

DEPiCTed

The annotation tool for the DEPiCT project

User manual (for Mac users)

Damaris Neuhof, Bianca Widlitzki & Viveka Velupillai

Content

1	Introduction.....	1
2	About DEPiCTed.....	1
2.1	What is DEPiCTed?.....	1
2.2	What does DEPiCTed do?.....	1
2.3	Why is there a need for DEPiCTed?	2
2.3.1	Principles of mark-up by tags.....	2
3	Download and installation	4
4	The components of DEPiCTed	6
4.1	The Menu bar	6
4.1.1	File	6
4.1.2	Edit.....	7
4.1.3	Find	9
4.1.4	Help.....	9
4.2	The Editing toolbar	10
4.3	The Status bar	11
4.4	The Tagging toolbar	12
5	Importing .txts, saving .xmls, opening .xmls and exporting projects.....	12
5.1	Importing .txt files.....	12
5.2	Saving .xml files	12
5.3	Opening .xml files	13
5.4	Exporting projects.....	13
6	Contextual and sociolinguistic annotation in DEPiCTed.....	14
6.1	Excerpts	14
6.2	Situation.....	15
6.3	Setting.....	15
6.4	Attitude	16
6.5	Description	16
7	Language Data	17
8	Making corrections and manual changes in DEPiCTed	17
8.1	Removing tags in DEPiCTed.....	17
8.2	Making manual changes in DEPiCTed	18
9	Comments, footnotes and general formatting.....	18
9.1	Comments	18
9.2	Footnotes.....	19
9.3	General formatting: Page break and line break.....	19
10	Editing annotations	20

10.1	Grouping data elements.....	20
10.2	Viewing and deleting Data tags	22
11	Entering bibliographical information in DEPiCTed.....	22
12	Known issues.....	25
13	Error reports	25
14	Sources	26


1 Introduction

This manual will introduce you to DEPiCTed, the annotation tool which was specifically designed by Magnus Nissel for the **Database of Early Pidgin and Creole Texts** (DEPiCT) project.

In this manual you will learn how to download and execute the tool and you will be introduced to the tool's various functions. This manual will give step-by-step instructions on how to prepare your texts, how to import your texts into DEPiCTed, as well as how to add sociolinguistic and linguistic annotations to your texts.

The manual is supplemented by 22 tutorial videos which you can find here:

<https://www.youtube.com/playlist?list=PLXRYh7O8yWqBri-vkAmXiGr7lxd1JnERq>.

You will also find a direct link to the respective videos in this manual. For those functions where there is a tutorial video, you will find the following symbol in the right margin: 

By clicking on the symbol you will be led directly to the relevant YouTube tutorial.

Examples of correct versus incorrect annotations will be given throughout. Examples of correct annotations are indicated with a green tick (✓) while examples of incorrect annotations are indicated with a red cross (✗).

At this stage there are still some issues that specifically relate to the Mac version of the tool. We are aware of these issues and are working to solve them as soon as possible. Please refer to the [Known issues](#) section for further information.

2 About DEPiCTed

2.1 What is DEPiCTed?

DEPiCTed is the annotation tool specifically developed by **Magnus Nissel** for the DEPiCT project. It is an easy to use and intuitive tool which allows each contributor to annotate text for meta-linguistic and linguistic information.

2.2 What does DEPiCTed do?

DEPiCTed will transform your normal text files (ending in .txt) into XML files (ending in .xml). XML stands for "Extensible Markup Language" and is an international standard markup language which serves to describe data. Transforming your text files into XML files will have the advantage that they can be easily moved and translated into other environments. To describe it more in linguistic terms, XML can be described as a "lingua franca" for encoding documents.

The created XML files will contain the original text and those tags which are needed to make the document machine readable. In addition, DEPiCTed will create another XML file containing the bibliographical information entered in DEPiCTed (ending in _bib.xml). Storing the data in these formats will ensure that the data can be easily integrated into the final database.

2.3 Why is there a need for DEPiCTed?

The tool will ensure a consistent annotation of texts among all contributors so that each text becomes maximally cross-comparable with all other texts in DEPiCT and can be easily integrated into the final database. The consistency is ensured by the fact that DEPiCTed provides pre-defined sets of tags.

For the contributors, DEPiCTed has the advantage that

- they are provided with an intuitive menu-driven Graphic User Interface (GUI), which guides them through the annotation steps so that
- no prior knowledge of mark-up language is required. The appropriate mark-up will be inserted by the tool; and
- the tool will automatically check the status of the XML and when relevant flag what type of error occurred.

2.3.1 Principles of mark-up by tags

Even if you do not need any prior knowledge of XML in order to use DEPiCTed, it might be helpful to have a brief look at the general principles to be better able to avoid and/or understand error messages that might appear in DEPiCTed.

- **tags come in pairs:** Every start-tag needs to have an end-tag; the end-tag is identical to the start-tag but has an / added to it, like this: <d> (start-tag) </d> (end-tag). For example:

✓ “<d>My people no do so bad; me wish`e me in Bidera</d>”
(**well-formed:** contains both start- and end-tags)

✗ “<d>My people no do so bad; me wish`e me in Bidera”
(**incorrect XML:** end-tag is missing)

The so-called empty-element tags indicating page break and line break form an exception in that they do not need both start- and end-tags, but are valid on their own in the form <pb/> (for page break) and <lb/> (for line break)

- **the characters “<” and “&” are not permissible in XML:** for instance, if “<” is placed in your text file it will create an error message, as it will be interpreted as the start of a new tag by the parser.

DEPiCTed is programmed in such a way that when you import a text into the editor, the tool will automatically replace the non-permissible characters with the appropriate pre-defined entity references:

&lt;	<	less than
&gt;	>	greater than
&amp;	&	ampersand

Using these pre-defined entity references will allow you to keep these non-permissible characters in an XML understandable format. You can also manually replace the characters by going to **Edit>Replace ampersands** and **Replace angle brackets**.

- **tags have to be nested inside other tags:** this means that if a tag element is opened inside another tag, it must be closed inside again. Look at the following example:

As the language data tag `<d>` opens inside the situation tag `<sit>`, it has to be closed inside the situation tag as well.

- ✓ `<sit>` [...] he looked as if his moral principles had received a shock, and said “`<d>`My people no do so bad; me wish`e me in Bidera`</d>`”`</sit>`
(**well-formed**: the `<d>` start- and end-tags are nested inside the `<sit>` start- and end-tags)
- ✗ `<sit>` [...] he looked as if his moral principles had received a shock, and said “`<d>`My people no do so bad; me wish`e me in Bidera”`</sit>` `</d>`
(**incorrect XML**: the `<d>` start-tag is inside the `<sit>` tags but the `</d>` end-tag is outside the `<sit>` tags)

- **no tags other than the predefined tags are valid:**
If your text already contains other tags not defined in DEPiCTed, the tool will automatically escape¹ the non-DEPiCT tags when importing your text into the editor
- **tags are case sensitive:** the tag `<Sit>` (with a capital *S*) is read as different from the tag `<sit>` (with a lower case *s*)
- **tags are classified as “parent” and “child” elements, whereby a parent tag cannot be nested into a child tag:**
XML documents contain one root element that is the “parent” of all other elements. For instance, in DEPiCTed `<ex>` (for *excerpt*) is the root element, while `<sit>`, `<set>`, `<lng>`, `<att>` (for *situation*, *setting*, *language description* and *attitude* respectively) represent “child” elements that have to be nested inside the parent element `<ex>`. In other words, child elements always have to be nested inside parent elements and parent elements cannot be nested inside child elements. This means that `<ex>` tags cannot be placed as child elements inside, for instance, `<sit>` tags; as `<ex>` is the parent element and `<sit>` the child element.

¹ In this context to ‘escape’ means to apply an alternate meaning to a character or set of characters, in this case the alternate meaning is to ignore the given character(s).

3 Download and installation

Please note that the tool is at this stage only possible to run via Python. We apologize for this inconvenience and are working as fast as we can to provide a proper .app for the tool.

Before being able to run DEPiCTed you will need to make sure that you have installed Anaconda Python 3 on your computer. If not, you can download Anaconda 3 for free by going to the following link: <https://www.continuum.io/downloads>.

The following links take you straight to the relevant file. To download the appropriate version for your computer, start by stating if your computer is running a 64-bit or 64-bit command line version:

- if your computer is running a 64-bit version:
https://repo.continuum.io/archive/Anaconda3-4.1.1-MacOSX-x86_64.pkg
- if your computer is running a 64-bit command line version:
https://repo.continuum.io/archive/Anaconda3-4.1.1-MacOSX-x86_64.sh

In case you have an older version of Python on your computer, please make sure that you uninstall it before downloading and installing Anaconda Python 3, as otherwise the two programs might interfere and cause problems when opening DEPiCTed. [To uninstall an older version of Python, simply open the “Finder” window, search the old version of Python, Control-click or right-click on the application’s icon and select “Move to Trash”.]

