Quantitative Text Analysis with MAXDictio

MAXDictio is a quantitative text analysis module that is part of MAXQDA Plus and MAXQDA Analytics Pro. It is not a part of MAXQDA Standard. MAXDictio offers tools for word-based analysis of text and PDF documents:

- Analysis of word frequencies and word combinations
- Analysis of Keyword-in-context (results tables and Interactive Word Tree)
- Quantitative content analysis with a dictionary

1. Exploration: Word Frequencies

Select MAXDictio > Word Frequencies to count and list words in all (or only in activated) text and PDF documents. Alternatively, you can restrict the search to the text currently displayed in the "Retrieved Segments" window.

The result table lists all words in the analyzed text by their absolute and relative frequencies. The columns "Documents" and "Documents %" indicate the number and percentage of documents in which a word occurs.

Word frequencies

In 2 documents (2834 words total) 672 Words (TTR = 0,2371)

Display top ranks   Frequency

Word	Word length	Frequency	%	Rank	Documents	Documents %
your	4	12	0,42	47	1	50,00
hand	4	11	0,39	53	2	100,00
horse	5	11	0,39	53	1	50,00
jorinda	7	11	0,39	53	1	50,00
jorindel	8	11	0,39	53	1	50,00
much	4	11	0,39	53	2	100,00
now	3	11	0,39	53	2	100,00
would	5	11	0,39	53	2	100,00
castle	6	10	0,35	60	1	50,00
goose	5	10	0,35	60	1	50,00
old	3	10	0,35	60	2	100,00
up	2	10	0,35	60	2	100,00
again	5	9	0,32	64	2	100,00
away	4	9	0,32	64	2	100,00

Word frequencies

Differentiation

☑ Only for activated documents ○ None
☐ Only in "Retrieved Segments" window ● By documents
○ By document groups top-level
○ By document sets
○ By codes

Characters to be cut off ...
Min.number of characters 2

☑ Apply stop list (Fairy Tales) ...
☐ Case sensitivity
☑ Lemmatize words English

OK   Cancel

Stop and Go Lists

Stop lists are a place to collect uninteresting words, for example to or and, that you wish to exclude from your analysis. You can send words to the stop list with a click on the green icon in result tables or open MAXDictio > Stop List to customize, export, import and merge stop lists.

Differentiate by Document, Document Group or Document Set

Result tables can hold additional columns to display word frequencies for documents or document groups. Columns can display how often a word occur, the rank of a word in this document (group) or if the word occurs at all (1, 0).
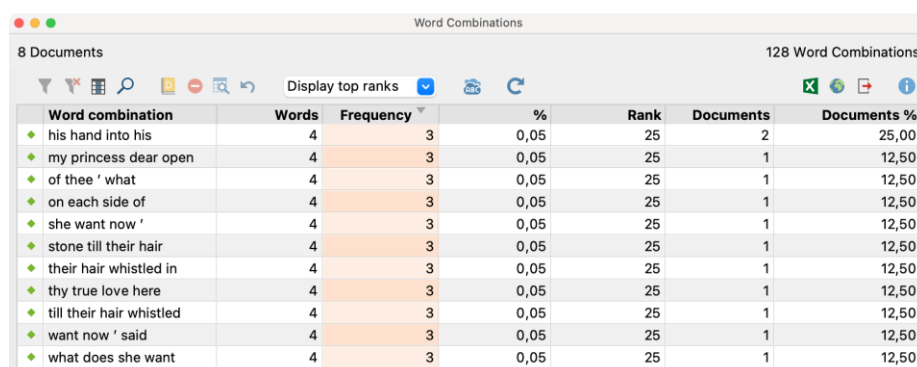
Lemmatization

Activate the lemmatization option to count words from the same stem together. For example, the frequency for "give" will include occurrences of give, gave, or given. Lemmatization is currently available for 14 languages (Bulgarian, Catalan, Czech, German, English, Spanish, Estonian, French, Hungarian, Italian, Polish, Portuguese, Swedish, Ukrainian).

2. Exploration: Word Combinations

Select MAXDictio > Word Combinations works similar to the Word Frequencies feature, but counts and lists word combinations rather than single word occurrences.

You can search for combinations of up to 5 words. You can decide if word combinations should occur within the same sentence or if word combinations should also be counted across paraphrase and full stops.

Tip: If you hover with a mouse over a row in the results table, all the lemmatized words or word combinations in this row will be displayed.

3. Exploration: Keyword-in-context

Select MAXDictio > Keyword-in-context to view selected keywords in a tabular view along with the words that appear before and after the word. The table also displays document information (name, paragraph).


Tip: This table is interactively linked to the original source data. Click into a table row to highlight the search item in the original text in the Document Browser.

3

4. Exploration: Interactive Word Tree

Select MAXDictio > Interactive Word Tree to visually explore word and word combinations in your data). Drag all the documents, document groups or sets into the dialog window that you want to use to create a Word Tree.

