

MAXDictio Reference Manual

Word Frequency Analysis,
Visual Text Exploration
and
Dictionary Based Content Analysis with
MAXODA



12

Support, Distribution:

VERBI Software. Consult. Sozialforschung. GmbH Berlin/Germany

<http://www.maxqda.com>

MAXQDA is a registered trademark of VERBI Software. Consult. Sozialforschung. GmbH, Berlin/Germany; SPSS is a trademark of SPSS Inc.; Adobe Acrobat is a trademark of Adobe Inc.; Microsoft, Microsoft Windows, Word, Excel, PowerPoint are trademarks of Microsoft Corp.

All other trademarks are the property of their respective owners, and may be registered in the United States and/or other jurisdictions.

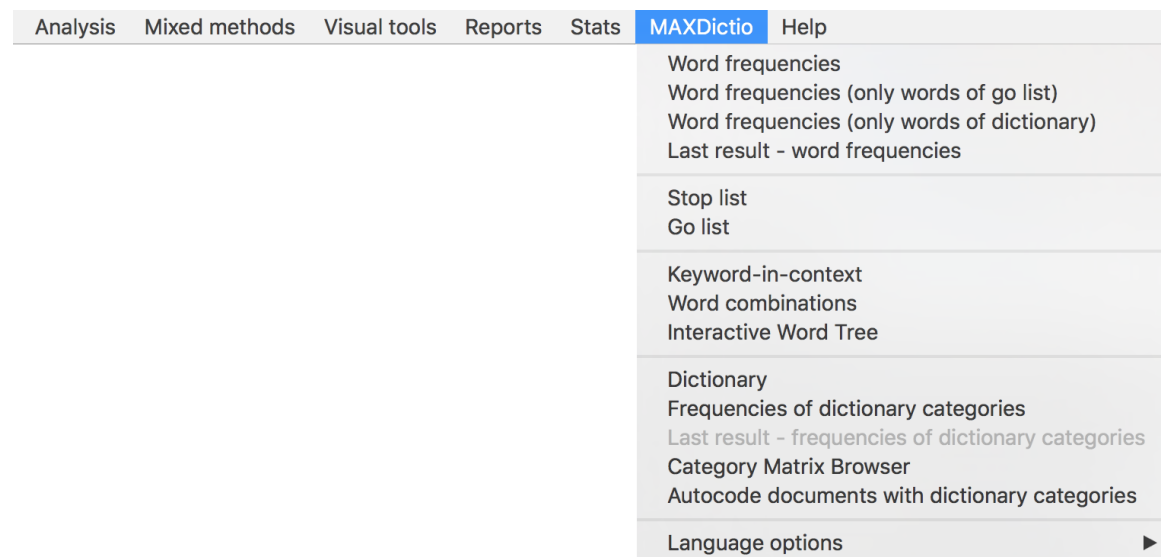
All rights reserved.

Content

1	Installation.....	5
2	What Does MAXDictio Offer?	6
3	Basic Terms	8
4	Word Frequencies	10
4.1	Analyze Word Frequencies	10
4.2	Word Frequencies: Table of Results	12
4.3	Word Frequencies for Words in the Go List	19
4.4	Word Frequencies for Dictionary Words	19
4.5	Viewing Last Frequency Count	20
4.6	Word Index: Search Results	21
5	Lemmatization	24
6	Stop Lists	26
7	Go Lists.....	33
8	Keyword-in-Context.....	34
9	Word Combinations.....	37
10	Interactive Word Tree.....	41
11	Managing Dictionaries	46
12	Dictionary Based Content Analysis	57
12.1	Schematic Course	57
12.2	Determine Frequency of Dictionary Categories	58
12.3	Result Table	61
12.4	Analyze Structured Texts.....	64
12.5	Further Processing of the Results with Excel or SPSS.....	68
13	The Category Matrix Browser	70
14	Autocode Documents with Dictionary Categories	72
15	Language Options.....	74
16	Limits of MAXDictio.....	76

1 Installation

MAXDictio can only be used in connection with MAXQDA. MAXDictio is not a separate software, it is accessed as a plug-in within MAXQDA. If you use [MAXQDA Plus](#) or [MAXQDA Analytics Pro](#), the main menu of MAXQDA contains an entry called "MAXDictio".



By clicking on MAXDictio, a pull-down menu will open and you can select one of the offered options

2 What Does MAXDictio Offer?

MAXDictio offers three essential functions:

1. Using the *word frequency function*, you can perform vocabulary analyses, for example to explore which vocabulary is present, which words occur in a text or text segment, and the differences between the vocabularies of two different texts within a project.
2. Using the *tools for text exploration* function, you can examine the occurrence of words and combinations of words in their respective contexts.
3. Using the *dictionary*, quantitative content analyzes can be carried out, in which the search terms from a dictionary are counted and the results are graphically displayed.

All functionalities can be understood and applied as independent methods, each with its own independent methodology. However, these functionalities are particularly suitable for complementing qualitative data analysis. The integration of these different methodological approaches is supported by the integration and interaction of MAXDictio and MAXQDA.

More precisely, MAXDictio provides following functions:

Interactivity and Integration with MAXQDA

- Complete integration with MAXQDA functionalities and thus the use of convenient functions such as automatic coding of text passages, building of groups by means of document groups and document sets, display of search results in their respective contexts, etc.
- Selection of texts according to specific criteria, for example, sociodemographic variables
- Editing of texts and changing of the assignment of units of analysis
- Automatic coding of multiple search terms in a single process using a words-based dictionary

General

- Analysis of text and PDF documents
- Exportability of results of frequency analysis to Microsoft Office programs (e.g. Excel, Word) and SPSS via the clipboard
- Lemmatization of words in results tables (in multiple languages)
- Definition of characters to be cut off

Word Frequencies and Word Combinations

- Counting, filtering and searching of word frequencies
- Output of word combinations of 1-5 words
- Differentiation of results by documents, document groups, document sets or codes - ideal for comparative analysis
- Results tables which display the rank frequencies and number of texts in which words occur



- List of sources for selected words with the possibility to jump from each list entry to the location of the word in the original text

Stop and Go Lists

- Restriction of evaluations using Go Lists (containing a list of the only words to be analyzed) and Stop Lists (containing a list of words to be ignored during the analysis)
- Saving of Stop and Go Lists within a specific project as global, cross-project lists
- Population of Stop Lists directly from the results table

Keyword-in-Context

- Searching of any desired words or words from the dictionary or Go Lists
- Presentation of keywords with a definable number of surrounding words
- Output of standardized properties, which are assigned to the original texts as variable values
- Jumping to the original location in the text with a single click

Interactive Word Tree

- Presentation and analysis of the most common phrases (word combinations) in an interactive tree structure
- Definition of any word as a "Root"
- Placement of desired words at the beginning, middle, or end of a phrase
- Presentation of phrases in the text with ability to jump directly to the location in the original text

Dictionary-based Functions (for Content Analysis)

- Construction of words-based dictionaries as simple or hierarchical dictionaries containing subcategories
- Simple transfer of words from word frequency tables to the dictionary
- Enabling and disabling of categories and search terms in the dictionary
- Export and import of dictionaries to and from Microsoft Excel
- Unconstrained definition of analysis units (with the possibility to overlap): documents, document groups, document sets
- Quantitative content analysis based on dictionary: determination of frequencies of the dictionary categories
- Interactive results tables, in which text segments included in the results can be jumped to directly
- Creation of validation file to verify text segments included in the results can be jumped to directly
- Easy processing of results using the MAXQDA statistics module "Stats," SPSS or Excel
- Visual display of results of the quantitative content analysis in the Category Matrix Browser

3 Basic Terms

In the following section, some important and frequently used terms will be explained:

Text

Here, “text” refers to text, PDF or table documents in the MAXQDA Document System.

Note: MAXDictio only analyzes the text itself, therefore it is irrelevant which font is being used, or if words appear in bold or italic.

Text Section/Text Unit

Similar to a book, texts can be subdivided into chapters and subchapters. Chapters may not overlap, just as chapters 3 and 4 may not overlap in a book. During import, a text can already be subdivided into sections (text units) with the text pre-processor. But it is also possible to make the text units in MAXQDA simply by coding text passages. The MAXDictio function “Frequencies for dictionary categories” can analyze text passages separately. Each text unit will appear individually in a line of the table of results.

Word

For MAXDictio, a word is any string of characters which are situated between two blanks or cut-off punctuation characters (such as , ; . :).

Word combinations

A word combination is understood by MAXDictio as a string of words which follow one another directly.

Word Stem

With a modification of the general terminology, in MAXDictio a word stem is to be understood as any part of a word. This means it is not only the left part or the stem form of the word; the “word stem” can also be situated in the middle of the word (so “go” would be a word stem of “going,” “bingo,” and “undergoes”).

Stop list

Stop lists contain collections of words that are supposed to be ignored in the word frequency analysis. Generally, these are indefinite and definite articles, numerals, and so on.

Go-List

A go-list functions pretty much conversely to a stop list. It contains words to which the frequency analysis should be confined, i.e. only words that are on the list will be considered in the analysis.

Index

An index is a list of words that includes their references, i.e. the passages where they can be found in the text(s).

Dictionary

A dictionary is composed of categories and words or character strings which are assigned to the categories. For instance, the words "Italy," "Great Britain," "Germany," "Belgium," "Netherlands," and "Spain" could be assigned to the category "Europe."

In principle, dictionaries are not project-dependent. They can be used again for other projects. For qualitative content analysis, the dictionary presents a classification scheme on the basis of which the frequency of the categories is determined.

Search Item

A search item is a string of characters, one or more words, which are attached to a category in a dictionary.

4 Word Frequencies

4.1 Analyze Word Frequencies

The simplest function of MAXDictio determines the vocabulary of all of a current project's texts.

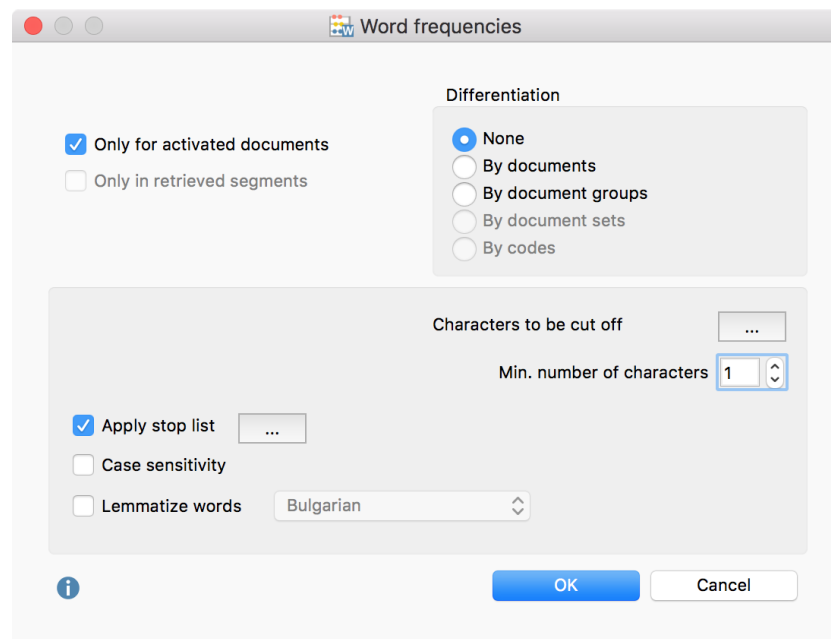
This function can be accessed by either:

- selecting the option **MAXDictio > Word Frequencies**, or
- just clicking on the corresponding quick button in the toolbar "MAXDictio".



Quick button "Word frequencies" in the MAXDictio toolbar

After starting the function, the following dialogue window appears. Here you may select all the options you need.



Before starting the "Word frequency" function you have to choose several options

Selection of texts to be analyzed

Only for activated documents – the frequencies procedure will be restricted to the activated text files

Only in retrieved segments – the frequencies procedure will be restricted to the coded segments actually displayed in the "Retrieved Segments" window

If neither option is selected, all text and table documents in the MAXQDA project will be analyzed.


Differentiation of results

None: The results table does not differentiate the results, providing only the totals over all analyzed texts.

By documents, document groups, document sets: The results table contains additional columns that can be used to compare word frequency within individual documents, document groups or document sets (see [Differentiation by Documents, Document Groups and Document Sets](#)). When the **Only for activated documents** option is selected, only activated documents within the document groups or document sets are taken into account, and only document groups or document sets containing activated documents will be analyzed.

By Codes: This option is available only if the analysis is restricted to the segments in the "Retrieved Segments" and a "Simple Coding Query" has been performed. The results table contains additional columns of recurring frequencies for each code that appears in the "Code System". This option is particularly helpful when texts have been divided into text units using codes for MAXDictio analysis, as it allows you to compare the word frequencies within different codes.

Further options

Characters to be cut off: By clicking the button with three dots , you will open a dialog box in which you can enter characters which are to be cut off from the words and ignored during the count. The setting of the characters to be cut off applies to all functions in MAXDictio.

How does MAXDictio define a "word"? A word, as showed above, is any sequence of characters between two delimiting characters. Delimiters can be, for example, blank spaces or punctuation marks. Take the example "work." As the last word of a sentence, this word is delimited by a space on its left and a period on its right.

The characters to be used as delimiters must be entered in the "Characters to be cut off" dialog box. Normally these characters include punctuation marks, question marks, etc. The selection of characters is stored in the respective project so the same results will be obtained for MAXDictio functions even if you open the file on another computer. By default, the following characters are entered automatically in new projects:

@ ! § \$ % & / () = ? ^ ° ' ^ ` " „ " " " { } [] \ # + * _ . : , ; < > ~ —

Some problems may occur with hyphens. If compound words should be counted as one word and not be split into their separate parts, the hyphen may not be declared as a delimiter. It is best to experiment a little with different possibilities. As the counting of the word frequencies can be repeated with no significant loss of time, it is advisable to look through the results for any conspicuities, change the options if needed, and then repeat the analysis.

Minimal number of characters – words with fewer characters will be skipped

Apply stop list – If a stop list is to be used, the corresponding box must be checked. Click on the button with the three dots  to open and edit the stop lists.

Case sensitivity: If this setting is activated, "Give" and "give", for example, will be counted as different words. If the setting is inactive, all words will be displayed in lowercase in the results list.

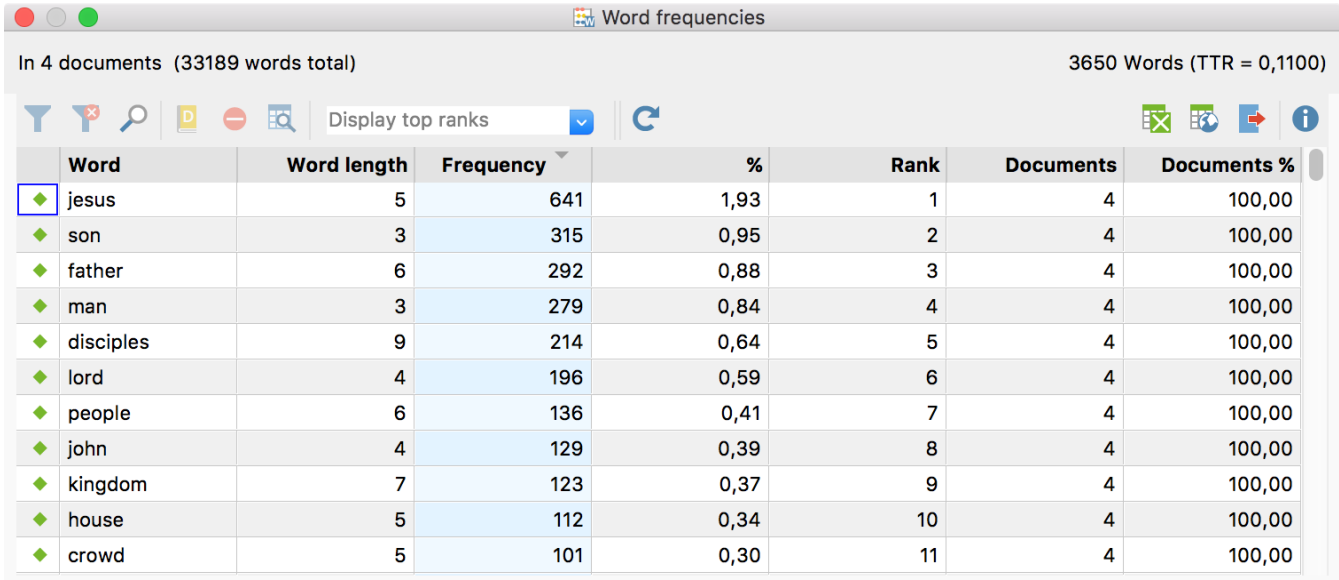
Lemmatize words – when this box is checked, the identified words in the texts will be simplified to their word stems (lemmas) by using a lemma lexicon in the chosen language. For example, if a text contains the words “gave”, “given”, and “gives”, MAXDictio will list the base form “give” in the results table only.

Click **OK**, to begin the analysis of word frequencies. Depending on the size of the texts, this process may take a few moments. A display informs you about the progress of the analysis.

Click **OK** again to begin the frequency procedure. Depending on the size of the texts, this process may take a few moments. A display informs you about the progress of the analysis.

4.2 Word Frequencies: Table of Results

The word frequency result table appears as follows:





The screenshot shows a window titled "Word frequencies" with a status bar at the top indicating "In 4 documents (33189 words total)" and "3650 Words (TTR = 0,1100)". Below the status bar is a toolbar with icons for filtering, search, and other functions. The main area displays a table with the following data:

Word	Word length	Frequency	%	Rank	Documents	Documents %
jesus	5	641	1,93	1	4	100,00
son	3	315	0,95	2	4	100,00
father	6	292	0,88	3	4	100,00
man	3	279	0,84	4	4	100,00
disciples	9	214	0,64	5	4	100,00
lord	4	196	0,59	6	4	100,00
people	6	136	0,41	7	4	100,00
john	4	129	0,39	8	4	100,00
kingdom	7	123	0,37	9	4	100,00
house	5	112	0,34	10	4	100,00
crowd	5	101	0,30	11	4	100,00

The “Word frequency” table

The table depicted above contains the following information:

- number of analyzed texts (top left, here = 4)
- total number of counted and separated words (= “tokens,” here = 33,189)
- number of different words in the texts (= “types,” here = 3,650)
- The first column serves to define a word as a stop-word. By double-clicking in the first row, you can change its status. At first, the default for each word is that it is counted. The symbol depicted then is green . After double-clicking on it, it changes into a red stop sign . The word will then be sent to the active stop list automatically.
- Column **Word**: the words appear as they appear in the text. Whether capitalization will be displayed or not depends on the option selected before starting the procedure.
- Column **Word length**: contains the word length counted in characters.

- **Column Frequency:** contains the absolute frequency of each word in the processed texts.
- **Column %:** contains the percentage of the word referring to the tokens, i.e. the total number of counted words in all texts. In our example, the word “Jesus” was counted 641 times, which is 1.93 percent of the total of 33,189 words.
- **Rank:** Indicates the ranking of a word. The most common word is assigned rank 1, the second most common word rank 2, etc. Words occurring with the same frequency are all assigned the same rank.
- **Documents:** Indicates the number of documents in which the respective word occurs.
- **Documents %:** Indicates the percentage of documents in which the respective word occurs, based on the number of documents that were taken into account during the counting process.

In quantitative content analysis, the so-called type-token-ratio is a measure for the language elaboration of a text. The type-token-ratio (TTR) is the quotient of different words (“types”) divided by the total number of words (“tokens”). It can be a maximum of 1. MAXDictio displays the TTR at top right of the window.

The results can be sorted by the individual columns, namely:


- by alphabetical wordlist
- by word length
- by absolute frequency of occurrence in the texts
- by percentage of the word in the texts (word frequency divided by tokens)
- by rank
- by number of texts, in which the words occur

Tip: If you hover over a row with your mouse pointer a tooltip will appear which displays the word that is listed in the row. If the option “Lemmatize words” has been chosen, the tooltip enlists all lemmatized words.

The Toolbar in the Table of Results

 **Filter** – Switches the currently set filter on and off.


 **Reset all filters** – Clears filters for all columns.

 **Search** – Opens an input field with which you can search within the entire result table.

 **Add to dictionary** – Adds the words of the selected rows to the active category in the dictionary.


 **Add to stop list** – Adds the words of the selected rows to the stop list.


 **Search results** – Creates a list of the search results for the selected rows.


 **Display top ranks** – In the drop down menu, select the level of the top ranks to be displayed. You can also enter a specific value in the field. Only the top most common words with the maximum entered rank will be displayed.

 **Refresh and apply stop list** – refreshes the table of results for the active stop list.

These standard MAXQDA functions are available as export options:



 **Open as Excel table** – MAXQDA creates a temporary Excel file and opens it in the program that you use for such files. If you have selected rows, only they are exported. If you want to hold on to this table for later use, choose to save it somewhere on your hard drive.

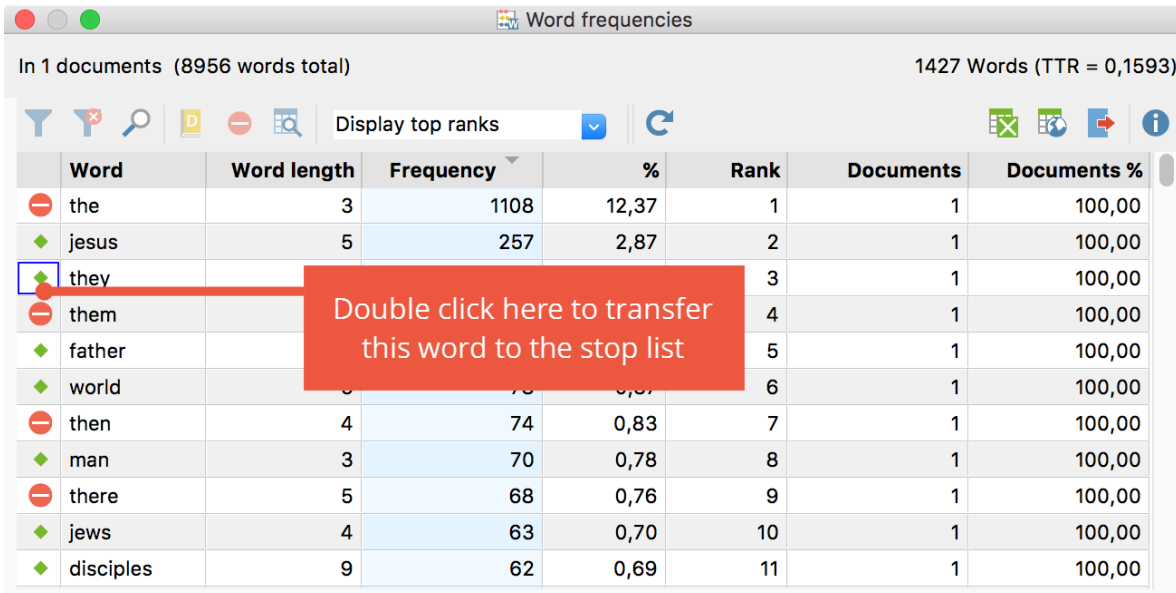
 **Open as HTML table** – MAXQDA creates a temporary table in HTML format and opens it in your internet browser. If you have selected rows, only they are exported.

 **Export** – MAXQDA opens a dialog window where you can enter a name for the file, the location you want it saved, and the format you want to save it in: Excel (XLS/X), Website (HTML), or Plain Text (TXT tab-delimited). The document will then be opened after the export is complete. If you have selected rows, only they are exported.

Transferring Words to the Stop List

To transfer words from the results table to the stop list, proceed as follows:












1. To open the stop lists, select the menu function **MAXDictio > Stop list** to confirm that the desired stop list is currently selected.
2. To transfer a single word to the stop list, double-click the green symbol in the first column of the results table . The symbol will change to a red stop sign , indicating that the word has been successfully added to the stop list. Alternatively, you can right-click the line containing the desired word and select **Add to stop list**, or click the corresponding button in the toolbar.
3. To remove a word from the current stop list, double-click the red stop sign in the first column of the results table.
4. To transfer multiple words to the stop list, hold down the Ctrl key (Windows) or Cmd key (Mac) while selecting the desired words, then select **Add to stop list** from the context menu or click the corresponding button in the toolbar.