Clicking on one of the above links will start the download of Anaconda Python 3. Anaconda will be saved in your “Downloads” folder. If you go to the “Downloads” folder and double click on the respective file, the installation of Anaconda Python 3 will start. Please consider that downloading Anaconda 3 might take roughly ten minutes and that you might need another 10 minutes for its set-up. After having installed Anaconda 3 restart your computer.

Get and install DEPiCTed:

In order to download DEPiCTed go to the following link:

<https://www.uni-giessen.de/faculties/f05/engl/ling/staff/professors/mhuber/depict/depicted/download>

This takes you to the Download page on our project website from which you can download the latest version of the tool. To download the appropriate version for your computer, select the version “DEPiCTed for Mac”. After clicking on **Download**, DEPiCTed will be downloaded into your “Downloads” folder.

An alternative way to install DEPiCTed is to have the program sent to you by email. You will then receive a .zip folder from the DEPiCT team. Once you receive the email with the .zip folder from us, go to the email and download the appropriate .zip folder.

The files will be saved in your “Downloads” folder on the hard disk. Go to the “Downloads” folder and search for the folder “DEPiCTed_for_Mac”. Your Mac will

automatically unzip the folder. Move the folder to the location where you want to keep it.

The folder will contain several files and folders. As a user, only the following will be of importance to you:

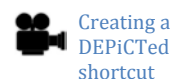
- **the folder “import”**: this is where you place those files that you want to import to DEPiCTed
- **the folder “documents”**: this is where the transformed XML files will be stored
- **the folder “export”**: this is where the final exported projects will be stored and
- **the file “depicted.py”**: which is the file you need in order to execute the program

To facilitate your work with DEPiCTed, it might be helpful to create a shortcut on your desktop.

To create a shortcut:

1. Go to the location where you have stored the DEPiCTed folder
2. Control-click or right click on the file ‘depicted.py’ and select **Make Alias**
→ A shortcut of the file depicted.py will be created and appear next to your original depicted.py file
3. Click on the shortcut which you have created and drag it to the desktop
→ The shortcut to DEPiCTed will appear on your desktop

If you prefer to have a shortcut to the whole DEPiCTed folder, follow steps 1-3 for the folder instead of the file.



To run DEPiCTed:

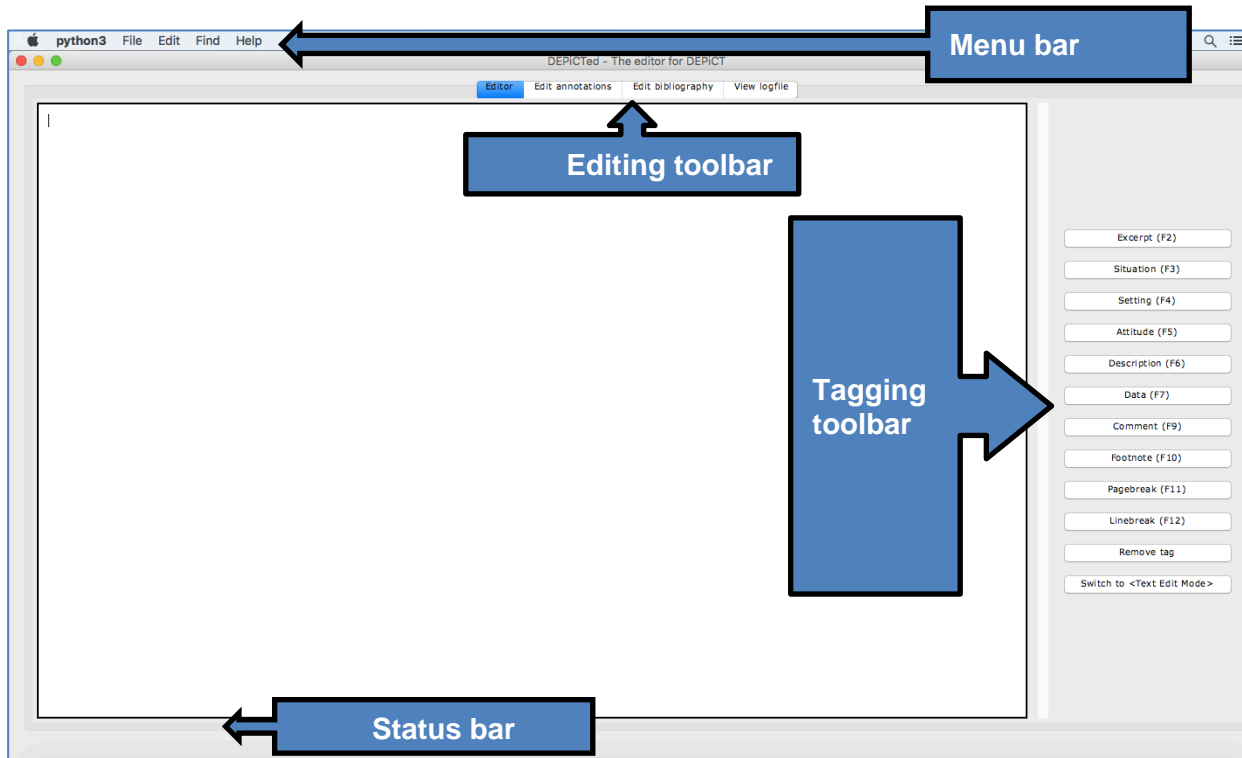
1. Click on **Go** in the Menu bar and select **Utilities**
2. Double click on “Terminal.app”
3. Type **cd desktop** and press **Enter** in case you have stored the depicted.py file on your desktop. In case you have stored it somewhere else, please indicate the path where you have stored it after cd.
4. Type **python3** and press **Enter**
5. Type **import idlelib.idle** and press **Enter**
→ The Python Shell will open on your screen.
6. Click **File>Open**, select the file ‘depicted.py’, and click **Open**
→ The depicted.py script will open
7. Click **Run>Run Module**
→ DEPiCTed will open on your computer



It might be that you receive some error messages in the Python Terminal or in the Python Shell when you open DEPiCTed. You can simply ignore these error messages if DEPiCTed runs on your Mac, as they are based on layout issues and will not prevent you from annotating your data. We are aware of this issue and are working to find a solution for it.

4 The components of DEPiCTed

When you first open DEPiCTed, you will be presented with the basic user interface of DEPiCTed as in the following screenshot. The tool consists of a [Menu bar](#), an [Editing toolbar](#), a [Tagging toolbar](#) and a [Status bar](#).



4.1 The Menu bar

On the menu bar you have the items [File](#), [Edit](#), [Find](#) and [Help](#).

4.1.1 File

In **File** you have the following options:

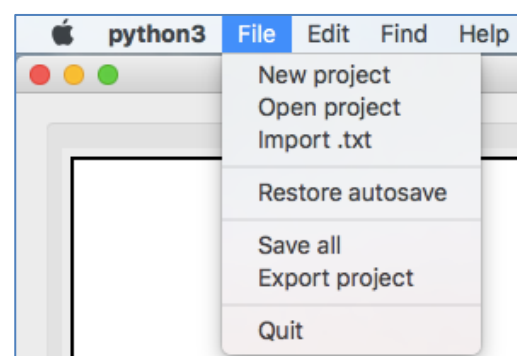
New project: This allows you to start a new project.

Open project: This allows you to open files that you have already worked on before and which were saved as .xml files.

Import .txt: This allows you to import a text file into the main editor.

Restore autosave: This allows you to retrieve any desired autosave, which will override the latest version.

Save all: This allows you to save your annotated data as XML.



Export project: This allows you to export your annotated data and to save it as a zip file after completing the annotation.

Quit: This allows you to shut down the DEPiCTed tool.

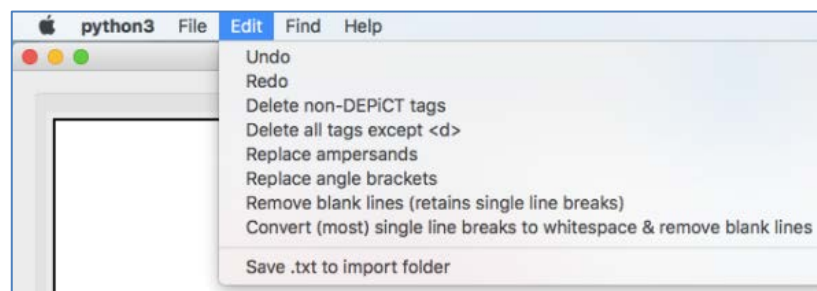
See [5. Importing .txts, opening .xmls, saving .xmls and exporting projects](#) for further information.

4.1.2 Edit

Text documents need to be well-formed and valid in order to save them as XML files. This means that they have to comply with certain syntactic rules. As explained in [2.3.1](#) the tool will automatically check the imported texts for various problematic characters, such as ampersands, brackets and non-DEPiCT tags and will replace these. In other words, the program will automatically transform the text into well-formed XML.

In case you inserted a problematic character by accident only after the importation of the text into the editor, the options under **Edit** are there to give you extra opportunities for manual clean-up:

Undo: This allows you to erase the last change you made to your document and will revert it to an older state. This function will only work when you are in the text edit mode.



