

Read the Damn Manual!

VOGON MANUAL



Julia Damerow
v1.0, November 2011

Table of Contents

Introduction	2
The Short Story	2
The Long Story	2
Starting Vagon	2
Working with Texts	3
Loading Texts	3
Saving Texts	4
Add Vocabularies and Concept Lists	4
Adding Vocabulary / Concept List Entries	5
Text Metadata	6
Creating Appellations	6
Theory	7
Praxis	7
Creating Relations	9
Theory	9
Praxis	10
Export Quadruples (Triples)	10

Introduction

This is a step-by-step description of how to use Vagon to annotate texts with quadruples. Since Vagon is still in development this manual can only describe half the workflow. There will be an updated version as soon as Vagon can do more.

The Short Story

Here is a brief outline of the annotation workflow.

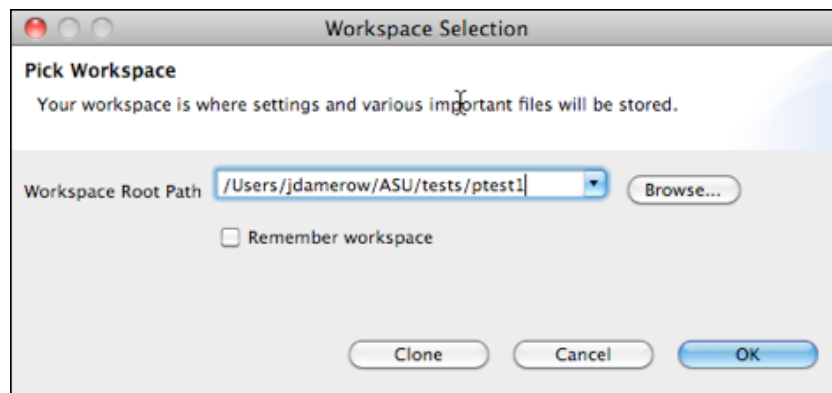
1. Open Vagon and choose workspace.
2. Load text into Vagon.
3. If not yet added, add vocabularies (aka dictionaries) and concept lists to Vagon.
4. Create appellations.
5. Create Relations.
6. Export quadruples. [NOTE: This feature is not yet implemented!]

No idea what you have to do? Then keep on reading the long story!

