alphabetical order as mentioned above.

**Important note:** When making this type of index, InDesign’s page numbering preference (under general preferences) should be set to *View: Section Numbering*. If this is not done, the page numbers for each file will always begin with page one.

An example of an index without chapter/section names appears in figure 3.4.



Canal plant, 35-38  
COMINCO, 46-48, 51-55, 57, 117-119, 121, 123, 128  
Deane, 35-103, 105-130  
Hydro, 36, 40-42, 44, 46, 49  
Macintosh, 49, 71, 79

*Fig. 3.4*

## Indexing With Chapter/Section Names

If the document being indexed has a page numbering scheme which includes a chapter/section reference and a page number (such as “1-2,” “6-12,” etc.), then each section start in the document should have an appropriate InDesign section marker such as “1-” or “6-”. Section markers are set using InDesign’s *Numbering and Section Options* in the *Pages submenu*. If the document consists of more than one file, then the names of the files making up the document should be in alphabetical order as described earlier in this section.

An example of an index with chapter names is shown in Figure 3.5.

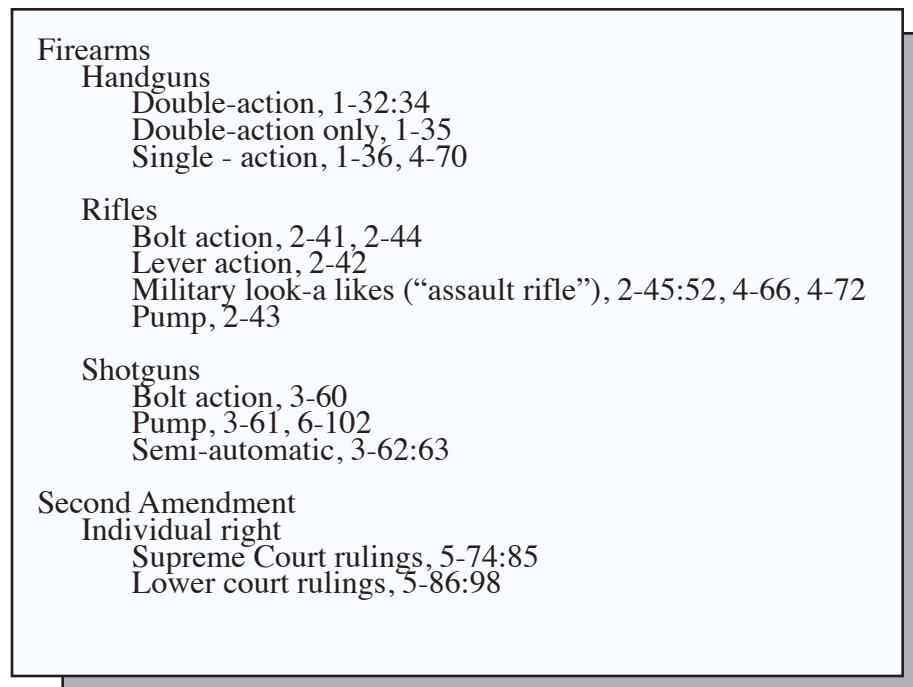


Fig. 3.5

## Section IV Creating a Word/Phrase List

**Note:** This section only applies to making an index. This section can be skipped if making a table of contents.

In order to make an index, there must first be a word/phrase list containing the items to be in the index. The word/phrase list can be made in a variety of ways as discussed next.

### Word-Frequency List

A word-frequency word/phrase list is based on how often words appear in a document. The entries are sorted alphabetically. This type of word/phrase list assumes that a word's significance decreases as its frequency increases.

In a word-frequency list, both numbers and words are selected. Numbers appear before letters; so numeric values appear first in the list and are followed by words beginning with the letter "A." Note: By redefining Sonar Bookends InDex Pro's alphabetic characters to make the characters "0123456789" non-alphabetic, numbers will not be indexed. For a list solely of numbers, see "Subject List" later in this chapter.

The word-frequency list can be edited using InDesign. The edited list can then be used as a word/phrase list to produce the final index. Before proceeding to make a word-frequency list of a document consisting of multiple files, make sure the instructions in "Document Organization" on page 14 in Section III have been followed.

### Creating a Word-Frequency List

To begin generating a word-frequency list, **open all files that are to be indexed or open a book** and then select *Build Word Frequency List...* in the *Word-Phrase List* menu, as shown in figure 4.1. The dialog box shown in figure 4.2 will then appear.

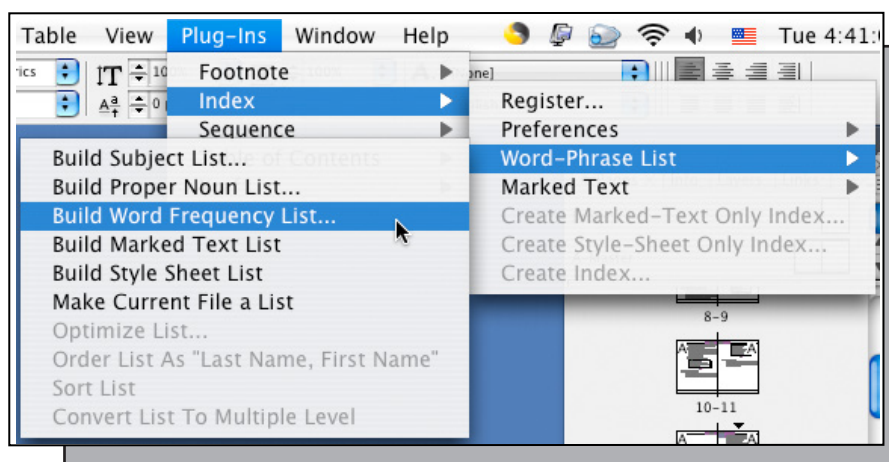
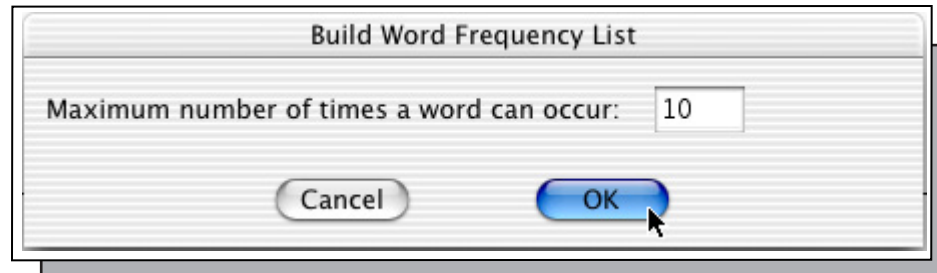
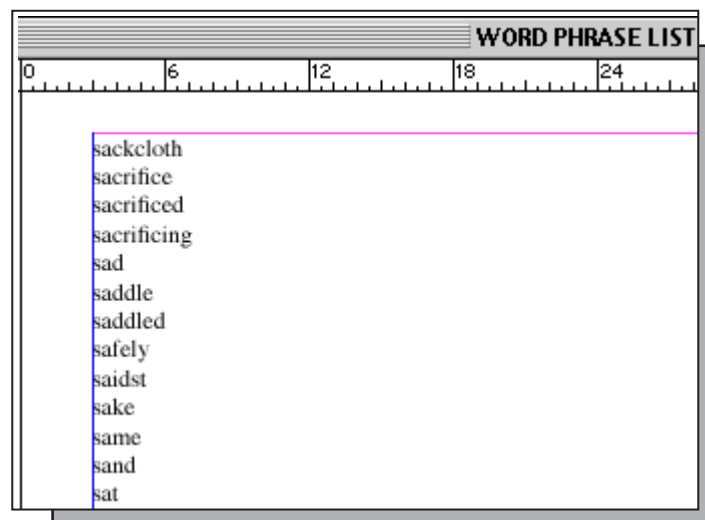


Fig. 4.1

*Fig. 4.2*

Refer to figure 4.2. Set the “Maximum number of times a word can occur:” to the cut-off frequency for a word to appear in the list. E.g., if a word appears more frequently than the specified value, it is not put in the list. This value should be greater than zero.

After clicking *OK*, Sonar Bookends InDex Pro will then process the document’s file or files and begin to build the word/phrase list. When completed the list appears in a window as shown in figure 4.3. **The demo version will only show words that start with “s”.**

*Fig. 4.3*

## Subject List

The subject list capability of Sonar Bookends InDex Pro is one of its most unique and powerful features. A subject list is a list of people, places, and things that are mentioned in a document. Sonar Bookends InDex Pro has a proprietary English-language parsing engine which analyzes the contents of a document to automatically create a subject list, taking into account sentence structure and the part of speech of each word. This contrasts with the proper noun detection capabilities of both the regular and professional versions of Sonar Bookends InDex which use capitalization and punctuation rules alone. A subject, unlike a proper noun, does not need to be capitalized to be detected.

To begin making a subject list, **open all files that are to be indexed or open a book** and then select *Build Subject List...* in the *Word-Phrase List* menu as shown in figure 4.4. Figure 4.5 will appear. **For the demo version, only the first 25 subjects will appear.**

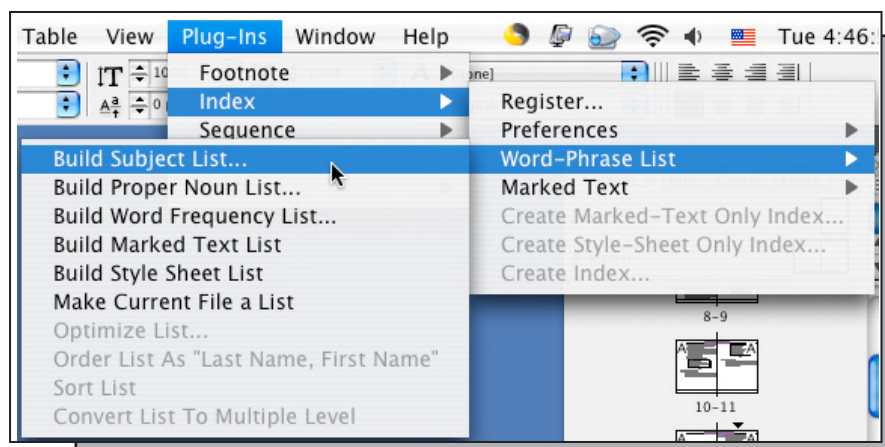


Fig. 4.4

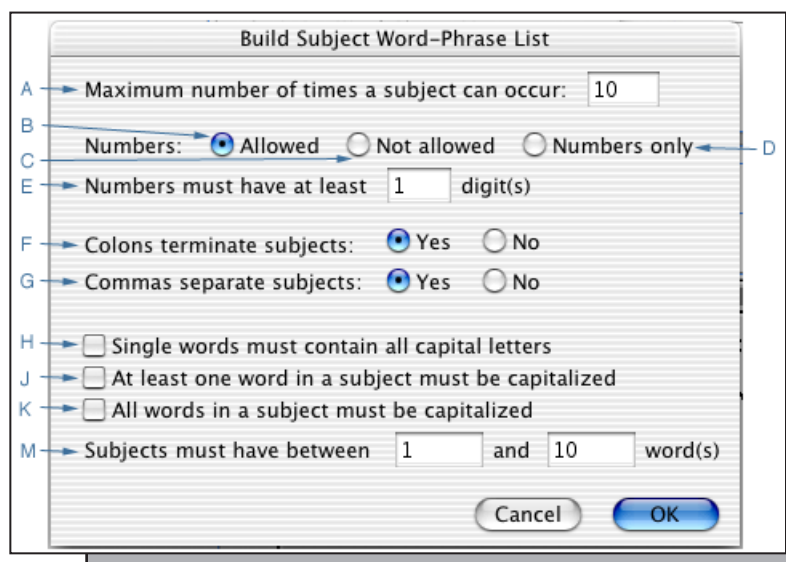


Fig. 4.5

Refer to figure 4.5. [A] eliminates any subjects that occur too frequently to be of any real value. The frequency can be set anywhere from 1 (for unique subjects) to 65535 (keeping every subject).

Numbers can be used to filter subjects. A number is any word that contains at least one value from the range 0-9 and optionally, other alphabetic characters. A digit is any character that makes up a number.

Select [B] if numbers are allowed to be part of a subject:

1234 West Broad Street  
34th Airborne  
Glock 27

If numbers are not allowed, select [C] and any subject that contains a number is removed. If subjects are to only contain single numbers, select [D]. This is useful for extracting part numbers from a document.

As a final numeric filter, the total digits required for a number to be considered valid is set by [E]. Any subject containing a number with less than this number of digits is removed.

Normally when a sentence is parsed by Sonar Bookends InDex Pro, colons and commas mark the end of a subject. This behavior can be changed with [F] and [G], respectively. For example, by changing both [F] and [G] to “No”, biblical references, like Eph. 1:22-23 can be found and put into a subject list. For this kind of application, unchecking Smart hyphens in the Special menu is also useful (see page 32). The resulting subject list may require some additional “cleanup” when changing these options.

A subject can be a single word. Each letter of a single-word subject can be required to be capitalized by selecting [H]. This limits single-word subjects to acronyms like NRA, GOA, USAF, etc.

Subjects don’t necessary need to have any capitalized letters. Selecting [J] requires that at least one word in a phrase be capitalized. [K] requires that every word (except articles, conjunctions, and prepositions) be capitalized. To find proper nouns only, select [C] and [K].

As a final filter, phrases can be eliminated based on the number of words in the phrase as defined by [M].

Click the *OK* button when you have made your choices. A subject list will appear in the Word/Phrase List window as shown in figure 4.6.



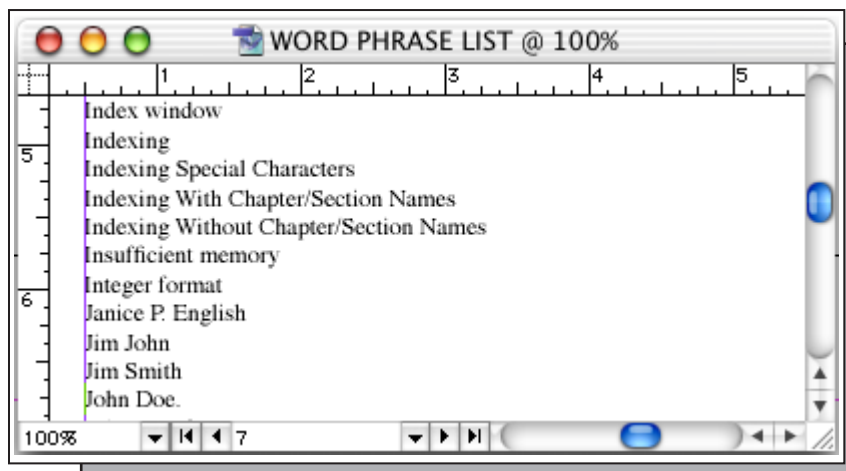


Fig. 4.6

## Proper Noun List

Sonar Bookends InDex Pro can create a list of proper nouns found in a document. The proper nouns are found using rules of capitalization and punctuation. Proper nouns which appear too often (above a user defined maximum number of times) are eliminated. **For the demo version, only the first 25 proper nouns will appear.** To generate a proper noun list, **open all files that are to be indexed** and then select *Build Proper Noun List...* in the *Word-Phrase List* menu as shown in figure 4.7 and the dialog box shown in figure 4.8 will appear.

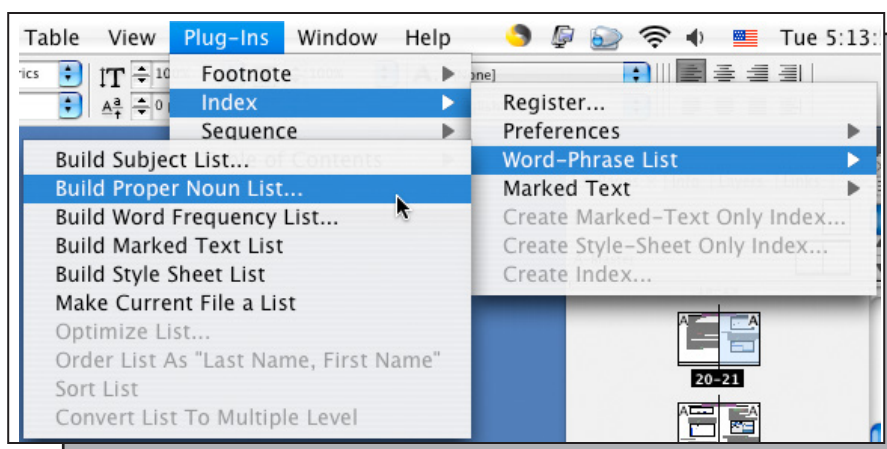


Fig. 4.7

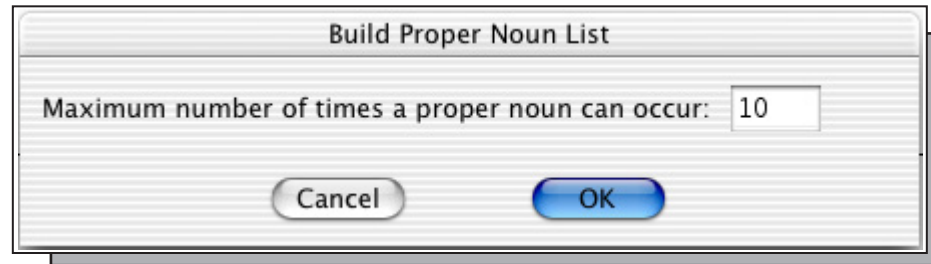


Fig. 4.8

Set the maximum number of times that a proper noun can occur. Setting this frequency acts like a filter, eliminating any proper nouns that occur too frequently to be of any real value. The frequency can be set anywhere from 1 (for unique proper nouns) to 65535 (keeping every proper noun). Sonar Bookends InDex Pro compiles and sorts the proper noun list when the *OK* button is clicked. This list is automatically displayed in a new “Word Phrase List” window as shown in figure 4.9. The word/phrase list can then be edited or saved, as desired.