Redo: This allows you to redo the last change you made to your document. This function will only work when you are in the text edit mode.

Delete non-DEPiCT tags: This allows you to delete tags which are not the ones predefined in DEPiCTed. Notice: Once you have clicked on **Delete non-DEPiCT tags** you cannot undo the function.


Delete all tags except <d> : This allows you to delete all annotation tags which were entered into DEPiCTed at once, with the exception of data tags which will have to be deleted via **Edit annotations**. Notice: Once you have clicked on **Delete all tags except <d>** you cannot undo the function.

Replace ampersands: This option allows you to replace ampersands (“&”) from your text. Notice: Once you have clicked on **Replace ampersands** you cannot undo the function.

Replace angle brackets: This option allows you to replace angle brackets (“<”) from your text. Notice: Once you have clicked on **Replace angle brackets** you cannot undo the function.

Remove blank lines (retains single line breaks): This option allows you to remove blank lines from your text. Notice: Once you have clicked on **Remove blank lines (retains single line breaks)** you cannot undo the function.

Original text, with blank lines:



```
Der Bukajunge Hellwigs schnitt diesem mit der Maschine die
Haare, Makiror sah zu. Plötzlich sagte er zum Buka:

„You no cut him gras belong master too much,
bymby coconut belong him look allsame bottle!"

„Du nicht schneiden es Haar gehörig zum Herrn zu sehr,
schließlich Kopf gehörig ihm sieht aus (blank) wie Flasche!"
```

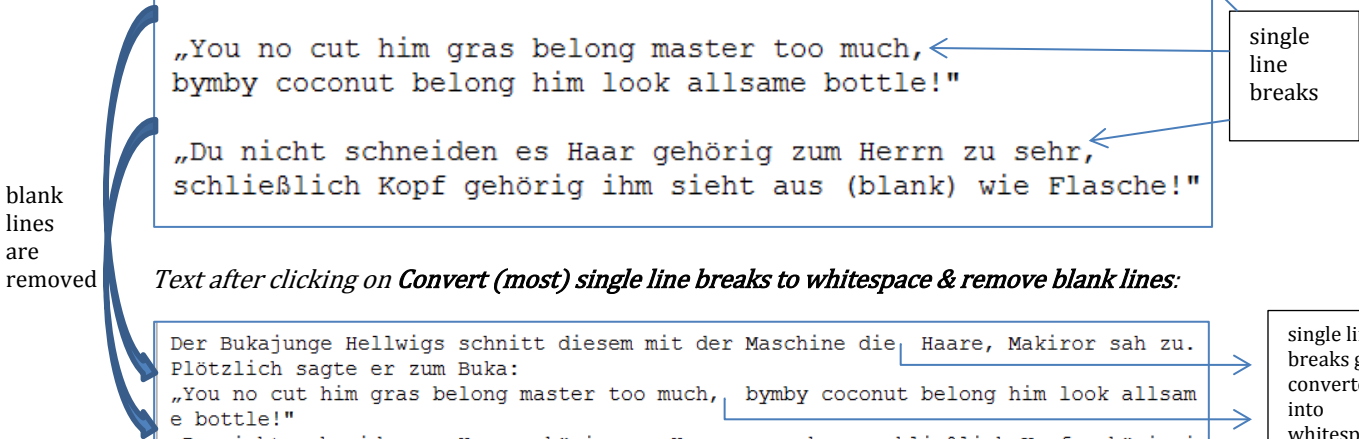
*Text, after clicking on **Remove blank lines (retains single line breaks)**. The blank lines are removed:*

```
Der Bukajunge Hellwigs schnitt diesem mit der Maschine die
Haare, Makiror sah zu. Plötzlich sagte er zum Buka:
„You no cut him gras belong master too much,
bymby coconut belong him look allsame bottle!"
„Du nicht schneiden es Haar gehörig zum Herrn zu sehr,
schließlich Kopf gehörig ihm sieht aus (blank) wie Flasche!"
```

(Vogel 1911: 25)

Convert (most) single line breaks to whitespace & remove blank lines: This option allows you to remove blank lines from your text and at the same time to convert single line breaks to whitespace. Clicking on this button will replace 50% of the line breaks in your imported text. This means that the more times you click on the button, the more line breaks are replaced. Notice: Once you have clicked on **Convert (most) single line breaks to whitespace & remove blank lines** you cannot undo the function.

Original text, with blank lines and single line breaks:



```
Der Bukajunge Hellwigs schnitt diesem mit der Maschine die
Haare, Makiror sah zu. Plötzlich sagte er zum Buka:

„You no cut him gras belong master too much,
bymby coconut belong him look allsame bottle!"

„Du nicht schneiden es Haar gehörig zum Herrn zu sehr,
schließlich Kopf gehörig ihm sieht aus (blank) wie Flasche!"
```

*Text after clicking on **Convert (most) single line breaks to whitespace & remove blank lines**:*

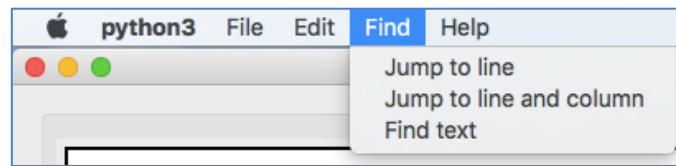
```
Der Bukajunge Hellwigs schnitt diesem mit der Maschine die  Haare, Makiror sah zu.
Plötzlich sagte er zum Buka:
„You no cut him gras belong master too much, bymby coconut belong him look allsam
e bottle!"
„Du nicht schneiden es Haar gehörig zum Herrn zu sehr, schließlich Kopf gehörig i
hm sieht aus (blank) wie Flasche!"
```

(Vogel 1911: 25)

Save .txt to import folder: This option allows you to save changes you make to your text prior to annotating it.

4.1.3 Find

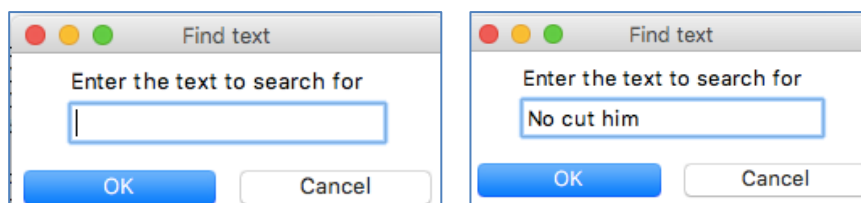
The item **Find** will be helpful in case you are searching for a specific sentence in your .txt file. You can select from the following options:



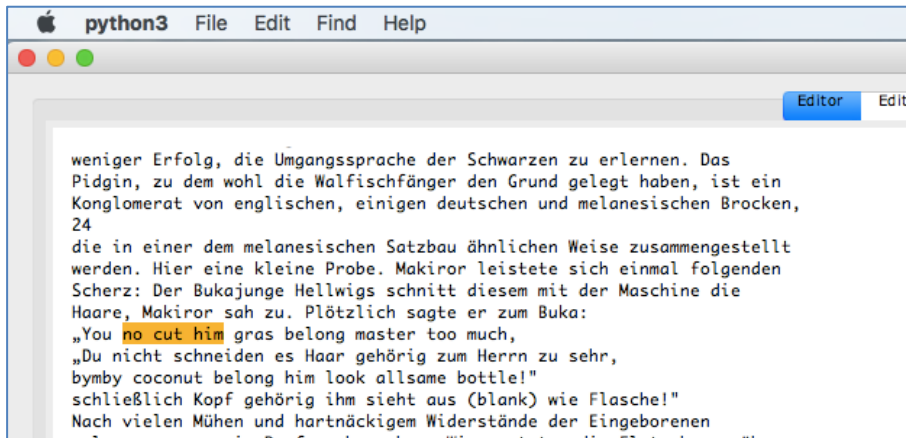
Jump to line: This allows you to trace back to places where the XML mark-up is incorrect (see Section [4.3. Status bar](#)).

Jump to line and column: This allows you to trace back to places where the XML mark-up is incorrect (see Section [4.3. Status bar](#)).

Find>Find text: This allows you to search for a specific sentence in your document. If you select **Find>Find text** a window will pop up in which you can type the text you are searching for.



By clicking **OK** the tool will jump to the sentence.

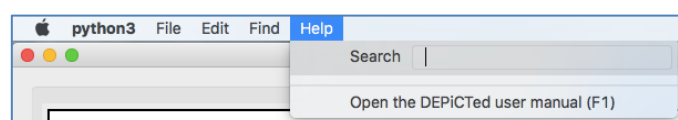


(Vogel 1911: 25)

Please make sure that you type in at least three words or are very precise. If you go to **Find>Find text** and just type in "I" it will take a lot of time until the tool has scanned through the fulltext and identified all instances of "I".

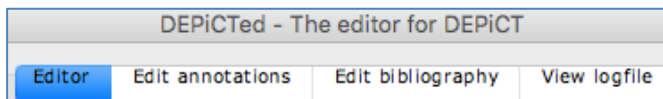
4.1.4 Help

The item **Help** will lead you to a version of this User manual. If you click on **Help>Open the DEPiCTed user manual**, or, alternatively, if you press **F1** on your keyboard, the user manual will open.