Word frequencies

In 1 documents (8956 words total) 1427 Words (TTR = 0,1593)

Display top ranks

	Word	Word length	Frequency	%	Rank	Documents	Documents %
	the	3	1108	12,37	1	1	100,00
	jesus	5	257	2,87	2	1	100,00
	they				3	1	100,00
	them				4	1	100,00
	father				5	1	100,00
	world				6	1	100,00
	then	4	74	0,83	7	1	100,00
	man	3	70	0,78	8	1	100,00
	there	5	68	0,76	9	1	100,00
	jews	4	63	0,70	10	1	100,00
	disciples	9	62	0,69	11	1	100,00

Transfer words from the results table to the stop list

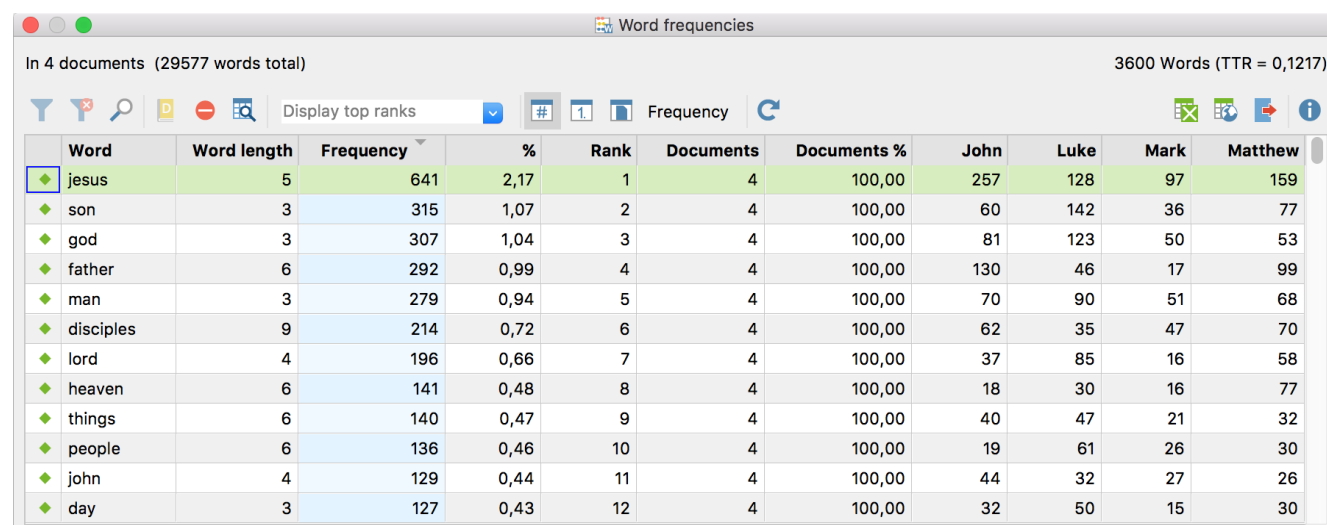
Note: If you transfer a lemmatized word into the stop list, not only the displayed word, but also the original words before lemmatization will be added to the stop list. Let's say the words "given" and "gave" have been simplified to their lemma "give" and only this word is displayed in the word frequency results. If you then transfer "give" into the stop list, also "given" and "gave" will be added to the stop list.

Differentiation by Documents, Document Groups and Document Sets

If you check the box to set your frequency count to be differentiated **By document set** in the dialog field, the counts will be separated into different columns in the results table. Each column will include the counts for the documents in each of the document sets you have created in the Document System. If you use this option in combination with the **Activation by document variable** function, you can do comparisons across specific groupings based on socio-demographic information, for example.

The results are shown in the same form and with the same meaning as in the results table for all documents.

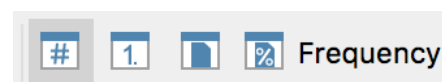
The difference is that new columns have been added at the right, so you can see exactly how many times each word showed up in each document, document group or document set.



Word	Word length	Frequency	%	Rank	Documents	Documents %	John	Luke	Mark	Matthew
jesus	5	641	2,17	1	4	100,00	257	128	97	159
son	3	315	1,07	2	4	100,00	60	142	36	77
god	3	307	1,04	3	4	100,00	81	123	50	53
father	6	292	0,99	4	4	100,00	130	46	17	99
man	3	279	0,94	5	4	100,00	70	90	51	68
disciples	9	214	0,72	6	4	100,00	62	35	47	70
lord	4	196	0,66	7	4	100,00	37	85	16	58
heaven	6	141	0,48	8	4	100,00	18	30	16	77
things	6	140	0,47	9	4	100,00	40	47	21	32
people	6	136	0,46	10	4	100,00	19	61	26	30
john	4	129	0,44	11	4	100,00	44	32	27	26
day	3	127	0,43	12	4	100,00	32	50	15	30

Word frequency table showing the word frequencies for each document


If you have selected a differentiation by document, document group or document set, the toolbar in the results table will contain additional icons that can be used to control the content of the additional columns. The icons can be switched on and off and only one of the icons can be selected at a given time. In addition to the symbols, the currently selected display is shown.




Icons for control of current content, in which "Frequency" is currently selected.

Show frequency of words – The number in the corresponding column indicates how often the word appears in the document, the document group, or the document set.

1. Show rank of words – The number in the corresponding column indicates the rank of the word in the respective document, document group, or document set.

 **Show presence of word in document** – The number in the corresponding column indicates the occurrence of the word in the document group or document set. In the differentiation by document, the column reflects whether or not the word appears in the document. Accordingly, in this case the function appears as **Show number of documents in which word occurs**.

 **Show percentage of documents, in which word occurs** – The number in the corresponding column indicates the percentage of the evaluated documents (per document group or document set) in which the word occurs. If a document group contains 10 documents and the word occurs in 7 of these documents, this column will display 70%. This option is only visible if a differentiation by document group or document set has been selected.

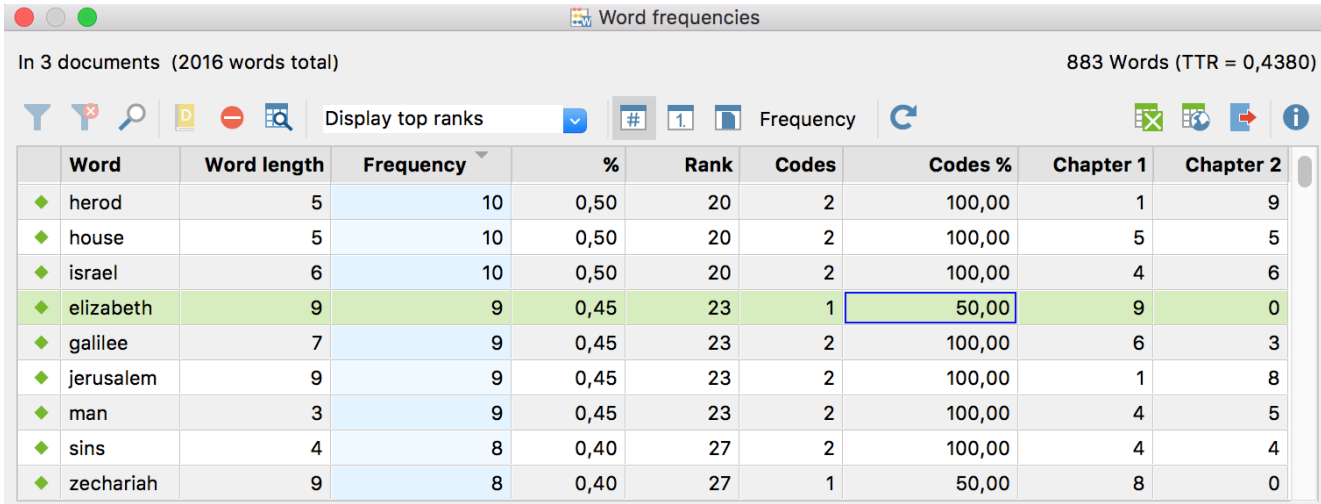
Differentiation by Codes

Differentiation **By codes** is an option if one of more texts have been divided into sections by means of coded segments. In the following example, the various chapters of the three gospels, Matthew, Mark, and Luke, were coded with the codes "Chapter 1", "Chapter 2", etc., in order to carry out word counts for the individual chapters.

To perform a separate word count for the words of coded segments, proceed as follows:

1. Activate the texts you wish to include in the analysis in the "Document System".
2. Activate the desired codes in the "Code System".
3. Select **MAXDictio > Word frequencies**.
4. Select the options **Only for activated documents**, **Only in retrieved segments** and differentiation **By codes**.

The results table that appears is similar to the previously presented results tables:



Word	Word length	Frequency	%	Rank	Codes	Codes %	Chapter 1	Chapter 2
herod	5	10	0,50	20	2	100,00	1	9
house	5	10	0,50	20	2	100,00	5	5
israel	6	10	0,50	20	2	100,00	4	6
elizabeth	9	9	0,45	23	1	50,00	9	0
galilee	7	9	0,45	23	2	100,00	6	3
jerusalem	9	9	0,45	23	2	100,00	1	8
man	3	9	0,45	23	2	100,00	4	5
sins	4	8	0,40	27	2	100,00	4	4
zechariah	9	8	0,40	27	1	50,00	8	0

Results table for word frequencies differentiated by codes

However, instead of the "Documents" and "Documents %" columns, the table contains the columns "Codes" and "Codes %". These columns indicate in how many codes the respective word occurs. For the example above, the two codes "Chapter 1" and "Chapter 2" have been activated, therefore the maximum number in the Codes column is 2. The percentages in the "Codes %" column always refer to the number


of codes that were activated in the “Code System” at the time the table was created. In the example, the values could therefore be 50% (one code) or 100% (both codes).

The two columns with the code names “Chapter 1” and “Chapter 2” contain the following information. All text segments from the “Code System” to which the code “Chapter 1” is assigned are analyzed and the respective word frequencies are counted. The same procedure is undertaken for the second column. The word “and” occurs 113 times in Chapter 1 and 99 times in Chapter 2. The sum total of 212, representing the word frequency for all evaluated text segments, appears in the “Frequency” column.


When differentiation “By codes” is selected, the following three icons allow you to control the display of the individual code columns:



Icons for managing content of additional code columns. In this example, “Frequency” is selected.

 **Show frequency of words** – The value in the additional columns indicates how often the word occurs within the codes.

 **Show rank of words** – The value indicates the rank of the word in the column.

 **Show number of codes, in which word occurs** – The value indicates the number of codes in which the word occurs. The maximum value corresponds to the number of activated codes at the time the frequency analysis began.

Maximum Rank: Show Only the Most Common Words

The “Rank” column indicates the ranking of a word based on the frequency of its occurrence. The higher the ranking of the word, the more frequently it occurs. In the table below, the word “house” holds the position 11 when all texts included in the analysis are taken into account.

The results table can easily be restricted to display the most common words:

- From the **Display top ranks** drop down menu, indicate to what rank words should be displayed, or,
- Enter the desired value directly in the drop down menu.

Word frequencies

In 3 documents (22580 words total) 3269 Words (TTR = 0,1448)

Display top ranks

Display top ranks

Word	Word length	Frequency	%	Rank	Documents	Documents %	Luke	Mark	Matthew
jesus	5	384	1,65	1	3	100,00	2	1	1
son	3	255	1,10	2	3	100,00	1	6	3
god	3	226	0,97	3	3	100,00	3	3	9
man	3	209	0,90	4	3	100,00	4	2	6
father	6	162	0,72	5	3	100,00	9	23	2
lord	4	159	0,70	6	3	100,00	5	29	7
disciples	9	152	0,67	7	3	100,00	11	4	5
heaven	6	123	0,54	8	3	100,00	16	29	3
kingdom	7	118	0,52	9	3	100,00	10	13	8
people	6	117	0,52	10	3	100,00	6	9	15
house	5	105	0,47	11	3	100,00	7	10	10
things	6	100	0,44	12	3	100,00	8	12	11

Select the number of most commonly occurring words

Once the maximum rank has been chosen, the display is restricted to only the most commonly occurring words (in the example, the top 10). If, as in the example, the word frequencies from individual documents are displayed, it is now easy to compare the frequency occurring words between documents. For example, while the word "people" is in position 10 within the text "Matthew", the same word is in position 6 in the text "Luke".

Word frequencies

In 3 documents (23274 words total) 10 Words (TTR = 0,0004)

10

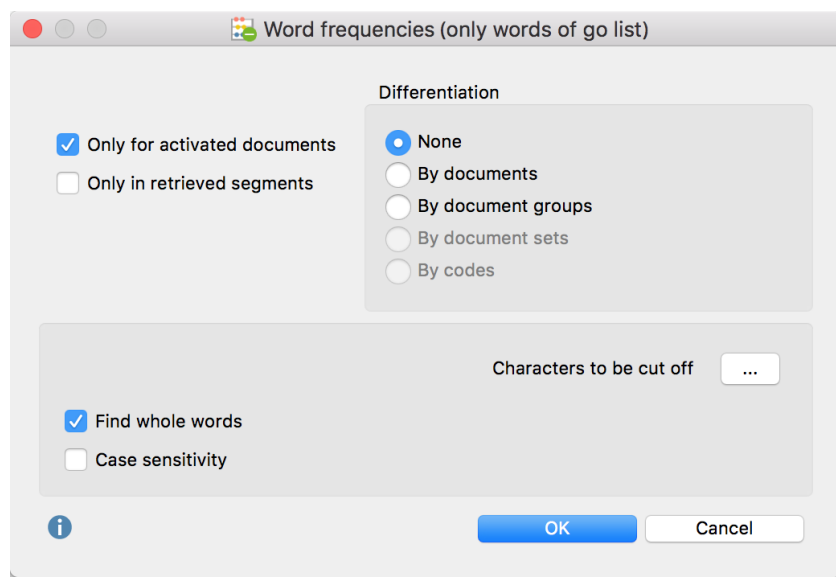
Word	Word length	Frequency	%	Rank	Documents	Documents %	Luke	Mark	Matthew
jesus	5	384	1,65	1	3	100,00	2	1	1
son	3	255	1,10	2	3	100,00	1	5	3
god	3	226	0,97	3	3	100,00	3	3	9
man	3	209	0,90	4	3	100,00	4	2	6
father	6	162	0,70	5	3	100,00	7	8	2
lord	4	159	0,68	6	3	100,00	5	9	7
disciples	9	152	0,65	7	3	100,00	9	4	5
heaven	6	123	0,53	8	3	100,00	10	9	3
kingdom	7	118	0,51	9	3	100,00	8	7	8
people	6	117	0,50	10	3	100,00	6	6	10

Only the 10 most frequently occurring words are displayed

Note: The calculation of rank order always refers only to rows currently displayed. If, as in the example, the view is restricted to the 10 most frequently occurring words (overall) the 10 ranks within each differentiation column will distributed as if only these 10 most frequent words were evaluated.

4.3 Word Frequencies for Words in the Go List

The calculation of word frequencies can be restricted to words appearing in the Go list (see [What are Go Lists?](#)). To access this function, select the menu option **MAXDictio > Word frequencies (only words of go list)**. The following dialog window will appear, which essentially corresponds to the standard dialog for word frequencies:



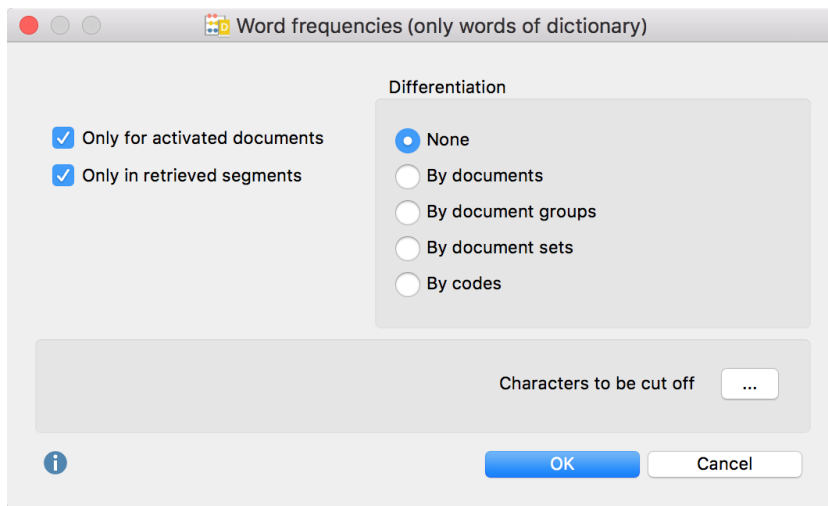
Options dialog for Word frequencies (only words of go list) function

The setting options are described in the section on Word frequencies (see [Analyze Word Frequencies](#)), however neither lemmatization nor stop lists can be used.

The results table will naturally contain only words present in the go list.

4.4 Word Frequencies for Dictionary Words

If you have created dictionaries with search terms (see [Creating and Managing Dictionaries](#)), you have the option to call up the frequencies for these words. Simply choose **MAXDictio > Word frequencies (only words of dictionary)** from the main menu. The following dialog window will appear, which essentially corresponds to the standard dialog for word frequencies:



Option dialog “Word frequencies – Count only dictionary words”

The setting options are described in the section on Word frequencies (see [Analyze Word Frequencies](#)), however neither lemmatization nor stop lists can be used.

Note: The options set for each category and search item in the dictionary will be taken into account for the count. If you have switched off a category in the dictionary its search items won’t be listed in the word frequency table.

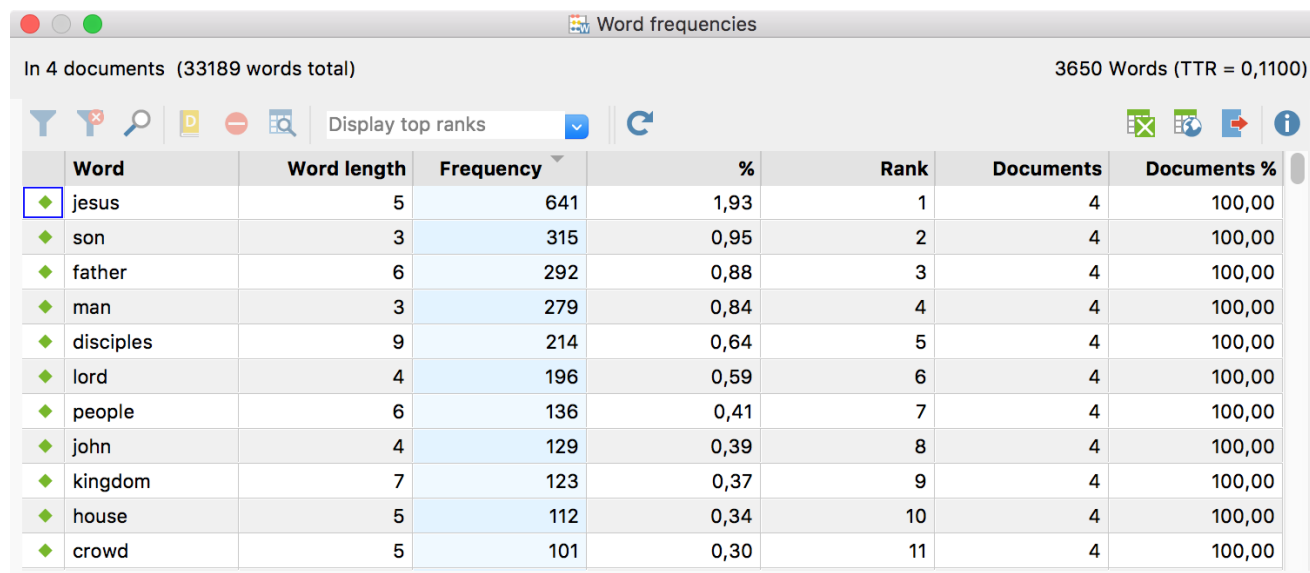
The count begins as soon as you click on **OK**. The results table naturally contains only dictionary words.

4.5 Viewing Last Frequency Count

By selecting **Last result – word frequencies** in the **MAXDictio drop-down menu** or clicking on the appropriate quick button (below) the last counting results of a word frequency function will be displayed anew, without the need to carry out the word frequency analysis again, thus saving time.



Icon for calling up the most recent frequency count



Word frequencies


In 4 documents (33189 words total) 3650 Words (TTR = 0,1100)

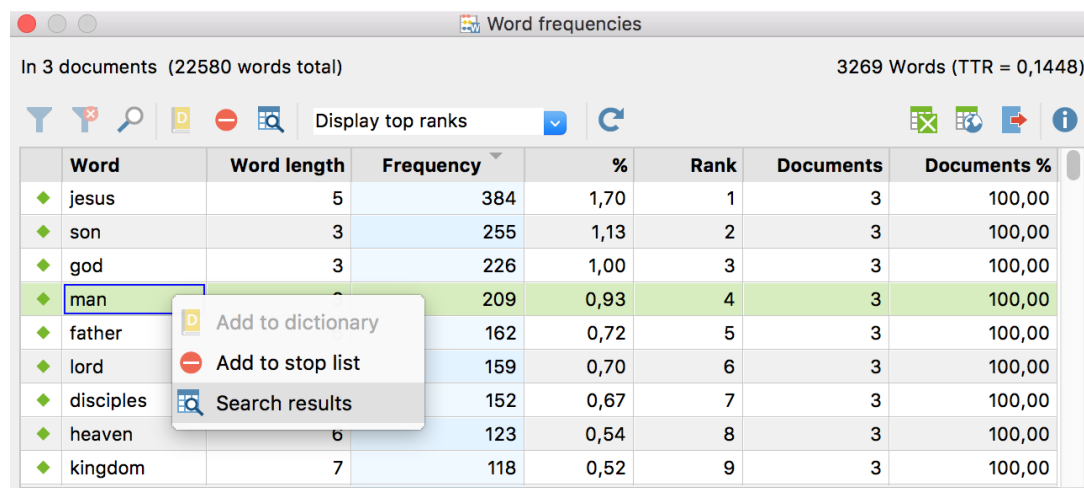
Display top ranks

Word	Word length	Frequency	%	Rank	Documents	Documents %
jesus	5	641	1,93	1	4	100,00
son	3	315	0,95	2	4	100,00
father	6	292	0,88	3	4	100,00
man	3	279	0,84	4	4	100,00
disciples	9	214	0,64	5	4	100,00
lord	4	196	0,59	6	4	100,00
people	6	136	0,41	7	4	100,00
john	4	129	0,39	8	4	100,00
kingdom	7	123	0,37	9	4	100,00
house	5	112	0,34	10	4	100,00
crowd	5	101	0,30	11	4	100,00

Example for the last frequency table

4.6 Word Index: Search Results

An index of the located positions in the text can be created for each word in the list of word frequencies. Click on the desired word with the right mouse button; then select **Search results** from the context menu (see illustration below). Alternatively, you can click the button with the same name  in the toolbar.



Word frequencies

In 3 documents (22580 words total) 3269 Words (TTR = 0,1448)

Display top ranks

Word	Word length	Frequency	%	Rank	Documents	Documents %
jesus	5	384	1,70	1	3	100,00
son	3	255	1,13	2	3	100,00
god	3	226	1,00	3	3	100,00
man	3	209	0,93	4	3	100,00
father	6	162	0,72	5	3	100,00
lord	4	159	0,70	6	3	100,00
disciples	9	152	0,67	7	3	100,00
heaven	6	123	0,54	8	3	100,00
kingdom	7	118	0,52	9	3	100,00

Context menu for 'man':

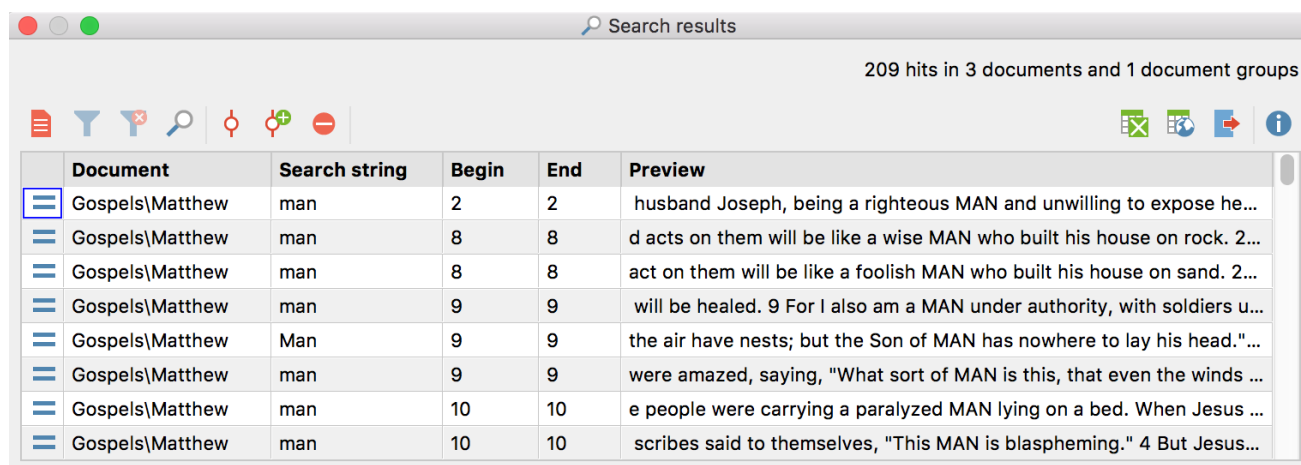
- Add to dictionary
- Add to stop list
- Search results

Right-clicking on the word opens the context menu in which you can find the "Search results" option

Tip: You can also select more than one word by holding the Ctrl-key (Win) or cmd-key (Mac). The list of search results will be created for all selected words.

In the screenshot above, you can see that the word "man" was found 209 times, making up 0,93 percent of the total vocabulary.