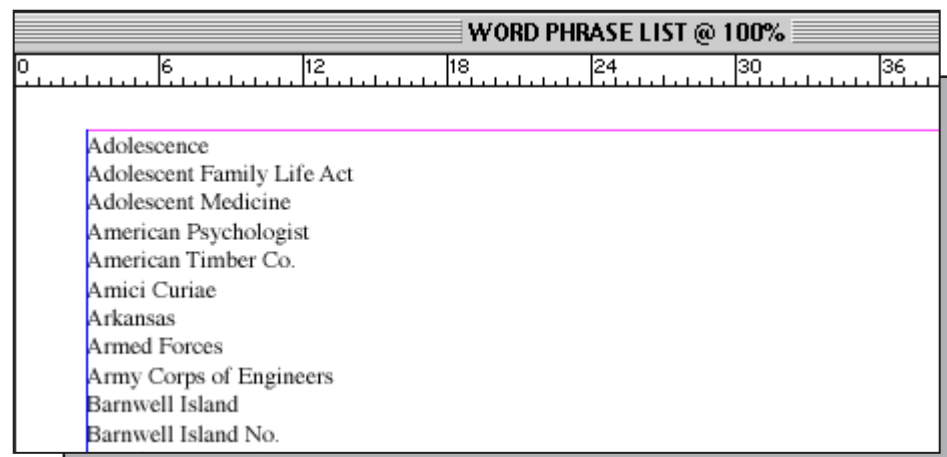


Fig. 4.9

## Generic Word/Phrase List

Besides using Sonar Bookends InDex Pro to create word-frequency, proper noun, and subject lists, a word/phrase list can be generated manually using InDesign or imported from a database, spreadsheet or a word processor file.

A word/phrase list can contain both individual words, combinations of words (phrases), and entries which incorporate advanced indexing techniques such as Boolean operations and wild cards. The word/phrase list can be either single-level or multiple-level, depending upon the type of index to be made. See Section VI, “Special Features” on page 41 for using Boolean operators, wild cards, reversing first and last names, optimizing plurals and overlapping phrases, and automatically converting single-level word/phrase lists to multiple-level.

Each word or phrase needs to be on a separate line. By default, Sonar Bookends InDex Pro indexes the words or phrases exactly as they are found in the word/phrase list, ignoring punctuation and spacing. To change the default values for word order, proximity, and punctuation, see Section VI, “Special Features” on page 41.

There is no practical limit set on the size of a word/phrase list, but InDesign may need more memory for extremely large lists.

To use the generic word/phrase list, create an empty document in InDesign, “place” the word/phrase list in the document, and, with that document as the front-most document, mark it as a word/phrase list using *Make Current File a List* as shown in figure 4.10.

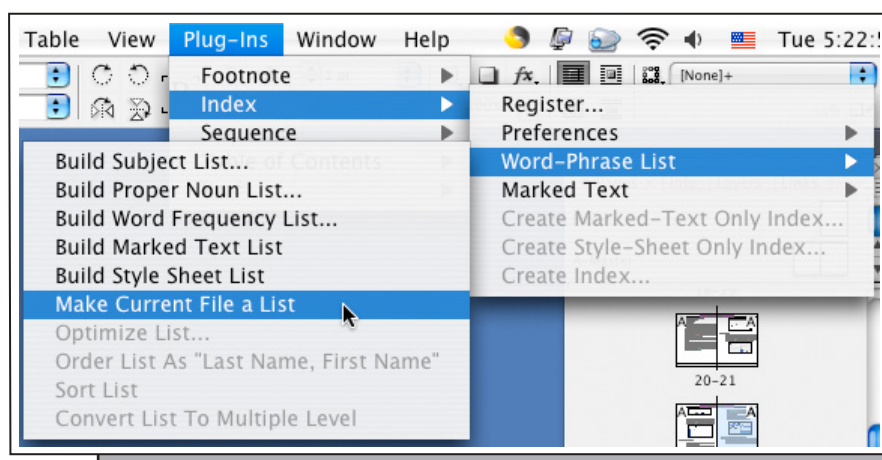
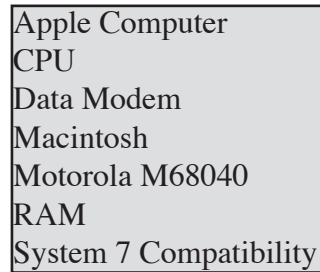


Fig. 4.10

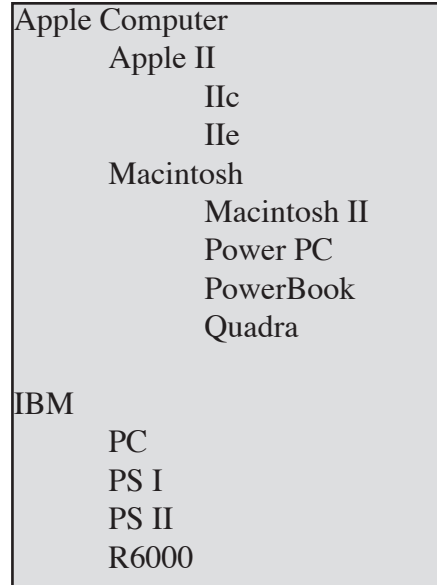
## Single- and Multiple-Level Word/Phrase Lists

Part of a single-level word/phrase list is shown in figure 4.11. Each entry is against the left margin, and is on a separate line. A multiple-level word/phrase list is shown in figure 4.12. Each entry in a multiple level word/phrase list is on a separate line and optionally indented with one or more tab characters. Only the most indented entries at any given level will receive page numbers when the list is indexed.



Apple Computer  
CPU  
Data Modem  
Macintosh  
Motorola M68040  
RAM  
System 7 Compatibility

Fig. 4.11



Apple Computer  
Apple II  
IIc  
IIe  
Macintosh  
Macintosh II  
Power PC  
PowerBook  
Quadra  
IBM  
PC  
PS I  
PS II  
R6000

Fig. 4.12

## Creating a Word/Phrase List Using a Database

For some applications, like indexing a catalog, the items to be indexed may be contained in a database. In this case, the word/phrase list can be generated directly by exporting the desired database fields into a text only file, one item per line. Separate indices for part names, part numbers, and part descriptions can be generated easily by creating three different word/phrase lists. Depending on available memory, Sonar Bookends InDex Pro can handle huge indices. For more information on exporting database files, refer to the user's guide provided with the database.

For catalogs, it is often advantageous to have Sonar Bookends InDex Pro look-up the page numbers based on a part number. The Translation operator is used for this type of application. (See "Using the Translation Operator" in Section VI, "Special Features," on page 41) For example, a word/phrase list entry like this:

[Computer Keyboard, Mac;00234-4521A](#)

would put the description "Computer Keyboard, Mac" into the index, but would find the page numbers by looking for "00234-4521A." The resulting index entry would look like this:

[Computer Keyboard, Mac, 56](#)

To use the database generated word/phrase list, make sure that the word/phrase list is the front-most document, then mark it as a word/phrase list using *Make Current File a List* as shown in figure 4.10 on page 23.

## Using Marked Text to Create a Word/Phrase List - Manual

Words and phrases in an InDesign document may be ‘marked’ manually for inclusion in an index. To mark text, drag over the desired word or phrase and select *Mark Text* in the *Index* menu. Marked text is highlighted in blue. Note that InDesign allows you to assign a function key to this (or any other) menu item. See figure 4.13.

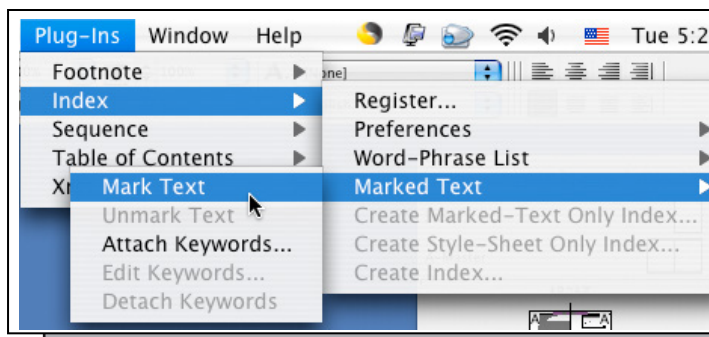


Fig. 4.13

**Note:** There are two marking modes. Regular mode marks all of the selected text as a single mark, but will not mark past a carriage return. List mode lets you mark a list of objects and splits each item in the list into separate marks automatically - a real time saver. The items in a list are separated either by commas or by carriage returns. For example highlighting:

John Smith, John Doe would produce John Smith, John Doe after selecting *Mark Text* with list marking enabled.

To enable list marking, check *Mark Lists* in the *Preferences* menu.

Once all of the text is marked, the actual word/phrase list is made using *Build Marked Text List* in the *Word-Phrase List* menu as shown in figure 4.14.

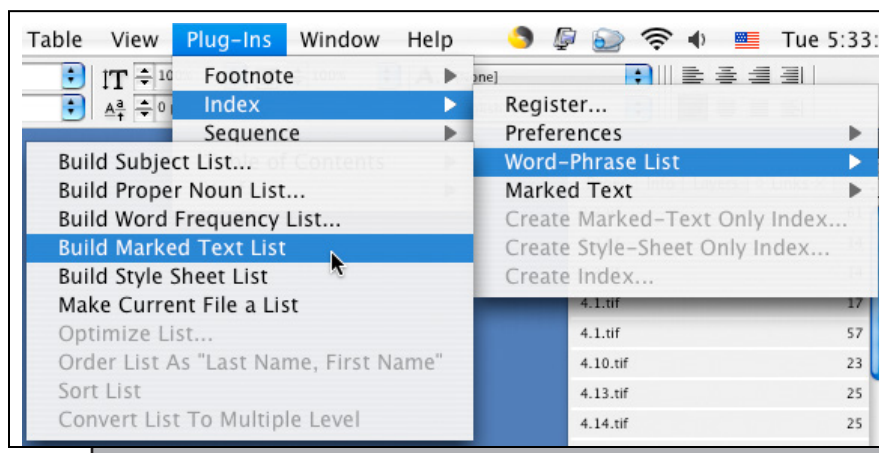


Fig. 4.14

The highlighting of marked text can be enabled or disabled using *Show Marks* in the *Preferences* menu as shown in figure 4.15.

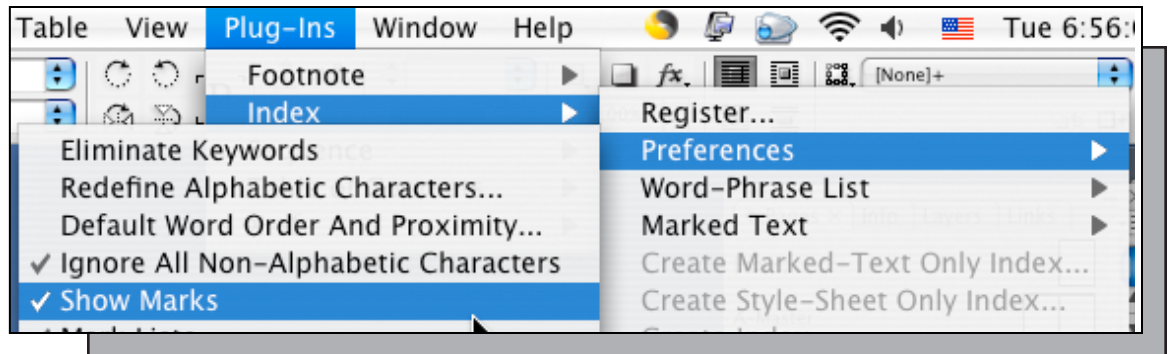


Fig. 4.15

## Using Marked Text to Add Hidden Keywords to a Document

Hidden text, called *keywords*, can be attached to text in a document and indexed. This facilitates indexing subjects which don't actually appear in a document. The hidden keywords can be put in a word/phrase list using the *Build Marked Text List* command, just like normal marked text. However, hidden keywords are put in the list with square brackets on both sides of the words. A semicolon allows the keywords to be split into separate lines arbitrarily.

For example, if the hidden keyword string was:

United States;United Kingdom

The marked text list would have the following entries:

[United States]  
[United Kingdom]

To attach hidden keywords to text, highlight the appropriate text and then selected *Attach Keywords...* in the *Marked Text* menu as shown in figure 4.16.

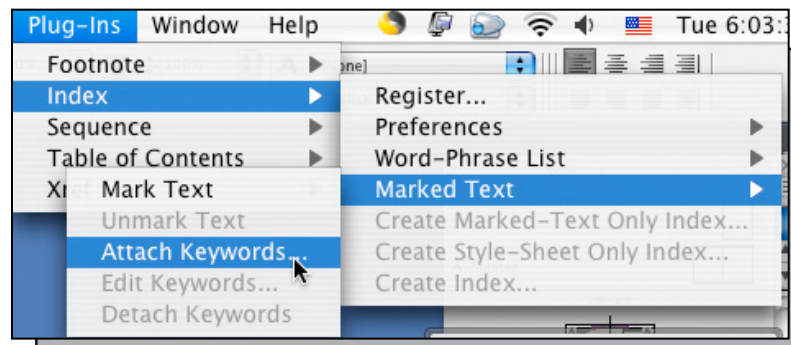


Fig. 4.16



The dialog box shown in figure 4.17 will appear to allow the keywords to be entered.

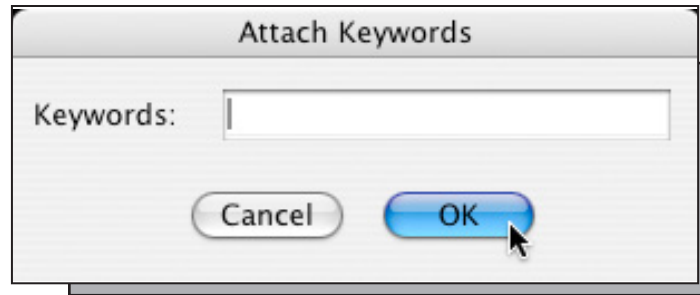


Fig. 4.17

## Using Marked Text to Create a Word/Phrase List - Style Sheets

Words and phrases in an InDesign document can be ‘marked’ for indexing using one or more style sheets. The style sheets can be either paragraph or character style sheets and must have “WPL0” as the first 4 characters of the style sheet name. For example, a valid style sheet name might be “WPL0 Part Numbers.” Existing style sheets can be used to mark text by simply changing their name to add “WPL0” to the front of the name.

Once all of the desired style sheets are properly named and the text is marked, the actual word/phrase list is made using *Build Style Sheet List* in the *Word-Phrase List* menu as shown in figure 4.18.

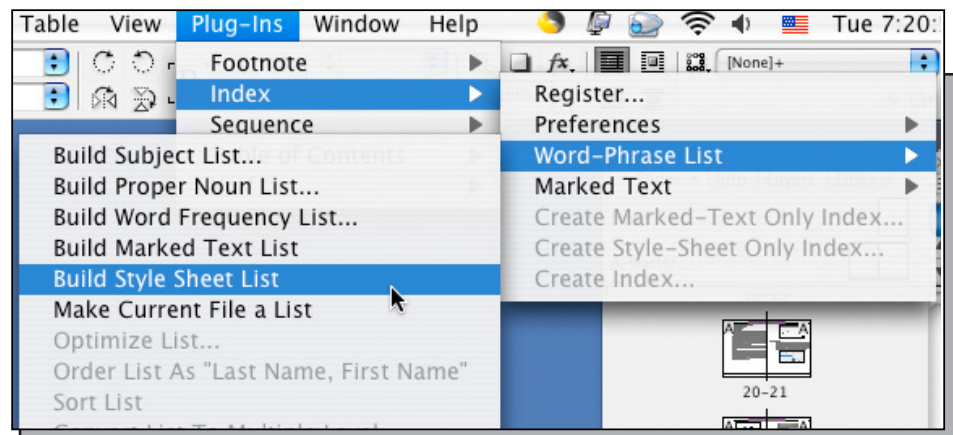


Fig. 4.18

## Sorting a Word/Phrase List

By default, Sonar Bookends InDex Pro assumes that word/phrase lists are already sorted and does not waste time resorting them. If a word/phrase list (either single-level or multiple-level) needs to be sorted, it must be sorted before making the index. **Note that sorting a single-level list removes duplicate entries.** To sort the word/phrase list, **make sure that the word/phrase list is the front-most window** and then select *Sort List* in the *Word-Phrase List* menu as shown in figure 4.19.