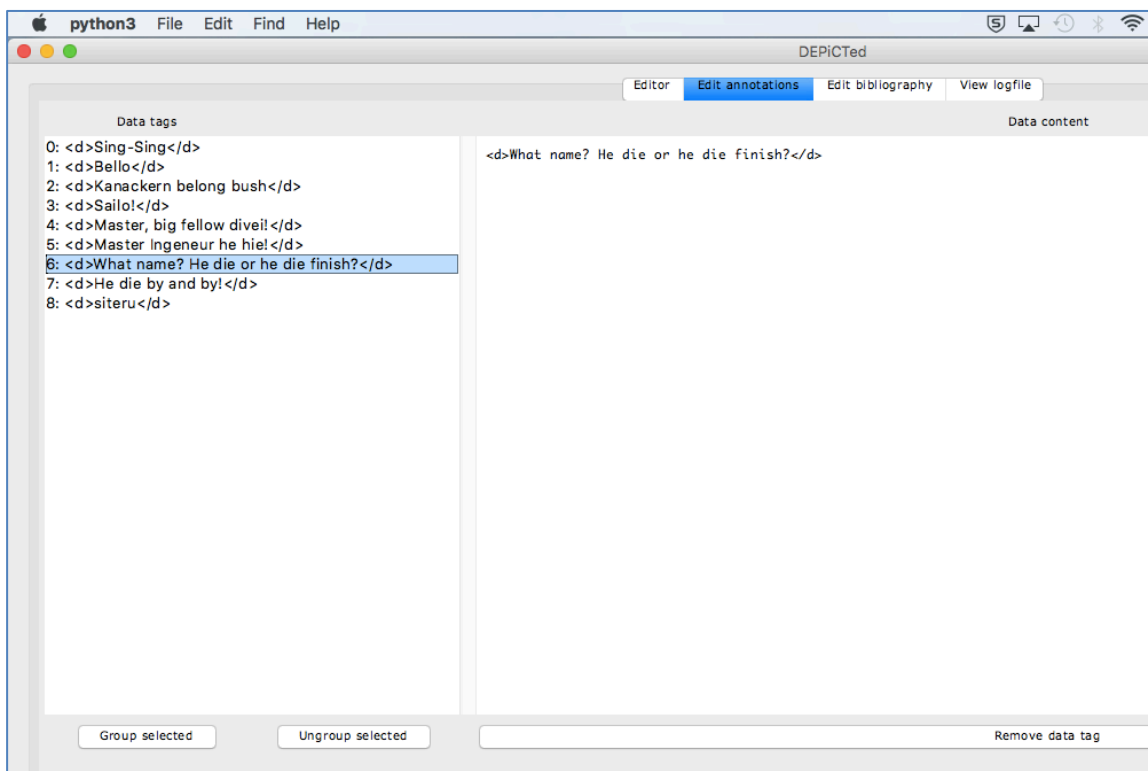
4.2 The Editing toolbar

The **Editing toolbar** is located below the **Menu bar** in which you have the following options:

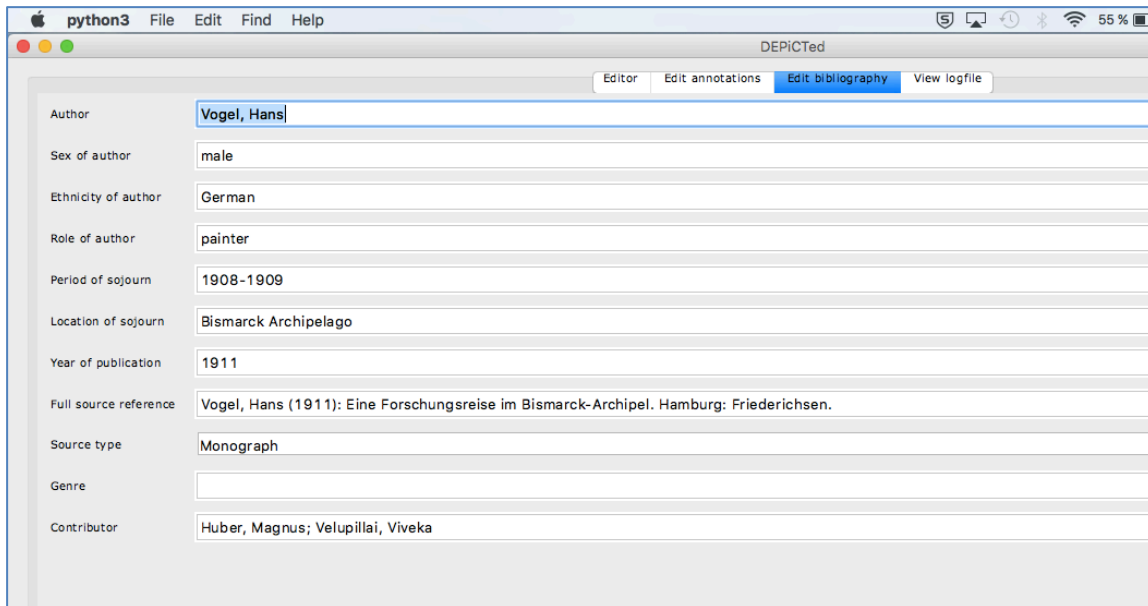


Editor: In the general **Editor** you will annotate your data for linguistic and metalinguistic information.

Edit annotations: In this tab you are able to view and delete your inserted data tags, as well as to group/ungroup data tags (see [Section 10, Editing annotations](#)).



Edit bibliography: The tab allows you to enter any kind of bibliographical information regarding the text you imported into the editor (see [Section 11, Entering bibliographical information](#)).



python3 File Edit Find Help

DEPiCTed

Editor Edit annotations **Edit bibliography** View logfile

Author: Vogel, Hans

Sex of author: male

Ethnicity of author: German

Role of author: painter

Period of sojourn: 1908-1909

Location of sojourn: Bismarck Archipelago

Year of publication: 1911

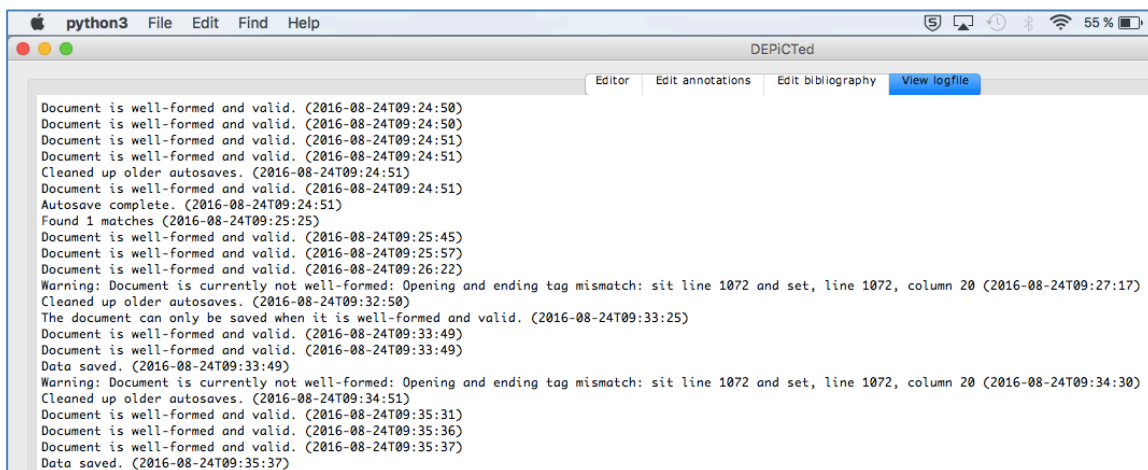
Full source reference: Vogel, Hans (1911): Eine Forschungsreise im Bismarck-Archipel. Hamburg: Friederichsen.