The Long Story

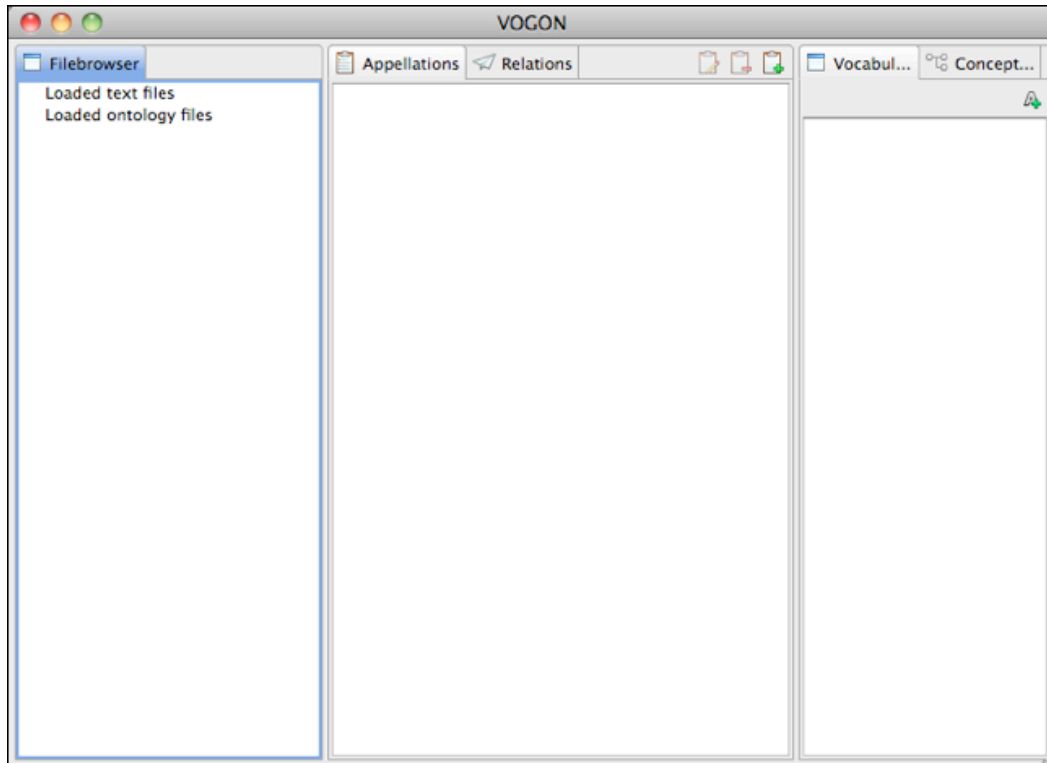
Starting Vagon

First of all, you need to start Vagon. Do that by double-clicking on its application symbol. First thing Vagon will do is to ask you which workspace you want to use. A workspace is a folder on your computer where Vagon keeps all files you load into Vagon. The first time you use Vagon you have to create a workspace folder somewhere on your computer, after that you can reuse this folder. You can create a new empty workspace folder at anytime if you for example start a new project. The input dialog you will see when you start up Vagon will look like that:



Use the "Browse..." button to select a workspace folder. Then click "OK." The pull-down menu lists all the workspaces you used before. If you're asked if you really want to create a workspace in this location, check if the folder is really the folder that you want to use and then click "OK." [NOTE Mac OSX users: although you normally don't see it but your user folder is located in /Users/yourusername/ and this path will appear in the dialog.]

Once Vogon is started up it will look somehow similar to this:



Vogon Environment

Working with Texts

This section describes how to work with texts in general.

Loading Texts

Let's start by loading a text into Vogon. From the main menu (the menu at the top of the screen for Mac OSX users or the menu at the top of the application window for PC users) choose "Analysis >> Analyse Texts..."; a dialog will open. This dialog has a button labeled "Choose..." next to the text "Choose input directory." Click that button and choose the folder in which your text is located (Vogon can read plain text or MEDLINE abstracts; NOTE: if you use plain text make sure the title is not part of the text!). Click "Open."

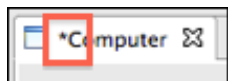
In the "Analyse Texts" dialog in the upper list there should now appear a list of all the files contained in the folder you chose. Select the checkbox corresponding to the text you want to annotate. If you have already ontologies loaded

into your workspace you can also select the ontologies you want to use to analyze the text. If you do that Vagon will try to find the terms that are in these ontologies in your text and highlight them. **NOTE: I use ontologies and concept list in this document interchangeably.** Click "OK."

On the left side of your application window there is the "filebrowser." You can now click on the arrow (or plus sign on Windows) next to "Loaded text files." There will be an entry with the path to the text file you've just added. Again click on the arrow next to the path to find the contained text. If you added a text plain file (with ending "txt") it should say "No title." Double-click on the text. An editor with the text will open in the center of your application window.

Saving Texts

When you create appellations, relations or modify the metadata of a text, you need to save your changes to persist them. You can find out if a text contains changes that haven't been saved yet by examining its tab. If the text in the tab (the title of the displayed text) has a little star in front of it, you made changes that haven't been saved yet.



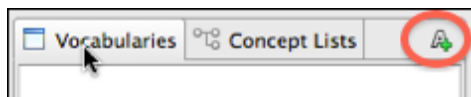
To save your text select "File >> Save" or the appropriate shortcut (on Mac OSX: cmd+S).

Add Vocabularies and Concept Lists

In Vagon vocabularies and concept lists are very similar but have very different purposes. A vocabulary is like a dictionary. It defines what terms exist, how they are spelled and what meanings they have. A concept list is more like thesaurus. It has an entry for every meaning of a word. We call these meanings "concepts." To annotate a text with Vagon you need both: how a word is spelled (its normalization) and what it means (the concept it refers to). For example, in the sentence "I ran up a hill," "ran" is the past form of the verb "to run" and we want to capture the information. We also want to this word refers to, the concept it refers to. In the example, that would be the "concept" of walking very fast.

Vagon lets you use remote vocabularies and concept list. This means that all the terms you need are stored on a server and you just "copy" them to your workspace when you need them. This way we can make sure, that everybody uses the same vocabulary and concept list.

To add a new vocabulary or concept list (it works the same way for both) click on the button with the little plus sign in the right upper corner of the vocabularies/concept list view. This view is by default on the right side of your application window. A view is a part of the application window that has a tab. But you can rearrange views by drag them at their tab and dropping them wherever you want.