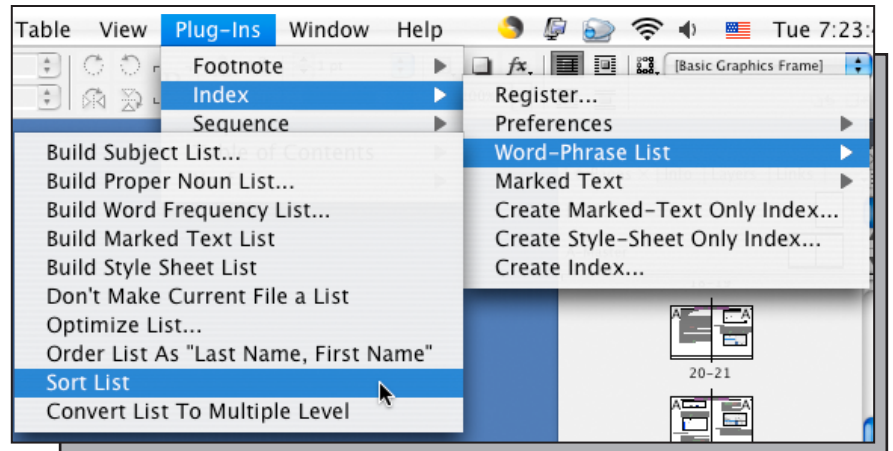


Fig. 4.19

Certain words, such as articles, do not affect the sorting order. For example, “The United States” is sorted under “U” and not “T.” Which words are ignored during sorting is controlled by the file named “SBISORTIGNORE.TXT”. The file is Unicode text-only and can be edited using InDesign. To edit the file, create a new InDesign document with an empty text box. Next, “place” the file into the text box. The text can now be edited as necessary. To save the modified text, “export” it as “text only” and, when prompted, save it with an encoding of “Unicode.” For the changes to take effect, quit and restart InDesign.

The collating sequence, or the order for sorting letters, is controlled by another Unicode text-only file, “SBISORTSEQUENCE.TXT”. The order that letters appear in the file is the order in which they are sorted. Letters that are on the same line together are sorted interchangeably. For example,

Aa  
Bb  
Cc

defines the normal first three letters of the alphabet. Since “B” and “b” are on the same line, they are sorted as if they are the exact same letter and both would appear interchangeably between “A” and “C.”

To edit the sorting sequence file, create a new InDesign document with an empty text box. Next, “place” the file into the text box. The text can now be edited as necessary. To save the modified text, “export” it as “text only” and, when prompted, save it with an encoding of “Unicode.” For the changes to take effect, quit and restart InDesign.

Word lists with entries that use the translation feature (see page 47) receive special treatment when sorted: if the word list has multiple entries which are duplicates except for their translation values, then the duplicate entries are removed and their translated values are combined using a Boolean “or” so that the single remaining entry will have all the page numbers, including those for the removed duplicates.



---

For example, if a word list looks like this (product name;part number to search for):

Computer; 2408  
Computer; 8102  
Computer; 10231  
Desk lamp; 1023  
Desk lamp; 1044

If left unsorted, the list word produce an index like this:

Computer, 86  
Computer, 92  
Computer, 93  
Desk lamp, 12  
Desk lamp, 14

If sorted, however, the list would have duplicates removed and the part numbers merged:

Computer; 2408 | 8102 | 10231  
Desk lamp; 1023 | 1024

Producing an index that would look like this:

Computer, 86, 92, 93  
Desk lamp, 12, 14



## Section V Making an Index and Table of Contents

To begin making an index, make sure that all the files that are to be indexed are open. *If all of the file names are in an InDesign “book,” just open the book.* Only open files or books are indexed. Also, if the word/phrase list is not already open, open it and make sure it is the **front-most window**. Several word/phrase lists can be open at one time, but only the front-most is used to make the index, the others are ignored.

Select *Create Index...*, *Create Style-Sheet Only Index...*, or *Create Marked-Text Only Index...* in the *Index* menu, as shown in figure 5.1. *Create Index...* indexes all of the text on all of the pages. *Create Style-Sheet Only Index...* only indexes the text on each page that has been marked using one or more style sheets whose name starts with the letters “WPL0”. *Create Marked-Text Only Index...* only indexes the text on each page that has been marked using *Mark Text* in the *Index* menu.

**Note:** If *Create index...*, *Create Style-Sheet Only Index...* and *Create Marked-Text Only Index...* are dimmed, then the front-most window is not a word/phrase list or the word/phrase list has not been marked as such using *Make Current File a List* in the *Word-Phrase-List* menu as shown in figure 4.10 on page 23.

The Index dialog box shown in figure 5.2 controls the indexing process. It controls whether the index is single-level or multiple-level, what type of symbols are to be used to separate page numbers, how many blank lines should appear between entries, etc.

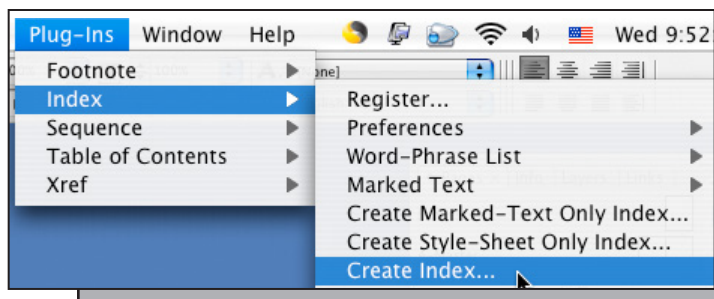


Fig. 5.1

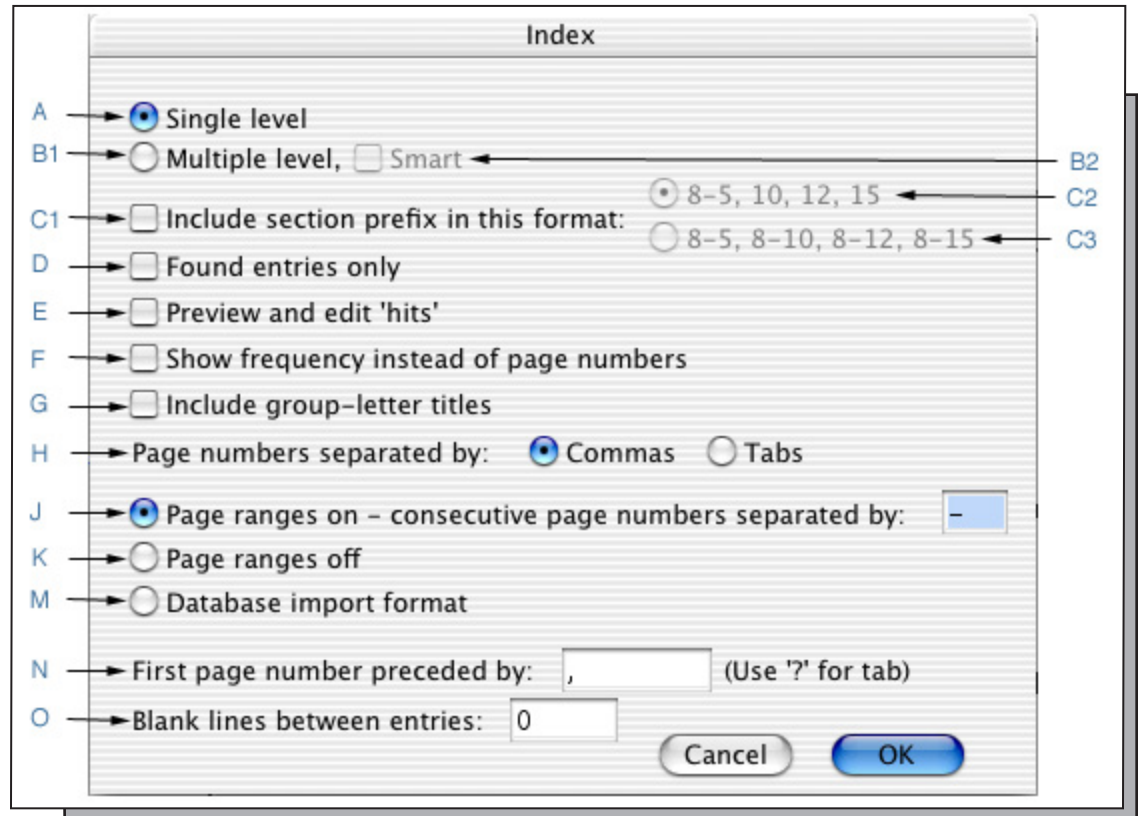


Fig. 5.2

## Handling Multiple Levels

Item [A] in figure 5.2 creates a single-level index and item [B1] a multiple-level index. If the word/phrase list is a single-level, then selecting item [B1] does not matter. For a multiple-level word/phrase list, selecting [B1] causes page numbers to appear next to each item, unless that item has one or more lower-level entries below it. If there are lower-level entries, only the lowest level entries will have page numbers. However, page numbers can be placed next to all entries in a multiple-level word/phrase list by simply not checking [B1]

If item [B2], “Smart,” is checked, then higher-level entries are used to help qualify page numbers. For example, with [B2] unchecked, the page numbers for *Papillon* and *Pekinese*, below, would be found by only looking for the words *Papillon* and *Pekinese*, respectively.

Dogs  
     Toy  
         Papillon  
         Pekinese

Checking [B2] would cause the page number for *Papillon* to be found by looking for *Dogs* on the same page as *Toy* on the same page as *Papillon*. Page numbers for *Pekinese* would be found by looking for *Dogs* on the same page as *Toy* on the same page as *Pekinese*.

**Note:** If you used *Convert List to Multiple Level* in the *Word-Phrase List* menu, do not check [B2].

Normally higher-level entries are followed by at least two lower-level entries. If there is only one lower-level entry (which is called an *orphan*), that entry is automatically combined with the higher-level entry at index time.

This example, which contains the orphaned line *Intel*:

Computers,  
Intel

would look like this after being indexed:

Computers, Intel, 105, 110

Single-level and multiple-level indices are discussed in more detail on page 5.

## Specifying Section Prefixes

If an index is to contain both section prefixes and page numbers, such as “2-1,” “6-12,” etc., then the *Include section prefix in this format* option in item [C1], figure 5.2, must be checked. Otherwise, the index will contain a series of consecutive page numbers without section prefixes.

There are two options for formatting section prefixes:

1. Option [C1] in figure 5.2 - each section prefix is placed at the beginning of a line, with all of the page numbers for that section following the section prefix:

1-6, 8, 10, 22  
2-10, 55

2. Option [C2] in figure 5.2 - the page numbers are strung together and each page number is prefixed with the associated section prefix:

1-6, 1-8, 1-10, 1-22, 2-10, 2-55

## Handling Non-Existent Entries

Normally if an item in the word/phrase list is not found, it is flagged with five asterisks (\*\*\*\*\*) in the index. However, if item [D] is checked then any entries which cannot be found are not placed in the index at all.

Use caution with this option. Normally it is useful to know that an item is not found so that the situation can be fixed. Many simple things can cause an item to not be found:

- One or more of the words in the entry are misspelled.
- The words in the entry are not in the same order that they appear in the document being indexed. This can be fixed by indicating that word order is not significant using *Default Word Order And Proximity...* in the *Preferences* menu. See page 13 for more information on changing the default word order and proximity.
- The entry truly does not exist.
- The entry contains special symbols that are used by Sonar Bookends Professional:

&:~|/\^+\*{}()!="#;

This can be fixed by checking the *Ignore All Non-Alphabetic Characters* item in the *Preferences* menu or by replacing the symbols with spaces.

If item [D] is not checked, then all entries that are not found are summarized at the beginning of the index, so that they can be more easily identified:

-----Index entries that were not found-----

Smithfield, John  
Young, Michael

-----End of entries that were not found-----

## Previewing and Editing ‘Hits’

Checking item [E] causes each occurrence of each entry in the word/phrase list to be displayed in one line of context. Each occurrence can then be either included or excluded from the index by selecting or deselecting the occurrence, respectively. This is a very powerful indexing tool that is discussed in detail in Section VI, “Special Features,” on page 41.

## Displaying Frequency Instead of Page Numbers

Caution: By checking item [F], each item in the index will be followed by its frequency instead of by the page numbers where the item is found. This option is used by people doing linguistic studies. Unless you have a specific reason to use this option, do not use it!

## Including Group-Letter Titles

Item [G] in figure 5.2 is used to break index entries into separate groups. The entries in each group have the same first letter, and that letter is put at the top of the group as a title:

A
Aardvark
Albatross
Arctic winters
B
Bavaria
Basic weather prediction
“Big-Bang” theory
The bust of Augustus
...

Notice that in the last two entries, above, leading punctuation and articles are ignored when breaking the index into groups.

## Separating Non-Consecutive Page Numbers

Item [H] in figure 5.2 affects the separation of non-consecutive page numbers. Either commas or tabs can separate page numbers. Examples of the two methods of separation are shown in figure 5.3.

Thunderbird, 112, 189, 201, 256	Thunderbird,   112   189   201   256
<i>Comma Separators</i>	<i>Tab Separators</i>

Fig. 5.3

## Abbreviating Consecutive Page Numbers

Consecutive page numbers (45, 46, 47, 48, 49) are abbreviated (45-49) by default. Item [J] in figure 5.2 enables abbreviation and identifies the single character that separates consecutive page numbers. A hyphen is the default, but a colon or any other single character can be used. A colon or a tilde (~) are good choices if the index is to contain hyphenated chapter names (“1-,” “2-,” etc.). This avoids confusing chapter names with consecutive page numbers:

1-2, 4, 6-12, 15
------------------

The above example shows chapter 1, pages 2, 4, 6 through 12, and 15. It would be clearer as:

1-2, 4, 6~12, 15

Selecting item [K] in figure 5.2 disallows any consecutive page number abbreviations:

1-2, 4, 6, 7, 8, 9, 10, 11, 12, 15

Page numbers can be further abbreviated by checking *Maximum Consecutive Page Number Abbreviation* in the *Preferences* menu. This will cause the second page number in a range to have any redundant leading digits dropped.

For example, the following page ranges:

101-104, 126-131

would be further abbreviated to:

101-4, 126-31

## Using Database Import Format

Item [M] in figure 5.2 causes each entry in the index to be formatted for easy importation into a database. The database format consists of one line for every occurrence of every word/phrase being indexed. Each line consists of the word/phrase being indexed and one page number. A word/phrase is repeated if it is on more than one page:

Computers, 5  
Computers, 12  
Computers, 13  
Computers, 14  
Dragging, 19  
Dragging, 46

## Separating Index Entries and Page Numbers

The characters between each word or phrase and its first page number are controlled by item [N] in figure 5.2. The default value is a comma and a space, but any sequence of characters can be used. With the default, the index entry for “Nuclear energy” would look like this:

Nuclear energy, 104, 110-114, 203



Changing item [M] to be “===” would change the index entry to look like this:

Nuclear energy === 104, 110-114, 203

The question mark (?) is a special character which is used to represent a tab character. Thus, if item [M] was set to “??” there would be two tabs between each word or phrase and its first page number:

Nuclear energy        104, 110-114, 203

## Setting Blank Lines Between Entries

Item [O] in figure 5.2 determines how many blank lines are inserted between index entries. A value of zero will put one entry immediately below the other:

Nuclear energy, 104, 110-114, 203  
Reactor, 33, 45

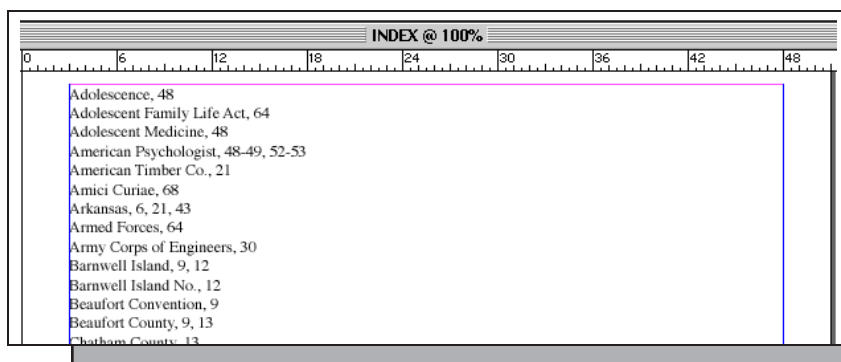
A value of 2 will place two blank lines between each entry:

Nuclear energy, 104, 110-114, 203  
  
Reactor, 33, 45

When all desired options in figure 5.2 have been chosen, click the *OK* button to begin generating the index.

When the index is complete, it appears in the Index window as shown in figure 5.4.

**Note:** If there are any mistakes discovered in the final index, the word/phrase list should be corrected and the index remade. This is important because if only the index is corrected, then each time the index needs to be made again, so do the changes!



*Fig. 5.4*

## Formatting the Index

The ability to easily format an index is new with version 7.0 of Sonar Bookends InDex Pro.

When an index is made, certain style sheets are dynmamicly created and automatically used to mark various parts of the index. Formatting the index is accomplished by modifying these style sheets to have the properties that you desire.

The only style sheets that are generated are those that are necessary to format a given index.

Here is a list of the style sheets that can be generated and what they mark:

SonarIndexNumbers	Marks page numbers and associated punctuation
SonarIndexBodyLevel0	Marks index terms that are not indented
SonarIndexBodyLevel1	Marks index terms that are indented with one tab
SonarIndexBodyLevel2	Marks index terms that are indented with two tabs
SonarIndexBodyLevel3	Marks index terms that are indented with three tabs
SonarIndexBodyLevel4	Marks index terms that are indented with four tabs
SonarIndexBodyLevel5	Marks index terms that are indented with five tabs
SonarIndexBodyLevel6	Marks index terms that are indented with six tabs
SonarIndexBodyLevel7	Marks index terms that are indented with seven tabs
SonarIndexBodyLevel8	Marks index terms that are indented with eight tabs
SonarIndexBodyLevel9	Marks index terms that are indented with nine ore more tabs
SonarIndexTitle	Marks Group-Letter Titles

Below is a sample index showing **Group-Letter Titles**, **index terms that are not indented**, **index terms that are indented with one tab**, and **page numbers**. Color is used to show which text is marked and which style sheet is marking it.