Source type: Monograph

Genre:

Contributor: Huber, Magnus; Velupillai, Viveka

View logfile: The logfile records a history of all events taking place while you are working with DEPiCTed.



python3 File Edit Find Help

DEPiCTed

Editor Edit annotations Edit bibliography **View logfile**

```
Document is well-formed and valid. (2016-08-24T09:24:50)
Document is well-formed and valid. (2016-08-24T09:24:50)
Document is well-formed and valid. (2016-08-24T09:24:51)
Document is well-formed and valid. (2016-08-24T09:24:51)
Cleaned up older autosaves. (2016-08-24T09:24:51)
Document is well-formed and valid. (2016-08-24T09:24:51)
Autosave complete. (2016-08-24T09:24:51)
Found 1 matches (2016-08-24T09:25:25)
Document is well-formed and valid. (2016-08-24T09:25:45)
Document is well-formed and valid. (2016-08-24T09:25:57)
Document is well-formed and valid. (2016-08-24T09:26:22)
Warning: Document is currently not well-formed: Opening and ending tag mismatch: sit line 1072 and set, line 1072, column 20 (2016-08-24T09:27:17)
Cleaned up older autosaves. (2016-08-24T09:32:50)
The document can only be saved when it is well-formed and valid. (2016-08-24T09:33:25)
Document is well-formed and valid. (2016-08-24T09:33:49)
Document is well-formed and valid. (2016-08-24T09:33:49)
Data saved. (2016-08-24T09:33:49)
Warning: Document is currently not well-formed: Opening and ending tag mismatch: sit line 1072 and set, line 1072, column 20 (2016-08-24T09:34:30)
Cleaned up older autosaves. (2016-08-24T09:34:51)
Document is well-formed and valid. (2016-08-24T09:35:31)
Document is well-formed and valid. (2016-08-24T09:35:36)
Document is well-formed and valid. (2016-08-24T09:35:37)
Data saved. (2016-08-24T09:35:37)
```

4.3 The Status bar

Situated on the bottom of the screen is the **Status bar** which will inform you about the status of the mark-up in your text. The mark-up in the texts needs to be well-formed and valid, which means that it has to adhere to the rules for XML files and to the rules set for DEPiCTed XML files. If the mark-up in your text is correct, you will find the following information in the **Status bar**:

Document is well-formed and valid.

If your mark-up is incorrect, an error message will appear in the **Status bar**. Look, for instance, at the following status from the toolbar. To see the error in context and how to correct it, see the YouTube video.

Warning: Document is currently not well-formed: Opening and ending tag mismatch: sit line 1072 and set, line 1072, column 20

You are told that the opening and ending tag mismatch due to the fact that the tags have not been nested inside each other. The lines show the user where exactly the mismatch happened. To find the line and column choose **Find>Jump to line and column** from the Menu bar and type in the line and column which appear in the Status bar. You can then correct your XML error and carry on annotating. Once you have corrected the error, notice that the error message will only disappear once you save your data again or once you set a new tag.

Important: Please do not forget to check continuously the content of the Status bar. It will facilitate your work with DEPiCTed as invalid mark-up can easily be detected.

4.4 The Tagging toolbar

Located on the right side of the DEPiCTed User Interface you will find the **Tagging toolbar**. This is the toolbar you will be using to actually annotate your text.

The Tagging toolbar contains all of the predefined tagging operations which are provided in DEPiCTed. The tags in the Tagging toolbar can be divided into:

- general tags for **contextual and sociolinguistic annotation**
- language **data** tags for linguistic annotation
- **formatting** tags

In addition, the button **Remove tag**, which serves to delete tags, and the button **Switch to <Text Edit Mode>**, which enables you to make manual changes to your uploaded text (such as, for instance, adding comments and correcting typos), can be found.

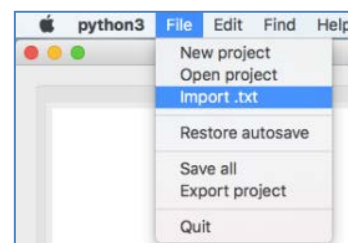
For an exact explanation what each tags signifies, see below ([Section 6 ff.](#)).

5 Importing .txts, saving .xmls, opening .xmls and exporting projects

As DEPiCTed is only able to import files which are saved as .txt files, make sure that you save your data in the appropriate format. Word and also other text editing programs will allow you to change the format of your file by clicking on **Save as**. In the window that opens you will have the opportunity to select the file format. In order to ensure that all characters are saved and displayed correctly, it is important to choose “Unicode (UTF-8)” as a text encoding. Save or move the .txt files which you want to annotate in the folder “import” of your extracted DEPiCTed folder.

5.1 Importing .txt files

Go to the menu bar and click on the item **File**. A submenu will open and will allow you to import, open and save files. Choose **Import .txt**, select a text and click on **Open**. The text will be loaded into the editor window.



 Importing .txt files

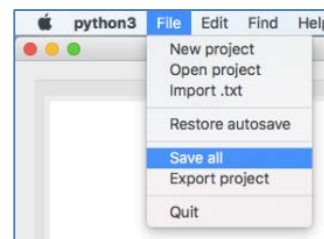
5.2 Saving .xml files

Remember that it is important that you ***continuously save your data!***

 Saving .xml files

In order to save all annotations and bibliographical information you have entered go to the Menu bar and click on **File > Save all**.

Your data will be saved in the folder “documents” as a DEPiCT XML file ending in `_text.xml`. The bibliographical information will be saved in the folder “bibdata” in an XML file ending in `_bib.xml`.



Notice that the option **Save .txt to import folder** which you find under **Edit** in the Menu bar should not be used to save your annotated data. If you chose **Save .txt to import folder** your annotations will be stored in your original .txt file, however all other information, such as the bibliographical information, will not be saved. The option **Save .txt to import folder** is restricted to exclusively save those changes you might have made to your text prior to annotating it.

5.3 Opening .xml files



Opening
.xml files

If you have already imported a text file and saved it as an .xml file once, you will have to go to the Menu bar and select **File>Open project** to select and open the file again. See the YouTube video for further guidelines on how to proceed.

5.4 Exporting projects



Exporting
your projects

Under **File** you will also be able to find the option **Export project**. You will only need to export a project once you have finished the annotation of a text. By clicking on it, an archive (similar to a .zip file) will be created and stored in your export folder which contains the original import file, the transformed .xml file, the bibdata (bibliographical data) and the log file. This archive will be saved with the ending **.tar.gz**.

Notice that **Export project** will export the last successfully saved version of your project files.

Please send the **.tar.gz** file to the DEPiCT team (depict@anglistik.uni-giessen.de) so that your annotated work can be incorporated into DEPiCT. If you want to open the .tar.gz file as well, you will need a zip program such as 7-Zip to extract the data (cf., for instance, <http://www.7-zip.de/>).

6 Contextual and sociolinguistic annotation in DEPiCTed

Contextual and sociolinguistic annotation tags are those indicating what kind of information your text contains (e.g. on the grammar of a language, sociobiographical speaker information, the domains of language use, language attitudes, etc.).

6.1 Excerpts

Your source text needs to be divided into excerpts before you can start tagging and glossing it. An excerpt should be a passage of text that can stand on its own (e.g. because it represents one particular event or situation). ***No other annotation can be made*** unless a passage has been tagged as an excerpt.

The reason for dividing your source text into excerpts is that you, as a contributor, are allowed to submit a whole book as an excerpt, but also just a paragraph or sentence. Furthermore, you might submit an un-annotated excerpt. Marking sections as excerpts will help identify which of the submitted sentences belong together.

Excerpts are tagged with `<ex>`; `</ex>`. For instance:

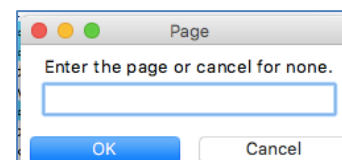
Several were permitted to come on board where they made some alterations in toilet, by transferring the kerchief from the head to the loins, and thus equipped, proceeded to present their testimonials and letters of recommendation, which they carried in leathern or tin cases suspended from the neck, or folded in the head-dress, to the commander, and to ask for employment.

`<ex>``<set>` *These are the Kroomen so frequently mentioned by African voyagers, and so favorably known to our traders and cruisers. They generally speak a little broken English, and from their acquaintance with the tribes and harbors of the coast, are very useful as pilots and interpreters. They occupy several villages along the coasts of Sierra Leone and Liberia, where, though mixed with other tribes, they preserve their own forms of government and religion.*`</set>``</ex>`

(Thomas 1860: 105)

To tag a text passage as an **Excerpt**:

1. Highlight the text that should be marked as an excerpt
Be aware that you will not always see the highlight. However, the function will work nonetheless. Cf. [Known issues](#).
2. Click on **Excerpt** in the Tagging toolbar or alternatively use the hotkey **F2**
→ A pop-up window will open and will ask you to enter the page number on which your excerpt starts or to press **Cancel** if no page number exists.
Please ALWAYS enter page numbers if you know them! This makes the citation of the excerpt more exact in the database.
3. Enter the page number in case your source has page numbers and click **OK**.
→ The tool will add `<ex>` in front of and `</ex>` at the end of the text passage you highlighted



→ If a page number is entered this will be shown in the start tag; for instance if the page number is 12, the opening tag will display `<ex p="12">`. Now the database will be able to give the user the exact reference for the excerpt.

6.2 Situation

The tagging option **Situation** is chosen to tag testimonies surrounding actual language data, meaning statements about the situation in which an utterance was made.

The tags used are `<sit>` and `</sit>`. Notice that situation tags have to be nested inside excerpts tags. For instance:

`<sit>` *They took us by the hand, and, looking in our faces with their mellow black eyes, accompanied by what were meant for witching smiles and a peculiar singing voice, recognised us as old acquaintances, whether they had seen us before or not, in their broken English, "<d>My chin-chin you; me savee you lass voyage; my washe you muche good; can do all the same now; muchee good sweetmeats you savee. </d>"* `</sit>`

(Jacobs 1844: 347)

To tag a part of the text as **Situation**:

1. Highlight the text in question
Be aware that you will not always see the highlight. However, the function will work nonetheless (see video [Tagging Situations](#)). Cf. [Known issues](#).
2. Click on the button **Situation** in the Tagging toolbar or use the hotkey **F3**
→ `<sit>` will be added in front of and `</sit>` at the end of the tagged data. The `<sit>` and `</sit>` tags appear within the `<ex>` and `</ex>` tags.