Add vocabulary



Add concept list

After you clicked on the add button a dialog opens that asks you for the URL of the vocabulary service/concept list service. So far, Vagon can only talk to Wordpower/Conceptpower. The URLs you need to enter for these services are

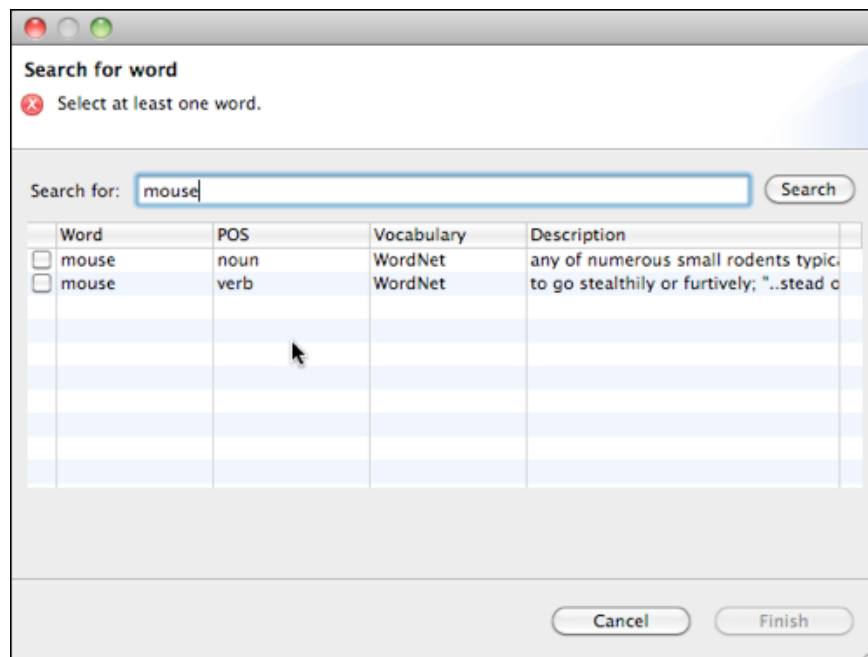
<http://myserver.uni.edu/wordpower> for Wordpower

and

<http://myserver.uni.edu/conceptpower> for Conceptpower where you replace myserver.uni.edu with the address of the server you want to use. Then click "OK." A new vocabulary / concept list will appear.

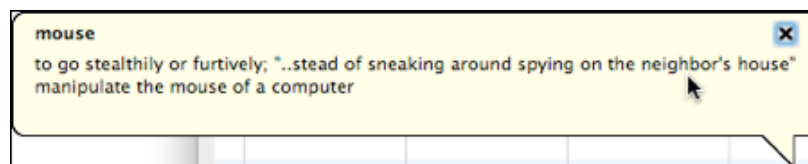
Adding Vocabulary / Concept List Entries

Once you have vocabulary / concept list you probably want to add entries to it (otherwise it wouldn't make much sense). To do that right-click on the vocabulary / concept list you want to add an entry to. A context menu will pop up. Select "Retrieve Word" (for a vocabulary) or "Retrieve Concept" (for a concept list) from that menu. A search dialog will open. Enter the word you're searching for and click "Search." If the remote vocabulary / concept list contains the word you're looking for the search results will be displayed in the list below the search field.



Search Results

If you're adding a term to a vocabulary then there can maximal be 4 search results (one for each type of word: noun, verb, adverb and adjective). If you're adding a concept there can be as many results as there are meanings of that word. To see the full description of the word double click on an entry. A little information balloon will open displaying the description.



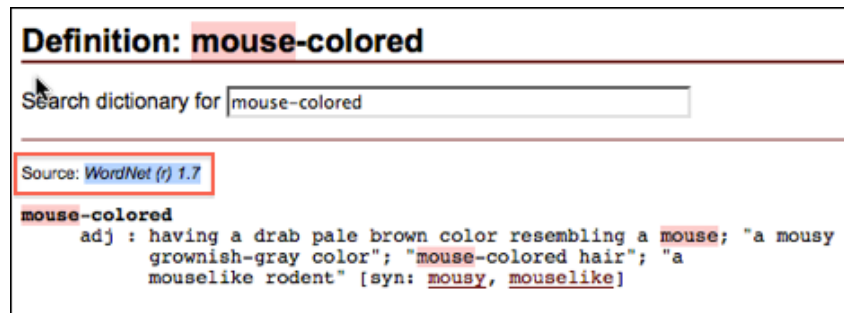
Information Balloon

If the word you're looking for is among the search result select its checkbox in the result list and click "Finish." A new term/concept will be added to your vocabulary/concept list.