A
Active
photosynthesis, 759
osmosis, 753, 757
Additive properties, 755
B
Black Plague, 444-446, 602,
903, 1024

## Making a Table of Contents

To make a table of contents, you must indicate which parts of a document make up the titles, subtitles, etc. This is done using one or more paragraph style sheets. The style sheet's name is what tells Sonar Bookends InDex whether or not a block of text is to be part of the table of contents. The name can be anything you want, but it must start with the characters "TOC" (Table Of Contents) followed by a single digit number from 0 through 9. For example "TOC0Header style" would be a valid table of contents style sheet name.

The digit following the letters "TOC" tells Sonar Bookends InDex how many tab characters to place in front of the text marked with that style sheet. Thus titles would be prefixed with "TOC0" and subtitles with "TOC1". You can have as many table of contents style sheets as you want, including multiple style sheets for any given level.

The table of contents is generated in the same order as the titles and subtitles appear in the document. Each line contains only one page number. The page number is preceded by a tab character and, optionally, the "Section marker" text (referred to as the section prefix).

Once the style sheet names are correct, use either *Make Table of Contents (With Section Prefix)* or *Make Table of Contents (Without Section Prefix)* in the *Table of Contents* menu as shown in figure 5.5. The table of contents is automatically placed in a new window.

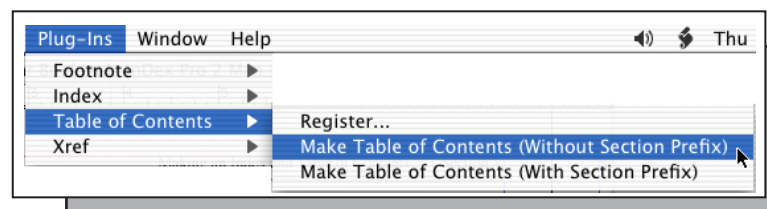


Fig. 5.5

Set some tabs stops, perform any other desired formatting operations and the table of contents is done!



## Section VI      Special Features

Sonar Bookends InDex Pro provides many advanced features for creating an index:

- Each occurrence of each phrase being indexed can be viewed in context to determine if that occurrence should be in the index.
- Wild cards and Boolean expressions can be placed in a word/phrase list to more precisely control indexing.
- “Translation” can “hide” the Boolean expressions and wild card characters so that they do not appear in the final index. It can also be used to add cross-references and notes to an index.
- First and last names in a word/phrase list can be reversed automatically taking into account titles and two-word last names.
- “Noise” words can be eliminated from an index.
- Singular and plural entries for the same subject can be combined.
- Redundant page numbers for subjects that overlap can be removed.
- A single-level word/phrase list can be converted into a multiple-level word/phrase list automatically.
- The list of characters that are considered alphabetic can be redefined.
- Text marked with certain style sheets can be ignored.
- Terms split across two pages can be detected.

### Previewing and Editing ‘Hits’

Without Sonar Bookends InDex Pro, determining which occurrences (hits) should be indexed is a very time consuming chore. The determination is usually made by manually finding each occurrence of each word/phrase in the document and then studying how it is being used. Sonar Bookends InDex Pro has a powerful feature that greatly cuts down on the time to do this kind of verification. Refer to figure 6.1. If item [A], Preview and edit ‘hits,’ is checked, then for each word/phrase being indexed, Sonar Bookends InDex Pro displays a dialog box containing one line of context for every occurrence of that word/phrase. Occurrences can then be selected or deselected for indexing quickly and easily, with all the information needed to make the decision at hand.

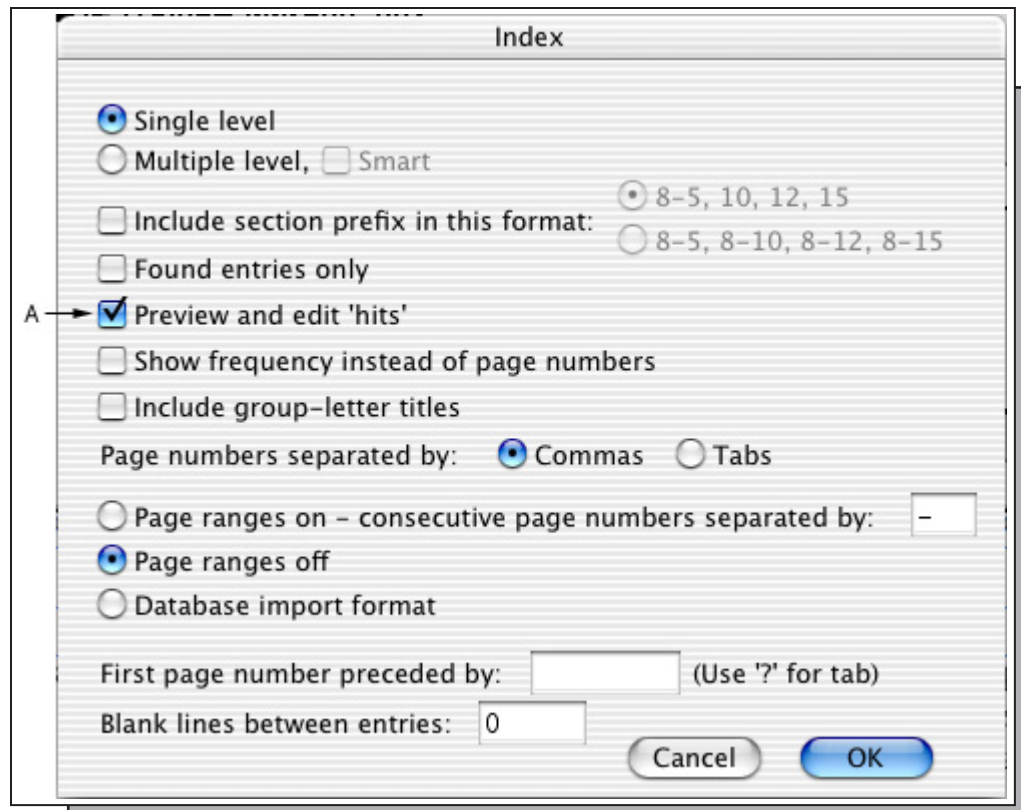


Fig. 6.1

Figure 6.2 shows the occurrences of the phrase “Active Window” (the current phrase is shown in the title bar [A]). [B] points to the four lines of context found for the phrase. The phrase is centered if possible. All highlighted occurrences are to be indexed.

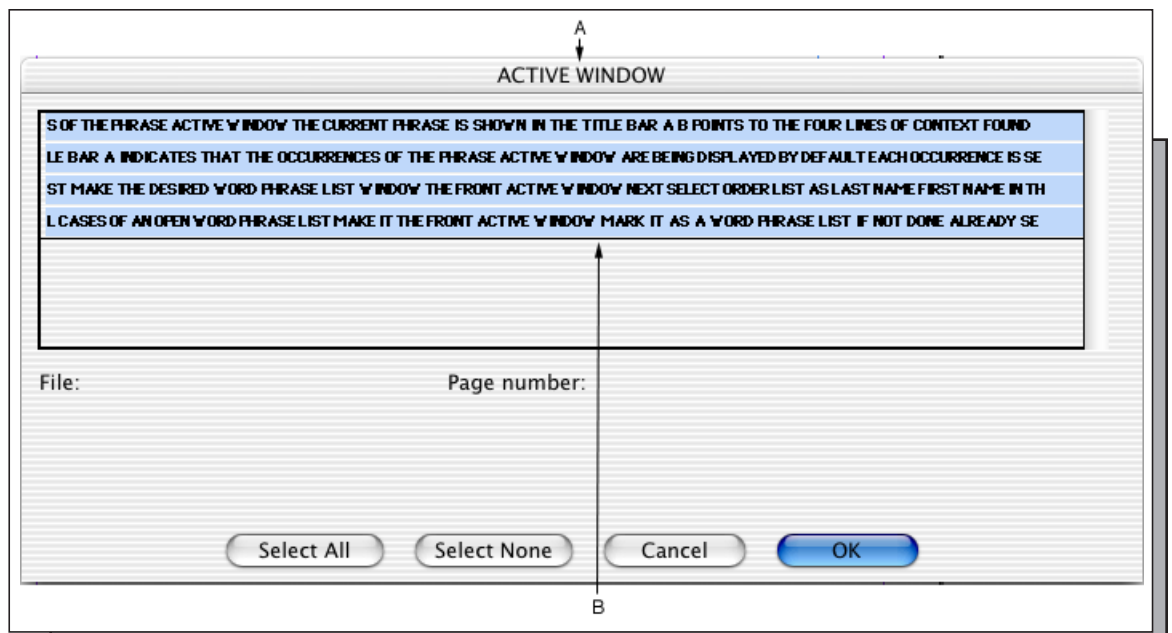


Fig. 6.2

In the dialog box shown in figure 6.2, the title bar [A] indicates that the occurrences of the phrase “Active Window” are being displayed. By default, each occurrence is selected (highlighted) and if possible, the phrase is centered on each line [B]. Clicking the *OK* button at this point will index all of the occurrences.

**Mac users:** To alternately select/deselect an occurrence, hold down the command key while clicking on the occurrence.

**Windows users:** To alternately select/deselect an occurrence, hold down the option key while clicking on the occurrence.

If the majority of the entries are to be deselected, click the *Select None* button and then select the desirable entries. *Select All* will select each occurrence again. Remember: only the occurrences which are highlighted when the *OK* button is clicked will be indexed.

To see more context for a given occurrence, simply click on the occurrence — an expanded context will appear in the window at the bottom of the dialog box. In figure 6.3 the expanded context, [A], is for the second line. To help in finding the occurrence manually, the associated file name and page number are shown at [B] and [C], respectively. If an entry to be viewed is already highlighted, you will need to deselect and then reselect that entry.

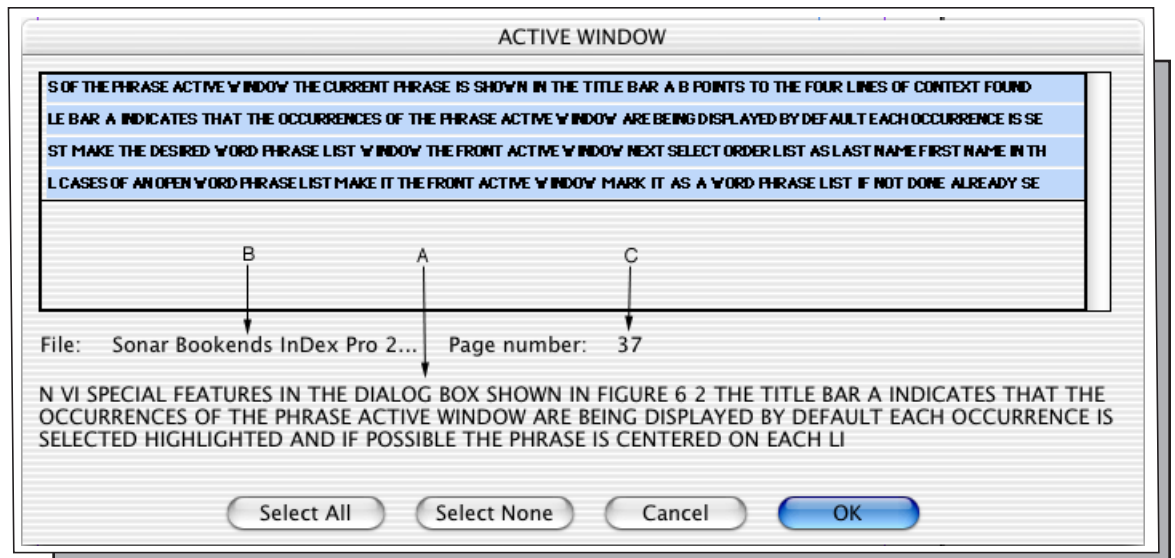


Fig. 6.3

When the *OK* button is clicked, Sonar Bookends InDex Pro displays the next word or phrase. This repeats until all words and phrases have been processed.

If a word or phrase appears too often (more than 50 times), then that entry is skipped. All page numbers for the offending word/phrase are placed into the index. For a word/phrase to appear so many times that it cannot be previewed, it is questionable if it should be included in the index anyhow.

When the last word/phrase has been processed, the selections are saved to disk for use with future indices. The previewed index appears in the Index window.

## Interrupting a Preview and Edit ‘Hits’ Session

Sonar Bookends InDex Pro allows a preview session to be interrupted. To interrupt, simply click the *Cancel* button in the dialog box shown in figure 6.3. A warning shown in figure 6.4 will appear.

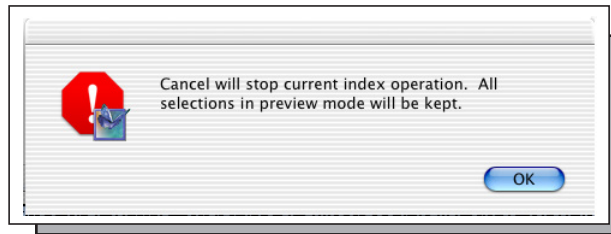


Fig. 6.4

All of the selections made up to the point of interruption are saved to disk. Whenever the index is made again, the preview session will pick-up at the point of the interruption.

## Rebuilding a Previously Previewed Index

As mentioned earlier, when a preview session is complete, the occurrence choices are saved for future use. This comes in handy if an index must be remade. To apply the previewed choices to a new index, make sure that Preview and edit ‘hits’ is selected as shown in figure 6.1. Upon detecting the saved preview information, the dialog box shown in figure 6.5 appears.

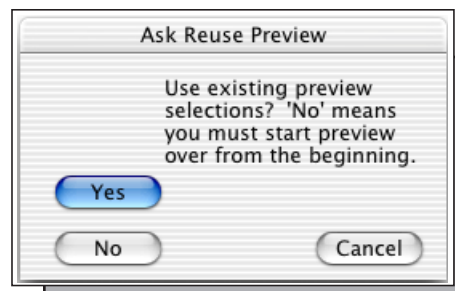


Fig. 6.5

By selecting *Yes*, the same choices made for each occurrence are applied again automatically, even if the page numbers have changed. However, if the number of occurrences of a word/phrase has changed or the word/phrase is new or modified or if the preview is continuing after a previous interruption, then a dialog box like the one shown in figure 6.2 appears to allow choices for that word/phrase.

**Caution:** Clicking *No* in figure 6.5, causes all of the saved choices to be discarded permanently and selection starts over with the first word/phrase entry.



## Wild Cards in Word/Phrase Lists

The asterisk (“\*”) is used as a wild card character. When making an index, a word is considered a match to a wild card if all characters up to the asterisk match exactly. For example, to index all words that start with son, enter son\*: son, SONAR, and song would all be matches, but not reason or resonant. There can be any mixture of wild cards and non-wild cards in a phrase. For example, *computer disk\* file\** would index *Computer diskette file* or *computer disks filed*.

See “Using the Translation Operator” on page 47, for information on how to “hide” wild card characters so that they do not appear in the index.

## Boolean Expressions in Word/Phrase Lists

Elaborate indices can be created using a simple but powerful index syntax to create a Boolean expression. A Boolean expression can contain any combination of phrases, operators, modifiers and parentheses. The basic syntax of a Boolean expression is defined as:

Phrase [operator [modifiers] Phrase]

(Bracketed ([]) items are optional)

### Phrases

A phrase consists of one or more words. Generally, punctuation in a phrase, such as periods and commas, are ignored. However, some non-alphabetic characters require special treatment (discussed later in this section). Here are some valid phrases:

John  
Jonathan Doe  
Doe, John  
United States of America

### Operators

Operators are used to perform Boolean arithmetic. The operators are:

John	&	Jim	Index each occurrence of John <u>and</u> Jim on the same page
John		Jim	Index each occurrence of either John <u>or</u> Jim
John	~	Jim	Index each occurrence of John only if it is <u>not</u> in the same file as Jim
John	^	Jim	Index each occurrence of John <u>only</u> if on the same page as Jim

### Modifiers

By default, the phrases on either side of an operator can appear anywhere in a file and can appear in any order. Modifiers are used to modify this behavior of operators:

John	&:15	Jim	Index John and Jim within 15 words of each other
John	&=	Jim	Index John and Jim on the same page, John must appear first
John	&:15=	Jim	Index John and Jim within 15 words of each other, John must appear first
John	~:5	Jim	Index John if not within 5 words of Jim
John	^:15	Jim	Index John <u>only</u> if within 15 words of Jim

## Notes on the ‘&’ and ‘^’ Operators

When unmodified, both the ‘&’ and ‘^’ produce the same result - the page numbers of the pages where the left and right arguments both appear. Thus, unmodified they can be used interchangeably. However, when modified with a proximity, they behave very differently and are not interchangeable.

Looking at how ‘&’ behaves with a proximity, the following two expressions would both produce the same page numbers:

John Doe &:100 Jim Smith  
Jim Smith &:100 John Doe

The page numbers are the same, because Sonar Bookends Professional looks for all locations where John Doe and Jim Smith are within 100 words of each other. Which phrase is found first does not matter. For each John Doe and Jim Smith that are within 100 words of each other, the page number is determined by the location of the first name that is found. If both names are on the same page, it makes no difference. However, if John Doe is on page 16 and Jim Smith is on page 17, the result of both expressions will be page 16 because that is where the first name, John Doe, is found. Putting Jim Smith on page 16 and John Doe on page 17 instead still produces page 16 as the result.