6.3 Setting

The tagging option **Setting** is chosen to mark testimonies regarding the languages spoken in a contact situation, as well as general information on the political and economic situation.

The tags used are `<set>` and `</set>`. They have to be nested inside the excerpt tags. For instance:

`<set>` *All the races along the coast, in their dealings with the white man, appear to speak a common dialect known as "Kru English."* `</set>`

(Newland 1916: 12)

`<ex>` `<set>` *Lange Jahre hindurch bezog die Firma Godefroy ihre Plantagenarbeiter von Malayta; auch nach Queensland liessen sich viele Wilde* `<pb>` *anwerben.* `</set>` `</ex>`

(Ribbe 1903: 43-44)

To tag a part of the text as **Setting**:

1. Highlight the text in question
Be aware that you will not always see the highlight. However, the function will work nonetheless (see video [Tagging Settings](#)). Cf. [Known issues](#).



2. Click on the button **Setting** in the Tagging toolbar or alternatively press **F4** on your keyboard
 → `<set>` will be added in front of and `</set>` at the end of the tagged data.
 The `<set>` and `</set>` tags appear within the `<ex>` and `</ex>` tags.

6.4 Attitude

The tagging option **Attitude** is chosen to mark testimonies indicating the attitude (either positive or negative) of the author or the society towards the language.

The tags used are `<att>` and `</att>`. They have to be nested inside the excerpt tags. For instance:

`<att>` *Pidgin English is probably the stupidest linguistic device ever invented; it has absolutely no point. It is not easier to learn than English ('chop' is just as difficult to pronounce as 'food', 'win' as 'find', 'pickin' as 'child', etc.), it sounds silly, it has little connection with negro languages and it is terribly ambiguous. Its only possible advantage – the moral one of making negroes seem childish, by talking to them childishly – is completely overbalanced by the endless muddles it leads to.* `</att>`

(Gorer 1935: 167)

To tag a part of the text as **Attitude**:

1. Highlight the text in question
 Be aware that you will not always see the highlight. However, the function will work nonetheless (see video [Tagging Attitudes](#)). Cf. [Known issues](#).
2. Click on the button **Attitude** in the Tagging toolbar or alternatively use the hotkey **F5**
 → `<att>` will be added in front of and `</att>` at the end of the tagged data.
 The `<att>` and `</att>` tags appear within the `<ex>` and `</ex>` tags.



6.5 Description

The tagging option **Description** is chosen to mark a part of a text as language description. Examples of sections to be tagged as descriptions are such sections that describe the language in such a way that we would not gloss it, such as grammatical descriptions.

The tags used are `<lng>` and `</lng>`. They have to be nested inside the excerpt tags. For instance:

`<lng>` *One of the most difficult English idioms is the redundant negative in questions (i.e. Didn't you see him? No, I didn't) and the use of pidgin has made this always crucial point in interrogations needlessly complicated. If a negro is asked whether he has done something by two different questions he will deny apparently contradictorily: Did you do it? No. (I did not do it.) Didn't you do it? Yes. (That is correct; I did not do it.) This apparently trivial point is the cause of endless muddles, lawsuits, apparent injustices and numberless memoranda.* `</lng>`

(Gorer 1935: 167)

To tag a part of the text as **Description**:

1. Highlight the text in question
Be aware that you will not always see the highlight. However, the function will work nonetheless (see video [Tagging Descriptions](#)). Cf. [Known issues](#).
2. Click on the button **Description** in the Tagging toolbar or alternatively use the hotkey **F6**
→<lng> will be added in front of and </lng> at the end of the tagged data.
The <lng> and </lng> tags appear within the <ex> and </ex> tags.

7 Language Data

The tag **Data** serves to mark actual language data. As you can see in the following example, language data tags are represented by <d> being inserted at the beginning and </d> being inserted at the end of the language data. Text passages tagged as Data are those sections that will have to be glossed at a later point in time. Data tags have to be nested inside the excerpt tags. Furthermore, they can appear inside of <sit>, <set>, <att> and <lng> tags but it is not obligatory.

*<ex><sit> They took us by the hand, and, looking in our faces with their mellow black eyes, accompanied by what were meant for witching smiles and a peculiar singing voice, recognised us as old acquaintances, whether they had seen us before or not, in their broken English, "**<d>**My chin-chin you; me savee you lass voyage; my washe you muche good; can do all the same now; muchee good sweetmeats you savee. **</d>**"</sit></ex>*

(Jacobs 1844: 347)

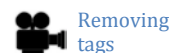
Similar as with tags for contextual and sociolinguistic annotation, in order to mark a passage in your file as language **Data**:

1. Highlight the language data
→Punctuation marks should be highlighted as well; **quotations marks should NOT be highlighted**
Be aware that you will not always see the highlight. However, the function will work nonetheless (see video [Tagging Language Data](#)). Cf. [Known issues](#).
2. Go to the Tagging toolbar and select **Data** or use the hotkey **F7**
→The tool will add the tags <d> to mark the beginning of the language data and </d> to show where it ends. Punctuation marks should be **inside**, quotation marks **outside** of the tags. The <d> and </d> tags appear within the <ex> and </ex> tags.

8 Making corrections and manual changes in DEPiCTed

8.1 Removing tags in DEPiCTed

If you notice that you have used the wrong tag or if you want to remove a tag for some other reason, you can delete your tag by using the **Remove tag** option in the Tagging toolbar.

To remove a tag:

1. Place your cursor in the start tag of the tag pair which you want to delete
→ Though you are not able to see the cursor in the main Editor, you can place the cursor wherever you want to place it
2. Select **Remove tag** from the Tagging toolbar
→ A pop-up window will appear on the screen and will ask you if you want to delete the highlighted tag
3. Click **Yes**
→ The tag will be removed

This is the procedure for all tags other than **Data**. Data tags will be removed with the help of the [Edit annotations](#) function and not with the **Remove tag** button.

8.2 Making manual changes in DEPiCTed

If you import your text file into the editor, you are not able to make changes to your text per se. This is in order to avoid that malformed mark-up is accidentally inserted. Furthermore, tagging corrections can be done without going to the text edit mode with the help of the **Remove tag** button. However, in cases where manual editing is necessary, such as, for instance, to correct a typo or to add a comment, you are able to make manual changes to your text in DEPiCTed. To do so, click on **Switch to <Text Edit Mode>** in the Tagging toolbar. You will see that you are in the Text Edit Mode when the status bar shows the following information:

Switched to manual text edit mode. Be careful.

You will then be able to edit your text. Remember to make sure that your data stays well-formed and valid (check the Status bar). In the Text Edit Mode, you can also make use of the Undo button (hotkey combination Ctrl+Z) and Redo button (hotkey combination Ctrl+Y) which you will find under **Menu bar>Edit**.

9 Comments, footnotes and general formatting

9.1 Comments

Your data might contain personal comments which you want to share and include in the database. DEPiCTed allows you to include your comments and/or explanations by using the **Comment** button from the Tagging toolbar. Comments are represented by the tags `<com>`, `</com>` in DEPiCTed.

To classify a comment in the imported text as being a comment/explanation by the contributor:



1. Highlight your personal comment or explanation
Be aware that you will not always see the highlight. However, the function will work nonetheless. Cf. [Known issues](#).
2. Choose **Comment** from the Tagging toolbar or alternatively use the hotkey **F9**
→ DEPiCTed will place `<com>` before, and `</com>` after the comment



To add a new comment into the data and classify it as a comment:

1. Switch to the Text Edit Mode by selecting **Switch to <Text Edit Mode>** from the Tagging toolbar
2. Move your cursor to the position where you want to add a new comment and type your comment
3. Click **Exit <Text Edit Mode>** after having written your comment
4. In the main Editor, highlight your new comment and select **Comment** from the Tagging toolbar or alternatively use the hotkey **F9**
→DEPiCTed will place <com> before, and </com> after the comment

9.2 Footnotes

Footnotes cannot be placed in the footer but need to appear in the running text. To nonetheless be able to differentiate between the running text and the footnote, footnotes need to be tagged as such with the help of the button **Footnote** in the Tagging toolbar.

Please keep in mind to differentiate between footnotes and comments. Comments, explanations, ideas etc. that *you* insert into the text need to be tagged as **Comment**. However, **Footnote** refers to footnotes that appear in the original source texts.

To classify text as a footnote:

1. Highlight the text that you want to classify as a footnote
Be aware that you will not always see the highlight. However, the function will work nonetheless. Cf. [Known issues](#).
2. Click on **Footnote** in the Tagging toolbar or alternatively use the hotkey **F10**
→DEPiCTed will place <fn> before and </fn> after the footnote



9.3 General formatting: Page break and line break

Formatting in DEPiCTed does not work as in a word processor like MSWord or OpenOffice. All texts are basically stored as if they were one long line of plain text. Thus, even though you might have paragraphs or a poem displayed in separate lines in your original text file, this formatting will not be saved in DEPiCTed, unless the appropriate tags are inserted.