The image below shows the Interactive Word Tree for the "I have a dream" speech (Martin Luther King, Jr., 1963)
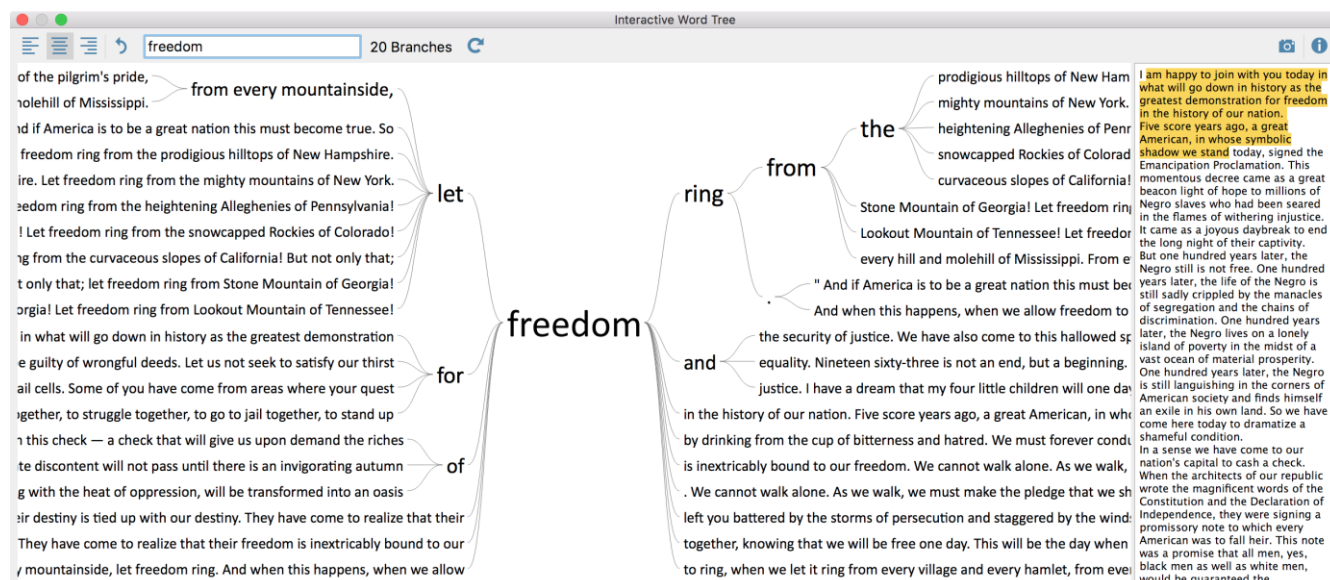

The word "freedom" is the root of the tree with branches to the left and right showing the words before and after that word. The root can be placed in the middle (as shown above) or as the starting or ending point. The window to the right contains the complete analyzed text. Yellow highlighted passages are currently displayed in the tree view.

5. Dictionary Based Context Analysis

To conduct a dictionary based quantitative context analysis, your first step is to create a dictionary via MAXDictio > Dictionary. A dictionary may contain any number

of categories with an unlimited number of search items. For example, a simple dictionary may look like this:
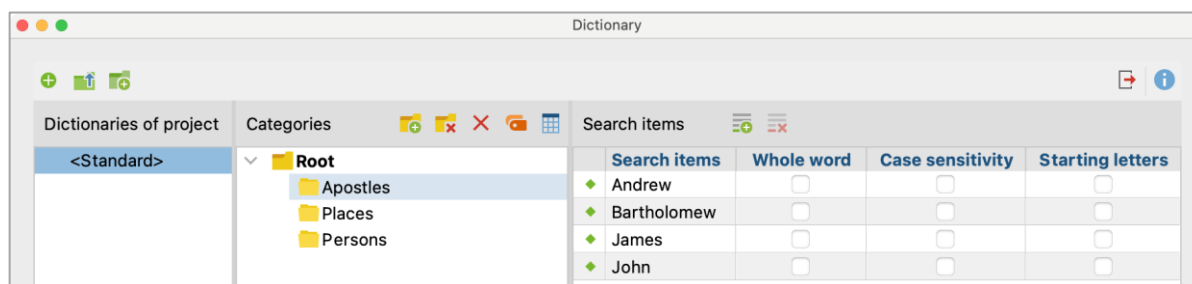
This dictionary contains three categories Apostle, Places and Persons, each of which holds a number of search items. A search item can also consist of multiple word.

Like stop-lists, dictionaries can be imported, exported, and merged. By default, dictionaries and stop lists are saved within a MAXQDA project, to ensure that if you share your project data with your team, everyone can use them.

Tip: You can transfer words from a word frequency table into a dictionary via drag & drop.

6. Frequencies of Dictionary Categories

Select MAXDictio > Quantitative Content Analysis to create a frequency table similar to the ones we discussed earlier. This time MAXQDA only searches for search items in active dictionaries.


Tip: You can also right click on a document or document group to start the quantitative content analysis for the selected document(s).

7. Visualize Frequencies of Dictionary Categories

Select MAXDictio > Category Matrix Browser to visualize the same type of frequencies in a visual matrix. The image blow visualizes how often search items

from the three categories occur in four different documents. You can also start this feature by clicking the matrix browser icon in the result table of a quantitative content analysis.
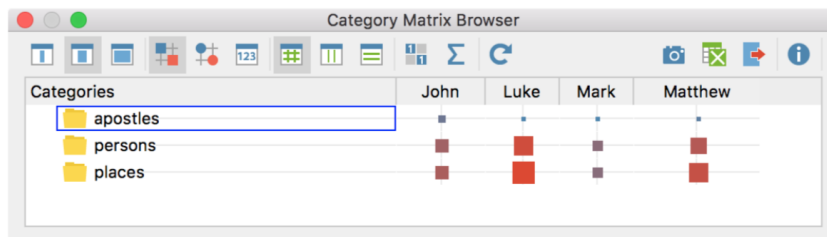
8. Autocode Documents with Dictionary Categories

Select MAXDictio > Autocode with dictionary to automatically search for occurrences of category search items. Each search result can be coded with the name of the category, that holds this search item. All the usual autocode options of MAXQDA are available, e.g. to code just the search item, or the (adjourning) sentence(s) or paragraph(s) it occurs in.