The above requested index appears as follows:



Document	Search string	Begin	End	Preview
Gospels\Matthew	man	2	2	husband Joseph, being a righteous MAN and unwilling to expose he...
Gospels\Matthew	man	8	8	d acts on them will be like a wise MAN who built his house on rock. 2...
Gospels\Matthew	man	8	8	act on them will be like a foolish MAN who built his house on sand. 2...
Gospels\Matthew	man	9	9	will be healed. 9 For I also am a MAN under authority, with soldiers u...
Gospels\Matthew	Man	9	9	the air have nests; but the Son of MAN has nowhere to lay his head."...
Gospels\Matthew	man	9	9	were amazed, saying, "What sort of MAN is this, that even the winds ...
Gospels\Matthew	man	10	10	e people were carrying a paralyzed MAN lying on a bed. When Jesus ...
Gospels\Matthew	man	10	10	scribes said to themselves, "This MAN is blaspheming." 4 But Jesus...

The "Search results" window for the word "man"

Users of MAXQDA are already familiar with the representation of search results in this form, since the same table is displayed for the MAXQDA search results.


In the table heading, information is displayed regarding how many texts the word has been found in and how many located positions there are all together in the text.

The column "Document" in the table indicates the name of the document in which the word has been found. In the column to the right, the search item (here: "man") is listed and the number of the relevant paragraph in the document is shown in the fourth column.

Tip: Clicking once on the column title will sort the table by column.

Going to Located Positions, Exporting, and Coding

By clicking on a **located position**, the text passage will be opened and loaded directly into the "Document Browser". The window "Search results" will stay in the front, so that for optimal working conditions, it is advisable to arrange the windows on your screen in such a way that this window won't cover the "Document Browser" window.

The function "Search results" is directly integrated into the lexical search functions of MAXQDA, which means that the located positions can be coded directly by clicking on the **Autocode search results** buttons in the toolbar located on top of the "Search results" window. More details about the autocode functions will be presented by clicking on the icon .

Search results

209 hits in 3 documents and 1 document groups

	Document	Search string	Begin	End	Preview
	Gospels\Matthew	man	2	2	husband Joseph, being a righteous MAN and unwilling to expose he...
	Gospels\Matthew	man	8	8	d acts on them will be like a wise MAN who built his house on rock. 2...
	Gospels\Matthew	man	8	8	act on them will be like a foolish MAN who built his house on sand. 2...
	Gospels\Matthew	man	9	9	will be healed. 9 For I also am a MAN under authority, with soldiers u...
	Gospels\Matthew	Man	9	9	the air have nests; but the Son of MAN has nowhere to lay his head."...
	Gospels\Matthew	man	9	9	were amazed, saying, "What sort of MAN is this, that even the winds ...
	Gospels\Matthew	man	10	10	e people were carrying a paralyzed MAN lying on a bed. When Jesus ...
	Gospels\Matthew	man	10	10	scribes said to themselves, "This MAN is blaspheming." 4 But Jesus...

The located positions can be coded automatically via the "Autocode" function

It is also possible to export all the „hits“ along with their surrounding text into a new file – either in RTF-format for text processing programs like Word or as a table in HTML or Excel format. All kinds of export files can be printed without problem.

Document	Search items	Search results	Begin	End
Gospels\Matthew	man	When his mother Mary had been engaged to Joseph, but before they lived together, she was found to be with child from the Holy Spirit. 19 Her husband Joseph, being a righteous man and unwilling to expose her to public disgrace, planned to dismiss her quietly. 20 But just when he had resolved to do this, an angel of the Lord appeared to him in a dream and said, "Joseph, son of David, do not be afraid to take Mary as your wife, for the child conceived in her is from the Holy Spirit.	2	2
Gospels\Matthew	man	23 Then I will declare to them, 'I never knew you; go away from me, you evildoers.' 24 "Everyone then who hears these words of mine and acts on them will be like a wise man who built his house on rock. 25 The rain fell, the floods came, and the winds blew and beat on that house, but it did not fall, because it had been founded on rock.	8	8
Gospels\Matthew	man	25 The rain fell, the floods came, and the winds blew and beat on that house, but it did not fall, because it had been founded on rock. 26 And everyone who hears these words of mine and does not act on them will be like a foolish man who built his house on sand. 27 The rain fell, and the floods came, and the winds blew and beat against that house, and it fell--and great was its fall!	8	8
Gospels\Matthew	man	8 The centurion answered, "Lord, I am not worthy to have you come under my roof; but only speak the word, and my servant will be healed. 9 For I also am a man under authority, with soldiers under me; and I say to one, 'Go,' and he goes, and to another, 'Come,' and he comes, and to my slave, 'Do this,' and the slave does it." 10 When Jesus heard him, he was amazed and said to those who followed him, "Truly I tell you. in no one in Israel have I found such faith.	9	9

An Excel table of the search results for the counted word "man" (and its surrounding paragraph)

5 Lemmatization

MAXDictio permits lemmatization of words in various languages for word frequency and word combination functions. When this option is activated, words are returned to their respective basic forms, so that words with the same meaning are combined regardless of declination or case. For example, only the word “give” will be counted for the various forms “give”, “gave”, and “given”.

In MAXDictio, lemmatization is performed using lemma lists which are available in the following languages:

- Bulgarian
- German
- English
- Estonian
- French
- Italian
- Catalan
- Polish
- Portuguese
- Swedish
- Spanish
- Czech
- Ukrainian
- Hungarian

The lists are in TXT format (UTF-8) and can be edited and expanded as required. It is recommended to create backup copies, because files may be overwritten during a new installation. You can find the lists in the installation folder of MAXQDA:

Windows: local installation

C: \ Programme (x68) \ MAXQDA12 \ Resources \ Lemmatization

Windows: portable installation on USB pen drive

USB drive \ MAXQDA 12 \

Mac: local installation

Program directory > Right-click on MAXQDA12 > Show package contents:

Contents / Resources / Lemmatization

Mac: portable installation on USB pen drive

USB drive > MAXQDA 12 Portable for Mac > Right-click on MAXQDA12 > Show package contents:



Contents / Resources / Lemmatization

Important license note: MAXDictio uses lemmatize lists for this function that have been published under the Creative Commons and the Open Database License. If you use this function for a publication you have to give appropriate credit in a short note like the following:

Lemma list for German: "A lemma list has been used that is based on "Deutsche Morphologie-Daten" by Daniel Naber (<http://www.danielnaber.de/morphologie/>), which is made available under the Creative Commons Attribution-ShareAlike 4.0 license (<http://creativecommons.org/licenses/by-sa/4.0/>)."

Lemma list for other languages: "A lemma list has been used that was originally provided by Michal Boleslav Měchura (<http://www.lexiconista.com/datasets/lemmatization/>), which is made available under the Open Database License (ODbL) (<http://opendatacommons.org/licenses/odbl/1.0/>)."

6 Stop Lists

Function of the Stop Lists

Stop lists are collections of “uninteresting” words, such as definite and indefinite articles, conjunctions, or numerals. Such words are usually sources of irritation when analyzing a text. For that reason, it is advisable to exclude them from the analysis. MAXDictio allows you to create as many stop lists as you’d like, as well as to edit and manage them.

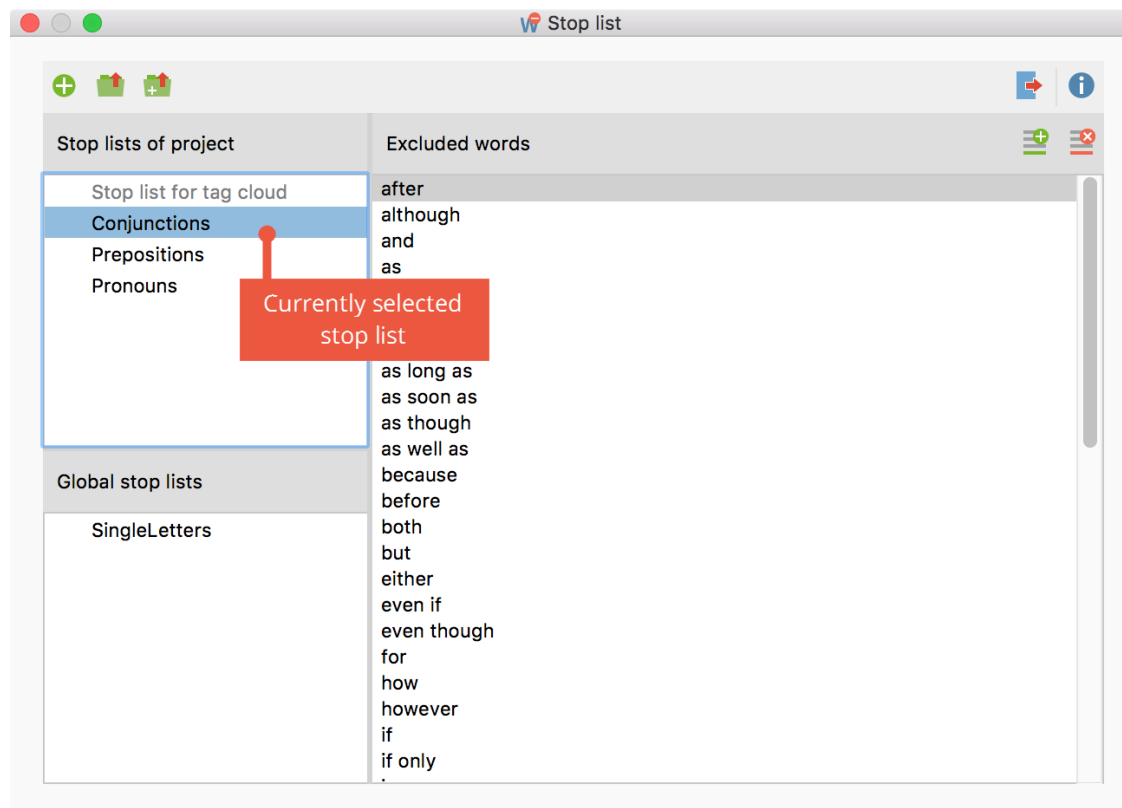
The window for Managing Stop Lists

To manage the stop list, you can either call up **MAXDictio > Stop list** or click the corresponding **Button** in the MAXDictio toolbar.



The “Stop list” quick button in the MAXDictio toolbar

The following window will appear:



The Stop list contains an overview of the excluded words

The window is divided into three sections:

1. Stop lists of project

This section lists all stop lists that are stored in the project file. If you open the project file on another computer in MAXQDA or send it to another person on a team by email, the lists displayed in this section will be available.

The "Stop list for tag cloud" is always visible as the first entry. It contains all stop words used to create word clouds in MAXQDA and can be edited just like any other stop list. It cannot be deleted. Below the "Stop list for tag cloud" at least one further stop list is always available, which also cannot be deleted.

2. Global Stop Lists

This section contains all stop lists that are not stored in the project file, but in the MAXQDA installation itself. These stop lists are available for each project opened with this MAXQDA installation.

If you open a MAXQDA project on another computer, the global lists may not be available.

3. Excluded words

This list displays all search terms that are included in the stop list. The list is always sorted in alphabetical order for user convenience.

Select Active Stop List

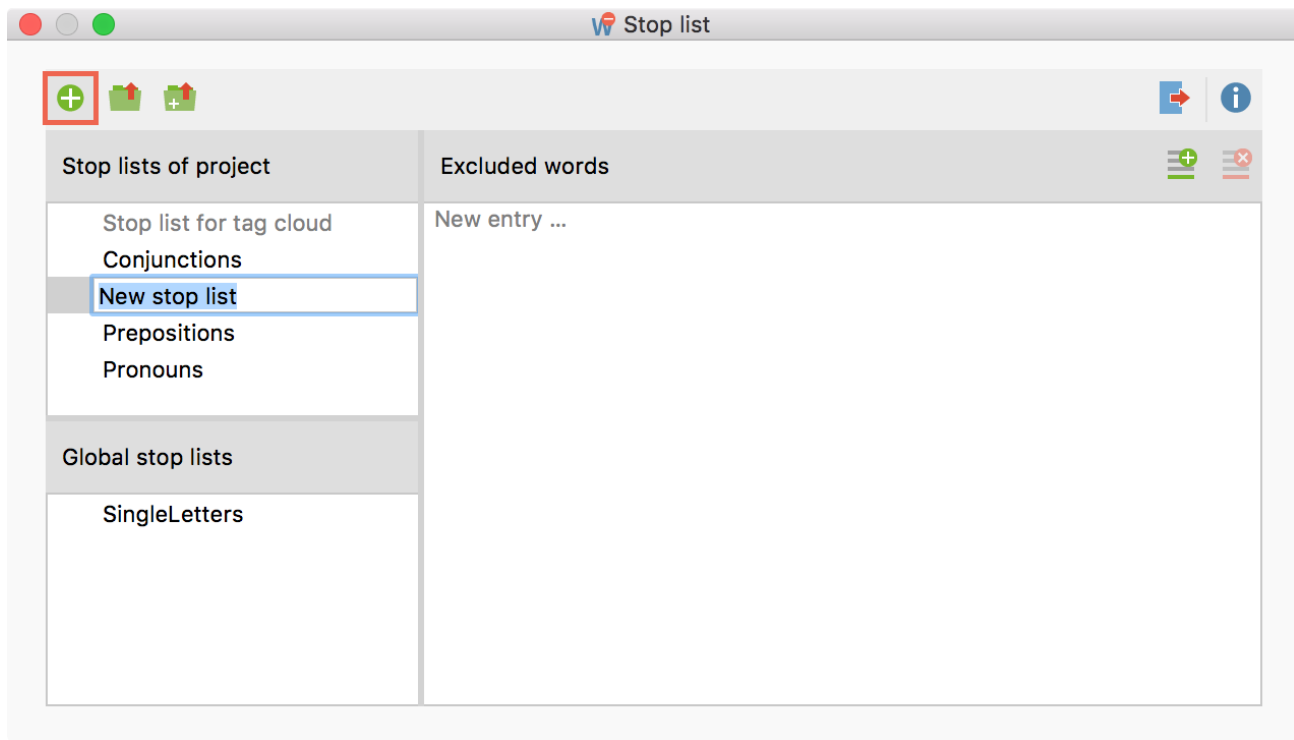
Click the name of a stop list to make it the active list. The stop words of the selected list will be displayed in the "Excluded words" section. MAXDictio always uses exactly these stop words, which are listed in the right pane. The selected stop list is always indicated with a blue or grey background.

Hint: If you have selected a global stop list and open the project on a computer where this global stop list is not available, or has been deleted, MAXDictio will automatically use the next stop list in the list.

Create New Stop List

By default, each MAXQDA project has at least two stop lists. The "Stop list for tag cloud" is used to create tag clouds and is always displayed at the top of the list. This stop list cannot be renamed or deleted. At least one additional stop list is always present.

To add a new stop list to the project, click the  icon.






Icon for adding a new stop list to a project

Thereupon a new list will be added to the project and opened. By default, this list is given the name "New stop list", and may include a sequential number to exclude duplicate names. The name of the list can be edited directly.

Add New Words to a Stop List

New words can be added to the stop list in two ways:

Firstly, you can enter new words by hand, by clicking the **New**  button above the Excluded words pane and typing the word. Multiple words can be added by hitting the Enter key after each entry.

Secondly, you can add words directly from a word frequency list by double-clicking the green symbol  in the first column, thus sparing the effort of typing in the word. Multiple words from a word frequency list can be transferred at once by highlighting them first, then clicking the Add to stop list  icon on the toolbar.

When using stop lists, the following must be observed:

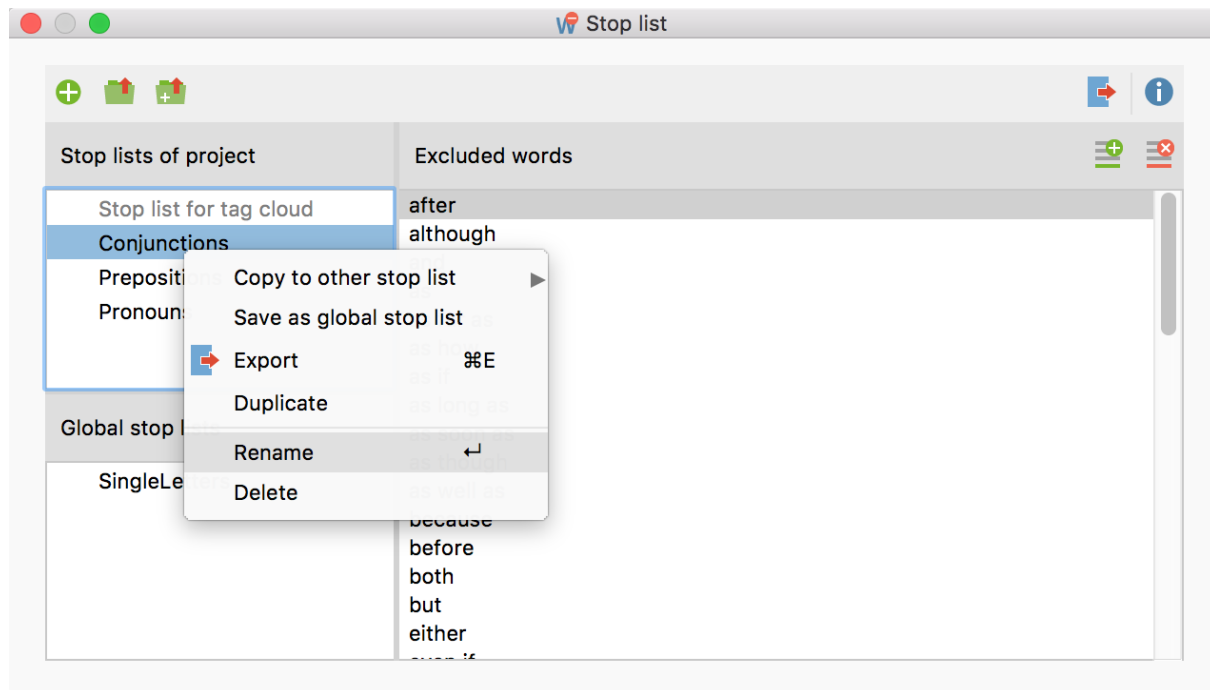
- The words of a stop list are automatically sorted in alphabetical order. This order is predetermined and cannot be modified. If a word already exists in the stop list, it will not be inserted.
- Stop lists can include word combinations separated by spaces, such as "George W. Bush" or "Walt Disney".
- Stop lists are case-sensitive. This distinction is only applied if the corresponding option is selected in the options dialog for a function.

Duplicate, Rename and Delete Stop Lists

Stop lists can be duplicated, for example if you wish to modify the list while retaining the original version. **Right click on the stop list** and select **Duplicate**. Thereupon the list will be duplicated and inserted. The duplicate list will have the same name as the original list supplemented with a sequential number.

To rename a stop list, click on the list to highlight it, then click again to change the name. Alternatively, **right-click the name of the list** and select **Rename**.

The option for deleting a list can also be found in the context menu.



Stop list functions in the context menu

Applying Stop Lists in Other Projects

To save a project-internal stop list as a global list for use in other projects, right-click the stop list and select **Save as a global stop list**. The list will be duplicated and inserted with the same name in the lower window pane. If the list name already exists, a sequential number will be added.

The Global stop lists can be transferred to any project:

1. Right-click the list in the "Global stop list" window pane.
2. Select **Copy into project** from the context menu.

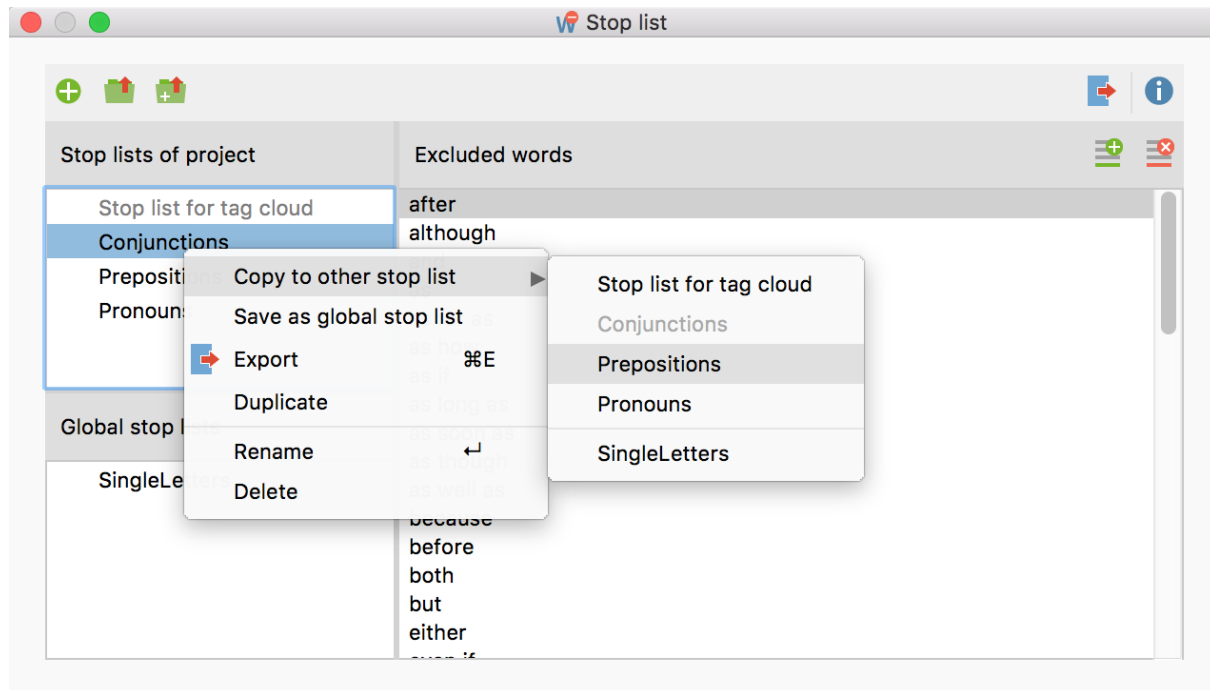
Tip: Stop lists can also be dragged and dropped from the "Stop Lists of Project" to the "Global stop lists" pane.

Hint: When you open a project that has already been opened in version 12.3 in an earlier version of MAXQDA, only an empty stop list will be displayed.

Merge Stop Lists

Many stop lists can be effectively used across projects. Modularization is recommended, meaning that specific types of stop lists (for example, definite and indefinite articles) are created, which can be universally applied. These stop lists can be combined into common stop lists as required.

1. **Right-click on a stop list.**
2. Select **Copy to other stop list** from the context menu.
3. Select the target stop list from the context menu.




Copy words from one stop list to another

After selecting the target list, a confirmation message appears to confirm whether you want to continue. If you confirm the request, the two stop lists are merged, that is, all excluded words from the source list are inserted into the target list, skipping already existing entries.

Export and Print Stop Lists


Stop lists can be exported as files to be edited, printed, or for documentation and archiving in other programs:


1. **Right-click on a stop list** and select **Export**, or click the **Export**  icon in the top toolbar, to export the currently displayed stop list.
2. Provide a file name and location in the dialog window.

The list will be exported in TXT Format (UTF-8) and can be edited and printed using a text editing or word processing program, or Excel.

Importing and Augmenting Stop Lists

MAXDictio allows you to load stop lists from TXT files, save them as new lists, or integrate the words into an existing list. For example, you can import lists that you have downloaded from the Internet, such as from the MAXQDA website: <http://www.maxqda.com/downloads/stoplists>.

Click the **Import list**  icon to create a new list from a TXT file.

Click the **Add to current list**  icon to combine the words from the imported stop list and the words of the currently displayed stop list.

Create and Edit Stop Lists in Office Programs

Stop lists are stored as normal text files in UTF-8 format; they can be opened and edited in any Office program. Excel is the best option for editing stop lists. Open the stop list by selecting Open File in Excel. The file type must be changed to "All files".

	A	B	C
1	after		
2	although		
3	and		
4	as		
5	as far as		
6	as how		
7	as if		
8	as long as		
9	as soon as		
10	as though		
11	as well as		
12	because		
13	before		
14	both		
15	but		

A stop list that has been opened in Excel

You can now add to or modify the list of stop words in Excel, then save them again. It is not necessary to sort the list as this will be done automatically when the list is opened in MAXQDA.

Location of Global Stop Lists

The global stop lists are stored by MAXDictio as files in TXT format (UTF-8) in the following file folder. They can also be edited here using external programs:

Windows: local installation

Documents / MAXQDA12 / Dictio / Stoplist

Mac: local installation

Documents / MAXQDA / MAXQDA12 / Dictio / Stoplist

Windows and Mac: portable installation on USB pen drive

USB pen drive \ MAXQDA 12 Files \ Dictio \ Stoplist

All TXT files in these folders are displayed as stop lists in MAXDictio.

7 Go Lists

What are Go Lists?

Go lists are essentially the opposite of stop lists. While stop lists contain a list of words which are disabled for the analysis and are intended to be excluded from the counting, go lists consist of words to which the analysis should be limited. This means that only these words will be counted.

Manage Go lists

To create a go list or customize existing go lists, select **MAXDictio > Go List**. The window that appears contains the same functionalities as the window for stop lists, the only difference being that there is no designated go list for tag clouds.

An explanation of all functions can be found in the [Stop Lists](#) section.

Use Go Lists for Analysis

MAXDictio allows the use of go lists in the following analyses:

MAXDictio > Word Frequencies (only words of go list) – Only the word frequencies of the currently selected go list are included in the results table (see [Word Frequencies for Words in the Go List](#)).

MAXDictio > Keyword-In-Context – Searches for words of the currently selected go list in all texts and displays them in their respective contexts (see [What is the “Keyword-in-Context” Feature?](#)).

