# Quantitative Text Analysis with MAXDictio
(Spotlight Session)

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
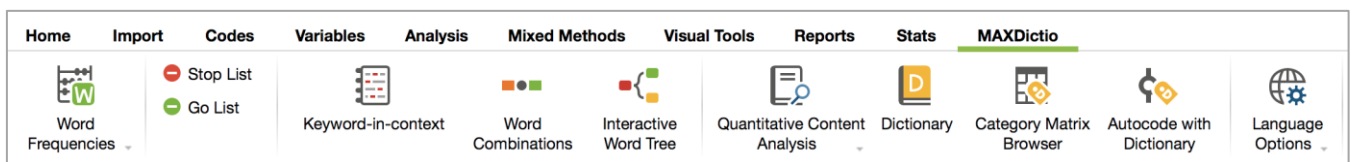
1. Exploration: Word Frequencies

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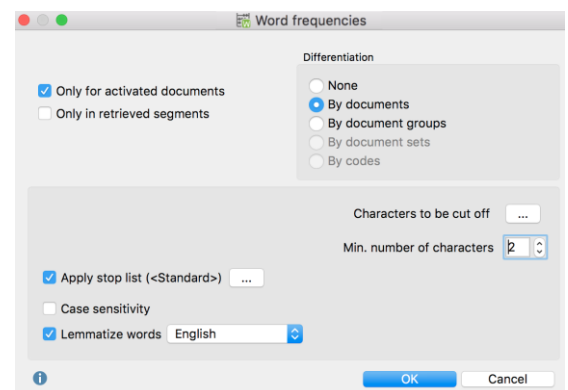
The result table lists all words in the analyzed text by their absolute and relative frequencies. The columns "Documents" and "Documents %" indicate the number and percentage of documents in which a word occurs.

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).

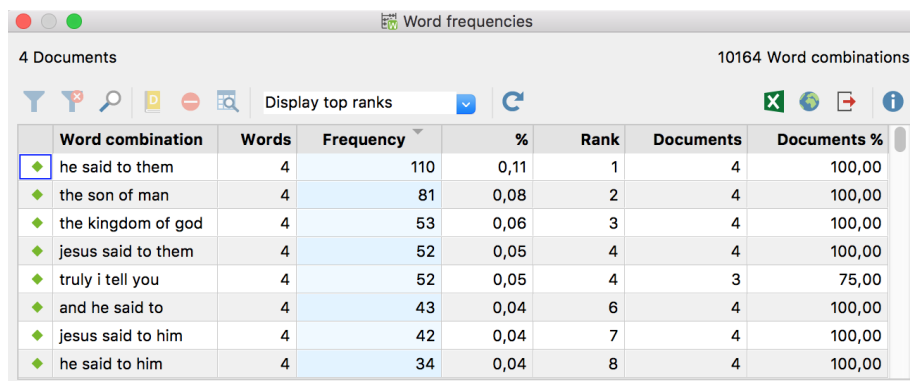### Lemmatization

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.
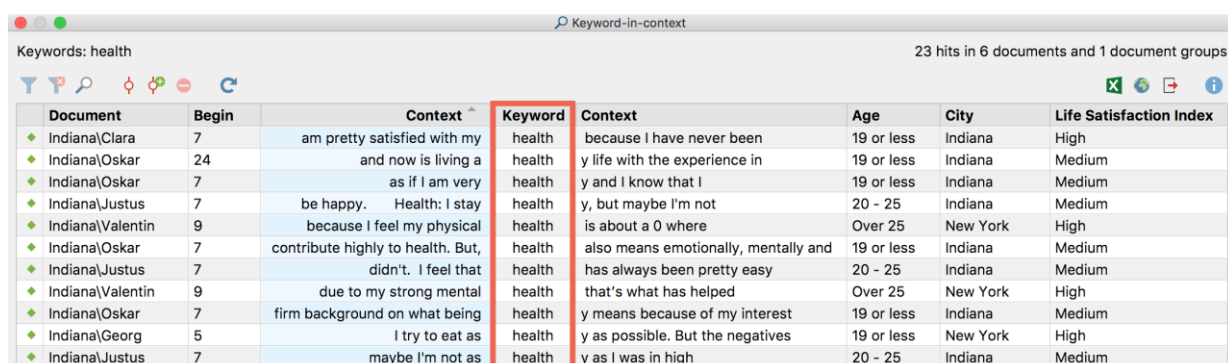