The ‘^’ operator behaves very differently. It returns the page numbers where the left argument appears as long as it is within the correct proximity to the right argument. Thus, the following two expressions can return very different results:

John Doe ^:100 Jim Smith  
Jim Smith ^:100 John Doe

Using the previous example of John Doe being on page 16 and Jim Smith being on page 17, the result of the first expression would be 16 and the second expression would be 17. Reversing the pages where John Doe and Jim Smith appear would reverse the results.

## Sample Boolean Expressions

John Doe	“John Doe” with “John” and “Doe” next to each other and in the specified order (standard default).
Reagan ~:15 (Ronald   Ron)	“Reagan” not within 15 words of either “Ronald” or “Ron.”
Adams &:20 Samuel	“Adams” within 20 words of “Samuel,” either word can come first. Same as: Samuel Adams:20!
Tab   Tabs   Tabbing	“Tab” or “Tabs” or “Tabbing.”
Convers* ~:1 conversation	Any word starting with “Convers” but not “conversation.”
John Doe:5! &:10 Jim Smith	“John” within 5 words of “Doe,” either word can come first, and both words within 10 words of “Jim Smith.”

See “Using the Translation Operator” on this page for information on how to “hide” Boolean expressions so that they do not appear in the index.

## Overriding Word Order and Proximity Defaults

Default word order and proximity values can not only be changed for all entries in a word/phrase list, as shown on page 13, but they can also be overridden for individual entries. Placing an exclamation sign (!) at the end of a phrase indicates that the words in the phrase can be in any order. The phrase “John Doe!” would match “John Doe,” “Doe, John,” etc. Adding a colon and a number to the end of a phrase sets the maximum number of words that can separate any two words in the phrase. For example, “John Doe:2” would match “John Doe,” “John E. Doe.” “John Doe:4” would match “John Edwin James Doe.” Both word order and proximity can be overridden at the same time: “John Doe:5!”

See “Using the Translation Operator,” below, for information on how to “hide” the word order and proximity override codes so that they do not appear in the index.

## Using the Translation Operator

### Hiding Text

The “translation” operator keeps Boolean expressions (on page 45), wild cards (on page 45), and default overrides from appearing in the final index. The symbol used for translation is the semicolon (;). Text appearing to the left of the semicolon appears in the final index, while text to the right of the semicolon is only used to find the page numbers. See the examples below:

Airplane(s);Airplane   Airplanes Boeing and Seattle;Boeing &:10 Seattle Cat;Feline Smith, John;Smith, John:15! Tacoma, Washington	Airplane(s), 6-7, 15, 33 Boeing and Seattle, 25 Cat, 22, 83 Smith, John, 92, 101-106 Tacoma, Washington, 59, 92
---	---

Word/Phrase List

Resulting Index

Note: In the above example, Tacoma, Washington does not require translation because Tacoma, Washington;Tacoma,Washington would be redundant. Also, translation does not require a Boolean expression on the right side. The phrase Cat;Feline places the word Cat in the index, but gets the page numbers by looking for the word Feline.

## Adding Cross-References and Blank Lines

Another use of the translation operator is to allow cross-references and blank lines to be inserted in a word/phrase list index. **Blank lines should not be added to an index which is going to be sorted.** To add a cross-reference to an index, place the comment on the left side of the semicolon and nothing on the right. To add a blank line, simply place the semicolon on a line by itself. Refer to the examples below.

Document Imaging Document Management (See Text Retrieval); ; Page Layout Word Processor	Document Imaging, 22 Document Management (See Text Retrieval)  Page Layout, 5, 7 Word Processor, 10-15, 19
---	--

Original List

After Indexing

## Reversing Names

Sonar Bookends InDex Pro can automatically reverse first and last names in an open word/phrase list. First, make the desired Word Phrase List window the front (active) window. Next, select *Order List As “Last Name, First Name”* in the *Word-Phrase List* menu, as shown in figure 6.6. Figure 6.7 shows a list before reversing names and figure 6.8 shows the same list after reversing names. Notice that a name that is already reversed is left unaltered, and that titles behind names and two-word last names are handled correctly. If any non-alphabetic characters, except period, are found in an entry, that entry is not reversed. If translation is being used, only the words up to the translation symbol (;) are reversed. *When reversing names, be sure and set the default word order to allow words in a phrase to appear in any order. Failure to do so causes most of the reversed entries to not be found. See page 13 for information on setting/ overriding default word order.*

**Note:** There is a shortcut for flagging names that are not to be reversed. Any name that has a ‘+’ as its first character is not reversed. Any leading ‘+’ characters are then stripped out of the reversed list automatically.

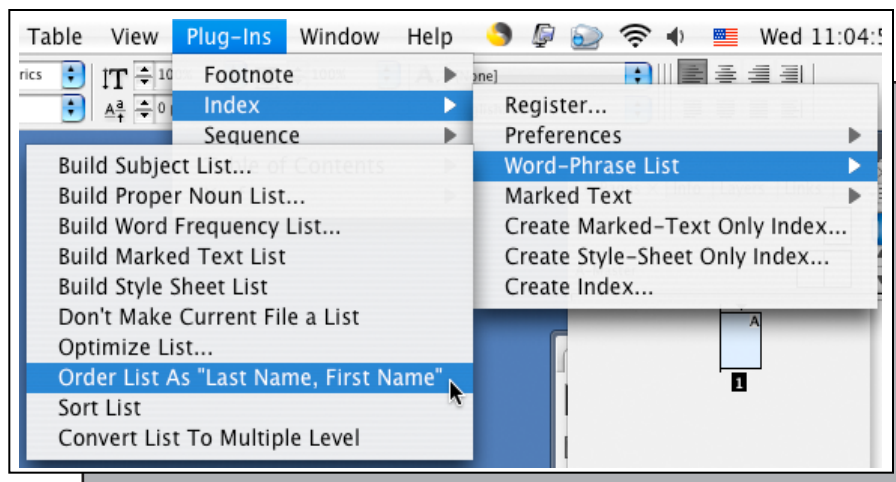


Fig. 6.6

John F. Doe Jr.  
Michael Del Gato  
Johnson, Michael  
Janice P. Jones  
Jill English, Ph.D.  
Greg Corbin  
Gabriel Lee  
David Lanz

Fig. 6.7

Doe, John F. Jr.  
Del Gato, Michael  
Johnson, Michael  
Jones, Janice P.  
English, Jill, Ph.D.  
Corbin, Greg  
Lee, Gabriel  
Lanz, David

Fig. 6.8

There are two files that control how Sonar Bookends InDex Pro reverses names. The first file, named “SBITITLES.TXT” and located in the System folder, contains a list of titles. The titles file list is only used to verify titles that are not preceded by a comma. In other words “John Doe, Jr.” would reverse correctly even if “Jr.” was not in the titles file. “John Doe Jr.” would only reverse correctly if “Jr.” is in the titles file.

The second file, named “SBIPREFIX.TXT” and located in the System folder, contains a list of words which are the first word of two-word last names. For example, if “Von” is in the prefix file, then all last names that have “Von” in them (like “Von Tree”) are considered two-word last names and reversed accordingly.

Both the “SBITITLES.TXT” and “SBIPREFIX.TXT” files are Unicode text-only and can be edited using InDesign. To edit a file, create a new InDesign document with an empty text box. Next, “place” the file into the text box. The text can now be edited as necessary. To save the modified text, “export” it as “text only” and, when prompted, save it with an encoding of “Unicode.” For the changes to take effect, quit and restart InDesign.

## Keyword Elimination

“Noise” words like “the,” “and,” “of,” etc. can be left out of a word frequency list by selecting *Eliminate Keywords...* in the *Preferences* menu. See figure 6.9. Keywords to be eliminated are contained in a text only file, one word per line. Any word frequency list created while *Eliminate Keywords...* is checked will not contain any of the keywords appearing in the elimination file. The file *SBIKWELIM.TXT* located in the *system* folder contains most of the common noise words.

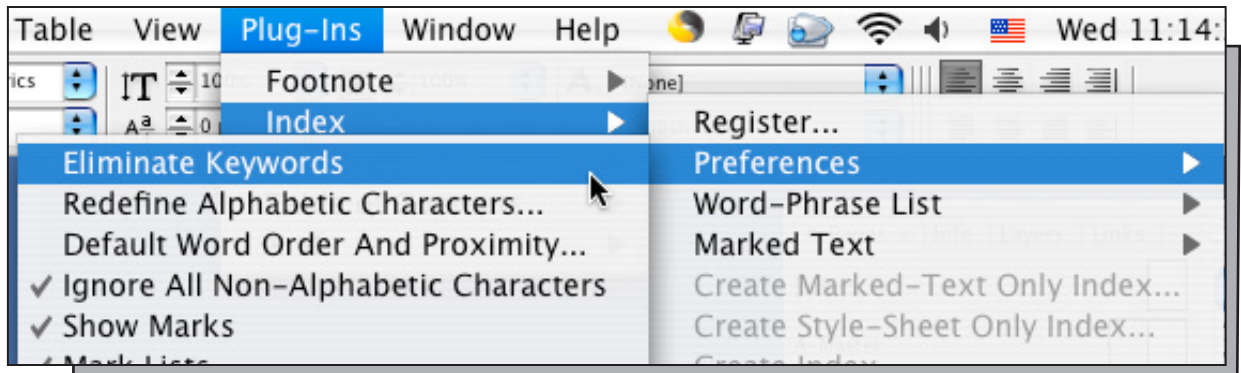


Fig. 6.9

Keyword elimination is only necessary when generating a word frequency index with a fairly large frequency. By specifying a small frequency, the high frequency noise words are eliminated anyway. In general, keyword elimination should not be used with word/phrase lists; as a phrase like “United States of America” would not be found if “of” was in the elimination file.

Selecting *Eliminate Keywords...* will alternatively activate and deactivate the eliminate keywords feature. Keyword elimination is activated when the menu item is checked.

The keyword elimination file is a Unicode text-only file and can be edited using InDesign. To edit the file, create a new InDesign document with an empty text box. Next, “place” the file into the text box. The text can now be edited as necessary. To save the modified text, “export” it as “text only” and, when prompted, save it with an encoding of “Unicode.” For the changes to take effect, quit and restart InDesign.

## Combining Singular and Plural Instances of a Phrase

Sonar Bookends InDex Pro has the ability to search an entire word/phrase list for entries that have both singular and plural cases, such as “Train” and “Trains”. Matching entries are combined “Train(s).” The redundant occurrences are then removed from the word/phrase list entirely. Indexing both the singular and plural cases is done automatically using translation and Boolean operations. Our example of train and trains would look like this after being combined:

Train(s);Train | Trains

The expression above puts the word “Train(s)” in the index, but gets the page numbers by looking for “Train” or “Trains.”

To combine singular and plural cases of an open word/phrase list, make it the front (active) window, mark it as a word/phrase list if not done already, select *Optimize List...* in the *Word-Phrase List* menu as shown in figure 6.10.

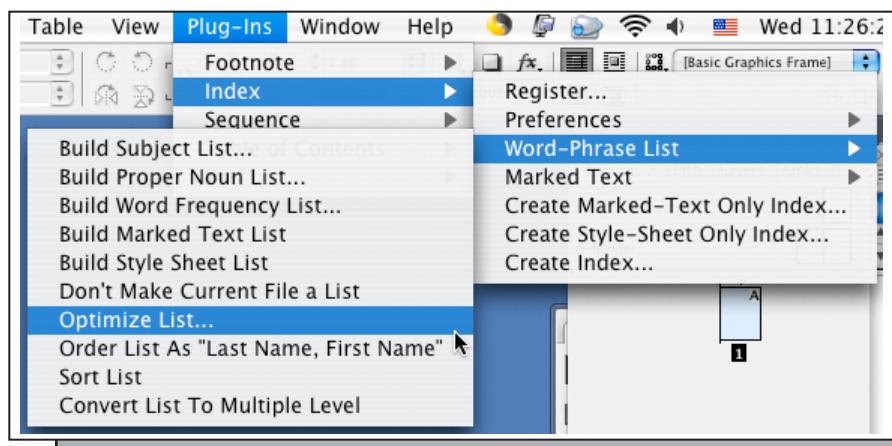


Fig. 6.10

When the dialog box shown in figure 6.11 appears, check only the *Singular/plural phrases* box and click the *OK* button. The optimized word/phrase list will then appear in the Word-Phrase List window.

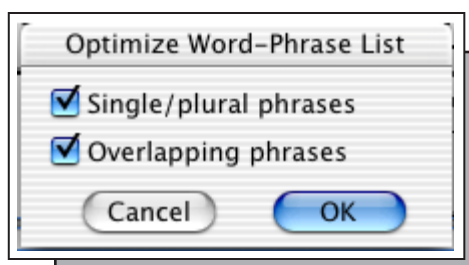


Fig. 6.11

Here is another example of how this optimization works:

#### Index of unoptimized list

Party, 44, 56  
Parties, 46  
Motor boat, 12  
Motor boats, 13-15

#### Index of optimized list

Party(ies), 44, 46, 56  
Motor boat(s), 12-15

**Note:** Optimization of word/phrase lists should be done prior to reversing first and last names or converting them to multiple-level.



## Removing Redundant Pages from Overlapping Phrases

Unless optimized for overlap, if one phrase (“Bear”) overlaps another phrase (“Black Bear” or “Bear Habitats”), then the page numbers for “Bear” will include the page numbers for the overlapped phrases as well. Here is a sample index with all overlaps underlined, except for “Brown Bears” where the overlap is highlighted in bold for clarity:

Bears, 14, 26, 33, 45, 90  
 Black Bears, 33  
**Brown Bears**, 14, 26  
Canadian Geese, 101, 126  
 Flight Patterns of Canadian Geese, 126  
 Mating Habits of **Brown Bears**, 26

After optimization to remove redundant page numbers, the index would look like this:

Bears, 45, 90  
 Black Bears, 33  
 Brown Bears, 14  
 Canadian Geese, 101  
 Flight Patterns of Canadian Geese, 126  
 Mating Habits of Brown Bears, 26

“Bears” now only shows page numbers where it is not part of the phrase “Black Bears,” “Brown Bears,” or “Mating Habits of Brown Bears.” Also “Brown Bears” only shows the page number where it is not part of the phrase “Mating Habits of Brown Bears.”

The overlap optimization is done automatically using translation and Boolean operations. Our example of “Bears” would look like this after being optimized:

Bears;Bears ~:1 (Black Bears|Brown Bears|Mating Habits of  
 Brown Bears)

The expression above puts the word “Bears” in the index, but gets the page numbers by looking for occurrences of “Bears” that are not part of the phrases “Brown Bears” or “Black Bears” or “Mating Habits of Brown Bears.”

To optimized overlapping cases in an open word/phrase list, select *Optimize List...* in the *Word-Phrase List* menu as shown in figure 6.12.



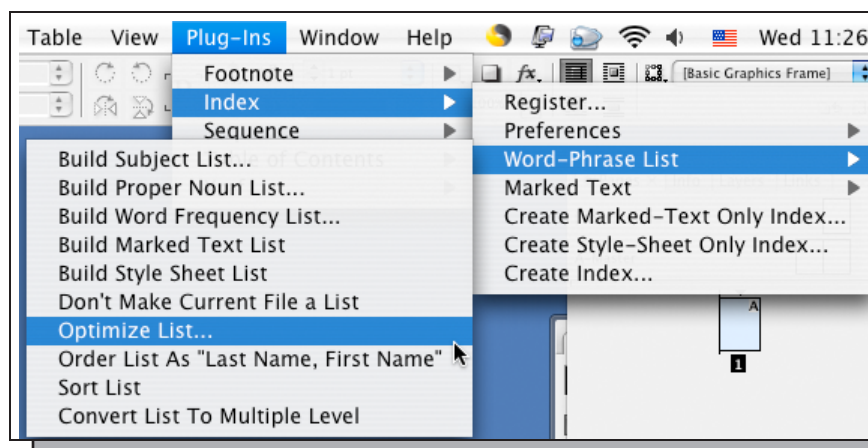


Fig. 6.12

When the dialog box shown in figure 6.13 appears, check the *Overlapping phrases* box and click the *OK* button. The optimized word/phrase list will then appear in the Word Phrase List window.

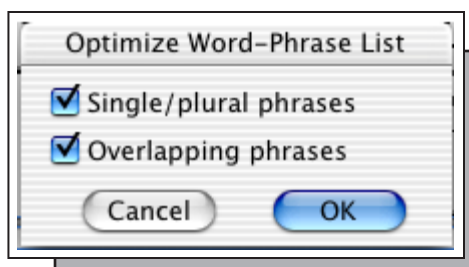
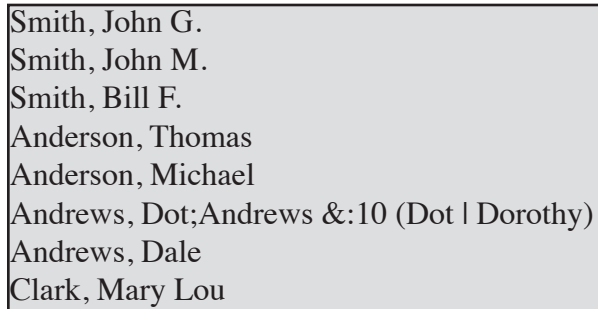


Fig. 6.13

Note: Optimization of word/phrase lists should be done prior to reversing first and last names or converting them to multiple-level.