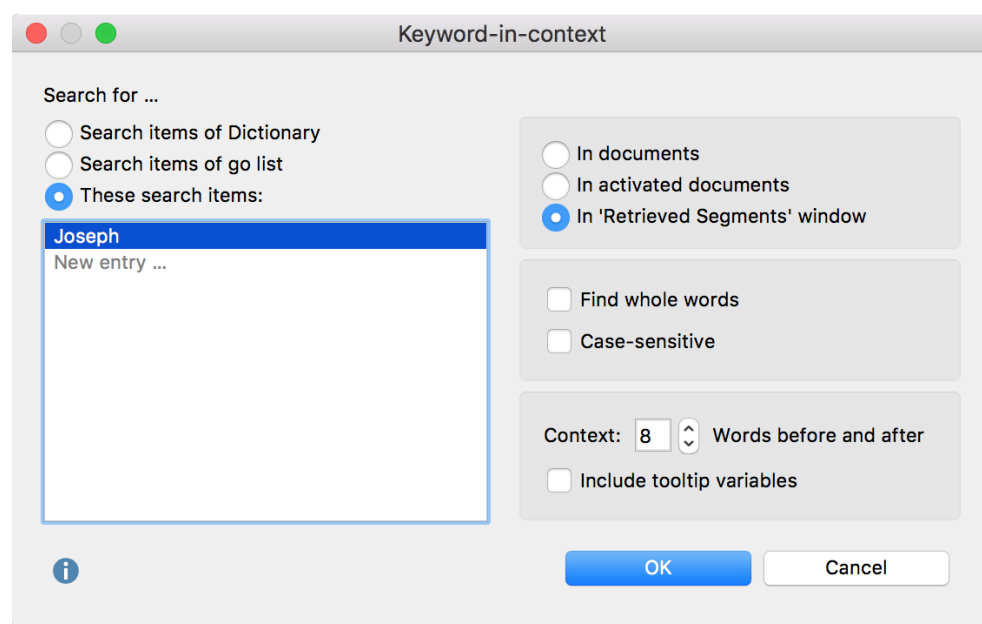
8 Keyword-in-Context

What is the “Keyword-in-Context” Feature?

The keyword-in-context feature allows you search for any number of terms relevant to the analysis (the “keywords”), and view them in a tabular overview along with the words that appear before and after (their respective contexts). The keywords are interactively linked to the original data, allowing you to jump directly to them in the original text.

Call Up the Function and Selecting Options

Begin the keyword search by selecting the menu option **MAXDictio > Keyword-In-Context**. The following dialog window will appear:



Set options for the keyword search

In the left window pane, select the keywords you would like to search for. Three options are available:

Search items of dictionary – All words in the currently selected dictionary will be searched for. The search options set in the dictionary are searched per keyword (case-sensitive, only whole word, beginning of the word). The corresponding options in the dialog window will be ignored.

Search items of go list – All words contained in the currently selected go list will be searched for.

These search items: – Enter the desired search terms in the list. The entries can contain spaces. Hit the Enter key to move to the next line.

In the right pane of the window you define the text to be included in the search. You can choose to search in all documents, in activated documents only, or in text segments that are currently displayed in the “Retrieved segments” window only.

The following further options are available:

Find whole words – If the option is enabled, the search for “or” will not find, for example, “Inventor” or “Elevator”. This option is available only if searching for words in the go list or if you enter specific search items.

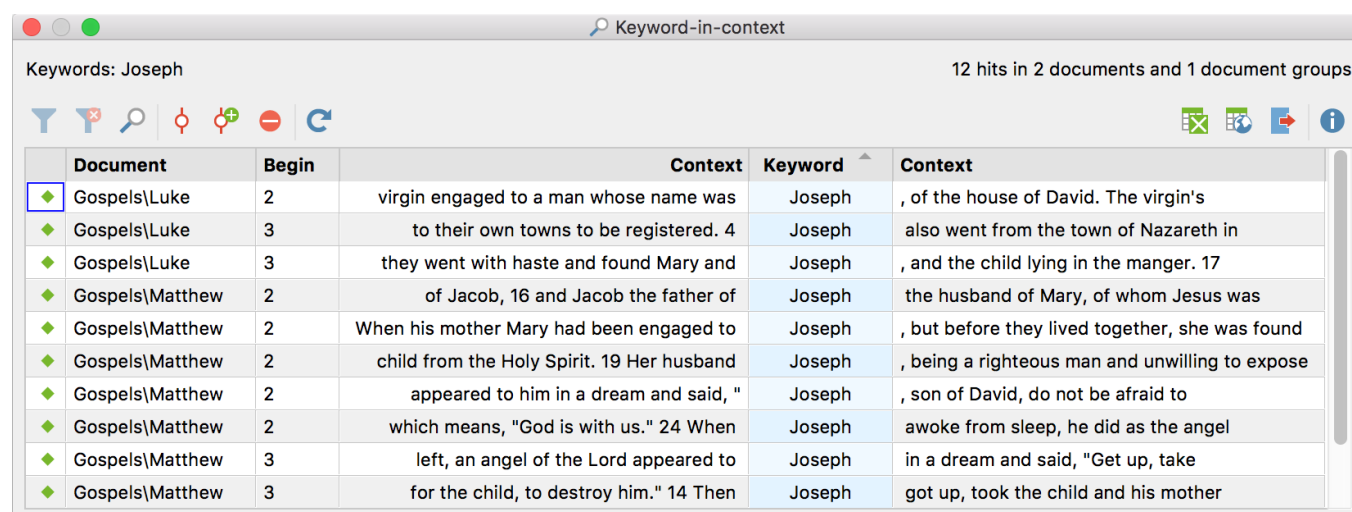
Case sensitive – If this option is enabled, the search for “You” will not find the lowercase form “you”. This option is available only if searching for words in the go list or if you enter specific search items.

Context x Words – Specify how many words to include before and after the keyword. The default setting is 5 words.

Include tooltip variables – Set this option to display document variables that are marked as tooltip variables under > **List of document variables** in their own columns. In the results table, right-hand columns are displayed with the variable values of their respective documents, so that when you look at the results, in addition to the verbal context in the form of the surrounding words, the content context can also be included in the form of variable information of the document in which the keyword appears.

When you click **OK** the following results table appears:

Results Table



Document	Begin	Context	Keyword	Context
Gospels\Luke	2	virgin engaged to a man whose name was	Joseph	, of the house of David. The virgin's
Gospels\Luke	3	to their own towns to be registered. 4	Joseph	also went from the town of Nazareth in
Gospels\Luke	3	they went with haste and found Mary and	Joseph	, and the child lying in the manger. 17
Gospels\Matthew	2	of Jacob, 16 and Jacob the father of	Joseph	the husband of Mary, of whom Jesus was
Gospels\Matthew	2	When his mother Mary had been engaged to	Joseph	, but before they lived together, she was found
Gospels\Matthew	2	child from the Holy Spirit. 19 Her husband	Joseph	, being a righteous man and unwilling to expose
Gospels\Matthew	2	appeared to him in a dream and said, "	Joseph	, son of David, do not be afraid to
Gospels\Matthew	2	which means, "God is with us." 24 When	Joseph	awoke from sleep, he did as the angel
Gospels\Matthew	3	left, an angel of the Lord appeared to	Joseph	in a dream and said, "Get up, take
Gospels\Matthew	3	for the child, to destroy him." 14 Then	Joseph	got up, took the child and his mother

Results table for “Keyword-in-context”

The upper section of the table displays which keywords were searched for, and how many occurrences were found in how many documents and document groups. The table contains one occurrence in each row and shows which document the respective keyword was found in along with its position.

Tip: The results table is interactive. Click on a row to highlight the search term in the original text in the Document Browser.


The table is sorted by keyword in ascending order each time it is opened.



Toolbar in the Results Table



A toolbar containing the following options is located in the upper section of the results table:


 **Filter** – Switches the currently set filter on and off.

 **Reset all filters** – Deletes filters from all columns.

 **Search** – Opens an input field in which you can search within the entire results table.


 **Autocode search results** – Autocodes search results with an existing code. All search hits that are not excluded from autocoding and marked with a stop symbol  in the first column will be autocoded.

 **Autocode search results with a new code** – Encodes the results list with a new code. All search hits that are not excluded from autocoding and marked with a stop symbol  in the first column will be autocoded.


 **Autocode/Export: Ignore hit** – Select one or more rows in the list and click this button to exclude these search hits during autocoding or export. Alternatively, you can double-click the green symbol at the beginning of a row to exclude the search hit.

 **Refresh** – Recalls the dialog box with the options settings.

Standard MAXQDA export options are available, meaning all rows excluding those with stop symbols will be exported (independent of row highlighting):

 **Open as Excel table** – MAXQDA creates a temporary Excel table and opens it with the program in which XLS/X files are saved.

 **Open as HTML table** – MAXQDA creates a temporary HTML table and opens it in the default browser.

 **Export** – MAXQDA opens a dialog window in which a filename and location can be defined. Format options include Excel (XLS/X), Website (HTML), Word (RTF) and Text (TXT, tab delimited). The file will open automatically upon export.

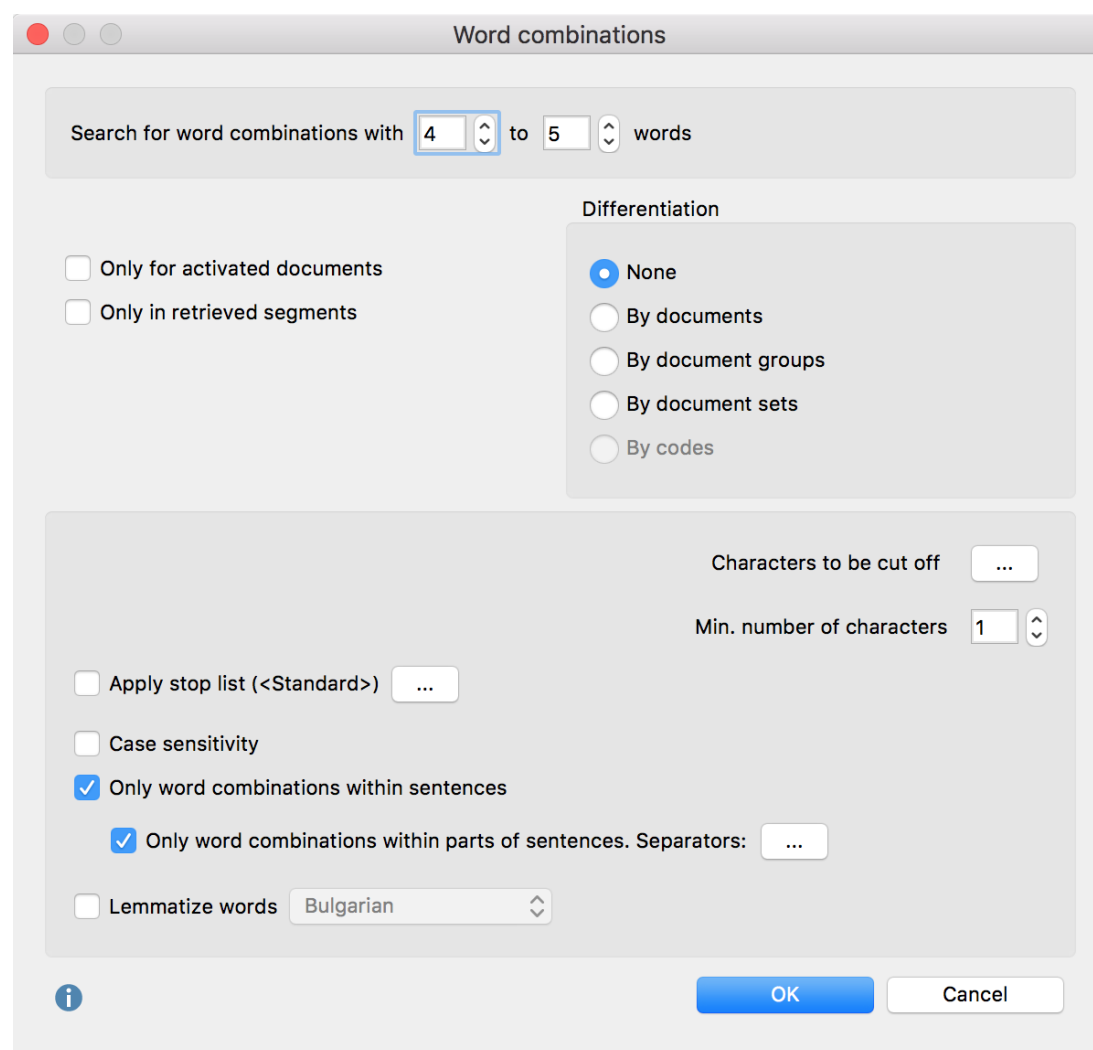
9 Word Combinations

What is the “Word Combinations” Function?

The word combinations function provides an overview of the most frequent combinations of up to 5 words, and word combinations which occur more rarely in selected texts. The function operates similarly to the word frequencies function (see [Analyze Word Frequencies](#)), in that not only individual words but combinations of up to 5 words can be listed.

Call Up the Function and Setting Options

Begin the search for word combinations via the menu option **MAXDictio > Word combinations**. The following dialog window will appear:



Set options for defining word combinations

Number of words

At the top of the dialog window, define how many words to include in the word combination. You can search for combinations of up to 5 words. The setting “Search for word combinations with 3 to 3 words”

will search for all word combinations with exactly 3 words. The setting “Search for word combinations with 2 to 4 words” will list all 2,3, and 4 word combinations.

Selection of texts to be analyzed

Only for activated documents: This option restricts the analysis to currently activated documents.

Only in retrieved segments: This option restricts the analysis to text segments that are currently displayed in the “Retrieved segments” window.

If neither option is selected, all text and table documents in the MAXQDA project will be analyzed.

Differentiation of results

None: The results table will not contain any differentiation of results, but only the totals across all analyzed texts.

By documents, document groups, document sets: The results table will contain additional columns that can be used to compare the frequencies of word combinations of the individual documents, document groups or document sets. With the option **Only activated documents**, only the activated documents are taken into account within the document groups or document sets, and only the groups and sets with documents that are activated are included.

By Codes: This option is only available if the analysis is restricted to the segments in the “Retrieved segments” and a “Simple Coding Query” has been performed. The results table contains additional columns of frequencies for each code that appears in the “Code System”. This option is particularly useful if you have divided texts into text segments using codes for analysis in MAXDictio, as it allows you to compare the frequencies of word combinations within different codes.

Further options

Characters to be cut off: By clicking the button with three dots (...), you will open a dialog box in which you can enter characters which are to be cut off from the words and ignored during the count. The setting of the characters to be cut off applies to all functions in MAXDictio.

How does MAXDictio define a “word”? A word is any sequence of characters between two delimiting characters. Delimiting characters may include, for example, spaces or punctuation marks. For example, the last word of the last sentence in this paragraph, “marks”, is delimited by a space to the left and a period to the right.

The characters to be used as delimiters must be entered in the “Characters to be cut off” dialog box. These normally include punctuation marks, quotation marks, etc.:

@ ! § \$ % & / () = ? ^ ° ' ' ` " „ " " " { } [] \ # + * _ . : , ; < > ~ —

Some problems may occur with hyphens. If compound words should be counted as one word and not be split into their separate parts, the hyphen may not be declared as a delimiter. It is best to experiment a little with different possibilities. As the counting of the word frequencies can be repeated with no significant loss of time, it is advisable to look through the results for any conspicuities, change the options if needed, and then repeat the analysis.

Min. number of characters: Words containing fewer characters than specified here will not be included in the analysis. The default option is set to 1 character. If this option is increased to 2 characters, words such as “he” and “it” will be excluded from the analysis and treated as words on the stop list.

Apply stop list: Check this box to apply a stop list. Clicking on the button with 3 dots (...) will open the stop list window where stop lists can be selected and edited.

Case sensitivity: When this setting is activated, the word combination “Go home” will be treated as a different word combination than “go home”. When the setting is not activated, all words will be displayed in lowercase in the results table.

Only word combinations within sentences: The definition of the word combinations in MAXDictio is such that all words of a text are written successively into a series. Up to 5 consecutive words then form the word combinations, ignoring paragraphs, points, exclamation marks etc. between words. Therefore, it is usually useful to turn this option on to ignore all word combinations that go beyond the end of a sentence.

Example: “It is warm. I am going home.” If the above option is not activated, the two-word combination “warm I”, which are unrelated, would be counted.

Hint: Sentences are defined by MAXQDA according to the following rules: A sentence always begins following a period, question mark, exclamation mark, or colon. The following exceptions apply:

A number that is not four digits appears before a period (e.g. 1. or 2.).

A single character appears before a period (to exclude abbreviations).

Two identical characters appear directly before a period (e.g. ff. or pp.).

Literal speech in quotation marks belonging to the sentence itself.

First letter after a sentence is in small letters.

A number appears directly following the end of a sentence.

Quotation marks appear immediately after the end of a sentence.

In text or table documents, a new sentence invariably begins after a paragraph.

Only word combinations within parts of sentences. Separators ...: Often, it is not logical to count combinations of words that are separated by, for example, a comma. It is therefore advisable to activate this option in order to obtain a breakdown of sentences into parts of sentences. Separators are defined by clicking the button with 3 points (...). By default, the following characters are defined as separators:

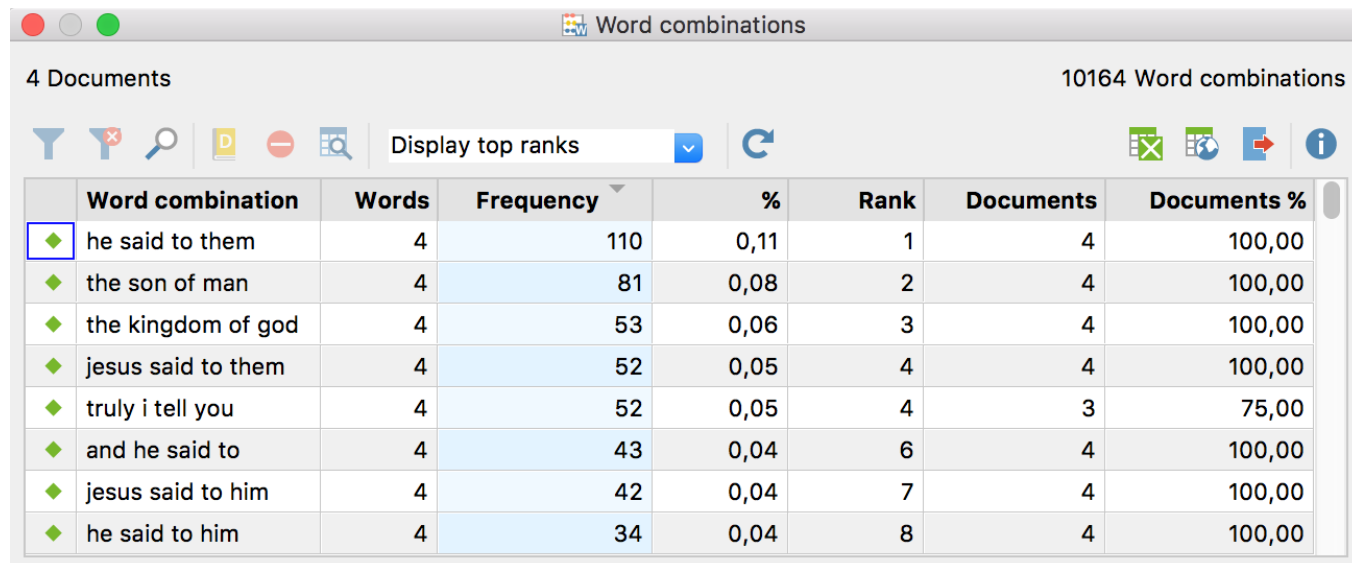
; , - () ... []

Example: “I’m tired, so I’m going home.” If the above option is selected, the sentence will be separated into two parts by the comma, so the 2-word combination “tired so” would not be counted.

Lemmatize words: If this option is activated, each word is returned to its basic form using a lemma lexicon in the selected language. For example, the words “give”, “gave”, and “given”, would be counted only as “give”.

When you click **OK**, the analysis of word combinations will begin. Depending on the length of the text, this process may take a few moments. A display will inform you about the progress of the analysis.

Results Table



4 Documents 10164 Word combinations

Display top ranks

	Word combination	Words	Frequency	%	Rank	Documents	Documents %
◆	he said to them	4	110	0,11	1	4	100,00
◆	the son of man	4	81	0,08	2	4	100,00
◆	the kingdom of god	4	53	0,06	3	4	100,00
◆	jesus said to them	4	52	0,05	4	4	100,00
◆	truly i tell you	4	52	0,05	4	3	75,00
◆	and he said to	4	43	0,04	6	4	100,00
◆	jesus said to him	4	42	0,04	7	4	100,00
◆	he said to him	4	34	0,04	8	4	100,00

Results table for "Word combinations"

The table above displays the number of documents included in the analysis and the number of word combinations found. The table is automatically sorted by frequency each time it is reopened, with the most frequent word combination appearing in the first row.

The columns and information displayed (including toolbars) are the same as those described here: [Word Frequencies: Table of Results](#), the only difference being that word combinations, rather than individual words, are displayed.

Lemmatization and Stop Lists

In the following section, the functioning of lemmatization and stop lists is described:

Stop Lists

- If a word within a word combination, or a word combination itself appears on the stop list, the word combination will be ignored.
- If part of a word combination appears on the stop list, the entire word combination will be ignored.
- All words will be lemmatized. The combinations of lemmatized words will be output.

Interaction of Lemmatization and Stop Lists

- If a lemmatized word appears on the stop list, the combination will also be ignored.
- If a lemmatized combination appears on the stop list, the combination will be ignored.
- If a lemmatized part of a combination appears on the stop list, the entire word combination will be ignored.

10 Interactive Word Tree

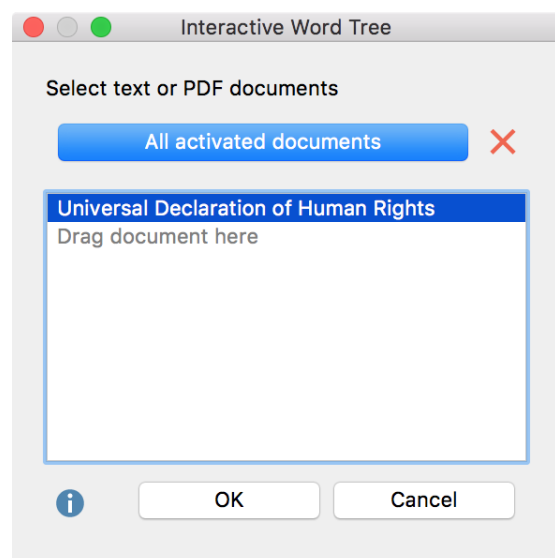
What is the Word Tree function?

Using the Interactive Word Tree, words and word combinations can be explored and analyzed visually in their respective contexts. The more frequently a word or word combination occurs, the more distinctly it appears in the tree. The Word Tree offers two-way interactivity: Firstly, it can be navigated via its individual “branches” in order to visualize words in their contexts. Secondly, the data is interactively linked to the original texts, so you can view words and phrases in MAXQDA's “Document Browser”. These functionalities make the Word Tree a visually supported form of keyword-in-context (KWIC) analysis.

WordTrees were originally developed by Wattenberg & Viégas (2008) (“The Word Tree, an Interactive Visual Concordance”) and have since formed part of the permanent repertoire of text exploration tools.

Call Up the Function and Select Text

To create a Word Tree, go to **MAXDictio > Interactive Word Tree**. The following window will appear, in which you can use the mouse to import text or PDF documents from the “Document System” window.



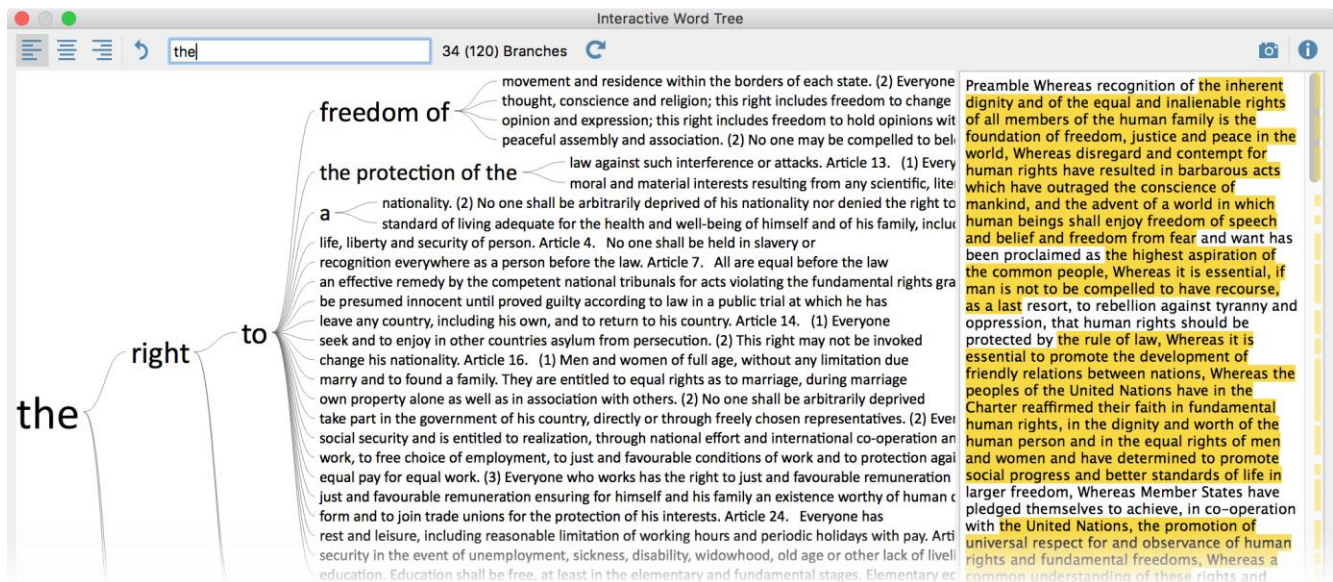
Selection of texts for the Word Tree

Alternatively, click the **All activated documents** button to add the currently activated text and PDF documents to the existing selection.