	Word combination	Words	Frequency	%	Rank	Documents	Documents %
◆	he said to them	4	110	0,11	1	4	100,00
◆	the son of man	4	81	0,08	2	4	100,00
◆	the kingdom of god	4	53	0,06	3	4	100,00
◆	jesus said to them	4	52	0,05	4	4	100,00
◆	truly i tell you	4	52	0,05	4	3	75,00
◆	and he said to	4	43	0,04	6	4	100,00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

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).

The option **Include Tooltip Variables** enables you to display document variables that are marked as tooltip variables in their own columns. This way, you can review the standardized information attached to the corresponding documents in addition to the verbal context of keywords. In the above example below, three columns on the right represent the variables *Age*, *City* and the *Life Satisfaction Index*.

Keywords: health   23 hits in 6 documents and 1 document groups

|	Document	Begin	Context	Keyword	Context	Age	City	Life Satisfaction Index
◆	Indiana\Clara	7	am pretty satisfied with my	health	because I have never been	19 or less	Indiana	High
◆	Indiana\Oskar	24	and now is living a	health	y life with the experience in	19 or less	Indiana	Medium
◆	Indiana\Oskar	7	as if I am very	health	y and I know that I	19 or less	Indiana	Medium
◆	Indiana\Justus	7	be happy. Health: I stay	health	y, but maybe I'm not	20 - 25	Indiana	Medium
◆	Indiana\Valentin	9	because I feel my physical	health	is about a 0 where	Over 25	New York	High
◆	Indiana\Oskar	7	contribute highly to health. But,	health	also means emotionally, mentally and	19 or less	Indiana	Medium
◆	Indiana\Justus	7	didn't. I feel that	health	has always been pretty easy	20 - 25	Indiana	Medium
◆	Indiana\Valentin	9	due to my strong mental	health	that's what has helped	Over 25	New York	High
◆	Indiana\Oskar	7	firm background on what being	health	y means because of my interest	19 or less	Indiana	Medium
◆	Indiana\Georg	5	I try to eat as	health	y as possible. But the negatives	19 or less	New York	High
◆	Indiana\Justus	7	maybe I'm not as	health	y as I was in high	20 - 25	Indiana	Medium

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.

Tip: Each column of a result table can be alphabetically sorted with a click on the column header. You can rearrange columns via drag & drop, and right click on a column header to filter for characters.