Note: The search functionality in Vogon is very limited. If you want to find out if a word is included in the remote dictionary (at the point of writing this tutorial this would be Wordpower/Conceptpower) try using the following website:

<http://dictionary.die.net/>

This website lets you use wildcards. When you found the word you were looking for on that website, check if there is source called "WordNet (r) 1.7." If there is then the word will most likely be included in Wordpower/Conceptpower. Try to search for exactly the term dictionary.die.net showed you in Vogon.



Result page of <http://dictionary.die.net/>

Text Metadata

Vogon lets you specify a minimal set of metadata: title, author and source URI. The source URI is the URI of the text in an repository. The idea is that the repository your text is stored in manages the text's metadata and Vogon only knows the URI of the text (if Vogon needs some metadata it can retrieve it using the URI). In DSpace such handle could look like that:

<http://hdl.handle.net/123456789/17>

To edit the title, author and source URI open the text you want to edit the metadata for. Then click on the button (blue with a white "i") in the upper right corner of the text editor.



An input dialog will open with text fields for the metadata. Note: author and title are just for your text organization in Vogon. Information in these fields won't be used later on.

Creating Appellations

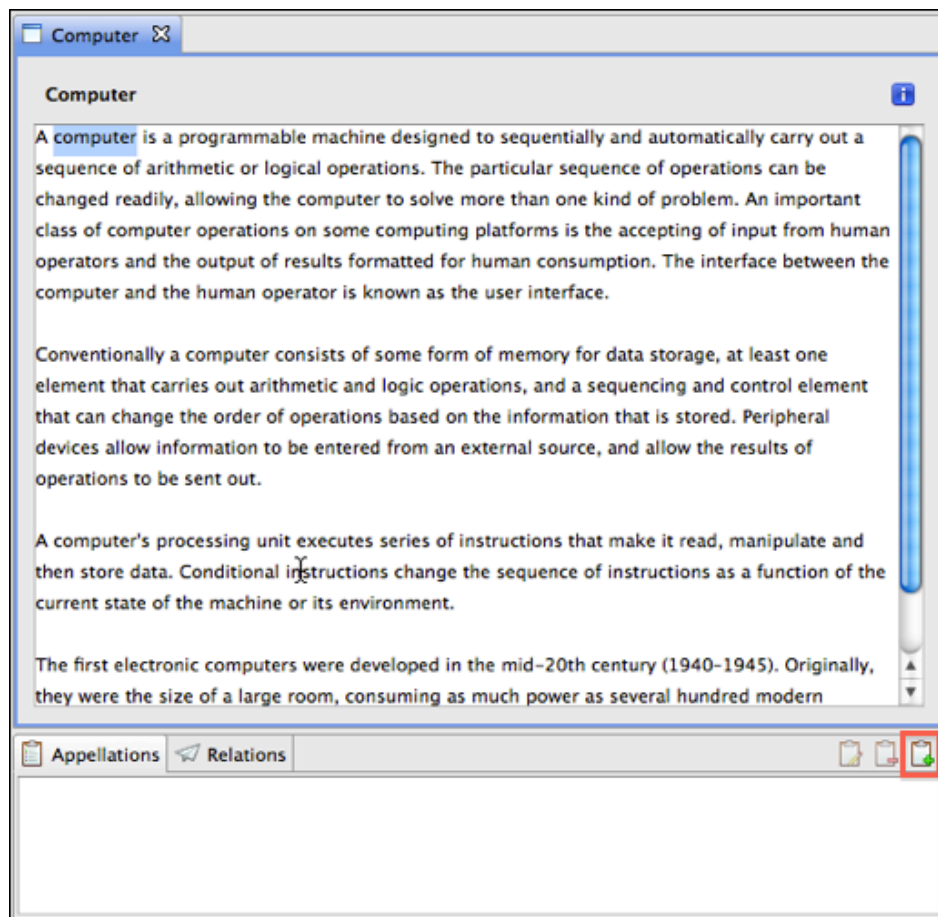
This section describes what appellations are in Vagon and how you can create them.