Click **OK** to begin creating your Word Tree. Progress of this operation is indicated in the display.

The Interactive Word Tree

The following figure shows a Word Tree for the UN Declaration of Human Rights created in MAXDictio.



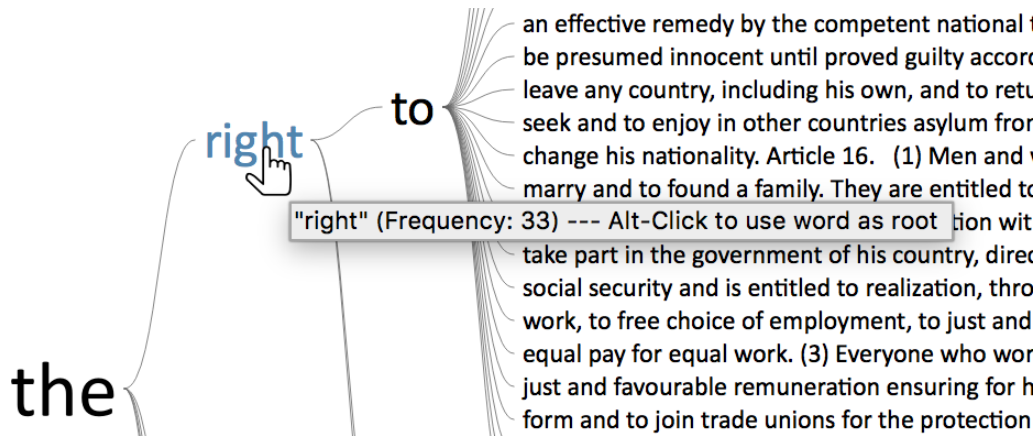
Word Tree for the UN Declaration of Human Rights

The Word Tree window is structured as follows:

- The actual Word Tree is displayed in the left pane.
- The right pane contains the original text without formatting. If several texts have been selected, they are displayed one after the other.
- In the upper window area, a toolbar allows you to adjust the view and to export the tree.

The Word Tree is created according to the following logic:

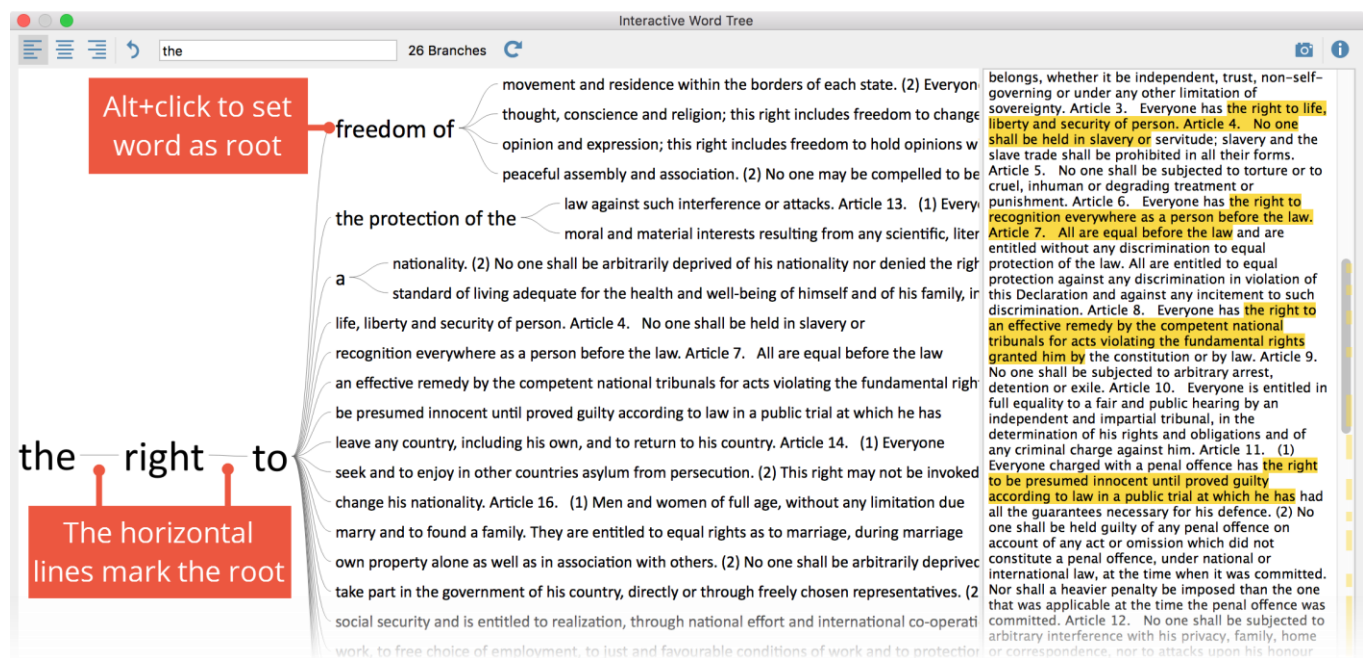
- The most frequent word is shown on the far left, in the example "the". This word forms the root of the Word Tree. The next most frequent word following the root word is displayed on the top branch, in the example, "right". The word "United" is the next most frequent word and is positioned at the second branch from the top. According to this logic, each branch of the tree grows longer, until only one phrase remains in the form of one branch per line.
- Stop words are not applied to avoid breaking the reading flow.
- Punctuation marks such as periods are considered single words.
- The visible portion of the word tree is displayed in the display pane. The number of currently displayed branches with different text lines is displayed in the upper toolbar. If there are more branches than space, the number of currently visible branches is displayed, followed by the total number of branches in parentheses. In the example above, 34 of a total of 120 branches are visible.
- The tooltip that appears when you hover over a word with the mouse shows how often this word appears in the analyzed texts. The following figure shows that the word "right" appears 33 times in the Declaration.



The tooltip shows the frequency of the highlighted word

Selecting Sentences and Extending the Root

Click on a word in the tree with the mouse to define all words from the previous root to the selected word as the new root of the tree. One can imagine observing a branch with binoculars and ignoring the other parts of the tree. The following screenshot displays the view of the tree after clicking the word "to" in the upper part of the tree:



Display of the updated Word Tree: The root of the tree is now "the right to".

Between the three words “the right to”, horizontal lines are displayed to show that “the” and “right” are followed by other words and phrases, but this word combination has been selected as the current root. As shown in the upper toolbar, the word combination “the right to” appears a total of 26 times in the text.

Hint: If there is no horizontal line displayed between two words (such as “freedom of” in the example), there are no further branches after the displayed branch. In this case, the first word in the text is always followed only by the second.

To shorten the root, click on one of the preceding words in the root. In the example above, clicking the word “the” shortens the root to this word.

Tip: Click the **Undo changes** icon  to restore the previous views of the tree step by step.

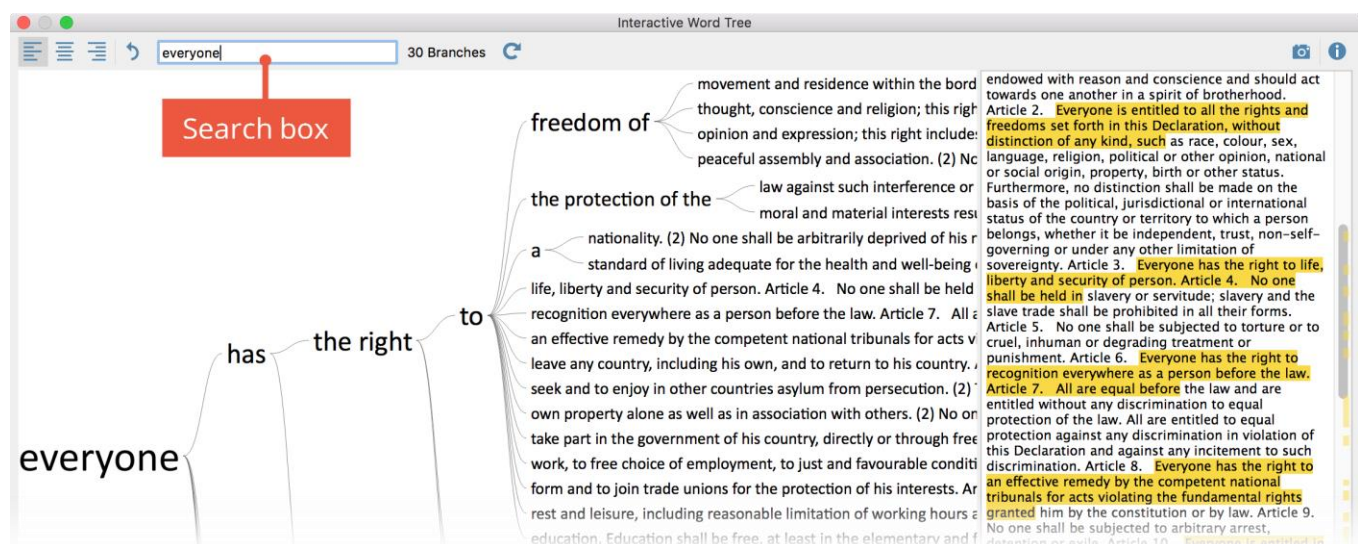
Using Words or Word Combinations as the Root

You can make any word or phrase in the tree the new root by holding down the **Alt** key and clicking the word. The clicked word immediately becomes the new root.

The words in the right pane containing the analyzed text are also interactive: click a word to make it the new root.

Search Words and Phrases

One of the main functions of the Word Tree is to search for interesting words within a text and to explore them in their respective contexts. To this end, a search field is located at the top of the window. Hit Enter key to begin the search process; if there is a hit, it will go directly to the new root of the tree:



Searching words within the text

Jump to Original Text

The right pane contains the complete analyzed text; if several text documents are selected, these are displayed one behind the other. The yellow-highlighted text corresponds to the phrases shown in the tree. The yellow highlighting on the scroll bar shows the positions of the phrases in the text to provide better orientation.

Right-click a yellow highlight in the text and select **View in Document Browser** to display the associated document in the “Document Browser” and highlight the corresponding text for further exploration.

Tip: Hover over the right pane of the display window to display the document name in the tooltip that appears.

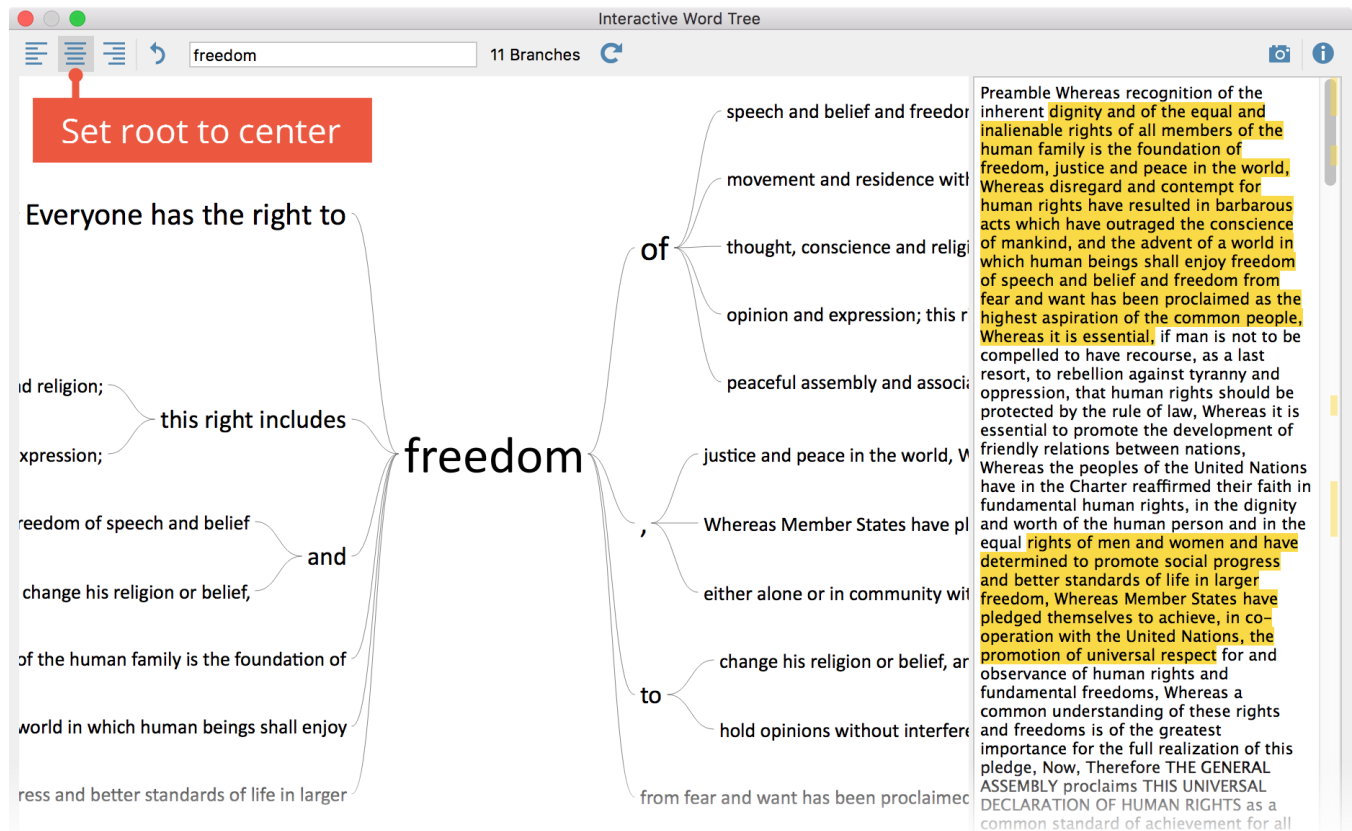
Placement of the Root: Start, Center, End

By default, the word of interest is displayed on the left, so you can examine which words follow. If, however, you want to analyze which words appear in the text before the word of interest, you can adjust the placement of the root using the three icons on the upper left of the screen:

Root: start – The root, that is the selected phrase, is left-justified in the window so you can examine which words follow.


Root: center – The selected phrase is displayed in the center of the window, so you can examine which words appear before and after. Clicking a word on the right half of the tree extends the root to the right; clicking a word on the left extends the root to the left.

Root: end – The selected phrase is right-justified in the window so you can examine which words appear before it.



The root "freedom" in the center position

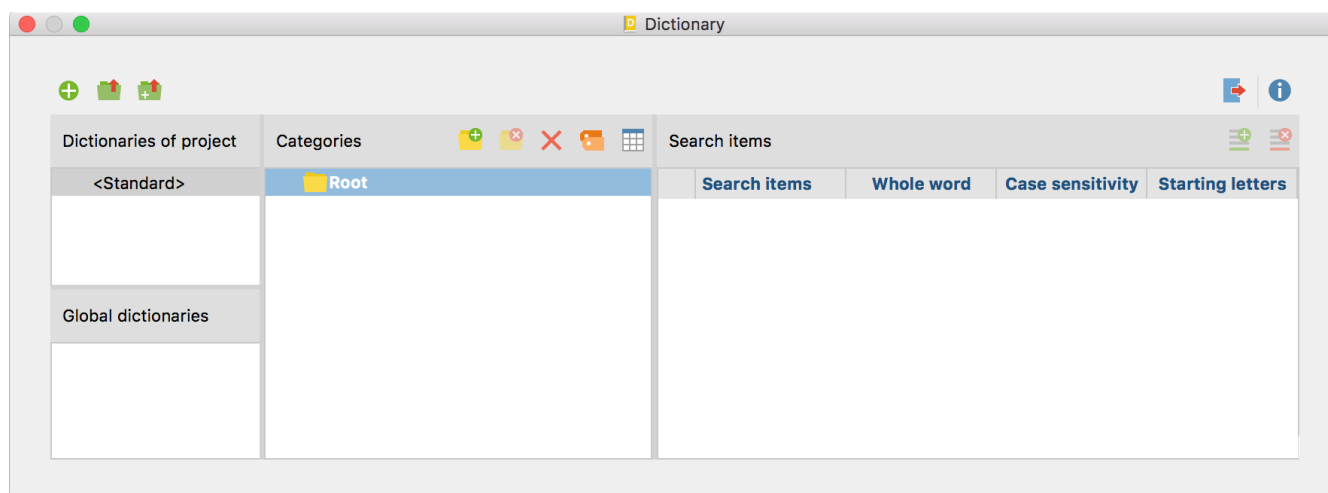
Export Word Tree

For documentation and presentation, you can copy the currently displayed Word Tree to the clipboard by clicking the **Copy current display as an image to clipboard** icon  to insert a graphic of the tree into a word processor such as Word or a presentation software such as PowerPoint. For better quality, the number of pixels is doubled relative to the view.

11 Managing Dictionaries

The Dictionary Function

MAXDictio allows you to create and manage dictionaries. A dictionary consists of one or more categories in which search terms can be defined.



Window for managing dictionaries

The window is divided into several sections:

1. Dictionary of project

This section lists all the dictionaries that are stored in the project file. If you open the project file on another computer in MAXQDA or send it to another person from a team by email, the dictionaries listed here will be available.

A new project contains a dictionary entitled <standard>, which is initially empty and can be renamed as you please.

Each project contains at least one dictionary; all dictionaries cannot be deleted.

2. Global dictionary

This section lists all the dictionaries that are stored not in the project file, but in the current MAXQDA installation itself. These dictionaries are available for each project opened with this MAXQDA installation.

If you open a project on another computer with MAXQDA, these dictionaries may not be available.

3. Categories and search items


In these sections, the currently selected dictionary can be edited. The “Categories” section contains categories in the form of a tree or linear list. The “Search terms” area contains the terms of the respective category and allows you to define the search options for the respective term.

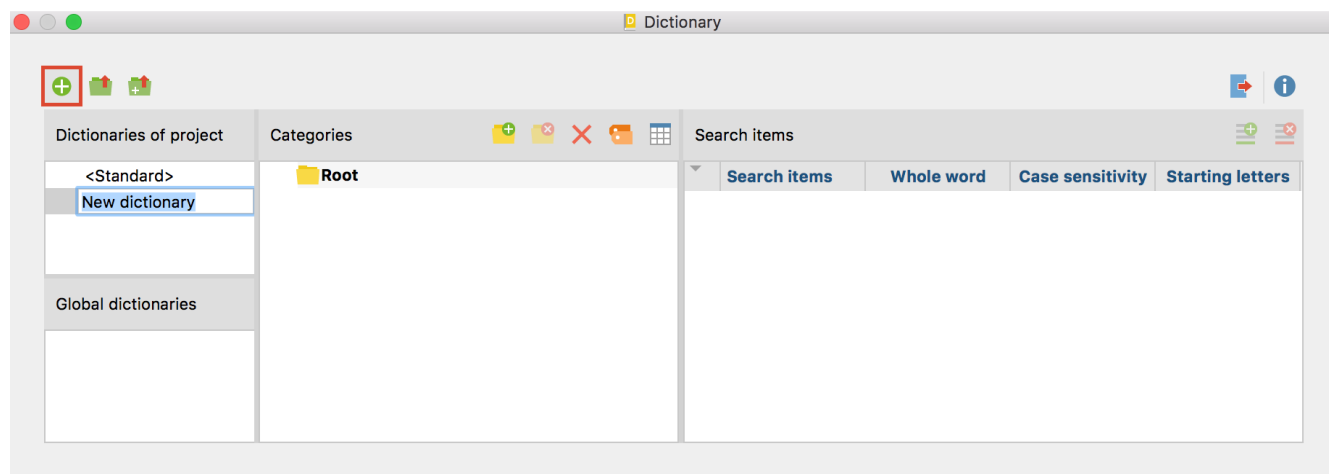
Select Active Dictionary

Click on the name of a dictionary to select it and make it the active dictionary. The corresponding categories and search terms will be displayed in the right pane. The currently selected dictionary can be identified by its blue or grey background.

Hint: If you have selected a global dictionary and open the project on a computer where this dictionary is not available, MAXDictio will automatically select the next dictionary on the list.

Create a New Dictionary

To add a new dictionary to the project, click the  icon.



Icon for adding a new dictionary to the project


A new dictionary will be added to the project and opened. By default, this dictionary will be named “New dictionary” with a sequential number as required in order to avoid duplicate names. The dictionary name can be modified directly.

Create Categories and Sub-categories

In MAXDictio basic and hierarchical types of dictionaries can be created. Hierarchical dictionaries are similar to the “Code System” of MAXQDA, as subordinated levels can be switched on or off as desired. There is no limit to the number of levels, but a dictionary normally doesn’t have more than three levels.

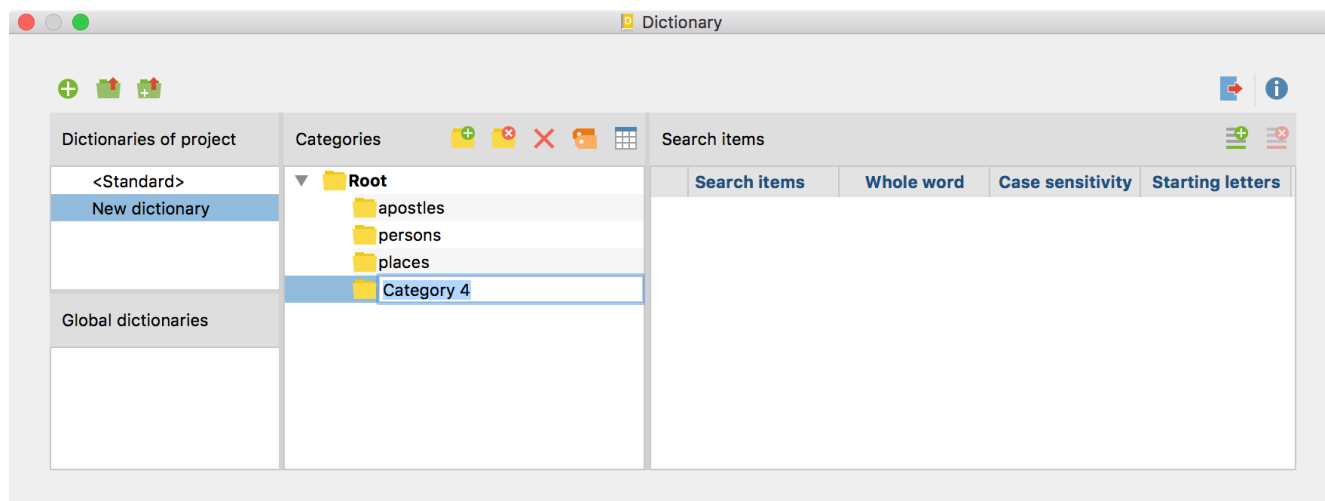
A dictionary consists of categories and sub-categories as well as a list of the search items that are allocated to each category.

Note: A dictionary in MAXDictio can contain any number of categories. Exactly one list of search items belongs to each category. The number of search items is unlimited as well.

To add a new category, click on the **New category** button  on the left side of the toolbar. Then a new category title can be entered.


Note: The title is not restricted to a certain length or form. It can also consist of several words, such as “old-age pension.” The title of a category can be modified later without effecting the dictionary itself.

After clicking on **New category**, MAXDictio will add a category named “Category #.” Here the character # stands for the sequential number of the category in the existing category system.



The illustration above shows that there are already three defined categories, namely:


- apostles
- persons
- places

After clicking on **New category** , MAXDictio will add a new category named “Category 4” at the bottom of the list.