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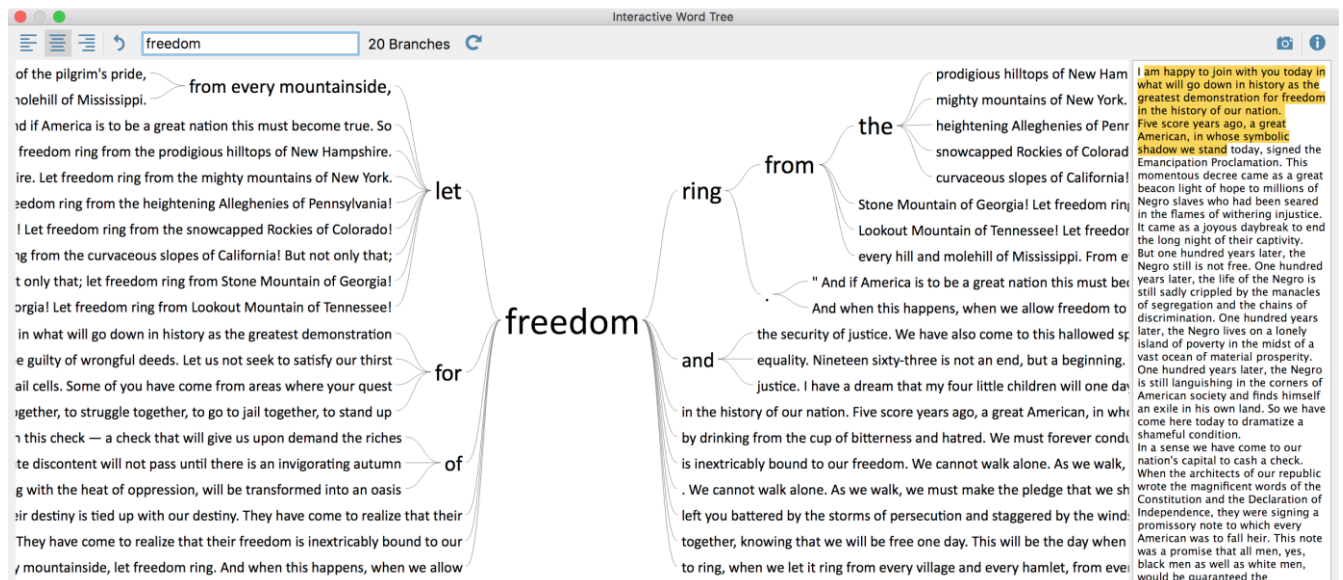
The word "freedom" is the root of the tree with branches to the left and right showing the words before and after that word. The root can be placed in the middle (as shown above) or as the starting or ending point. The window to the right contains the complete analyzed text. Yellow highlighted passages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

Tip: Click on another word to use it as the new root or search for word (combinations) in the search field.

5. Dictionary Based Context Analysis

Dictionary
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:
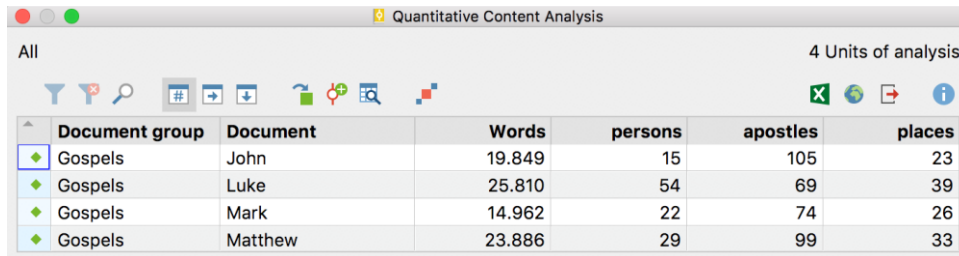

This dictionary contains three categories *Apostle*, *Places* and *Persons*, each of which holds a number of search items. A search item can also consist of multiple word.

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.

Document group	Document	Words	persons	apostles	places
Gospels	John	19.849	15	105	23
Gospels	Luke	25.810	54	69	39
Gospels	Mark	14.962	22	74	26
Gospels	Matthew	23.886	29	99	33

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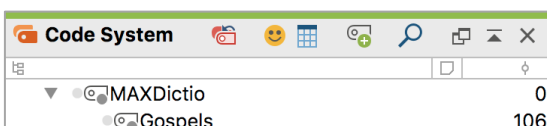
7. Visualize Frequencies of Dictionary Categories

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.
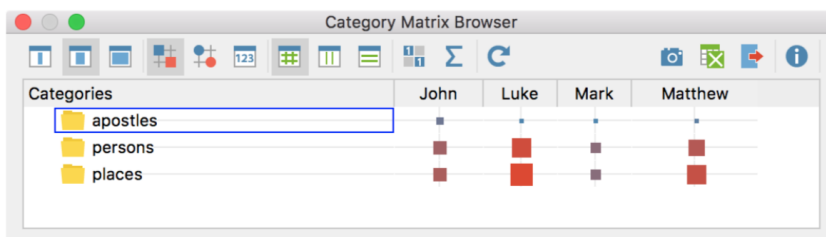
8. Autocode Documents with Dictionary Categories

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.