In principle, formatting should be kept to a minimum in the tool as the main focus is on the content of the texts. Nonetheless, DEPiCTed allows the insertion of page breaks and line breaks. The former is necessary especially in cases in which an excerpt extends beyond one page. The latter can be helpful for texts like letters, poems and songs which conventionally have a specific structure and where it might be helpful to keep the structure in order to get the meaning out of it.

Page breaks are represented by the tag <pb/> and line breaks by the tag <lb/>. Remember that these two tags do not need both start- and end- tags, but are valid on their own in the form <pb/> and <lb/>.

To insert a page break or line break:

1. Move your cursor to the position where you want to insert the page- or line break and make a click with the mouse
2. To insert a page break select **Pagebreak** from the Tagging toolbar or alternatively use the hotkey **F11**



- The tool will add `<pb/>` at the position at which the page break should be inserted
- 3. To insert a line break select **Linebreak** from the Tagging toolbar or alternatively use the hotkey **F12**
 - The tool will add `<lb/>` at the position at which the line break should be inserted



10 Editing annotations

The tab **Edit annotations** allows you to review and delete data tags, as well as to group data elements.

10.1 Grouping data elements

Not every new data tag (`<d>`) is necessarily a new utterance and sometimes a sentence is split by a comment. Imagine the following example:

```

small artificially made passage. This I did, and on
reaching it, sat down amongst them.
<ex p="136"><sit>Many women were now coming ashore at
another landing some distance away, and one of
the men pointed in their direction, saying:
"Tambu, tambu." I understood. them to mean
that it would be wrong for me to attempt to go
where the women were, and nodded assent. Some
of the men could speak a little bêche-de-mer, but
one was very fluent, and said that he had been to
Queensland. He asked if I had previously been to
"Maratta", and I told him that I had not.
"<d>Maratta not all-'e-same 'nother island,</d>" he
explained; "<d>you must look out</d> (take care)."
"<d>S'pose you try for take'm mary, man Maratt
must kill you,</d>" he added brightly.</sit></ex>
Very cheery information with which to greet a

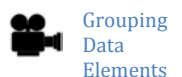
```

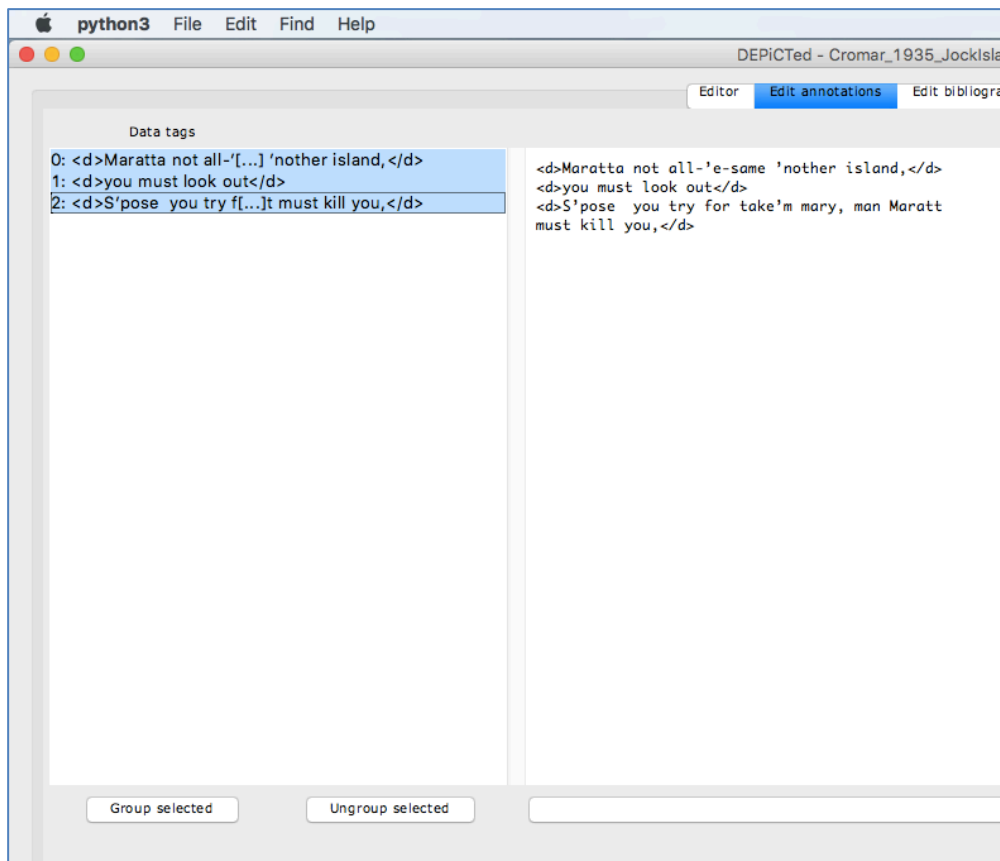
(Cromar 1935: 136)

In such a case, two or more data tags belong together. In the tab **Edit annotations** you have the opportunity to group those elements together.

In order to group two or more data tags together:

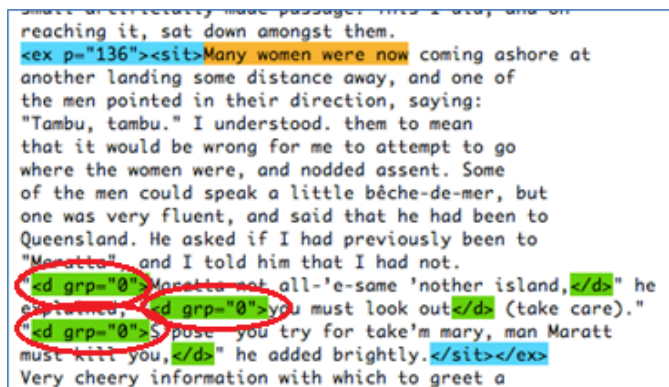
1. Go to the tab **Edit annotations**
 - On the left side of the screen, in the window titled Data tags, all language data which you have tagged will be listed
2. Use **SHIFT/CTRL+Click** to select at least two items from the Data tags list





3. Click on **Group selected**

→ The groups will be indicated in the text by **grp=""** appearing in the data tag bracket after the d



(Cromar 1935: 136)

Similarly, you can ungroup your data:

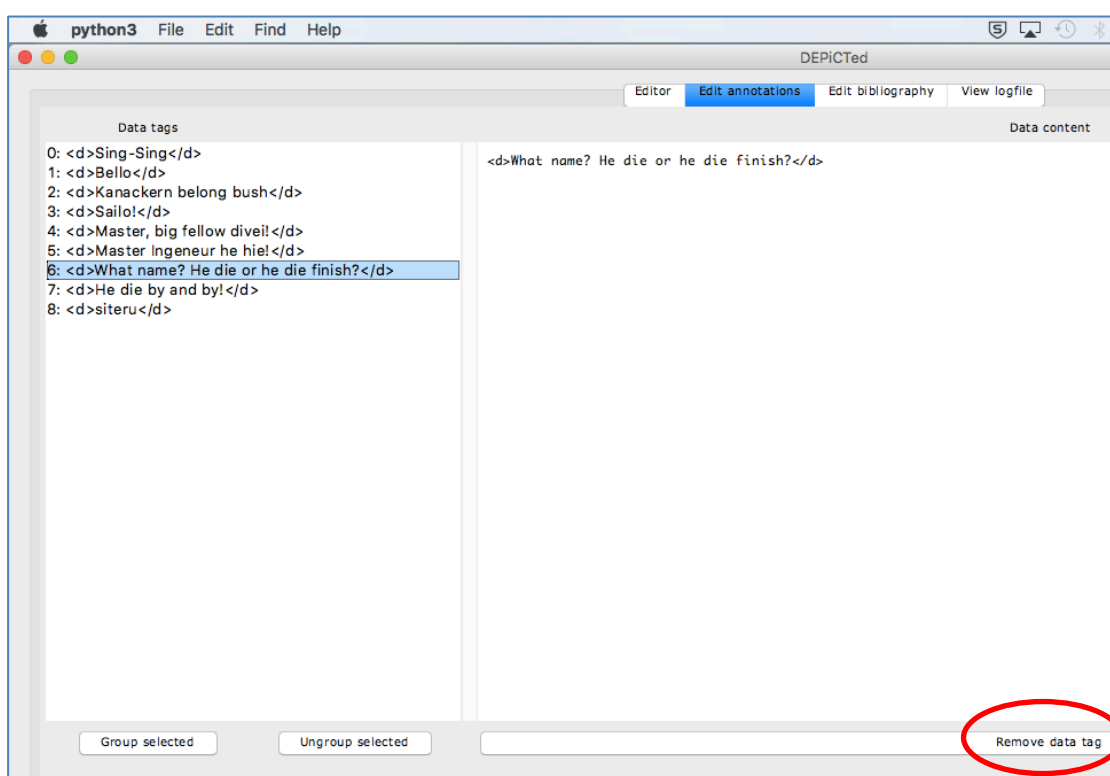
1. Go to the tab **Edit annotations**
→ On the left side of the screen, in the window titled Data tags, all language data which you have tagged will be listed
2. Use **SHIFT/CTRL+Click** to select at least two items from the Data tags list
3. Click on **Ungroup selected**

10.2 Viewing and deleting Data tags

The **Edit annotations** tab is also helpful if you want to review or delete your data tags at a later point in time.