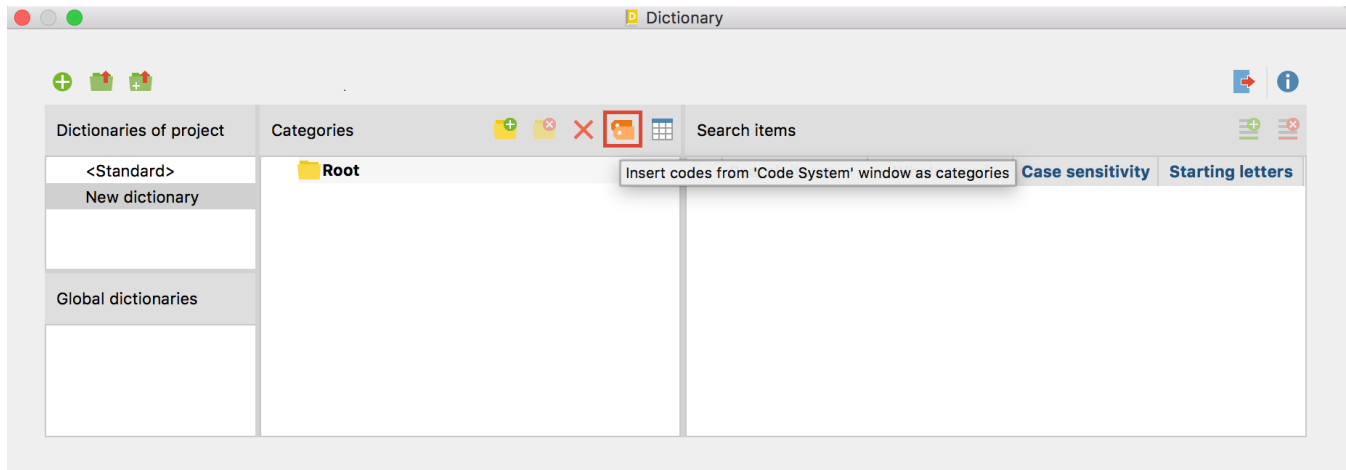
In most cases, you will want to change the title into a more meaningful title. To do so, double-click on “Category 4” and enter the new title.

Transfer Category Names From the “Code System”

The hierarchy and code names can be transferred directly from the “Code System” to the currently selected dictionary:

1. First, activate the desired codes in “Code System”.
2. Next, click the **Insert codes from “Code System” window as categories**  icon.

MAXDictio will ask if only the activated codes should be accepted and then inserts the new categories into the tree. Existing categories names on the same level will not be inserted twice, but rather their subcategories will be integrated.



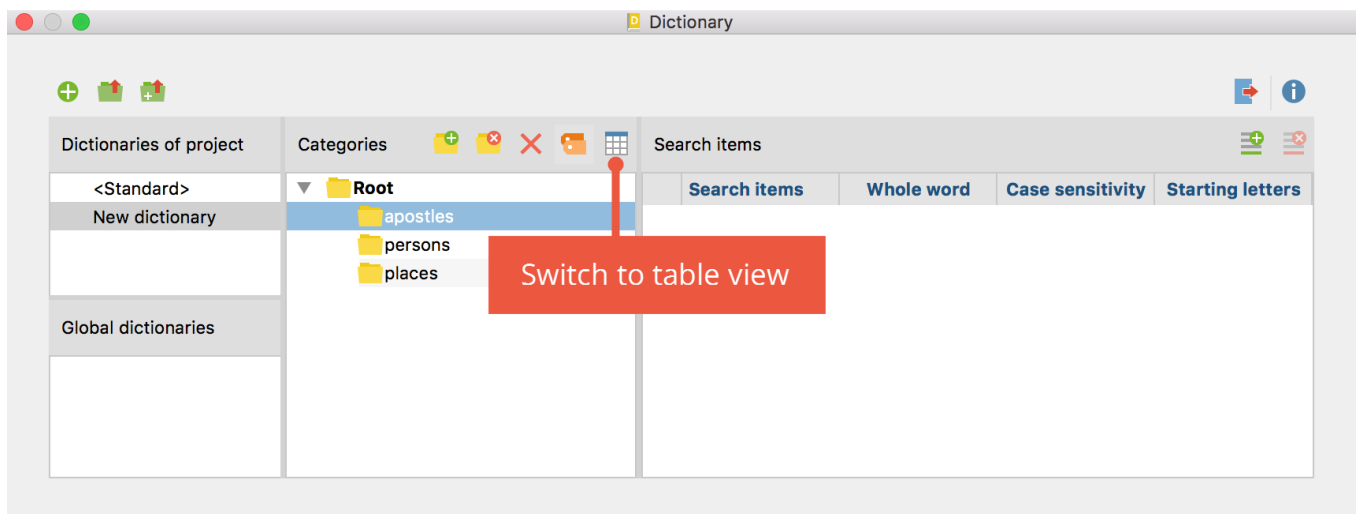
Insert categories from the “Code System” window

Tip: You can also drag and drop a code from the “Code System” into the category window. The code name and related subcode names will be inserted in the dictionary.

Sort and Organize the Category System

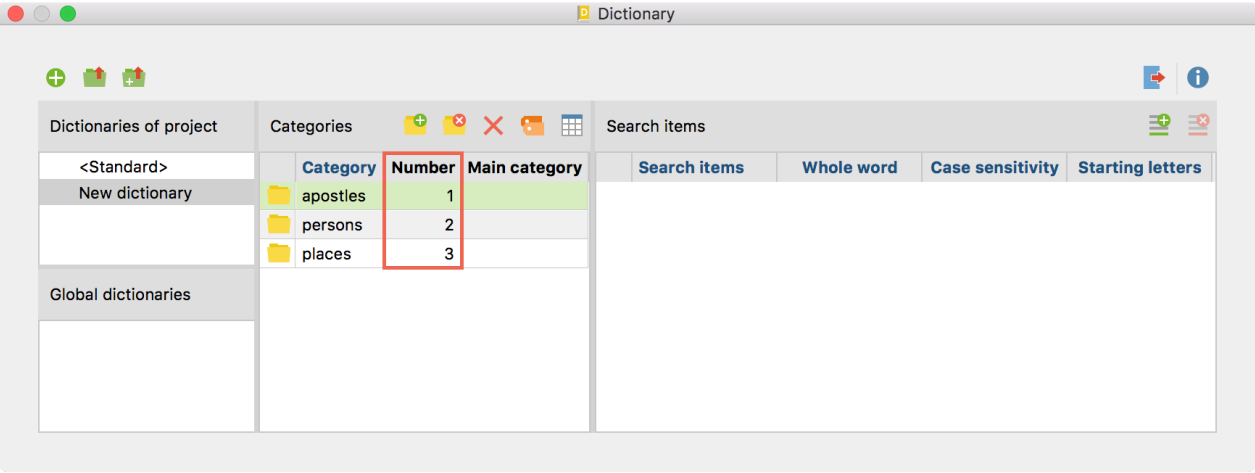
The category system can be displayed as a list or as a hierarchical tree.

When using the hierarchical view, the category system may be organized in the same way as the “Code System” in MAXQDA: by dragging the mouse, categories can be placed at the desired position.



The hierarchical view of the category system

When working in list view (see the following figure) you may sort the whole category system by clicking on the **column’s header**.



The list view of the category system

Note: The “Number” column contains an identification number that is created in the sequence the categories were created. Thus, the category defined first, is associated with the number 1.


Turn Categories On and Off

Categories may be switched on and off so that you can decide which categories will be counted by MAXDictio. In the left column in front of each category is a symbol which indicates whether a category is turned on or off for the counting. The initial state of a category is the on-state. This is indicated by the plain folder icon. Double-clicking on the icon will change this to a stop sign. This shows that the category will not be included in the analysis.

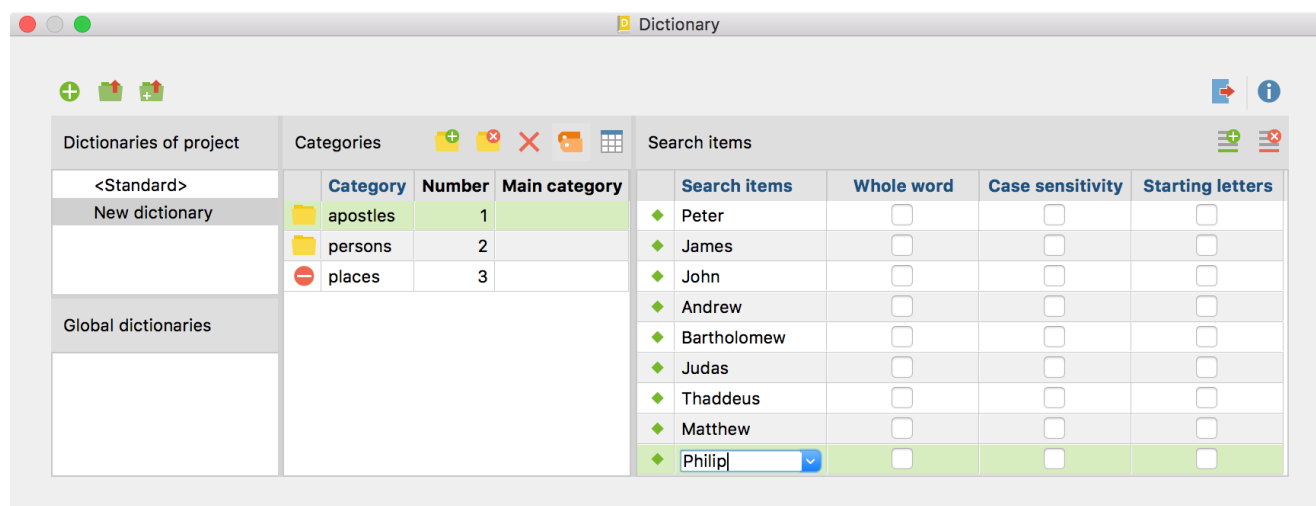
	Category	Number	Main category
⊘	apostles	1	
📁	persons	2	
📁	places	3	

The red cross in the first column indicates that the category “apostles” is turned off

Add and Delete Search Items

To add category search items, select the desired category by clicking on it with the left mouse button. Then click the **New search item button** in the toolbar . Now it is possible to enter a new search item. However, it is much easier to take a search item directly from a word frequency list and add it to the appropriate category (see description below).

An easier way to add terms it to simply drag them in from a word frequency list or right-click on them there and choose to add them to the dictionary.



A new search item is added to the category system

Individual or multiple search items can be deleted from the list by selecting them and then clicking on **Delete search item**.

Tip: Several words can be selected by using the combinations **Ctrl (Win) or cmd (Mac) + left mouse button** or **Shift + left mouse button** (to select all of the words between an already selected word and another).

Turn Search Items On and Off

Each search item of a category can be turned on or off individually for the analysis. The procedure is the same as with the categories. The symbol next to each word indicates its state: green point = on, stop sign = off.

Set Options for the Search Items

Options can be set for each search item. Search items are strings of characters whose occurrences are supposed to be investigated in the text. These can be individual words, parts of words (such as the stem of a word), or terms consisting of several words or strings of characters.

The following options are available:

Whole word – only words that match completely with the search item will be counted

Case sensitivity – case will be taken into consideration

Starting letters – the search item must appear at the beginning of a word

Examples:

"Whole word"

"president" will code President, but not presidential

"Whole word" and "Case sensitivity"

"president" will not code President

"Starting letters" activated

"tree" will code trees and treehouse

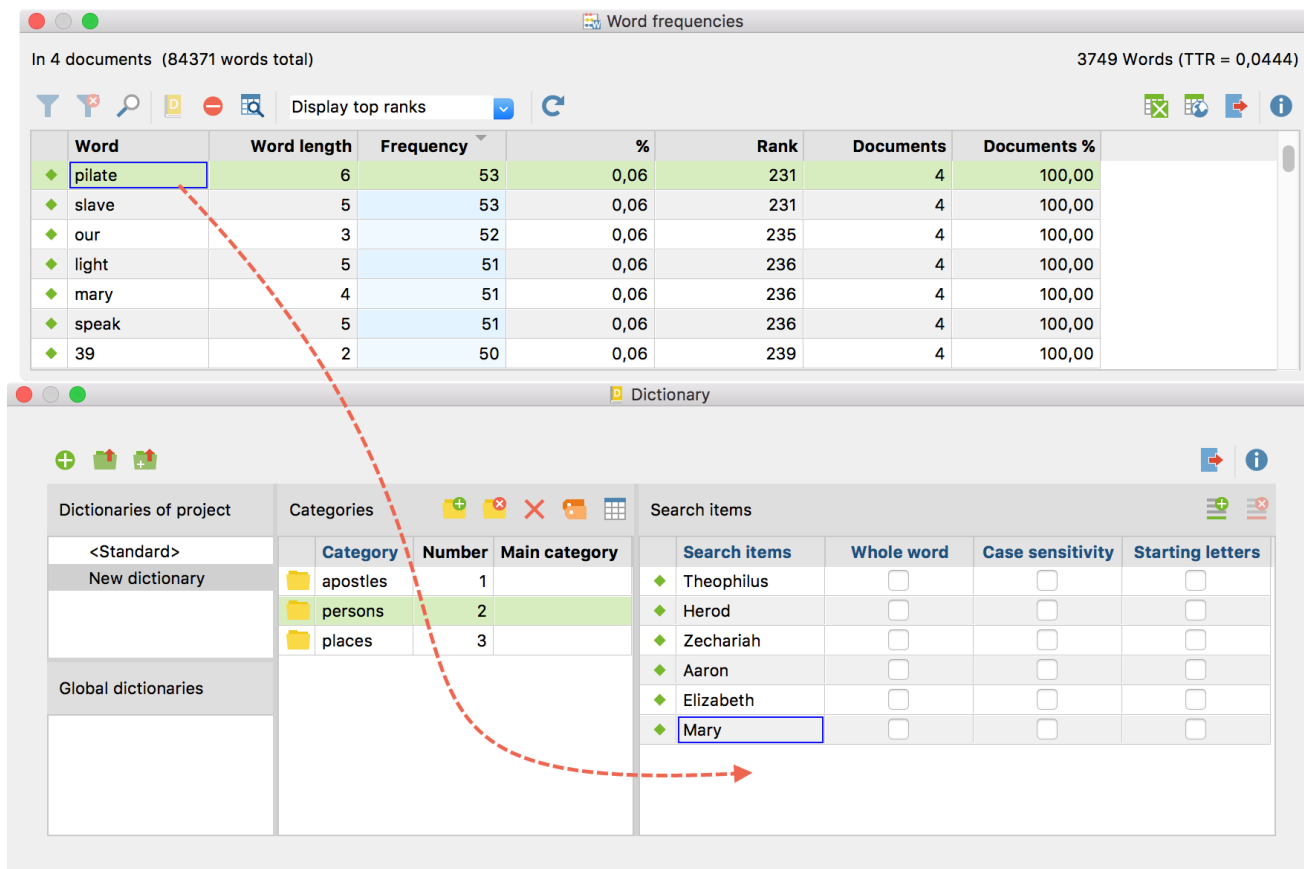
"Starting letters" deactivated

"tree" not only codes tree, but also street

Transfer Words from the Word Frequencies List into a Dictionary

Each word that is contained in a word frequency list can be sent directly to the dictionary:

1. The best way to do this is to open both lists and arrange the windows on the screen so that they are next to each other and can be seen at the same time.




The screenshot shows two windows side-by-side. The top window, titled 'Word frequencies', displays a table with columns: Word, Word length, Frequency, %, Rank, Documents, and Documents %. The bottom window, titled 'Dictionary', shows a list of categories (apostles, persons, places) and a search items table. A red dashed arrow originates from the word 'pilate' in the frequency list and points to the 'persons' category in the dictionary window.

Word	Word length	Frequency	%	Rank	Documents	Documents %
pilate	6	53	0,06	231	4	100,00
slave	5	53	0,06	231	4	100,00
our	3	52	0,06	235	4	100,00
light	5	51	0,06	236	4	100,00
mary	4	51	0,06	236	4	100,00
speak	5	51	0,06	236	4	100,00
39	2	50	0,06	239	4	100,00

Category	Number	Main category
apostles	1	
persons	2	
places	3	

Search items	Whole word	Case sensitivity	Starting letters
Theophilus	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Herod	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Zechariah	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Aaron	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Elizabeth	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Mary	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

2. Optimized view to transfer words from word frequency list into a dictionary
3. Click on that category in the dictionary window in which you would like to add words.
4. To transfer a word from the word frequency into the chosen category, you can either
 - a) double-click on the desired word in the frequency list, or
 - b) drag the desired word with the mouse on a category name or in the list of search items, or
 - c) use the button **Add to dictionary**  or select the equally called entry in the context menu.

Tip: Several words can be selected by using the combinations **Ctrl (Win) or cmd (Mac) + left mouse button** or **Shift + left mouse button** (to select all of the words between an already selected word and another).

Transfer Search Items from One Category to Another

It is easy to transfer search items from one category to another. Just click on the term and hold the left mouse button down as you move it into the desired category. This is also known as the drag-and-drop function.

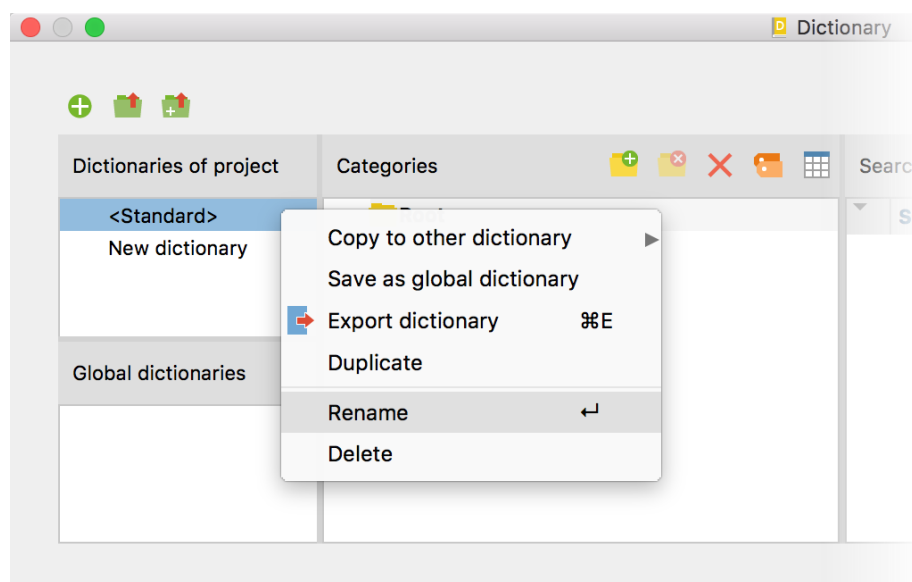
The search item will be removed from the old category immediately. The setting options for the given search item (such as “whole word”) accompany the word to the new category.

Duplicate, Rename and Delete Dictionaries

Dictionaries can be duplicated, for example if you wish to modify a dictionary while retaining the original version. Right-click the name of a dictionary and select **Duplicate**. Thereupon the dictionary will be duplicated and inserted. The duplicate will have the same name as the original dictionary, supplemented with a sequential number.

To rename a dictionary, highlight it with the mouse, then click again to enter a new name. Alternatively, you can right-click the name of the dictionary and select **Rename**.

The option for deleting a dictionary is also located in the context menu.



Functions for dictionaries in the context menu

Use a Dictionary in Other Projects

To save a project-specific dictionary as a global dictionary for use in other projects, right-click the name of the dictionary and select **Save as global dictionary**. The dictionary will be duplicated and inserted under the same name in the lower pane. If the name already exists, a sequential number will be added.

In turn, the global dictionaries can be transferred to any project:

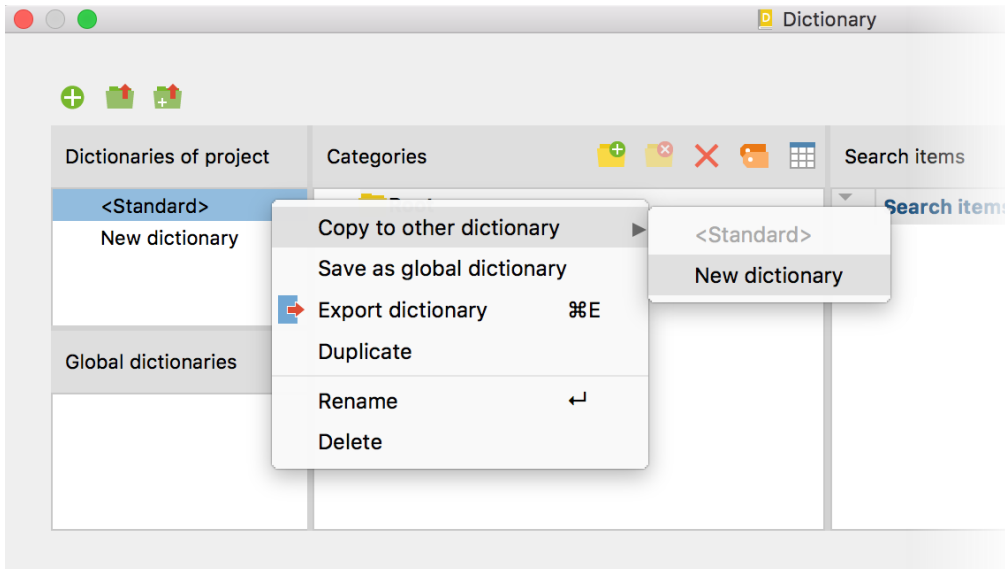
1. Right-click an entry in the lower pane, “Global dictionaries”.
2. From the context menu, select **Copy into project**.

Tip: Dictionaries can be conveniently dragged and dropped between the “Dictionaries of project” in the upper pane and “Global dictionaries” in the lower pane.

Merge Dictionaries

MAXDictio allows multiple dictionaries to be merged

1. Right-click on a dictionary.
2. Select **Copy to other dictionary** from the context menu.
3. Select the target dictionary from the menu that appears.




Copy search items from one dictionary to another

After selecting the target, a request appears to confirm whether you want to continue. If you agree, the two dictionaries will be merged, that is, all categories and search terms from the source list will be inserted into the target dictionary, ignoring already existing entries.

Export and Print Dictionaries

Dictionaries can be exported into other programs, for example, for editing, printing, or for documentation or archiving purposes.

1. Right-click the dictionary and select **Export**, or click the **Export**  icon in the toolbar to export the currently displayed dictionary.
2. In the dialog window that appears, enter a file name and location.


The exported file will be saved in TXT format (UTF-8) and can be opened, edited and printed in text editing programs, Excel, or word processing programs.


	A	B	C	D	E	F	G
1	Category	Search item	Whole word	Case sensitivity	Starting letters	Category activated	Search item activated
2	apostles	Peter	0	0	0	1	1
3	apostles	Andrew	0	0	0	1	1
4	apostles	James	0	0	0	1	1
5	apostles	John	0	0	0	1	1
6	apostles	Philip	0	0	0	1	1
7	apostles	Bartholomew	0	0	0	1	1
8	apostles	Thomas	0	0	0	1	1
9	apostles	Matthew	0	0	0	1	1
10	apostles	Judas	0	0	0	1	1
11	apostles	Thaddaeus	0	0	0	1	1
12	apostles	Simon	0	0	0	1	1

Exported dictionary in Excel

Add and Supplement Dictionaries

MAXDictio allows you to import dictionaries from TXT or Excel files as well as from the previous MAXDictio format, MXD, as new dictionaries. In addition, existing dictionaries can be supplemented with files in TXT or Excel format.

Click the **Import dictionary**  icon to create a new dictionary from a file.

Click the **Add Dictionary**  icon to add the words of the imported dictionary to the categories and search terms of the currently displayed dictionary.

Since the introduction of MAXQDA 11 for Mac, MAXDictio has created a new, simple import format for importing and exporting dictionaries from files. If you wish to import a dictionary from Excel, create an Excel file that contains the following columns:

Category	Search item	Whole word	Case sensitivity	Starting letters	Category activated	Search item activated
Category 1	Search Item 1	0	0	0	1	1
Category 1	Search Item 2	0	0	0	1	1
Category 1\Subcat. 1.1	Search Item 2	1	1	0	1	1
Category 1\Subcat. 1.2		1	1	1	0	
Category 1\Subcat. 1.1\Subc. 1.1.1		1	0	0	0	1

Tip: The easiest method is to next export an empty dictionary, because MAXQDA automatically creates an Excel file with the necessary column headings.

Each search term of the dictionary has its own row. In the "Category" column, specify the category where subcategories can be created using the backslash "\". The search string must be entered in the "Search

item" column. The other columns can be used to switch the available dictionary options on and off. A "1" stands for "yes" and a "0" stands for "no".

Tip: To create a backslash on a Mac, use the keyboard combination Alt + Shift + 7.

When importing, MAXDictio adheres to the following rules:

- Column order is irrelevant.
- The order of rows determines the sorting of categories in tree view.
- If a category is set to "on" somewhere in the file, it is switched on.
- If the "Category activated" or "Search item activated" columns are missing, the search terms and categories are all set to "on".
- All values different than "1", or missing values, will be interpreted as "0".
- If a word a category already exists, it will be ignored.
- Rows in which categories but no search terms have been defined will be inserted as empty categories.
- Rows in which categories but no search terms have been defined will not have entries for any of the three options.