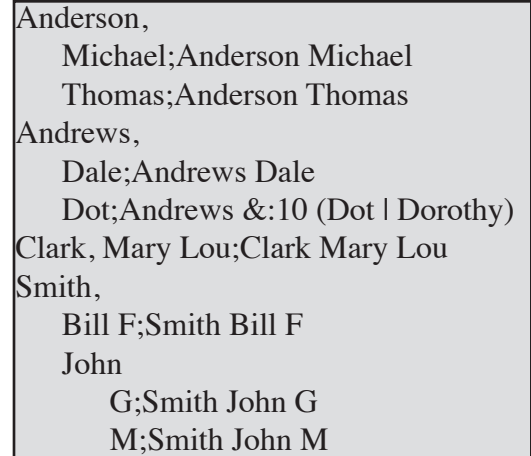
## Converting to a Multiple Level Word/Phrase List

Sonar Bookends InDex Pro can automatically convert a single-level word/phrase list into a multiple-level word/phrase list. As each entry is converted, a semicolon and a search phrase is generated and placed after every entry which does not already have a semicolon. If an entry already has a semicolon, then the semicolon and the text to the right of the semicolon is used instead. Figure 6.14 shows a single level word/phrase list and figure 6.15 shows the converted list.



Smith, John G.  
 Smith, John M.  
 Smith, Bill F.  
 Anderson, Thomas  
 Anderson, Michael  
 Andrews, Dot;Andrews &:10 (Dot | Dorothy)  
 Andrews, Dale  
 Clark, Mary Lou

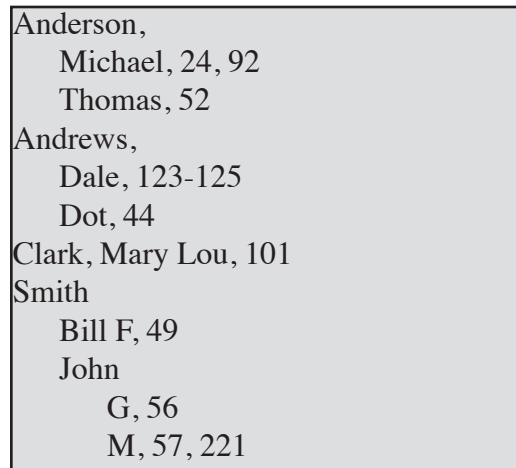
Fig. 6.14



Anderson,  
     Michael;Anderson Michael  
     Thomas;Anderson Thomas  
 Andrews,  
     Dale;Andrews Dale  
     Dot;Andrews &:10 (Dot | Dorothy)  
 Clark, Mary Lou;Clark Mary Lou  
 Smith,  
     Bill F;Smith Bill F  
     John  
     G;Smith John G  
     M;Smith John M

Fig. 6.15

Notice that the multiple level word/phrase list has been sorted. Figure 6.16 shows the index generated by using the word/phrase list in figure 6.15.



Anderson,  
     Michael, 24, 92  
     Thomas, 52  
 Andrews,  
     Dale, 123-125  
     Dot, 44  
 Clark, Mary Lou, 101  
 Smith  
     Bill F, 49  
     John  
         G, 56  
         M, 57, 221

Fig. 6.16

**Note:** if you do not want a group of words to be broken up during the single-level to multiple-level conversion, use an underline (    ) instead of a space between the words:

Van\_der\_Graff, John

The underlines will be left in the resulting word list and can then be removed manually.

To create a multiple level word/phrase list, a single level word/phrase list must already be open and be the front (active) document. Next, select *Convert List to Multiple level...* in the *Word-Phrase List* menu. Refer to figure 6.17.

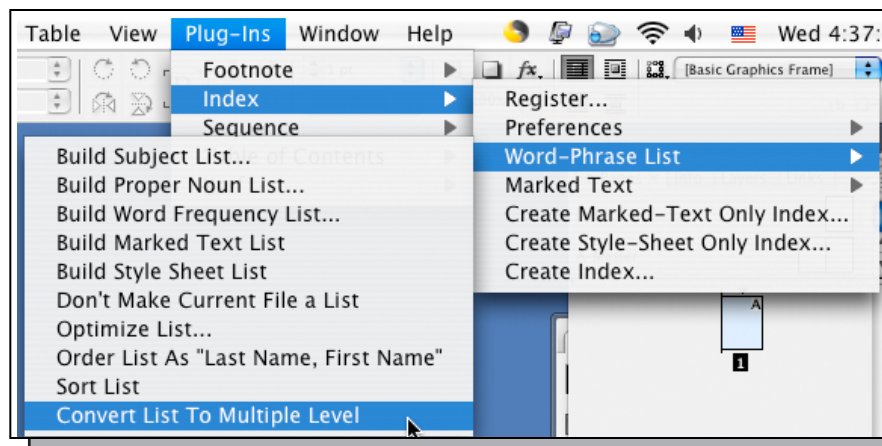


Fig. 6.17

A dialog box will appear allowing you to choose how many levels that terms can be split into (from 2 to 9). 2 is a good value if the list consists of proper names. Larger values are useful for more general lists, including product names.

The new multiple level word/phrase list is opened in a new Word Phrase List window.

## Redefining the Alphabetic Character Set

The advanced user can change Sonar Bookends InDex Pro's default alphabetic character set. Normally the letters from A-Z, the numbers from 0-9, non-English characters like *î* and *é*, hyphen and En dash are the only characters which are recognized by Sonar Bookends InDex Pro. Thus, "%test" is indexed as "test" and not "%test." By redefining the alphabetic characters set to include the character "%", "%test" would be indexed as "%test."

Refer to figures 6.18 and 6.19. Select *Redefine Alphabetic Characters...* in the *Preferences* menu.

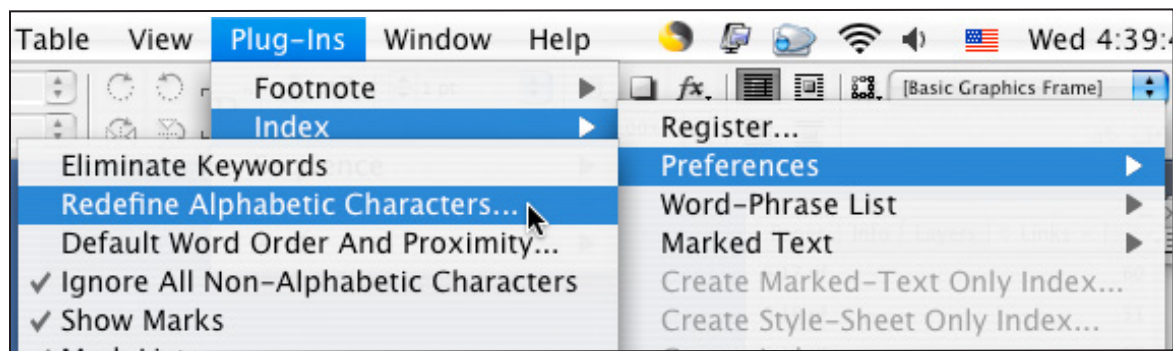
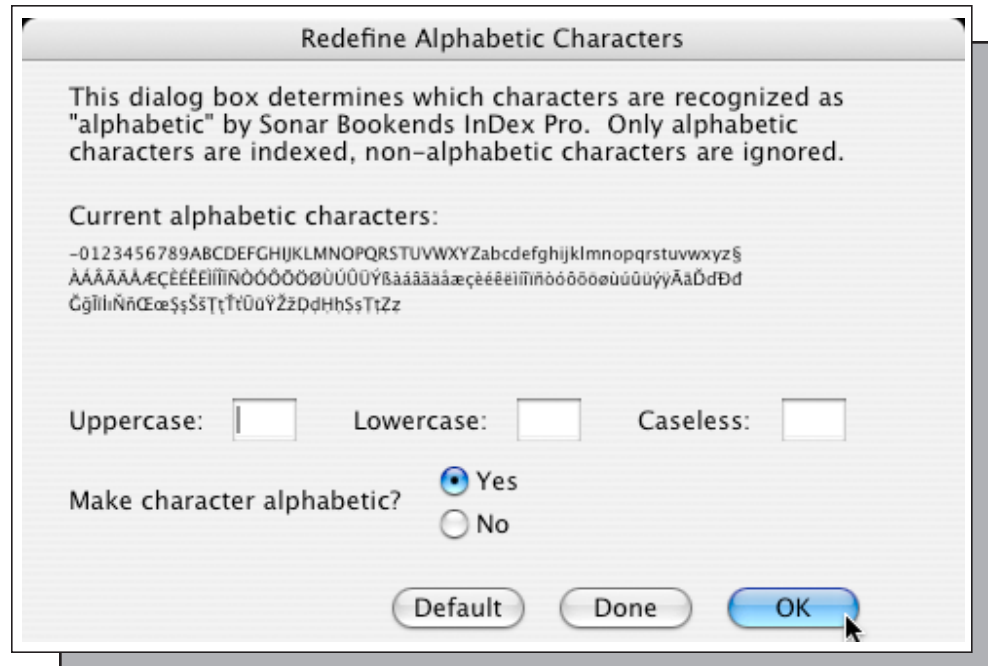


Fig. 6.18



*Fig. 6.19*

The currently defined alphabetic characters are listed in the dialog box shown in figure 6.19.

To add a new character or remove an existing one, enter its uppercase version and its lowercase version. For example, for the character **A**, you would enter **A** in the *Uppercase:* box and **a** in the *Lowercase:* box. If the character does not have an uppercase or lowercase representation (like **%**), then enter the character in the *Caseless:* box.

To make a character alphabetic, select *Yes* in the *Make character alphabetic?* area or select *No* to make a character non-alphabetic.

Click the *OK* button to make the change, which will appear under *Current alphabetic characters*. Repeat for the next character you want to change.

When you are finished, click the *Done* button.

To set the alphabetic characters back to the factory default, click the *Default* button.

## Ignoring Non-Alphabetic Characters

If a word/phrase list has any items that inadvertently contain characters reserved by Sonar Bookends InDex Pro, like the list is shown in figure 6.20, Sonar Bookends InDex Pro will attempt to interpret those characters as part of a Boolean expression. The result is unpredictable, but it usually causes the item not to be found or an error message to appear.

Phrase	Offending characters
John Doe (Junior)	()
Part #3245	#
temp = first + last;	= +
Smith, Jones & McClusky	&
First/last steps:	/:

Fig. 6.20

By selecting *Ignore All Non-Alphabetic Characters* in the *Preferences* menu, as shown in figure 6.21, all offending characters in the item are treated like spaces and effectively ignored.

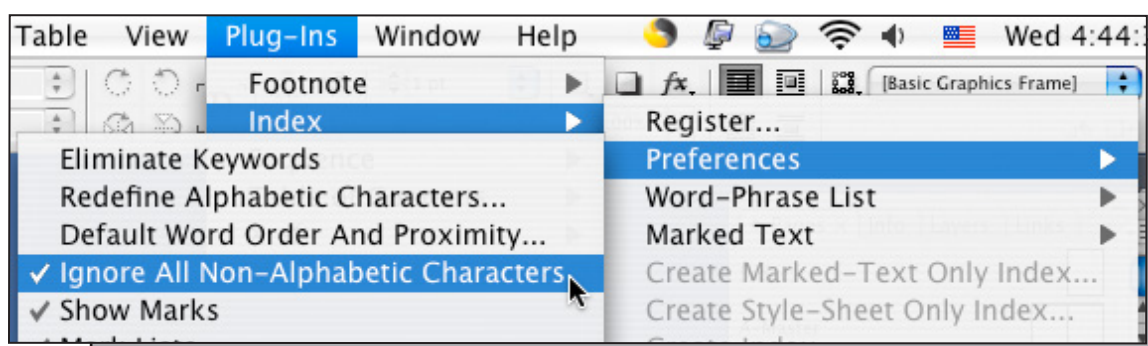


Fig. 6.21

If any Boolean operations are to be used in this mode, then they must be preceded by a semi-colon (;) :

John Doe;John &:5 Doe

See “Using the Translation Operator” on page 47 for details.

Note: Characters which are both non-alphabetic and not reserved by Sonar Bookends InDex Pro are treated like a space (ignored for all practical purposes).

`&:~|/^\^+;{}()!=#`

*Characters Reserved by Sonar Bookends InDex*

## Indexing Special Characters

If a character that is normally reserved by Sonar Bookends InDex Pro is to be part of a word, then it must be treated in a special manner. First, the character should be made alphabetic using *Redefine Alphabetic Characters...* in the *Preferences* menu (see page 55). Next, anywhere that the character appears as an alphabetic, it should be prefixed by a back-slash char-

acter (\). For example, if the phrase “Johnson & Johnson” is to be indexed and the ampersand is to be treated as an alphabetic character and not a Boolean “and,” then the phrase should look like this:

Johnson \& Johnson

When the index is complete, using InDesign simply search for the back-slashes and get rid of them. Figure 6.22 shows the steps.

Johnson, Smithers & Alfonso 50% Reduction Ratio (Part 6) Alpha/Beta Testing	<i>Part of an original word/phrase list containing several special characters.</i>
Johnson, Smithers \& Alfonso 50\% Reduction Ratio \ (Part 6\ Alpha\Beta Testing	<i>Word/phrase list with “\” characters inserted before all special characters.</i>
Johnson, Smithers \& Alfonso, 24, 52-56, 88 50\% Reduction Ratio \ (Part 6\ Alpha\Beta Testing, 2-5, 16, 22-28	<i>Resulting index.</i>
Johnson, Smithers & Alfonso, 24, 52-56, 88 50% Reduction Ratio (Part 6), 19, 24 Alpha/Beta Testing, 2-5, 16, 22-28	<i>Final index after removing all “\” characters.</i>

Fig. 6.22

## Smart Hyphens

By default a hyphen that is touching an alphabetic character on both sides is considered to be a part of a word and is kept. For example, the hyphens in the word “cul-de-sac” would be left untouched.

However, if a hyphen is followed by any combination of spaces or carriage returns on the right, by default that hyphen is not considered to be part of a word and is removed. For example,

alpha-  
betic

is considered to be the word “alphabetic” and not “alpha-betic.”

This default treatment of hyphens is called “smart hyphens.”

However, you can change this behavior so that hyphens are kept even if they are followed by blanks and carriage returns by unchecking *Smart Hyphens* in the *Preferences* menu. Refer to figure 6.23.

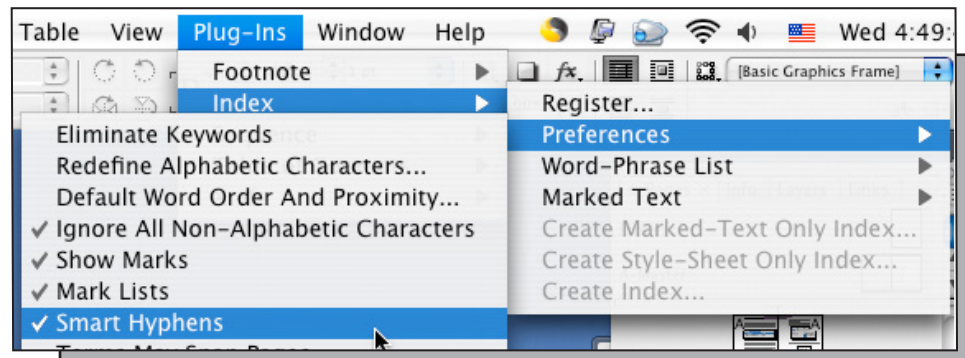


Fig. 6.23

With *Smart Hyphens* unchecked

123-ABC-  
456

is considered “123-ABC-456” and not “123-ABC456.”

## Filtering Proper Nouns

Sonar Bookends InDex Pro searches the system folder for a file named “SBIPN.TXT”. The filter file is a text only file with one word per line - the same as a keyword elimination file. Words in the filter file will not become part of a proper noun if the filtered word is the first word in a sentence. However, if a filtered word is not the first word in a sentence, but is capitalized, Sonar Bookends InDex Pro may consider it part of a proper noun.

Anyone building a non-English proper noun list will want to change or replace the supplied English file with one having equivalent words in the desired language.

## Using Style Sheets to Ignore Text

If a character or a paragraph style sheet name begins with the letters, **IGNORE**, then any text marked with that style sheet is not indexed nor is it used to generate a word/phrase list. This is useful for skipping over footnotes and endnote numbers, etc. Without this capability, footnote and endnote numbers become part of the word they are touching (**Smith<sup>1</sup>** becomes **Smith1**). If a style sheet named **Footnotes** marks footnotes, then simply change the name to **IGNOREFootnotes** or **IGNORE Footnotes**, etc.



## Controlling Treatment of Phrases That Span Pages

Depending on where a page break is located, a proper noun, subject, or any other multiple-word phrases can be split up over two pages. For example, the term “United States of America uses nuclear power” could be split up with “United States” on page 12 and “of America uses nuclear power” continuing on page 13. Sonar Bookends InDex Pro has two options for indexing such items, each with its own advantages and disadvantages.

Method one, which is the default, is to only index terms that are completely found on a single page. This method would **not** find “United States of America” in the above example because it wouldn’t see the entire phrase on either page. The phrase “nuclear power” **would** be found on page 13, because that entire phrase is on that page.

Method two treats the first several words of the continued sentence on the second page (up to 50 characters) as if those words are on both the first and second page. For a phrase that is split across the two pages, as in the example of “United States of America,” the index will show the phrase to be on the first page (page 12 in the example). The trade-off is that the second phrase, “nuclear power,” will appear in the index as being on both pages (12-13). That is because it is seen as being completely on page 12 and also completely on page 13.

Figure 6.24 shows what Sonar Bookends InDex Pro “sees” on each page when indexing using methods one and two.