To view and delete data tags:

1. Go to the tab **Edit annotations**
→ On the left side of the screen, in the window titled **Data tags**, all language data which you have tagged will be listed
2. Select a Data element and click on it
→ The chosen element will appear in the **Data content** window on the right side of the screen
3. Click on **Remove data tag** to delete the data tag
→ The data tags will be removed and will no longer appear in the editor



11 Entering bibliographical information in DEPiCTed

The tab **Edit bibliography** in the Editing toolbar serves to enter all kind of bibliographical information regarding the text you imported into the editor.

As the figure shows, the first four fields refer to the author of the text and serve to enter information such as the author's name, his/her sex, his/her ethnicity and his/her role (explorer, plantation owner, official, officer, etc.).

Following that are two fields which serve to enter the time (period) and location(s) the source refers to, for instance, time: *1908-1914*; location: *Bismarck Archipelago*.

The fields *Year of publication*, *Full source reference*, *Source type* and *Genre* will contribute to have an adequate reference to the source included in the database.

Please make sure to give the full reference of the source!

Under *Source type* you can select from a given drop-down list. The options to choose from are explained in the following table:

Source Type	Explanation	Example
Archive material	Documents which no longer have an everyday use but are stored due to their historical value in an archive should be assigned as Archive material . Note that the reference code for these items must be cited.	Letter from African House, 31 October 1721; PRO T/70/60: 32. Royal African Company 1678-1681. African Company Letterbook. Pro T 70/1. Letters received from Africa and West Indies.
Collected work	A collection of the work (articles, other written pieces) of a single author should be classified as Collected work .	Gandhi, Mahatma Karamchand. 1958. <i>The Collected Works of Mahatma Gandhi</i> . 1st ed. Delhi: Publications Division, Ministry of Information and Broadcasting, Government of India. Volume 1 (1884-1896).
Conference proceedings	A collection of works from a scholarly conference should be assigned as Conference proceedings .	Alders, Robyn Gwen ; Spadbrow, Peter B., and Young, M. P. (Eds.), <i>Village chickens, poverty alleviation and the sustainable control of Newcastle disease: Proceedings of an international conference held in Dar es Salaam, Tanzania, 5-7 October 2005</i> . Canberra, ACT: Australian Centre for International Agricultural Research.

Contribution in edited volume	A single specific article or chapter which was published in an edited book should be characterized as Contribution in edited volume .	Landtman, Gunnar 1920. Religious beliefs of the Kiwai-speaking Papuans. In Beaver, Wilfred N. (ed). <i>Unexplored New Guinea. A record of the travels, adventures, experiences of a resident magistrate amongst the head-hunting savages and cannibals of the unexplored interior of New Guinea</i> . London: Seeley, Service & Co. Ltd., pp. 301-321.
Edited volume	A collection of articles or chapters written by different authors and collected in one book should be classified as Edited volume .	Beaver, Wilfried N. (ed.) 1920. <i>Unexplored New Guinea. A record of the travels, adventures, and experiences of a resident magistrate amongst the head-hunting savages and cannibals of the unexplored interior of New Guinea</i> . London: Seeley, Service & Co. Ltd.
Interview material	An unpublished documentation of an interview should be characterized as Interview material .	Huber, Magnus. 2015. Early Pidgin and Creole Languages. Interview with Viveka Velupillai, 10/10/2015.
Journal article	An article which appeared in an academic or scholarly journal or in another type of periodical should be classified as Journal article .	Hocart, Arthur Maurice. 1937. Fishing in Eddystone Island. <i>Journal of the Royal Anthropological Institute</i> . 67. 33-41.
Lecture	A handout or notes from an unpublished lecture or talk should be classified as Lecture .	Clark, Ross. 1978. The rise and fall of New Zealand Pidgin. Paper presented at the second New Zealand Linguistics Conference, August 1978. New Zealand.
Manuscript	A written piece (monograph-length or article-length) which has not been published should be characterized as Manuscript .	Kelsey, Dudley E. 1938. An old man's legacy: A story of the Far North Australian Bush by Larrabah. Unpublished Manuscript. Mortlock.
Monograph	A single complete work written by a single author or a team of authors should be assigned as Monograph .	Zöller, Hugo. 1891. <i>Deutsch-Neuguinea und meine Ersteigung des Finisterre-Gebirges. Eine Schilderung des ersten erfolgreichen Vordringens zu den Hochgebirgen Inner-Neuguineas</i> . Stuttgart: Union Deutsche Verlagsgesellschaft.
Newspaper article	An article which appeared in a daily or weekly newspaper should be classified as Newspaper article .	Harwings, Joseph. Bakvisch en generaal. In <i>De Indische Courant</i> . 1932-29-04.
Personal communication	Unpublished notes from conversations, telephone conversations, letters and emails resulting out of personal communication should be classified as Personal communication .	Parkvall, Mikael. 2015. Français Tirailleur to Viveka Velupillai, 05/20/2015.

Radio or TV broadcast	The textualisation of a program broadcast on TV or radio should be assigned as Radio or TV broadcast .	Seke, Sam. 2015. PNG Polis na Difens joint-investigesen. ABC Radio Australia, 11/30/2015. Available online at http://www.radioaustralia.net.au/tokpisin/2015-11-30/png-polis-na-difens-jointinvestigesen/1520336 .
Special issue	An issue of a periodical which was published in addition to the normal regular issues should be classified as Special issue .	Buchner, Max. 1886. Kamerun-Englisch. <i>Beilage zur Allgemeinen Zeitung</i> . 16.November 1886.
Unknown source	Documents of which only incomplete bibliographic information is available and which were not collected in an archive should be assigned as Unknown source .	Aborigine Department. NA.: Correspondence received. South Australia.

In the field *Contributor* you are asked to enter your name as the contributor of the source.

12 Known issues

The tool is at this stage only possible to run via Python: We apologize for this inconvenience and are working as fast as we can to provide a proper .app for the tool.

Error messages in the Python Terminal and/or Shell when opening the tool: It might be that you receive some error messages in the Python Terminal or in the Python Shell when you open DEPiCTed. You can simply ignore these error messages if DEPiCTed runs on your Mac, as they are based on layout issues and will not prevent you from annotating your data. We are aware of this issue and are working to find a solution for it.

While you are highlighting a text passage, you are not able to see the highlight: Only after selecting the Tag button from the Tagging bar, which you want to insert, the appropriate tags will appear.

Sometimes you are able to write in the Editor window even though you are not in the Text Edit Mode: Please make sure that you nonetheless only edit your text in the Text Edit Mode.

To find an updated list of the **Known issues** click [here](#).

13 Error reports

In order to continuously improve the DEPiCTed tool, we are collecting error reports. If you have any problems or suggestions for improvements, please contact the DEPiCT team (depict@anglistik-uni-giessen.de).

14 Sources

- Behrmann, Walter (1922): *Im Stromgebiet des Sepik. Eine deutsche Forschungsreise in Neuguinea*. Berlin: August Scherl.
- Cromar, Jack (1935): *Jock of the islands: Early days in the South Seas*. London: Faber & Faber.
- Gorer, Geoffrey (1935): *Africa Dances: a book about West African Negroes*. New York: Alfred A. Knopf. First Edition.
- Jacobs, Thomas Jefferson (1844): *Scenes, incidents, and adventures in the Pacific Ocean; or, the islands of the Australasian seas, during the cruise of the clipper Margaret Oakley under Captain Benjamin Morrell*. New York: Harper & Brothers.
- Newland, H. Osman (1916): *Sierra Leone: its people, products and, secret societies. A journey by canoe, rail, and hammock, through a land of kernels, coconuts, and cacao, with instructions for planting and development*. London: J. Bale, Sons & Daniellson.
- Ribbe, Carl (1903): *Zwei Jahre unter den Kannibalen der Salomo-Inseln. Reiseerlebnisse und Schilderungen von Land und Leuten*. Dresden-Blasewitz: Hermann Beyer.
- Thomas, Chas. W. (1860): *Adventures and observations on the West Coast of Africa, and its islands. Historical and descriptive sketches of Madeira, Canary, Biafra, and Cape Verd islands; their climates, inhabitants, and productions. Accounts of places, peoples, customs, trade, missionary operations, etc., etc., on that part of the African coast lying between Tangier, Morocco, and Benguela*. New York: Derby & Jackson.
- Vogel, Hans (1911): *Eine Forschungsreise im Bismarckarchipel*. Hamburg: Friederichsen.