Saving Location of Global Dictionaries

Global dictionaries are stored by MAXDictio as files in TXT format (UTF-8) in the following file folder, and can also be edited there using external programs:

Windows: local installation

Documents / MAXQDA12 / Dictio / Dictionary

Mac: local installation

Documents / MAXQDA / MAXQDA12 / Dictio / Dictionary

Windows and Mac: portable installation on USB pen drive

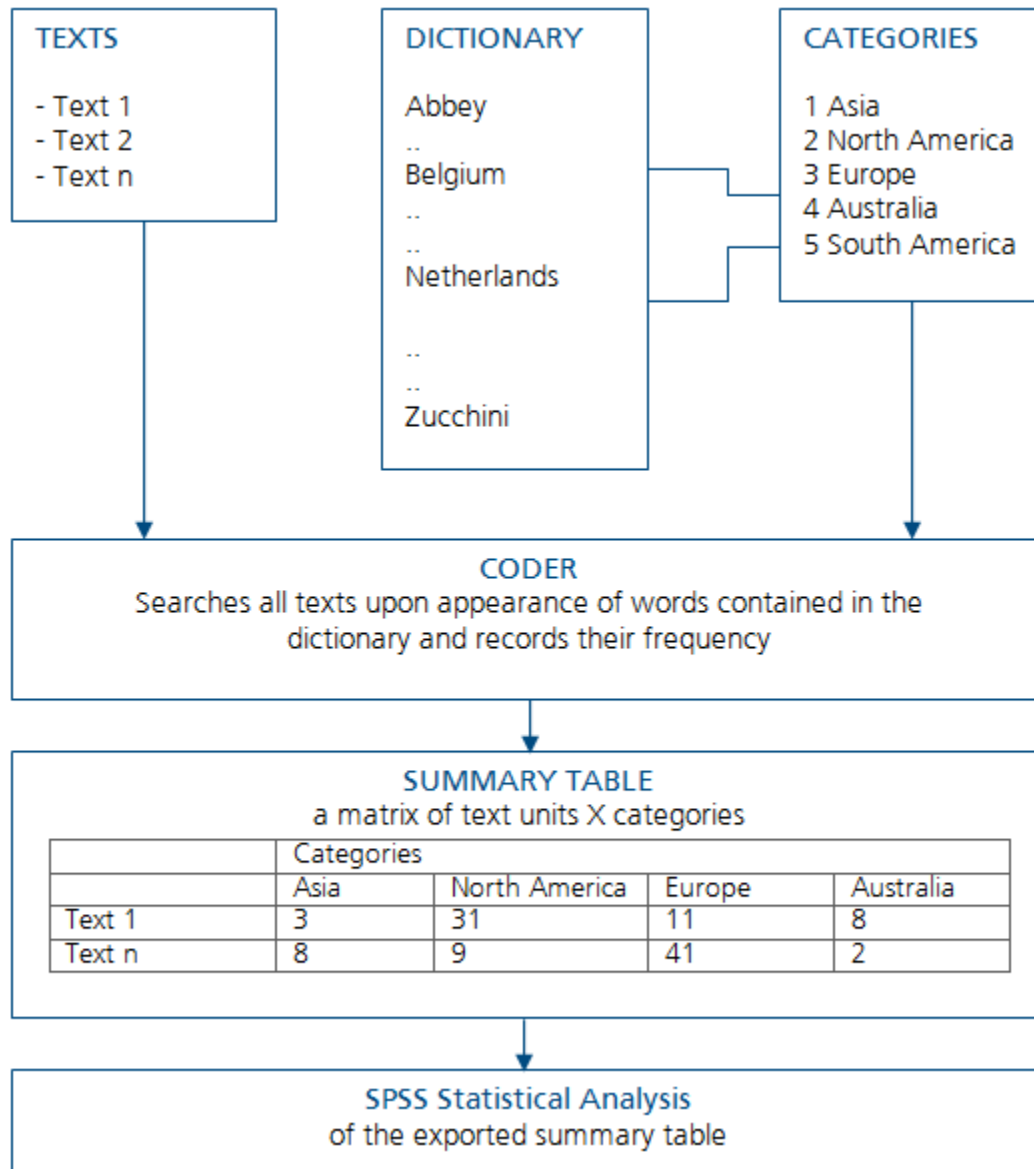
USB pen drive / MAXQDA 12 Files / Dictio / Dictionary

All TXT files in these folders will be displayed as dictionaries in MAXDictio.

12 Dictionary Based Content Analysis

12.1 Schematic Course

Based on well-known textbooks like KRIPPENDORF (2004), an outline of the following classical quantitative content analysis schema is possible:



Schematic view

The starting point of the analysis is a certain number of texts, often called the text corpus. If necessary, the texts can be subdivided into individual sections, such as a book is divided into chapters. These sections are

also called text units. They can also function as counting units, so that if you perform a dictionary based content analysis, each unit of analysis will be counted separately.

Independent from the texts, you work with a dictionary or encyclopedia. This dictionary consists, on the one hand, of an unlimited number of categories. On the other hand, it includes the words that are attached to these categories. In the example below, the words "Belgium" and "Netherlands" belong to the category "Europe."

The content analysis allows counting the category frequencies for texts or parts of text on the basis of the dictionary. It is as if a certain number of savings boxes are put up for each text – exactly as many as there are categories. The text editing starts with the first word of the text. Then, for each word that is also contained in the dictionary, a counting piece will be put into the savings box in the assigned category. After the whole text has been processed, the number of counting pieces in each savings box is determined. A statistical evaluation for this text will be calculated which contains, among others, the following information: total number of coded words, number of coded places for Category 2, and so on.

12.2 Determine Frequency of Dictionary Categories

With MAXDictio, a quantitative content analysis can be carried out to determine how frequently the search terms of the individual categories of a dictionary occur in a freely definable text corpus. The result of the material transport is a matrix containing the frequencies of the categories, e.g. the output is strictly quantitative.

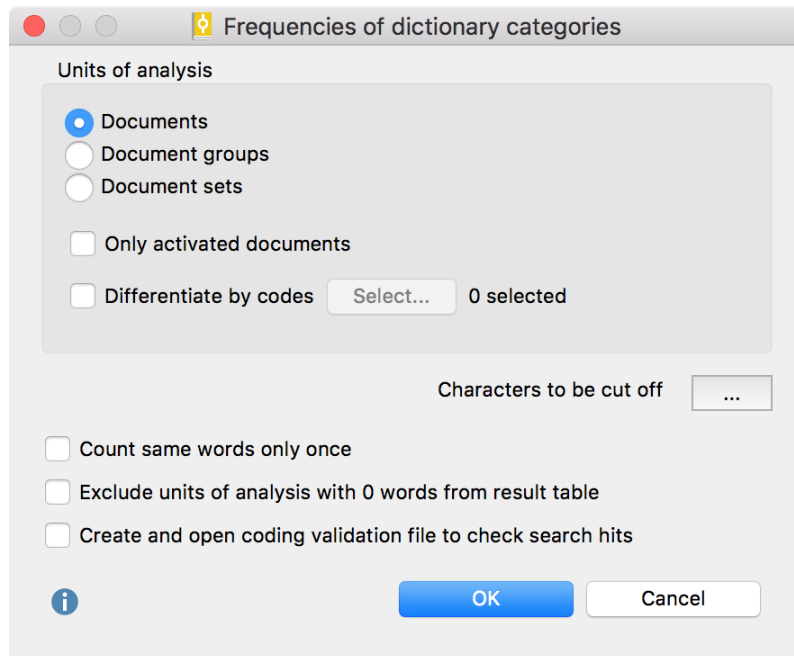
To run a quantitative content analysis, proceed as follows:

- Via the menu option **MAXDictio > Frequencies of dictionary categories**, or
- Using the quick button of the same name in the MAXDictio toolbar.



The quick button for the "Frequencies of dictionary categories" function in the "MAXDictio" toolbar.

The following dialog window will appear:



Options dialog for the “Frequencies of dictionary categories” function

Options

In the upper pane of the window, define the units of analysis, i.e. the breakdown of the results table:

Documents – The results table will contain a row for each document, so that the category frequency per document can be compared.

Document groups – The results table will contain a row for each document group, so the frequencies per group can be compared with one another.

Document sets – The results table will contain a row for each document set, so the frequencies per set can be compared with one another.

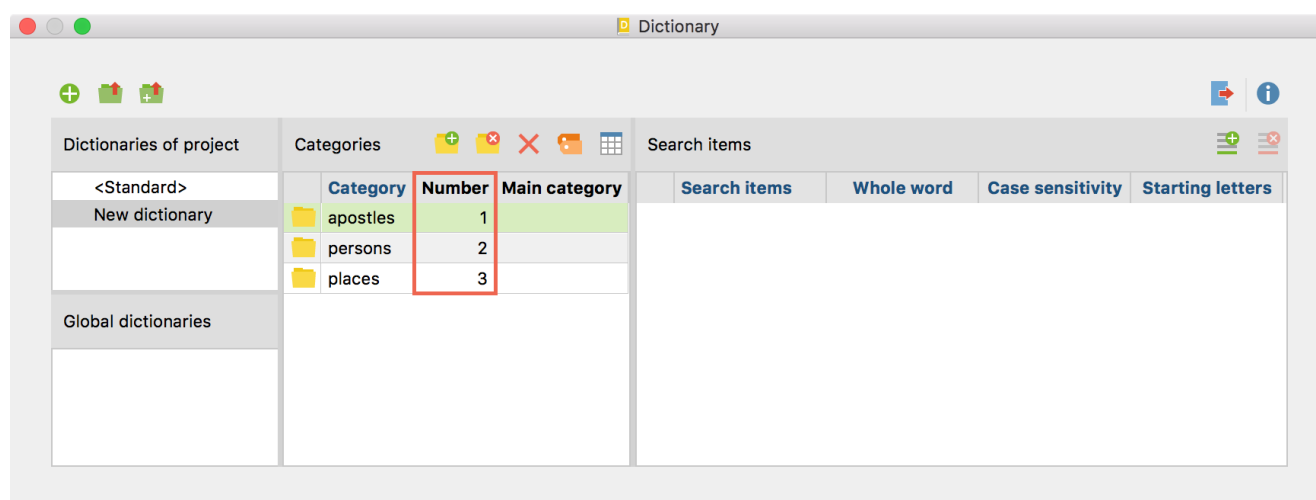
Only activated documents – Only activated documents are taken into account. If a document group or document set does not contain activated documents, the respective group or set will not be shown.

Differentiate by codes – The units of analysis (documents, document groups or document sets) selected in the dialog above will be subdivided into the codes assigned in the documents. The results table will contain one row for each code per unit of analysis, so if four texts have been analyzed and three codes have been selected, the results table will contain $4 \times 3 = 12$ lines. This option is recommended if text documents have been further subdivided by coded segments. You can use the **Select Codes** button to determine which codes are taken into account.

Count same words only once – Here you can determine how the words or search items should be counted that appear more than once in the same unit of analysis. If you only want the counter to count them once, select the appropriate option. If, for instance, a category contains the search item “Peter,” a unit will be counted once, no matter whether “Peter” is found one or more times in the given text unit. Suppose the category “apostles” contains the names of the 12 apostles as search items. The counting result 12 means that all 12 apostles have been found in this counting unit, given that the option has been selected. Otherwise, if this option hadn’t been selected, the counting result would not allow such a conclusion.

Exclude units of analysis with 0 words from result table – This option only concerns structured texts. Here you can set if empty units are displayed or not. Empty units may occur if an empty text has been analyzed or a code has not been applied to text. Please see [Analyze Structured Texts](#) for more information.

Create and open validation file to check search hits – Further, you have the possibility to create a validation file. Such a file allows you to carefully review the search hits. The validation file contains the original text of each unit of analysis. The locations where search items have been found are marked. Behind each search hit, you can find an identification number (ID-number) in brackets which points to the assigned category. The ID-number is determined when the category is created and is shown in the categories window. Categories can also be assorted by their numbers, and so you can create an assorted printout of the dictionary. This printout might be helpful when monitoring the validation file. At the head of a validation file is a legend which provides information about the assignment of the categories to the individual numbers.



Each category has its own ID-number

The following lines are a short extract from a validation file:

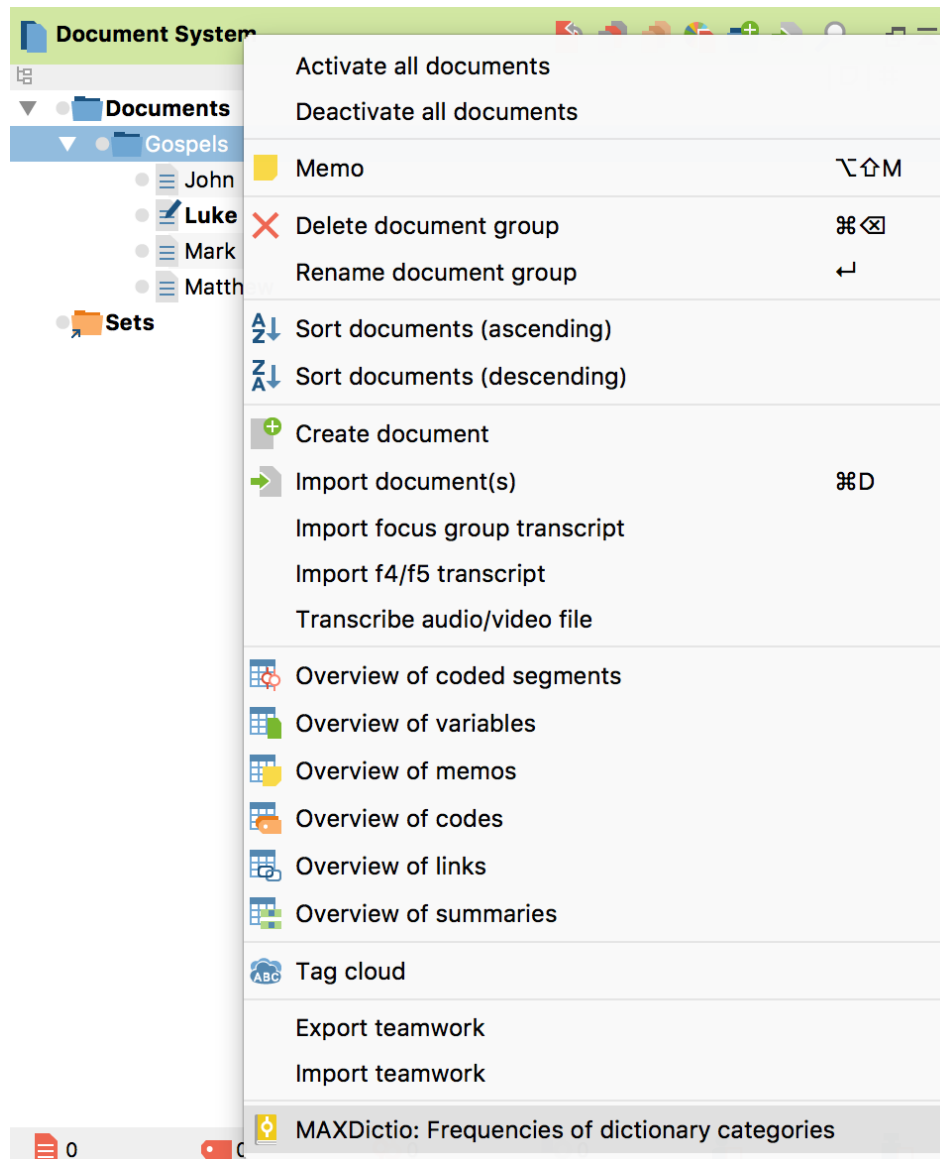
Now in the fifteenth year of the reign of Tiberius Caesar, Pontius Pilate[2] being governor of Judea[3], and Herod being tetrarch of Galilee[3], and Philip[1] his brother tetrarch of Iturea, and the country of Trachonitis, and Lysanias tetrarch of Abilina; Under the high priests Annas and Caiphas; the word of the Lord was made unto John[2], the son of Zachary, in the desert. And he came into all the country about the Jordan[3], preaching the baptism of penance for the remission of sins;

Begin Count from the "Document System"

The dictionary-based content analysis can be started directly from the "Document System" for the text documents of a selected level. Right-click one of the following levels and select **MAXDictio: Frequencies of dictionary categories**:

- Single document
- Document group
- Document set
- Root in the "Document System" for all documents

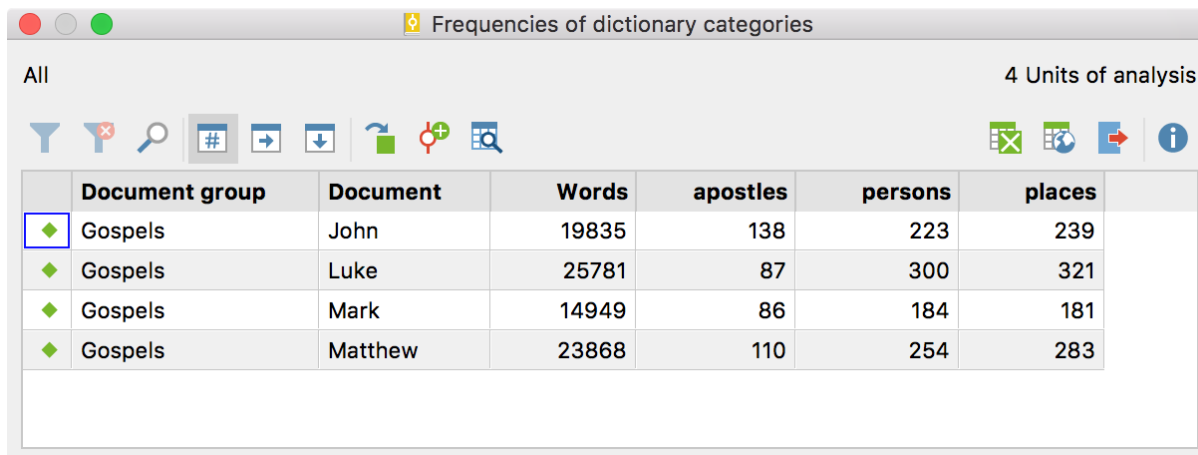
The same dialog window will appear, depending on the selected level, however, the selection of the units of analysis is limited. For example, a breakdown by document group cannot be obtained for a single document.



Opening the function “Frequencies of dictionary categories” via right-clicking on a document group

12.3 Result Table

The results table for the function “Frequencies of dictionary categories” is a data matrix, in which the units of analysis are represented in the lines and the categories in the columns:



	Document group	Document	Words	apostles	persons	places	
◆	Gospels	John	19835	138	223	239	
◆	Gospels	Luke	25781	87	300	321	
◆	Gospels	Mark	14949	86	184	181	
◆	Gospels	Matthew	23868	110	254	283	

Result table of the function "Frequencies of dictionary categories"



Such a matrix resembles the illustration above. Each line contains the results of one text and begins by naming the document group and document name. Then comes a column that contains the total number of counted words in the text. These numbers are followed by the frequencies of the categories – each in a single column.

There are numbers in the cells that specify how many times a given category has been found in the given unit of analysis.


This matrix is like other tables in MAXQDA – similar to a table in Excel. By clicking on a column title, the whole table will be assorted in ascending or descending order.



Note: Click on the Icon **Search results** to display a table of the origins of the search hits for selected or for all categories.

Display Row and Column Percentages

Use the symbols  **Row percentages** and  **Column percentage** in the upper toolbar to display the percentages for rows or columns. This makes it easy to compare the frequencies of categories for the different units of analysis.


Export Results

By clicking on the button **Export**  in the toolbar, the matrix can be saved as an Excel, HTML, RTF or plain text file.

The table of results can also be opened in HTML or Excel format by way of the **Open as HTML table**  button or **Open as Excel table** button  resp. This can easily be printed or imported into other programs.

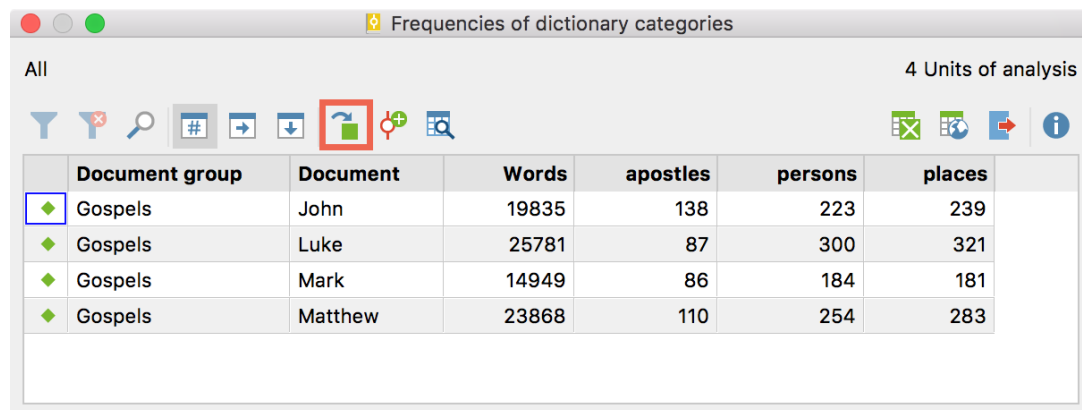
Convert Results to MAXQDA Document Variables

It is possible to save the columns with the category frequencies permanently in MAXQDA by converting them to MAXQDA variables.

Begin the transformation into variables using the **Transform into variable**  button in the toolbar of the results window. The button is visible only if the unit of analysis: documents, without differentiation by codes, has been selected.

The category frequencies are treated as individual variable types in MAXQDA. They appear in a different color (yellow) so that they are easier to recognize. The dictionary name serves as the variable name for the converted category frequencies. If, for example, a dictionary category is called “Legal Aspects,” this will serve as the variable name in MAXQDA’s variable list. The number of times the category appears in each of the texts on the list will be shown in the column of the variable matrix entitled “Legal Aspects.”

Note: The variable values of MAXDictio variables are not updated automatically if the origin data is changed; they represent the status of the time they were created. This type of variable cannot be updated by importing a variable matrix.



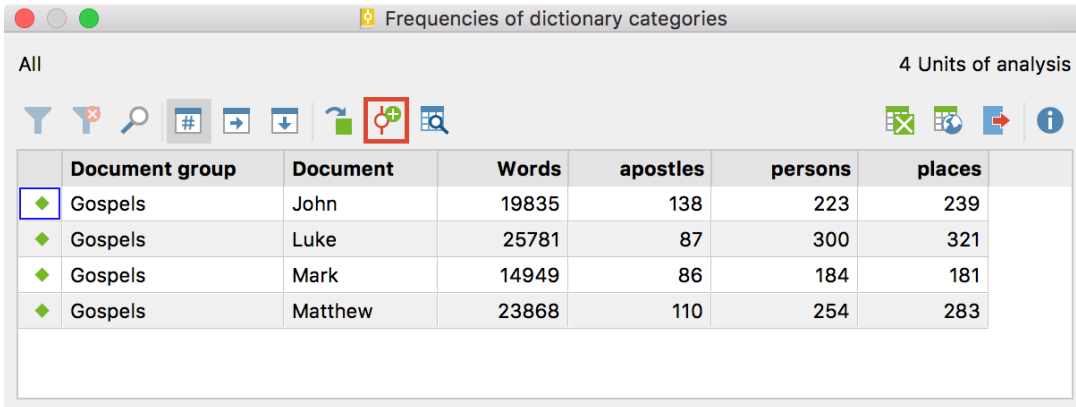
Document group	Document	Words	apostles	persons	places
Gospels	John	19835	138	223	239
Gospels	Luke	25781	87	300	321
Gospels	Mark	14949	86	184	181
Gospels	Matthew	23868	110	254	283

Symbol to transform the category names into document variables

Autocode Results with MAXQDA Codes

In MAXQDA you can use the standard and extended lexical search function to search for words and auto-code the results with a *single* code. MAXDictio extends this function and lets you autocode the search items of several dictionary categories at once – with the category name as a code label:


Click on the icon **Autocode search results** in the toolbar and choose, if you want to code the search item only or the surrounding sentence or paragraph, too. MAXQDA will create a new code “MAXDictio” in the “Code System” window, which has as many subcodes as there were categories defined in the dictionary. In the example the subcodes “persons”, “apostles”, and “places” will be created. In all analyzed documents the search items will be coded with the title of the category they belong to.

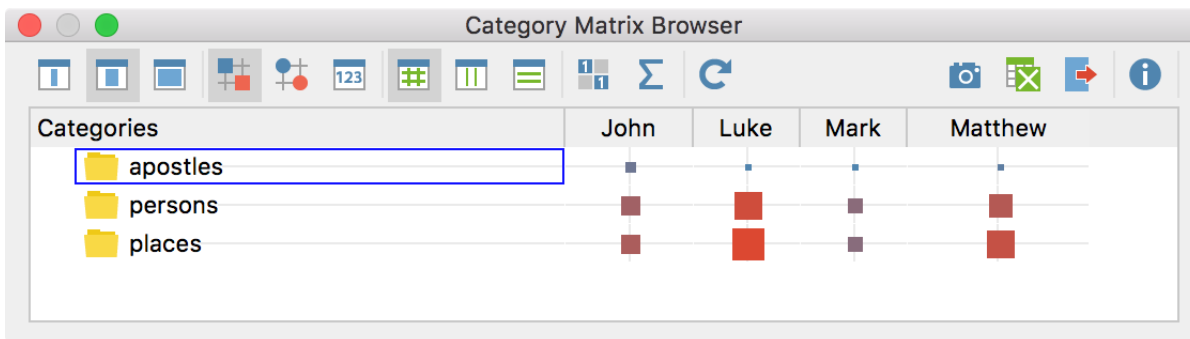


Document group	Document	Words	apostles	persons	places
Gospels	John	19835	138	223	239
Gospels	Luke	25781	87	300	321
Gospels	Mark	14949	86	184	181
Gospels	Matthew	23868	110	254	283

Symbol to autocode the search terms of individual categories with the category names

Visualize Results in the Category Matrix Browser

Using the Category Matrix Browser, the results of the count can be displayed graphically: each category of the underlying dictionary forms one row, and each unit of analysis forms a column. Click on the  icon in the toolbar of the results table to open it in the Category Matrix Browser.