End of page 12	Beginning of page 13
United States	of America uses nuclear power. This is
Method 1	
United States of America uses nuclear power	of America uses nuclear power. This is
Method 2	

Fig. 6.24

Method 2 can be enabled by selecting *Terms May Span Pages...* in the *Preferences* menu. Refer to figure 6.25.

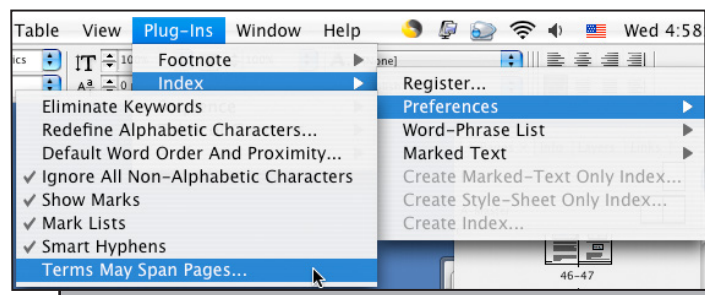


Fig. 6.25

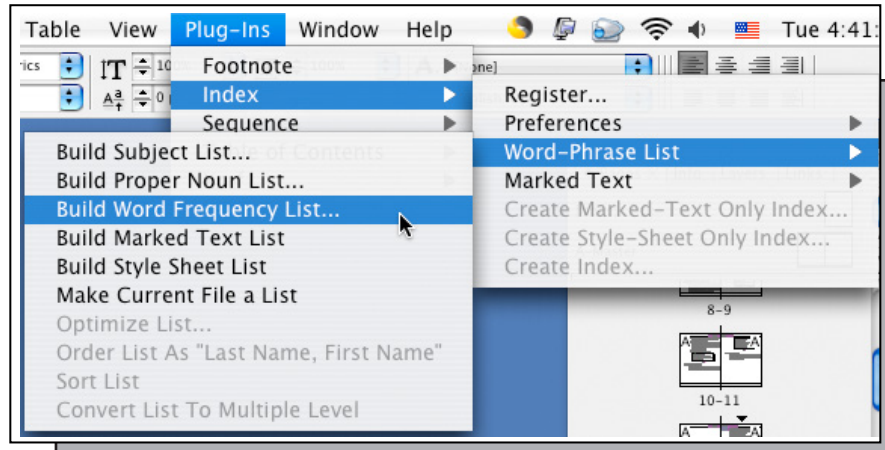


## Section VII Tutorials

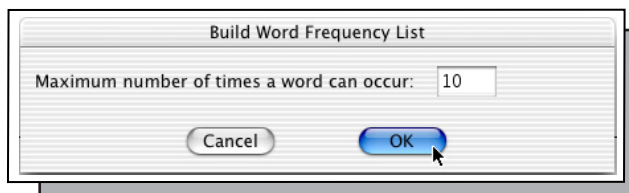
### Making a Word Frequency Index

After starting InDesign, open all of the files or books that make up the document to be indexed.

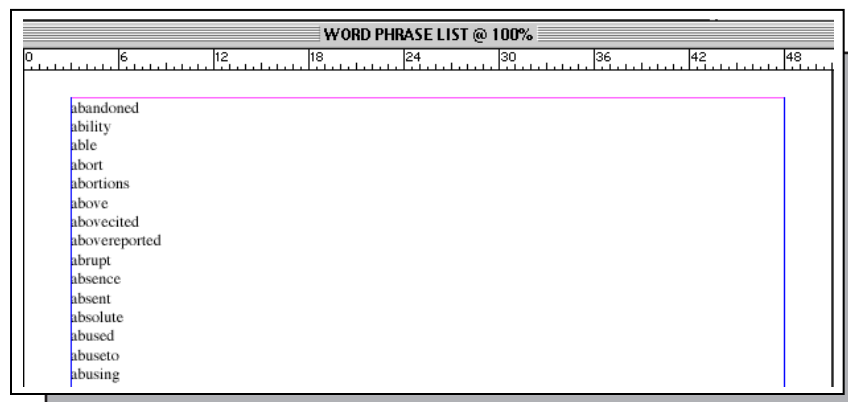
Select *Build Word Frequency List...* in the *Word-Phrase List* menu:



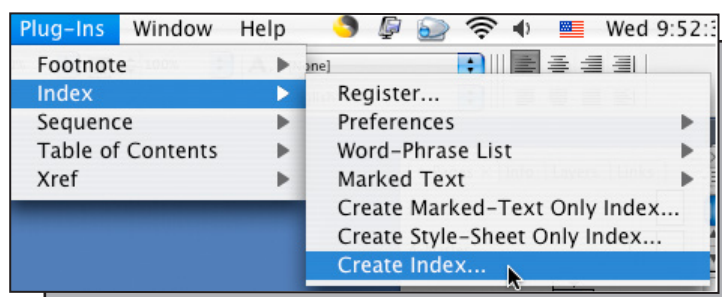
Fill in the desired frequency - the larger the number, the more words in the index. Click *OK*.



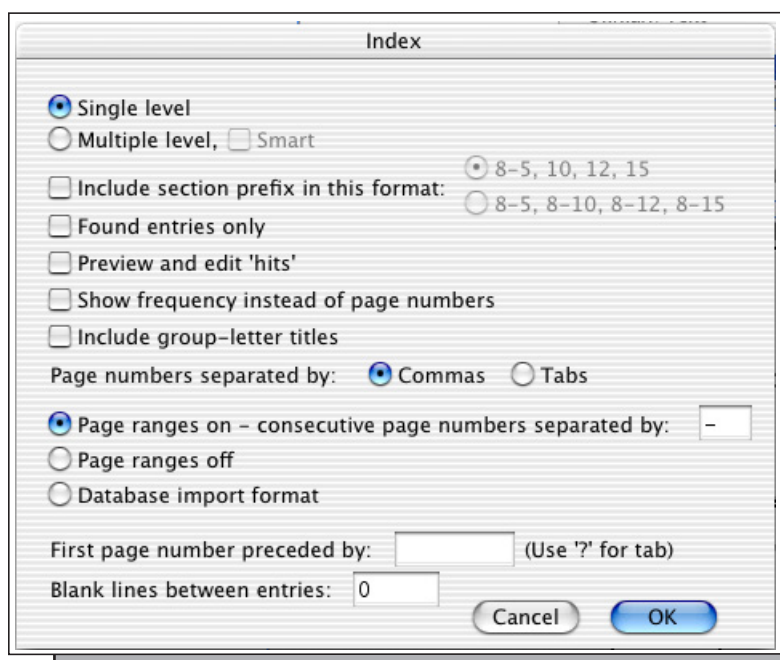
The Word Phrase List window will appear with the words to be indexed. **For the demo version, only words starting with the letter 's' will appear.** The word list can now be edited if necessary.



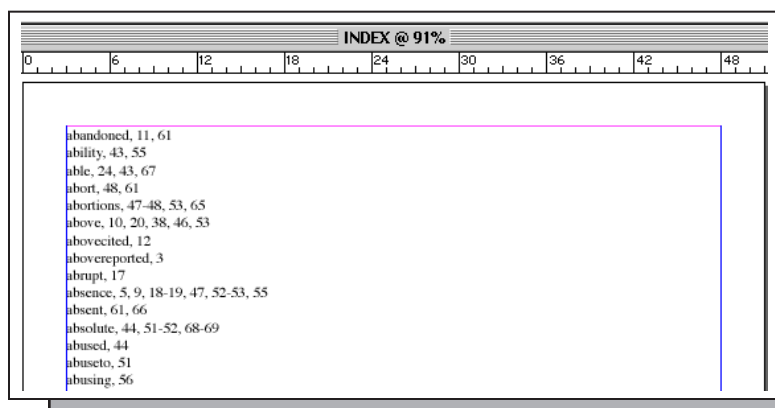
Select *Create Index...* in the *Index* menu:



In the Index dialog box make sure that *Show frequency instead of page numbers* is not checked. Set any other desired options and click the *OK* button:



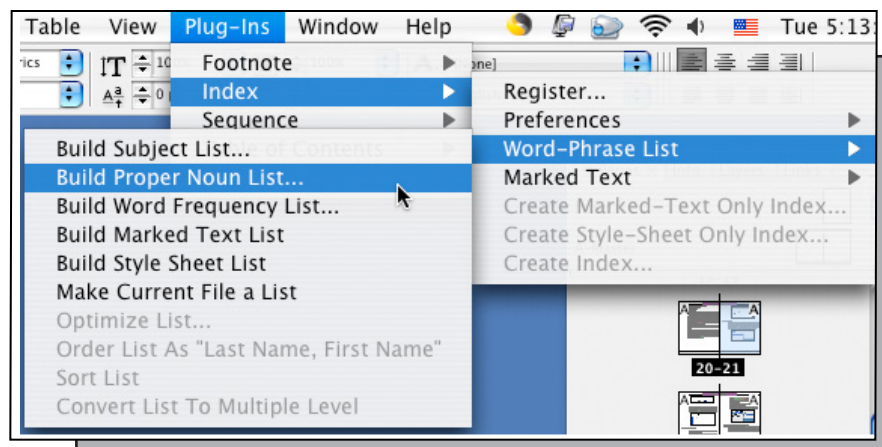
The index will now appear and can be printed and/or saved. **For the demo version, only the first two page numbers where each word is found will appear in the index.**



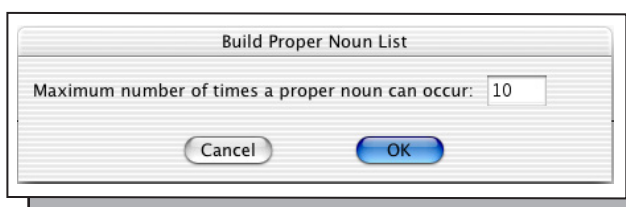
## Making a Multiple Level Proper Noun Index

After starting InDesign, open all of the files or books that make up the document to be indexed.

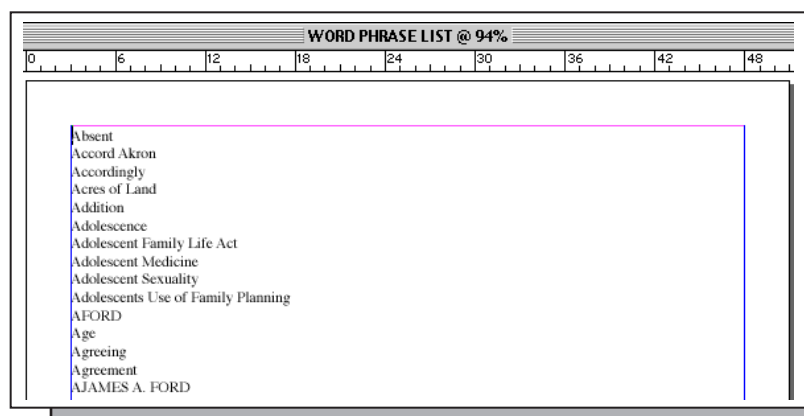
The first step in making a proper noun index is to find the proper nouns. Select *Build Proper Noun list...* in the *Word-Phrase List* menu:



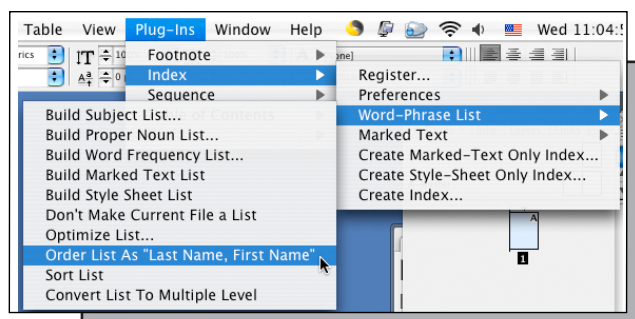
Set the maximum number of times that a proper noun can appear before it is disqualified and click *OK*:



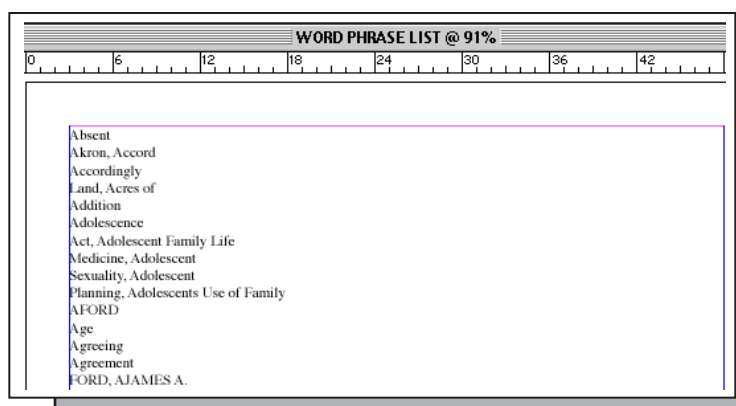
The Word Phrase List window will appear with the proper noun list. **For the demo version, only the first 25 proper nouns will appear.** The proper noun list can now be edited if necessary.



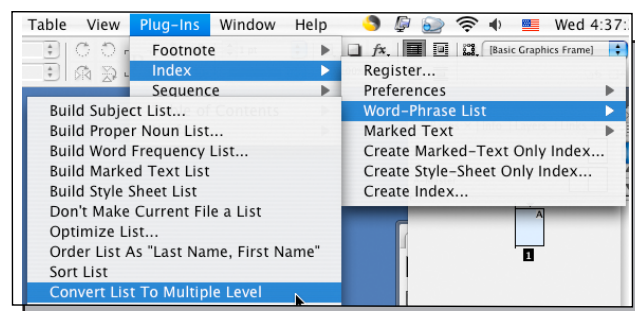
Next, reverse first and last names using *Order List As "Last Name, First Name"* in the *Word-Phrase List* menu:



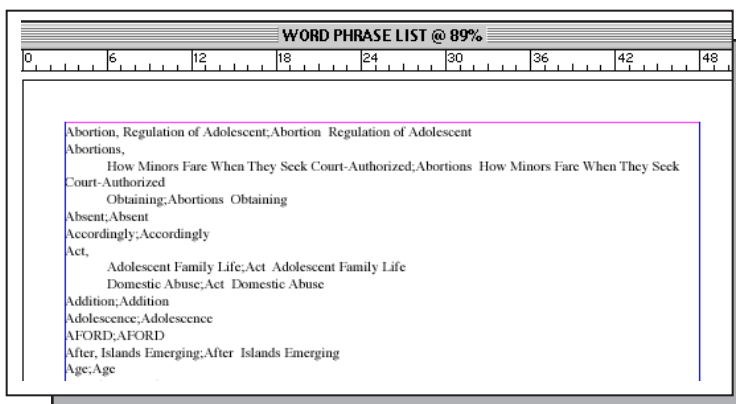
A new Word Phrase List window will appear with the reversed proper noun list.



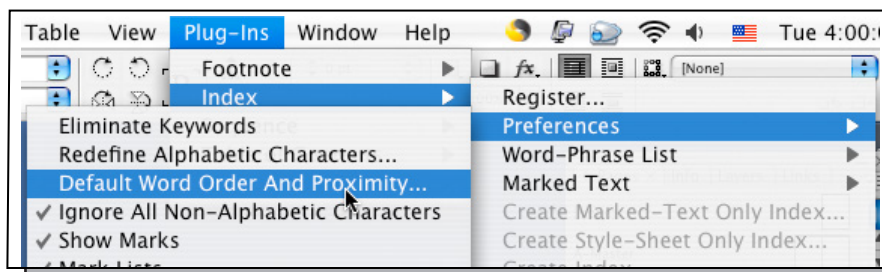
To make the single-level list into a multiple-level list, select *Convert List To Multiple Level* in the *Word-Phrase List* menu:



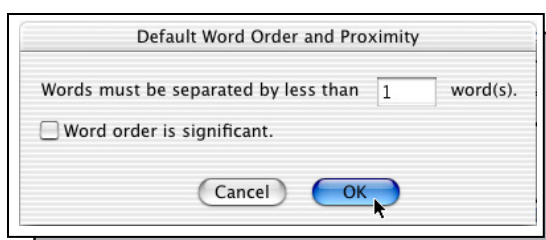
A new Word Phrase List window will appear with the multiple-level, reversed, sorted, proper-noun list.



