CECL Papers 6

The UCLouvain Error Editor User guide - Version 2.0

Sylviane Granger, Helen Swallow and Jennifer Thewissen



2023

CECL Papers 6

The UCLouvain Error Editor User guide - Version 2.0

Sylviane Granger, Helen Swallow and Jennifer Thewissen



2023

The UCLouvain Error Editor User guide - Version 2.0 – 2023

Sylviane Granger, Helen Swallow & Jennifer Thewissen

Centre for English Corpus Linguistics, UCLouvain

TABLE OF CONTENTS

Table of	f Cor	ntents	. 3
Table of	f Fig	ures	. 4
Introduc	ction		. 5
1.1.	Inst	allation	. 5
1.2.	Tag	set selection	. 5
1.3.	Mo	dification of tagsets	. 7
1.4.	Add	ling new tagsets	. 8
2. Me	nu B	Bar	. 9
2.1.	File		. 9
2.2.	Sett	tings	10
2.2	.1.	"Tags and colors"	10
2.2	.2.	No correction	10
2.3.	Edi	t	11
2.4.	Tag	s (Error tagging)	11
2.5.	Che	eck	13
2.5	.1.	"Check all"	14
2.5	.2.	Review specific tags ("Check only")	16
2.5	.3.	Nested error tags	17
2.6.	Ana	alyse	17
2.7.	Sea	rch	24
2.8.	Exe	rcises	24
2.9	Hel	p	27
2.9	.1.	The Louvain Error Tagging Manual (2.0)	27
2.9	.2.	"UCLEE Helpfile"	27
Appe	ndix	: Working with texts previously tagged with uclee-en-1.3	28

TABLE OF FIGURES

Figure 1: Tagset selection	6
Figure 2: Main interface	6
Figure 3: The hierarchical structure of a UCLEE-conformant tagset	8
Figure 4: Open several files (Merge) option	9
Figure 5: Settings	10
Figure 6: "Edit" menu functionalities	11
Figure 7: Highlighting erroneous words/strings of words in the text	12
Figure 8: Error code format	12
Figure 9: Error tagging missing and superfluous words	13
Figure 10: Error tag with no correction	13
Figure 11: "Check" menu	14
Figure 12: Error tagging in the "Check all" display	14
Figure 13: "Update tag" option	15
Figure 14: "Only update correction" option	15
Figure 15: Review specific tags only	16
Figure 16: "Check only" option after the selection of a specific tag	16
Figure 17: Nested error tags	17
Figure 18: "Analyse" menu	18
Figure 19: Overview of the UCLEE concordance tool	18
Figure 20: Exporting the concordance lines	20
Figure 21: Absolute frequencies in the concordance tool	21
Figure 22: Detailed percentages of errors sorted by tag	21
Figure 23: Detailed percentages of errors sorted by file	22
Figure 24: Relative frequencies of error tags	23
Figure 25: Exporting frequencies	23
Figure 26: Search tab	24
Figure 27: "Exercises" tab	25
Figure 28: Creating exercises based on multiple tags	25
Figure 29: Exporting tailor-made exercises	26
Figure 30: Illustration of exercise formats	27

<u>APPENDIX</u>

Figure A 1: Tagset selection	. 28
Figure A 2: Main interface	. 29
Figure A 3: Convert files from 1.3. tagset with double dollars	. 30
Figure A 4: Select files to convert	. 31
Figure A 5: Conversion completed	. 31
Figure A 6: File in 1.3 tagset with double dollars before conversion	. 32
Figure A 7: Dialog box after conversion	. 32

INTRODUCTION

The present document is the second version of the UCLouvain Error Editor (UCLEE) user guide. The software, which was developed at the Centre for English Corpus Linguistics (Université catholique de Louvain, Belgium), can be used with both raw and previously error-tagged texts and speeds up the insertion of error codes and corrections in learner texts. It also contains a concordancing function and an exercise generator.

1.1. Installation

Run the UCLEE.exe file in Windows, the UCLEE.dmg file in MacOS or the UCLEE file in Linux. This will open the "Tagset selection" window described in the following section.

1.2. Tagset selection

By default this version uses the Louvain Error Tagging Manual Version 2.0 (Granger, Swallow & Thewissen 2022^{1}). It also contains a built-in converter which allows users to work on texts tagged with the previous version of UCLEE (version 1.3^{2}). Alternatively, users can design their own tagset and use it to tag their data.

The "Tagset selection" window provides three tagset options (Figure 1):

- 1. Select uclee-en-2.0.tag if you wish to work with the new error typology described in Granger, Swallow & Thewissen (2022);
- 2. Select uclee-en-1.3.tag if you want to open texts previously tagged with the 1.3 version of the error annotation system;
- 3. Upload a new tagset: user-designed tagset.

Further guidance on Option 2 is given in **Appendix 1**.

¹ Granger, S., Swallow, H. & Thewissen, J. (2022). *Error tagging manual. Version 2.0.* CECL Papers 4. Louvain-la-Neuve: Centre for English Corpus Linguistic/Université catholique de Louvain. The manual is open access and can be downloaded from <u>https://cdn.uclouvain.be/groups/cms-editors-cecl/Granger%20et%20al._Error%20tagging%20manual_v2.0_2022.pdf</u>

² Dagneaux, E., Denness, S., Granger, S., Meunier, F., Neff, J., Thewissen, J. (2008). *The Louvain Error Tagging Manual Version 1.3.* Centre for English Corpus Linguistics, Université catholique de Louvain, Louvain-la-Neuve

It is very important that you read and implement the guidance set out in Appendix 1 before starting to work on texts previously error-tagged with version 1.3.