Theory

First of all, what does it mean to create an appellation with Vogon? To explain that, it is necessary to understand the problem we want to solve by creating appellations. When the computer "sees" a word (for example "house") it can tell you that the word consists of 5 characters and that these characters are h - o - u - s - e. It can also tell you how many bytes this word needs to be stored in a computer etc. but that's it basically. Everything else, like the meaning or if it's singular or plural, needs further computation. It's like someone would show you a finish or Chinese word (or whatever language you can't read). So, one of the big challenges in computer science is to make computers "understand" words and texts; from the grammar of words and sentences to the meaning of texts. By creating an appellation in Vogons you give a computer information about the grammar of a word and its meaning.

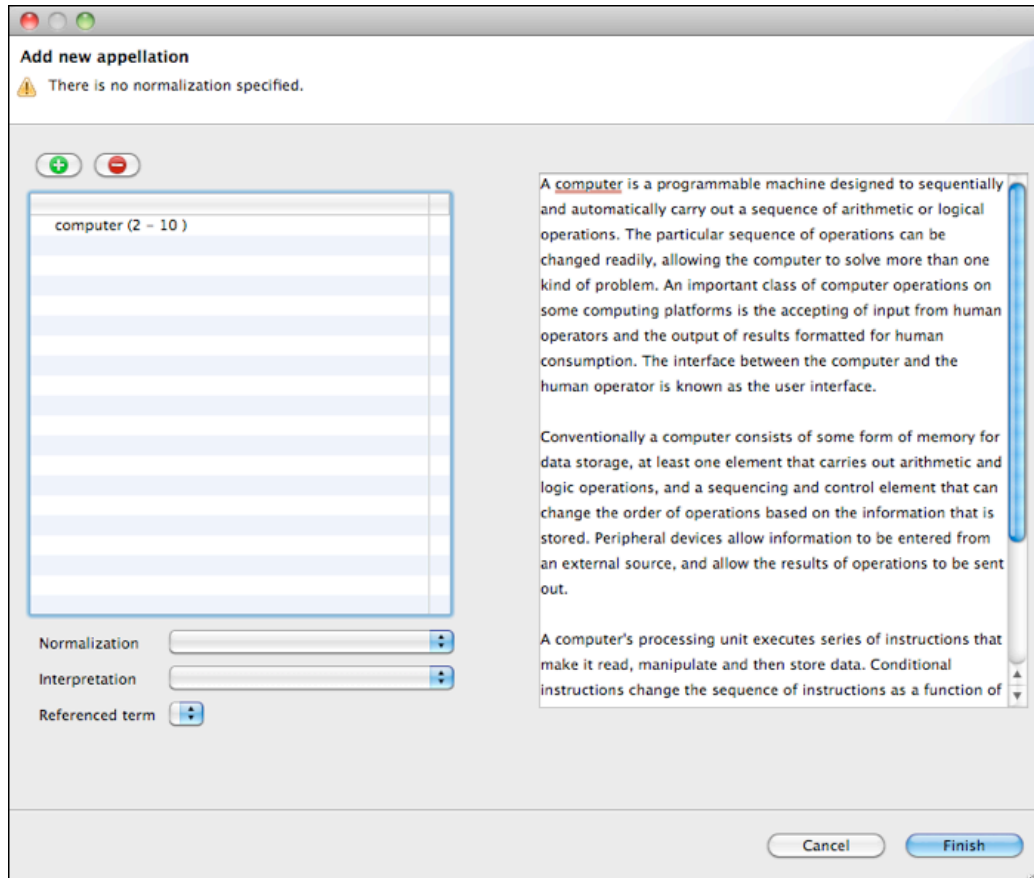
Praxis

To create an appellation first make sure that the term you want to create an appellation for is in one of your vocabularies and that its meaning is a concept in one of your concept lists (see section "Adding Vocabulary/Concept List Entries" for how to add terms). Then, select the word you want to create the appellation for. Make sure that the "Appellations" view is selected by clicking on its tab below the text editor. On the right upper side of that view there is a button with a clipboard sign and a little plus sign. Click this button.



Create an Appellation

A dialog will open that lets you specify the properties of the new appellation. First thing that needs to be defined are the parts of the term you want to create an appellation for. For instance, if you have the text "arithmetic or logical operations" and the term you are interested in is "arithmetic operations," you need to specify two parts "arithmetic" and "operations."



Create new Appellation Dialog

The list of the left side of the dialog shows all parts of the term you want to create an appellation for. You can add a new part by selecting it in the text field on the right side of the dialog and click the button with the plus sign above the list on the left side. Words that were added to list are underlined red in the text field on the right side. You can remove a part by selecting it in the list and click the button with the minus sign above the list.