Results displayed in the Category Matrix Browser

The Category Matrix Browser functions in the same way as the Code Matrix Browser in MAXQDA, except that instead of codes, the categories of the dictionary form the rows of the matrix. The larger a square at a node, the more search terms for these categories were found in the respective document. All settings and exports work exactly like the Code Matrix Browser.

Tip: Double click a square to list all related hits in a table.

12.4 Analyze Structured Texts

The functions for dictionary based quantitative content analyse in MAXDictio offers very complex possibilities of counting structured texts. Texts can be structured in many different ways. Suppose a text is divided into four sections (text units). This corresponds to a book with four chapters. This results in the following structure:

Chapter 1

Chapter 2

Chapter 3

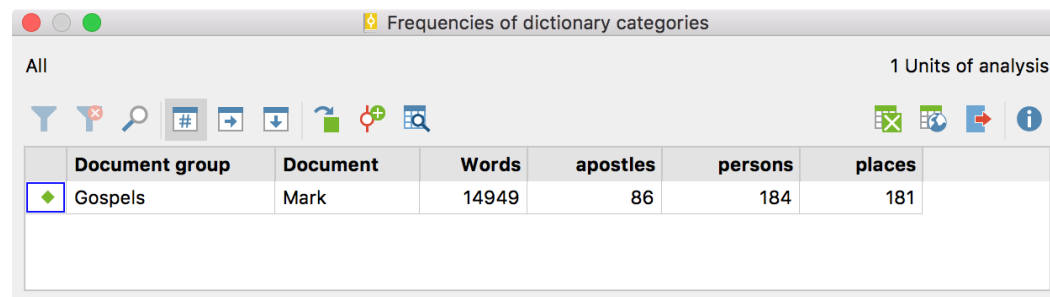
Chapter 4

The chapter names are managed by MAXQDA as codes. The subdivision of texts into sections can be achieved simply by coding the relevant paragraphs in the “Document Browser”.

Proceed as follows:

1. Create the codes “Chapter 1,” “Chapter 2,” “Chapter 3,” and “Chapter 4” in the MAXQDA “Code System” window.
2. Open the text in the “Document Browser”, select the first section with the mouse, and assign the code “Chapter 1” to it. Do the same with the sections two, three, and four. In the following example, this has been done with the first four chapters of the Gospel of Mark (the other chapters have been deleted for this example).

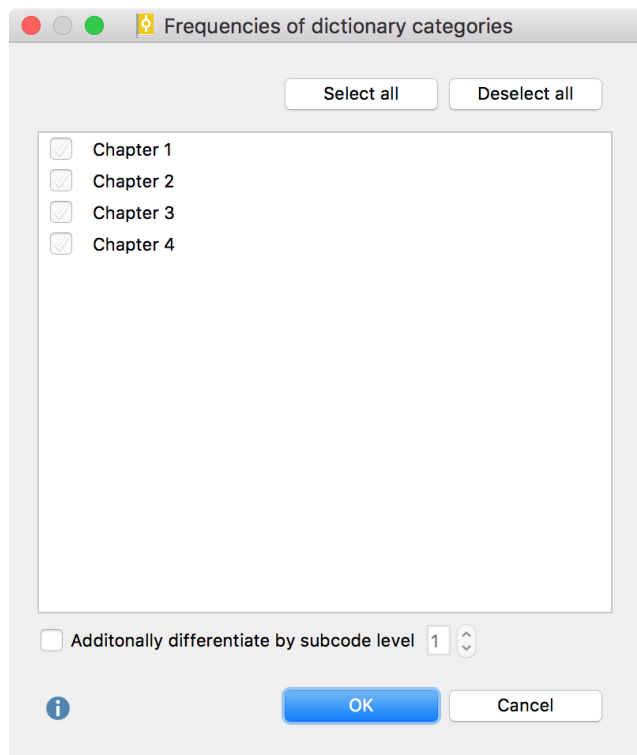
After starting the function **MAXDictio > Frequencies of dictionary categories** for a single text document without a differentiation by codes the table of results would only contain one line for the text:



Document group	Document	Words	apostles	persons	places
Gospels	Mark	14949	86	184	181

Only one line for the four sections of the text is shown if no differentiation by codes is selected

To differentiate by the chapters of the text select the option **Differentiate by codes** in the options dialog and click on the button **Select ...** to choose some codes. The following dialog window will appear:



Select codes to subdivide a text into more units of analysis

The window contains all top level codes from the MAXQDA "Code System" window. You can select one or more top level codes by ticking their selection box. Use the option **Differentiate by subcode level** to subdivide the texts by the subcodes of the selected top level codes. Level 1 selects the direct subcodes, level 2 additionally subdivides the texts by the subcodes of the subcodes.

The following table of results will be created:

	Document group	Document	Code	Words	apostles	persons	places
◆	Gospels	Mark	Chapter 1	918	14	11	11
◆	Gospels	Mark	Chapter 2	703	2	9	9
◆	Gospels	Mark	Chapter 3	659	14	6	8
◆	Gospels	Mark	Chapter 4	887	0	7	10

Result table for the text „Mark“, divided into the chapters 1 to 4

When adding the relevant category frequencies, their sum will be the same as the numbers contained in the overall table. However, this only applies when the option **Count same words only once** has not been selected. Otherwise, subdividing a text into sections can produce results that would differ from those of the overall analysis.

If you had only selected “Chapter 1” and “Chapter 2” in the options window, the result table for this text would only consist of two lines, one for each of the counted text units (Chapter 1 and 2 of the Gospel of Mark).

The following norms for structured texts should be taken into consideration:

- The structuring features must be level 0 codes on the highest level of the “Code System”.
- Overlapping of structuring features is allowed.
- The number of structuring features is unlimited.

Structuring texts can be simplified considerably by using the MAXQDA text pre-processor (**Documents > Import structured document (Preprocessor)**). With its help, structuring features can already be created while a text is being imported, meaning that they can be written directly into the imported text.

Differentiate by Subcodes

Many diverse possibilities result from subdividing texts on a more differentiated level. Analogue to a book, you can divide a text into:

Chapter 1 – Paragraph 1

Chapter 1 – Paragraph 2

Chapter 1 – Paragraph 3

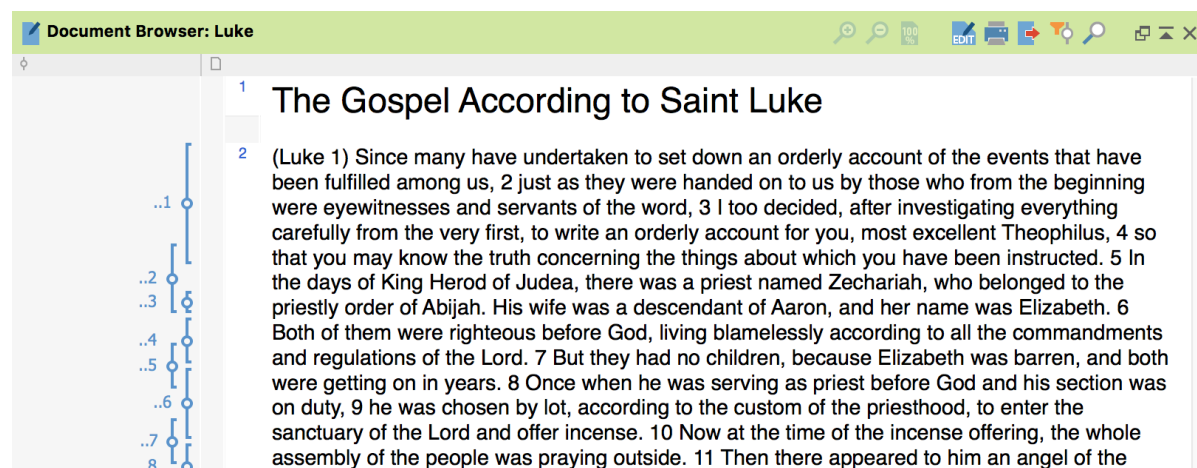
Chapter 2 – Paragraph 1

Chapter 2 – Paragraph 2

Now you can carry out the analysis on the lowest level of the text – the paragraphs. To do this, select the option **Additionally differentiate by subcode level 1**.

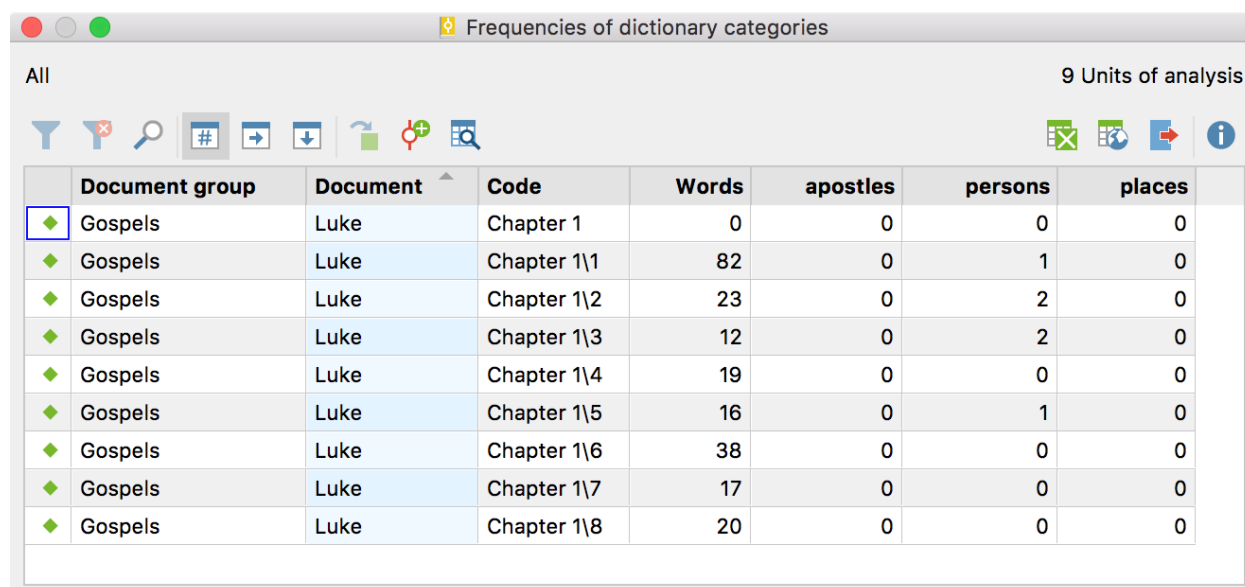
Example

The text below has been processed so that each sentence in the chapters has been coded individually, beginning with the codes “Chapter 1/1,” “Chapter 1/2,” and so on.



Visualization of the coded sentences in the “Document Browser”

The results for the first 8 sections that have been analyzed with subcode level 1, appears as follows:



The screenshot shows a software window titled "Frequencies of dictionary categories". It displays a table with 9 units of analysis. The table has columns for Document group, Document, Code, Words, apostles, persons, and places. The first row is selected, showing "Gospels" for the document group, "Luke" for the document, and "Chapter 1" for the code, with 0 words, 0 apostles, 0 persons, and 0 places.

	Document group	Document	Code	Words	apostles	persons	places
◆	Gospels	Luke	Chapter 1	0	0	0	0
◆	Gospels	Luke	Chapter 1\1	82	0	1	0
◆	Gospels	Luke	Chapter 1\2	23	0	2	0
◆	Gospels	Luke	Chapter 1\3	12	0	2	0
◆	Gospels	Luke	Chapter 1\4	19	0	0	0
◆	Gospels	Luke	Chapter 1\5	16	0	1	0
◆	Gospels	Luke	Chapter 1\6	38	0	0	0
◆	Gospels	Luke	Chapter 1\7	17	0	0	0
◆	Gospels	Luke	Chapter 1\8	20	0	0	0

Result table for differentiation by subcode level 1

You can see that each sentence has been counted individually and the results are represented by an individual line in the table. In the first sentence of Chapter 1, for instance, 82 words have been counted.


Note: Click on the Icon **Search results** to display a table of the origins of the search hits for selected or for all categories.

12.5 Further Processing of the Results with Excel or SPSS

The results matrix is an initial point for further statistical analysis. This can help answer further questions, such as:

- Which category is named most frequently in which texts or text units?
- In which text (or texts) is a category coded especially often? Where does it appear very seldom?
- How do the categories correlate?
- Can patterns be found in the category frequencies?

How to Transport the Matrix into Excel

To export the results to Excel, you can click on the icon **Open as Excel table** . Please note that only the selected rows will be displayed in Excel. If no row is selected, all rows will be exported.

The second possibility would be to create a text file of the matrix by clicking on the option **Export** and to choose the Excel format. Then you can decide, where to save the file and after export it will be opened automatically in Excel.

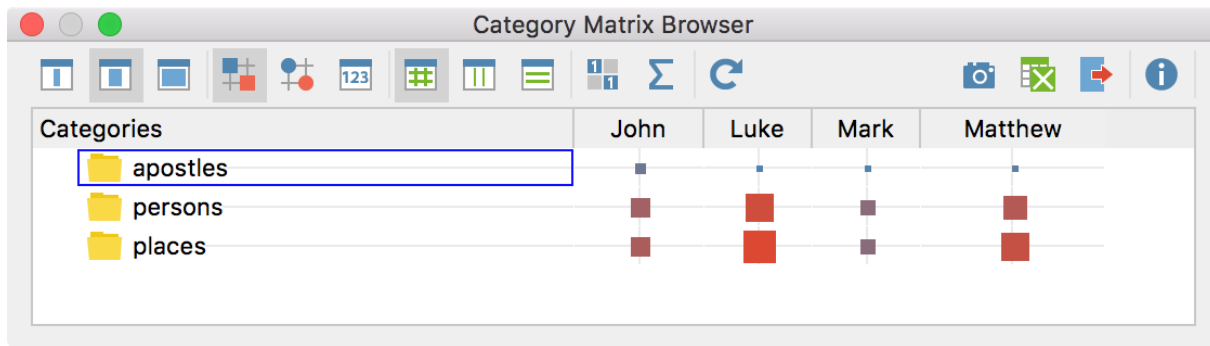


How to Transport the Matrix into SPSS

To export the matrix to SPSS, export the file in Excel format and import it in SPSS.

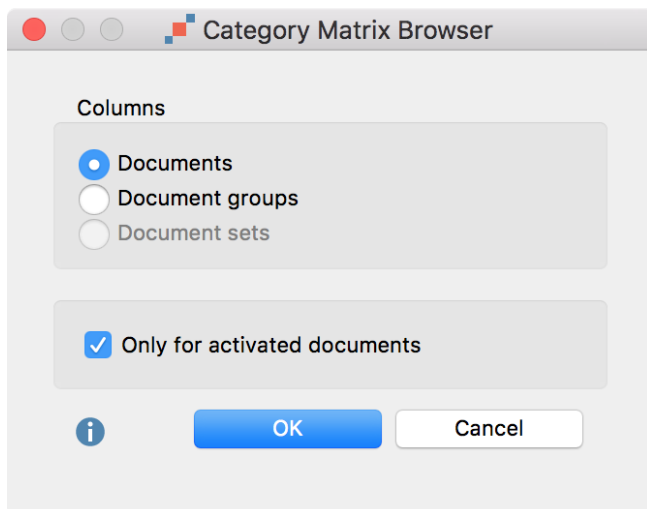
13 The Category Matrix Browser

The Category Matrix Browser graphically displays the frequencies of the dictionary categories per unit of analysis. The categories of the dictionary form the rows while the evaluated texts form the columns:



The Category Matrix Browser

Call up the function via the menu **MAXDictio > Category Matrix Browser**. The following dialog window will appear, in which you can set whether the documents, document groups or document sets will be displayed in the columns:



Setting options for the Category Matrix Browser


If you select the option **Only for activated documents**, only the documents activated in the “Document System” will be included in the display. Document groups and document sets without activated documents will not be displayed.

The nodes of the Category Matrix Browser contain the results of the count of the category frequencies. The larger a square, the more occurrences of the search terms of the respective dictionary category were found in the analyzed document, document group or document set.

Tip: The display is interactive. Double-click a node to display the search hits in the original text.

Customize View

Customize the view using the symbols on the left:

 **Names, columns: none, short, full** – Sets the column widths

 **Calculation of symbol size refers to all cells**


 **Calculation of symbol size refers to columns**


 **Calculation of symbol size refers to rows**

 **Display nodes as squares**

 **Display nodes as circles**


 **Display nodes as values**

 **Binarize view** – Regardless of how many hits per cell are present, it is always indicated whether at least one hit is present or not. The symbol size is standardized.


 **Sum** – Hides or shows a sum column and a sum row

Export Results

Using the icons on the right, the view can be exported as an image or a table:

 **Copy current display as an image to clipboard** –The currently displayed matrix will be copied to the clipboard and can be inserted with Ctrl + V (Win) or Cmd + V (Mac) into Word, PowerPoint and other programs.

 **Open as Excel Table** – Displays the matrix with frequencies in Excel.

 **Export** – Exports the displayed matrix as values in Excel, HTML or TXT format. It is also possible to export the matrix as a graphic file in PNG and SVG format.

Hint: When exporting as a table, rows and columns are reversed in order to allow easy evaluation using statistics programs. The rows will therefore contain the units of analysis (cases) and the columns will contain the categories (comparable to variables).

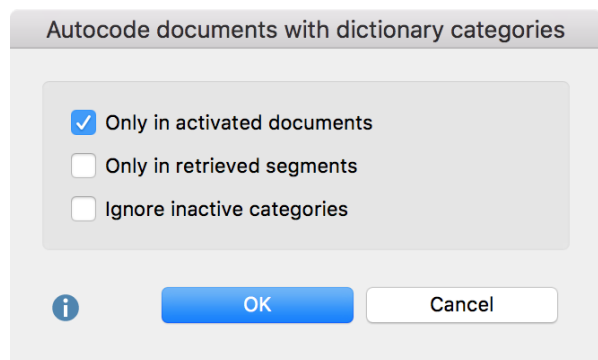


14 Autocode Documents with Dictionary Categories

The lexical search in MAXQDA allows you to search for words within text, PDF and table documents, and to encode the search results automatically with a code. MAXDictio extends this possibility enormously by making it possible to automatically encode the search terms of a dictionary category with the name of the category.

To automatically encode texts using the search terms of a dictionary, proceed as follows.


1. Open **MAXDictio > Dictionary**.
2. Create a new dictionary. Create the categories whose names are to appear later as codes in the “Code System” and add the desired search terms to each category. Alternatively, you can choose an existing dictionary.
3. Start the function **MAXDictio > Autocode documents with dictionary categories**.
4. In the dialog window that appears, select the appropriate settings.



Autocoding options

Only in activated documents – Autocoding is only performed in activated documents.

Only in ‘Retrieved Segments’ – Autocoding is restricted to text segments displayed in the “Code System”.

Ignore inactive categories – If a category in the dictionary has been switched off by with a double-click and displays a stop symbol , its search terms will be ignored and no code with the category name will be created in the “Code System”.

When you click **OK**, the codes for the dictionary categories will be created in the “Code System” and the following dialog box will appear, where you can define the usual options for the weight of the coded segments and the scope of the texts to be coded.

Code

PDF documents

☒ Only search string ☐ Context

☐ Sentence Words before

0 Words after

Text and table documents

☒ Only search string ☐ Context

☐ Sentence Words before


☐ Paragraph Words after



Cancel

Options for autocoding: set weight and text range

Click Autocode. A display will appear which indicates the progress of the operation until all hits are coded.

Tip: If you have used the function „Frequencies of dictionary categories“, you can also start autocoding directly from the results window using the icon  **Autocode documents with dictionary categories.**

⚙ Frequencies of dictionary categories

All

4 Units of analysis



	Document group	Document	Words	apostles	persons	places	
◆	Gospels	John	19835	138	223	239	
◆	Gospels	Luke	25781	87	300	321	
◆	Gospels	Mark	14949	86	184	181	
◆	Gospels	Matthew	23868	110	254	283	

Symbol to autocode the search words of each category with the category names

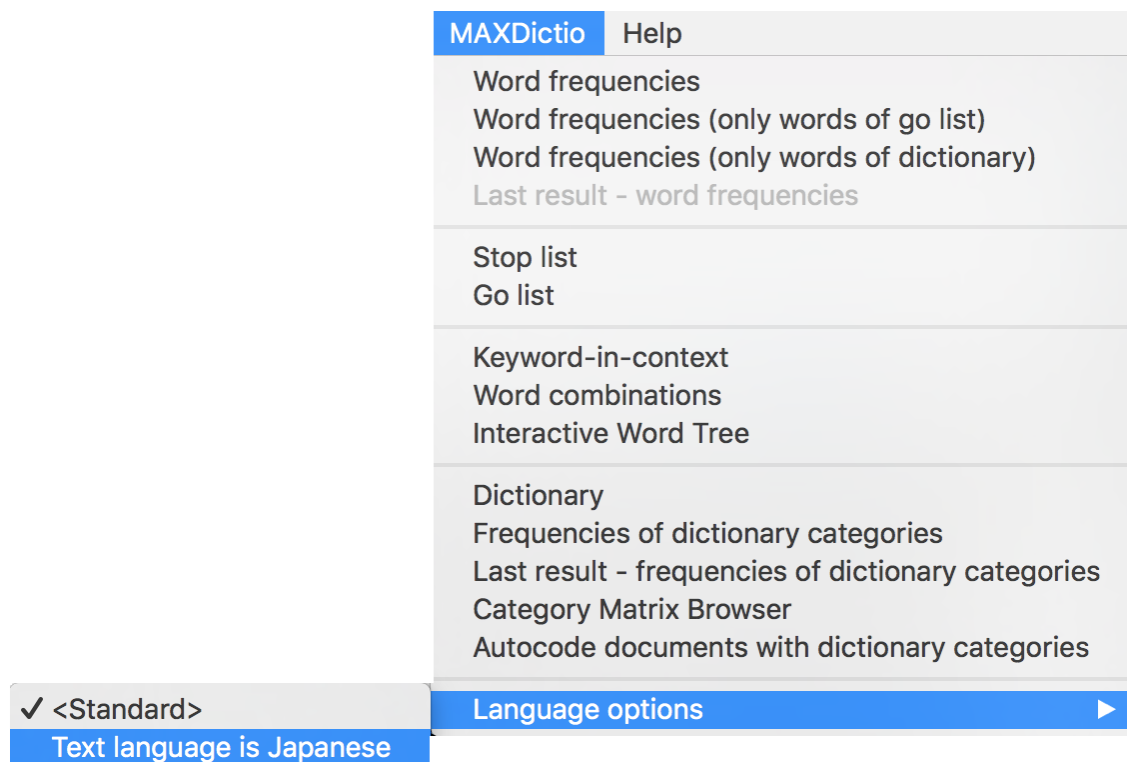
15 Language Options

The functions of MAXDictio are based on algorithms which are used to separate words from spaces, as is standard for Latin languages. Each new word begins after a space or at the beginning of a paragraph. In syllable-based languages, e.g. Japanese, spaces do not occur between words, therefore this differentiation is not possible.

Analyzing Japanese Texts

Beginning with MAXQDA version 12.3, it is nonetheless possible to analyze texts in the Japanese language. For this purpose, a separation of the Japanese words is made using “MeCab: Yet Another Part-of-Speech and Morphological AnalyzerMeCab” (<http://taku910.github.io/mecab/>).

To analyze a Japanese text with MAXDictio, go to **MAXDictio > Language options** and select **Text language is Japanese**.



Select option for Japanese text in the MAXDictio menu

It is recommended to use this option only if Japanese texts are currently being analyzed, because you might get wrong results for texts in other languages.

Tip: You should add the following characters to be cut off for functions such as word frequency, so these will not be counted as words. Please note that these characters are technically different from their latin equivalents although they look very similar:



@ ! \$ % & / () = ? ^ ' " " 「 」 『 』 [] [] < > < > 《 》 << >> { } 【
】 ¥ # + - - * _ : ; ~ 。 . 、 ,

All functions of MAXDictio can be used for Japanese texts, with the exception of the **Whole word** option in the dictionary and in the starting dialog for “Keyword-in-context function”.

16 Limits of MAXDictio

Neither the number or size of documents, nor the number of text units or words per text unit are technically so restricted as to be significant for a typical project in qualitative and mixed-method research.

Only the length of words is limited. Here, a maximum of 64 characters are permitted; longer words in a text are ignored. The representation in the results tables, e.g. for word combinations, is limited to 64 characters, and longer word combinations may be truncated, indicated by three points (...) at the end of the row.

To a high degree, working with MAXDictio depends on how fast the computer is and how many resources it has. If the number of texts to be counted at the same time reaches a certain quantum, the computer becomes noticeably slower. Projects of up to a million words are manageable for today's average computers. This should be sufficient for usual projects of qualitative social science.

The more text to be processed simultaneously, the more time the computer needs. All functions also have a bar that indicates the progress. Hence, it is visible how much text MAXDictio has already processed and, thus, possible to calculate how much time will still be needed to complete the entire analysis.