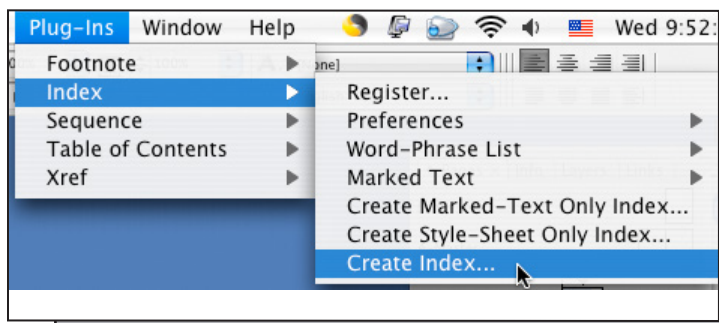
Since the names have been reversed, it is important to make sure that the reversed names will be indexed correctly. To do this, select *Default Order And Proximity...* in the *Preferences* menu:



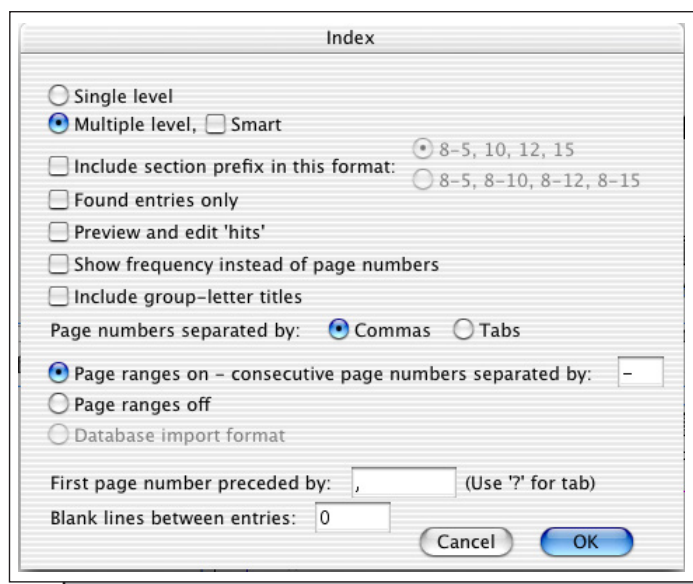
Make sure that *Word order is significant* is not checked and click the *OK* button:



Now that the word/phrase list is finished, make the index by selecting *Create Index...* in the *Index* menu:



In the Index dialog box make sure that *Show frequency instead of page numbers* is not checked and that *Multiple level* is selected. Set any other desired options and click the *OK* button:



Index

☐ Single level

☒ Multiple level, ☐ Smart

☐ Include section prefix in this format: ☒ 8-5, 10, 12, 15 ☐ 8-5, 8-10, 8-12, 8-15

☐ Found entries only

☐ Preview and edit 'hits'

☐ Show frequency instead of page numbers

☐ Include group-letter titles

Page numbers separated by: ☒ Commas ☐ Tabs

☒ Page ranges on – consecutive page numbers separated by:

☐ Page ranges off

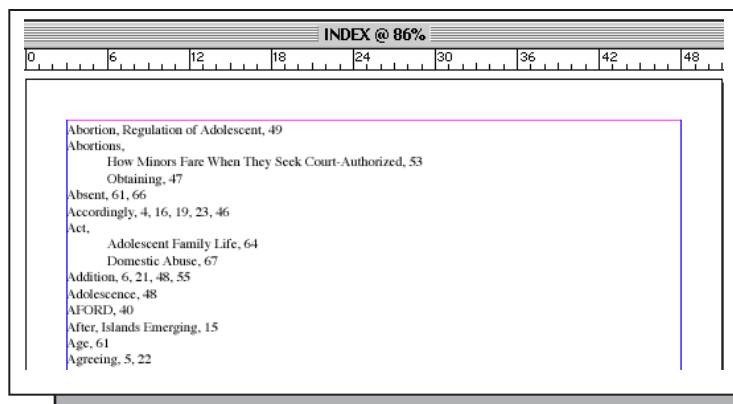
☐ Database import format

First page number preceded by:  (Use '?' for tab)

Blank lines between entries:

Cancel OK

The index will now appear and can be printed and/or saved. **For the demo version, only the first two page numbers where each word is found will appear in the index.**



INDEX @ 86%

0 6 12 18 24 30 36 42 48

Abortion, Regulation of Adolescent, 49

Abortions,

    How Minors Fare When They Seek Court-Authorized, 53

    Obtaining, 47

Absent, 61, 66

Accordingly, 4, 16, 19, 23, 46

Act,

    Adolescent Family Life, 64

    Domestic Abuse, 67

Addition, 6, 21, 48, 55

Adolescence, 48

AFORD, 40

After, Islands Emerging, 15

Age, 61

Agreeing, 5, 22

## Section VIII      Appendices

### Appendix A - Glossary

**Boolean** - A logical operation like *and*, *or*, or *not*. A Boolean expression would look like: “John or Jim but not Mary.” Using Sonar Bookends InDex, the phrase would be coded: “(John | Jim) ~ Mary”

**Chapter/section prefix** - Page numbers can be prefixed with the name of the chapter or section where the page numbers are contained. For example, if page 24 were in a chapter named “6-”, then the page number generated by Sonar Bookends InDex would be “6-24”. **The section prefix comes from InDesign’s “Section Marker” text in the Section dialog box.**

**Collating sequence** - The sequence into which the letters of an alphabet are sorted.

**Consecutive page numbers** - A series of page numbers that progressively increase by one.

**Decimal tab** - A type of tab stop where numbers are lined up on their decimal point, whether the decimal point is displayed or not.

**Dictionary** - A storage library of the words that exist in a document or folder. A dictionary is created when processing a file or folder.

**Keyword** - Treating a document as a series of words, ignoring punctuation, and spacing. By using keywords, Sonar Bookends InDex can index “John Doe” even though it may appear as “John, (Doe)” in the document.

**Keyword elimination file** - A list of words that are not to be placed into a word-frequency word/phrase list.

**Linguistic studies** - The study of languages. Such studies often include the frequency of words and phrases in documents.

**Multiple-level index** - An index where related entries are grouped together, with sub-entries indented with tab characters.

**Noise words** - High-frequency words that provide little or no content information. Noise words usually include words like “the” or “a.”

**Non-alphabetic characters** - Characters that are not used to make up a word. Just about any character can be made into an alphabetic character, including numbers.

**Phrase** - One or more words.

**Proper noun** - The name of a person, place or thing: United States of America, John Quincy Adams, Sonar Bookends InDex Pro, etc.

**Proximity** - The distance (measured in words) between two words or phrases. For example, “John within 10 words of Mary.” Using Sonar Bookends InDex Pro this would be written as either: “John Mary:10!” or “John &:10 Mary.”

**Single-level index** - An index where information is organized alphabetically and not by any kind of relationship.

**Special characters** - Characters that are reserved by Sonar Bookends InDex Pro. Special characters include those used to create Boolean expressions.

**Subject list** - A list of people, places, things, and concepts found by analyzing the English text in a document. Unlike a proper noun, capitalization does not play a major roll in detecting subjects for indexing.

**Text only file** - Free-text files that do not contain graphics or formatting information. Also known as ASCII files.

**Translation** - The “translation” operator keeps Boolean expressions, wild cards, and default overrides from appearing in the final index. It is also used for adding comments and blank lines to an index. The symbol used for translation is the semicolon: (;). Text appearing to the left of the semicolon appears in the final index, while text to the right of the semicolon is only used to find the page numbers. For example, *cat;feline | kitten* places the word *cat* in the index, followed by the page numbers where *feline* or *kitten* are found.

**Wild card** - Use of the asterisk as a special character to determine certain matching criteria during a search. Sonar Bookends InDex uses a suffix wild card. Suffix wild cards match the ending characters of a word. For example, *Mark\** finds *mark*, *marking*, *marker*, or *markers*.

**Word/phrase list** - A text only file containing the words and phrases that are to be in an index.



## Appendix B - Error Messages

**Dictionary damaged** - Contact technical support should this error message appear.

**Disk full** - This disk is full, use another.

**Insufficient memory** - There is not enough RAM to process the request. If Boolean expressions are not being used in the word/phrase list, check the *Ignore All Non-Alphabetic Characters* item in the *Preferences* menu and try making the index again.



## Appendix C - Menu Summary

Index ▶	Indexing menu items
Table of Contents ▶	Table of contents menu items
Register...	Register serial number to enable full indexing/TOC features
Preferences ▶	Set indexing preferences
Word-Phrase List ▶	Manipulate word/phrase lists
Marked Text ▶	Mark selected text for making a word/phrase list
Create Marked-Text Only Index...	Make an index of only where text is marked manually
Create Style-Sheet Only Index...	Make an index only where text is marked with style sheets
Create Index...	Make an index of all locations where words/phrases appear
Register...	Register serial number
Make Table of Contents (Without Section Prefix)	TOC without section marker prefix
Make Table of Contents (With Section Prefix)	TOC with section marker prefix
Eliminate Keywords	Eliminate certain words from index
Redefine Alphabetic Characters...	Change characters that make up words
Default Word Order And Proximity...	Set word order/proximity preferences
Ignore All Non-Alphabetic Characters	Non-alphabetic characters not used
Show Marks	Highlight marked text on/off
Mark Lists	Allow marking of items in a list as a single selection
Smart Hyphens	Only keep hyphens if they are at end of a line on/off
Terms May Span Pages...	Find phrases if they are split across two pages
Maximum Consecutive Page Number Abbreviation	Remove redundant digits from a page range
Build Subject List...	Create a list of all subjects of sentences
Build Proper Noun List...	Create a list of proper nouns (names)
Build Word Frequency List...	Create list of words based on frequency
Build Marked Text List	Create a list of terms marked manually
Build Style Sheet List	Create a list of terms marked with style sheets
Make Current File a List	Make/Unmake front file a word/phrase list
Optimize List...	Remove plurals and overlapping page numbers
Order List As "Last Name, First Name"	Reverse first and last names
Sort List	Sort the list alphabetically
Convert List To Multiple Level	Convert a single level list to multiple-level



## Appendix D - Trouble Shooting

**Problem:** The *Create Index...* menu item is dimmed and cannot be accessed.

**Solution:** The front (active) file is not a word/phrase list. Make sure the front document is a word/phrase list. If the front document is not titled “Word Phrase List”, use the *Make Current File a List* item in the *Word-Phrase List* menu to mark the front file as a word/phrase list.

**Problem:** When indexing a document consisting of multiple files, page numbers are correct, but out of sequence: Car, 1, 120, 40, 92

**Solution:** The file names being indexed are not in alphabetical order. Alphabetize the file names so that they are in the same order as they appear in the book.

**Problem:** When indexing a document, there are incorrect page numbers.

**Solution:** Only the files making up the document to be indexed and the Word Phrase List window should be open. Check to make sure that there are no other files open. If there are some extraneous files open, then they are being indexed.

**Problem:** Five asterisks (\*\*\*\*\*), indicating that Sonar Bookends InDex Pro cannot find a word/phrase list entry, follow entries that are definitely in the document being indexed.

**Solution:** This can be caused by four things:

1. If the offending entries contain any of the following special symbols, refer to “Ignoring Non-Alphabetic Characters” on page 56:

&:~|/^\^+;{}()!=#

2. If keyword elimination is being used, then make sure that the offending entries do not contain any of the eliminated keywords.
3. Make sure that the spelling of the entry is correct. It is easy to accidentally substitute a zero (0) for the letter O, for example.
4. Finally, if phrases have had their first and last names reversed, make sure that *Word order is significant* is unchecked in the *Default Word Order And Proximity* dialog box as described under “Setting/Overriding Default Word Order and Proximity” on page 13.

**Problem:** The word “Cat” is to appear in the index, but “feline” is what actually appears in the document being indexed.

**Solution:** Use “translation” (;) to resolve this situation. For more on translation see “Using the Translation Operator” on page 47. You can also put “cat” on the page using hidden keywords and index “cat” in the normal fashion. There is more information about hidden keywords on page 26.

**Problem:** The index is to include chapter references as well as page numbers, but the chapter references do not appear.

**Solution:** Check *Include chapter names in this format:* in the Index dialog box. Also, make sure that you have entered the section/chapter name in the *Section Marker* field in InDesign’s *Section Options...* dialog box. Refer to “Indexing with Chapter/Section Names” on page 16.

**Problem:** Sonar Bookends InDex Pro has reversed first and last names in the word/phrase list. However, when the index is made, Sonar Bookends InDex Pro is unable to find the reversed names.

**Solution:** Even though Sonar Bookends InDex Pro was used to reverse first and last names, Sonar Bookends InDex Pro still needs to be told that word order does not matter when indexing. This is done using *Default Word Order And Proximity...* in the *Preferences* menu and must be done before making the index. See “Setting/Overriding Default Word Order and Proximity” on page 13.

**Problem:** The page numbers for each section in a multiple section document always begin with page one, even though different starting page numbers had been manually specified.

**Solution:** In InDesign’s general preferences, select *View: Section Numbering*. Do not use *View: Absolute Numbering*.

# Index

## A

Abbreviation of consecutive page numbers, 35  
Adding comments, 68  
Alphabetic character set (defn, changing), 55  
ASCII, 68  
Asterisk, 8, 34, 45, 68, 73  
Attach keywords..., 26

## B

Blank lines, 2, 31, 37, 48, 68  
Books, 31, 61, 63  
Boolean expressions, 2, 41, 45, 47, 68-69  
Build proper noun list..., 21, 63  
Build style sheet list..., 27  
Build word frequency list..., 17, 61

## C

Chapter/section names, 1, 15-16, 74  
Characters reserved by Sonar Bookends, 56-57  
Collating sequence, 28, 67  
Colons terminate subjects, 20  
Comma, 8, 10, 35-36, 49  
Comma Separators, 35  
Commas separate subjects, 20  
Comments, 68  
Consecutive page numbers, 8, 15, 33, 35, 67  
Converting single level word/phrase lists to multiple level, 53  
Create Index..., 31, 62, 65, 73  
Create Marked-Text Only Index..., 31  
Create Style-Sheet Only Index..., 31  
Creating a word/phrase list, 17-28, 30  
Cross-reference, 1, 48

## D

Database import format, 36  
Default alphabetic character set, 55  
Default word order and proximity..., 13-14, 23, 34, 47, 65, 71, 73-74  
Demo version, 3, 18-19, 21, 61-63, 66  
Discretionary hyphens, 13

## E

Editing ‘Hits’, 34, 41  
Eliminate keywords..., 50  
Error messages, 69  
Exclamation sign (!), 47  
Exporting, 5, 24

## F

Filtering proper nouns, 59  
Five asterisks (\*\*\*\*\*), 34, 73  
Footnotes, 2, 59  
Free-text, 68

## G

Glossary, 2, 67  
Group-letter titles, 35, 38

## H

Hidden keywords, 1, 26, 74  
Hiding text, 47  
High-frequency noise words, 67  
Hyphenated chapter names, 10, 35  
Hyphenated words, 13

## I

Ignore all non-alphabetic characters, 34, 57, 69  
Ignoring specified text, 59  
Index dialog box, 31, 62, 65, 74  
Index formatting, 38  
Index with chapter names, 16  
Index without chapter names, 15  
Indexing special characters, 57  
Installing Sonar Bookends InDex Pro, 2

## K

Keyword elimination, 50, 59, 67, 73

## L

Linguistic studies, 34, 67



## M

Make index, 31, 44  
Make table of contents, 39  
Mark lists, 25, 71  
Mark text, 25, 27, 31  
Maximum consecutive page number abbreviation, 36  
Menu summary, 71  
Modifiers, 45  
Multiple-level word/phrase lists, 10, 23, 32, 41, 53  
Multiple level index, 6  
Multiple level proper noun index, 63

## N

Name reversal, 2, 48  
Non-alphabetic characters, 45, 48, 56, 67, 71, 73  
Non-consecutive page numbers, 35  
Non-English characters, 5, 55

## O

Operators, 22, 45-46  
Optimize word/phrase list, 22, 51-52  
Order list as last name, first name, 48, 63  
Orphaned multiple-level entry, 33  
Overlapping phrases, 22, 52-53  
Overriding default word order, 13, 47-48, 73-74

## P

Pagination, 13  
Parentheses, 45  
Preview mode, 2, 41, 44  
Proper noun index, 63  
Proper noun list, 1, 5, 7, 21-22, 59, 63-64  
Proximity, 13-14, 23, 34, 46-47, 65, 68, 71, 73-74  
Punctuation, 1, 7, 13, 18, 21, 23, 35, 38, 45, 67

## R

Redefine alphabetic characters..., 55, 57  
Redundant page number optimization, 22, 52-53  
Register, 3, 71  
Registration card, 3  
Reversing names, 2, 48

## S

Sample Boolean expressions, 47  
SBIKWELIM.TXT, 50  
SBIPN.TXT, 59  
SBIPPREFIX.TXT, 49  
SBISORTIGNORE.TXT, 28  
SBISORTSEQUENCE.TXT, 28  
SBITITLES.TXT, 49  
Semicolon (use of), 26, 47-48, 53, 68  
Separation of non-consecutive page numbers, 35  
Setting/Overriding Default Word Order, 13, 48, 73-74  
Setting Blank Lines Between Entries, 37  
Single-Level Index, 5, 32, 68  
Single-level word/phrase lists, 5, 10, 22-23, 41, 53  
Single/plural phrases, 2, 41, 50-51  
Single level index, 5  
Singular/plural optimization, 2, 41, 50-51  
Smart hyphens, 13, 20, 58-59  
Smart option, 7, 32  
Sonar Bookends InDex Pro's Alphabetic Character Set, 55  
Sort list..., 27  
Sorting, 2, 27-28  
Sorting order, 2, 28  
Special characters, 34, 37, 45, 57-58, 68, 73  
Subject word/phrase list, 1  
Suffix wild card, 68

## T

Tab character, 9-11, 13, 23, 35, 37-39, 47, 67  
Table of Contents, 1, 3, 11, 14, 17, 31-40, 71  
Tables, 2  
Technical support, 3-4, 69  
Terms May Span Pages..., 60  
Text only, 1, 24, 28, 49-50, 59, 68  
Text separating index entries and page numbers, 36  
Tilde, 10, 35  
Translation operator, 7, 24, 45, 47-48, 57, 68, 74  
Trouble shooting, 73  
Two-word last names, 2, 41, 48-49

## W

Wild cards, 1, 22, 41, 45, 47, 68

Word/phrase list index, 5-6, 48

Word frequency index, 5, 50, 61

Word Frequency Instead of Page Numbers, 34

Word order, 13-14, 23, 34, 47-48, 65, 71, 73-74