Once you added all parts of a term you have to specify its normalization and interpretation. The normalization of a term is the word that you would find in a dictionary. For instance, if the text says "he ran up a hill" and you're creating an appellation for "ran" the normalization would be "run;" for plural nouns the normalization would be their singular form, etc. You use the terms specified in your vocabularies to define normalizations. To define a normalization click on the pull-down list next to "Normalization" below the term parts list and select a term.

The interpretation of a term is its meaning. You use the terms that you defined in your concept lists for that. For example, if you have the word "desktop" you need to specify if the top of a desk is meant or the desktop "the area of the

screen in graphical user interfaces against which icons and windows appear."¹ Similar to normalizations you specify the interpretation of a term by clicking on the pull-down list next to "Interpretation" and select a term.

The last thing you can specify is "Referenced term." This property needs to be set if the term you're annotating is an anaphora. For example, if the text is "Tom bought bananas. He ate them all." and you want to create an appellation for "he," you would have to specify that "he" refers to Tom in the previous sentence. You define a referenced term by using the pull-down menu next to "Referenced term." **Note: You can only reference appellations. This means before you create an appellation for "he" you need to create an appellation for "Tom."**

Click "Finish." The term you've created the appellation for will be highlighted in the text editor and an appellation will be listed in the "Appellations" view below the text editor.

Creating Relations

This section describes what relations are and how you can create them.

Theory

In Vogon, relations represent subject - predicate - object statements. We believe that every statement can be decomposed into a set of such triples based on the following assumptions:

- Triples can be nested (a part of triple can reference another triple).
- It is not necessary to represent a phrase exactly but interpret its meaning and add triple parts if necessary (for instance, add a predicate that is only implied in the text).

Example

"A computer is a programmable machine designed to carry out a sequence of arithmetic or logical operations."

The following triples can be created (text in brackets represents the normalized/ concept of a term):

T1: machine - is (be) - programmable

T2: arithmetic operations - or (is in disjunction with) - logical operations

T3: T1 - designed to carry out (carry out) - T2

T4: computer - is (be) - T3

This example shows how the above mentioned assumptions are applied. For instance, the adjective-noun construction "programmable machine" can be translated to a subject - predicate - object triple by adding "is" (to be) to it (triple T1). The disjunction "arithmetic or logical operations" can similarly be translated. The phrase "designed to carry out" can be interpreted simply as "carry out." By nesting triples, triple T4 comprises the whole information of the sentence: a computer is a machine that is programmable that carries out arithmetic or logical operations.

Note: Keep in mind that you should only create triples for what you think is the relevant information of a text.

¹ WordNet 3.0 definition

Praxis

So, now you know what relations are. Let's learn how to create them with Vogon. First, make sure that all terms that you want to use to create a relation are annotated with an appellation. You can only use appellations to create relations. Then, activate the "Relations" view below the text editor by clicking on its tab (the tab labeled "Relations").

In the activated "Relations" view, click on the bubble symbol with the plus sign in the right upper corner of the view:



An input dialog opens that allows you to specify subject, predicate and object of the new relation. The pull-down lists on the left side (labeled "Subject", "Predicate", "Object") contain all appellations that were created for the text. If there aren't many appellations, the fastest way is to simply select subject, predicate and object from these pull-down menus.

However, if you've already created many appellations you can filter these lists. To filter a pull-down menu, click on the button with the tool symbol next to the pull-down menu (the filter button). The button will show that it is selected by being of darker grey color than before. Next, click on the (blue) highlighted word in the text field on the right side that you want to use in a new relation. **Note: the highlighted words in the text field represent words that are annotated with appellations. You can only use these words in your relation.** The pull-down menu will now only contain appellations that annotate the word you've just selected. Select the appellation you want to use from the, now filtered, pull-down menu. If you want to remove the filter from a pull-down menu, simply click the filter button again and select an appellation from the whole list of appellations. Once you've decided on an appellation for a pull-down menu, make sure that the filter button isn't selected anymore. You can deselect it by clicking it again.

Appellations that are selected in one of the pull-down lists are highlighted in orange in the text field. When you've a subject, predicate and object click the "Finish" button. A new relation will be listed in the "Relations" view. If you select a relation in the "Relations" view the words in the text that are used in the relations will be highlighted in orange.

Export Quadruples (Triples)

This feature is not yet implemented. This documentation will be updated when this is done.