Whichever tagset you choose to use, the tagged files will automatically be given an ERA file extension (e.g. ICLEFR001.era).

A LICIEE Tracet colortion			
Source - lagset selection			
Which tagset should we use for this session?			
uclee-en-2.0.tag			
uclee-en-1.3.tag			
uclee-en-2.0.tag			
upload a new tagset			

Figure 1: Tagset selection

Once you have selected a tagset, simply click on the "Open UCLEE" button at bottom right of the window. This will redirect you to the software's main interface (see Figure 2).

Note: If you want to select a different tagset, you need to close and reopen UCLEE.



Figure 2: Main interface

1.3. Modification of tagsets

Existing tagsets can be modified, i.e. researchers can choose to add new tags, delete unwanted tags or modify existing tags to create a tagset better suited to their research purposes.

Tagsets can be accessed by clicking on the "Show tagsets" button in the opening "Tagset selection" box, or in Settings, Tags and Colors. To modify a tagset, open the corresponding *.tag file (e.g. uclee-en-2.0.tag) with a text editor such as Notepad or Notepad++ or any default text editor, and update the file in line with your preferences, paying particular attention to file structure. Save the new tagset with a new name so as to be able to return to the original tagset if you want to. As shown in <u>Figure 3</u>, the hierarchical structure of the tagset (e.g. the fact that the two error tags FS and FM belong to the broader category denoted by F) is represented by [] and **indents**:

- A broad category is enclosed in [] and headed with an error tag that directly follows [.
- Indents indicate sub-categories.
- There should be only one tag per line.

Sub-categories can be named as desired and do not have to start the same way as their parent. Be careful not to use the same tag name twice, even in different categories.

Warning! In order to apply changes and use the modified tagset, you must restart the software.



Figure 3: The hierarchical structure of a UCLEE-conformant tagset

1.4. Adding new tagsets

You can also add new tagsets to UCLEE. To create a new tagset, open a new text file and structure your tagset with [] and indents as explained above. Give your new file a name and save it in *.tag format (manually change .txt to tag when you save the file) in UCLEE's *resources/tagsets* repertory. All tagsets stored in this repertory will be made available to you when the program is launched.

Another way of creating a new tagset is to use one of the existing tagsets as a template, having saved it under a new name before introducing your changes; this means that you can use the existing structure as a basis rather than creating the new one from scratch.

To access the existing tagsets in order to use it as a template, go to UCLEE's *resources/tagsets* repertory (accessible by clicking on the "Show tagsets" button in the opening "Tagset selection" box).

The new tagset you have created will appear in the drop-down menu in the opening "Tagset selection" window once you closed and reopened the UCLEE software.

2. MENU BAR

2.1. File

Once redirected to the main interface, you can upload two types of text: either a raw (.txt) or a previously tagged (.era) text file by selecting the "Open" option in the "File" menu.

File	Settings	Edit	Tags	Check	Analyse
	New windo	w			Ctrl+N
	Open				Ctrl+O
	Open sever	al files	(Merg	e) Ctrl-	+Shift+O
	Save				Ctrl+S
	Save as			Ctrl	+Shift+S
	Print to PD	F			
	Export			Ctrl	+Shift+E
	Close				Alt+F4
	Quit				

Figure 4: Open several files (Merge) option

Multiple files can be uploaded by clicking on the "Open several files (Merge)" option (Figure 4). This will merge the uploaded documents into one. Further files can only be added subsequently by performing a new upload.

Note: If you do *not* want to merge your files, select the "New window" option; this will open a new workspace in which you can work on texts separately.

Note. *.txt files should preferably be encoded in UTF-8 format.

2.2. Settings



Figure 5: Settings

2.2.1. "Tags and colors"

This submenu allows you to view your current tagset and customise the colours you are working with. Clicking on this option opens a box with two tabs: "Tags" and "Colors".

- The "Tags" tab contains the tag hierarchy. This is the list that is used to generate the menu-driven tag insertion function; it lists the tags that will be recognised by the parser. (For instructions on modifying tagsets or creating new ones, see <u>1.3</u> <u>Modification of tagsets.</u>)
- The "Colors" tab specifies the colours used by the tag-highlighting system in the "Check" and "Exercises" modes. Each line must be in the "T: RRGGBB" format, in which T corresponds to the tag that will be coloured in this way. **RRGGBB** corresponds to a certain colour in the RGB hexadecimal format (see http://www.javascripter.net/faq/colornam.htm for examples) The line enclosed by "_" defines the default colour for the tagset.

2.2.2. No correction

If you don't want to insert any corrections in the texts you are error-tagging, you can apply the "No correction" setting (see <u>Figure 5</u>). This will result in the following tag format: 'it's my great <FS>presure</FS> to share my college life with you.'

Do not use this setting if you only wish to omit corrections in the case of certain specific errors (see Section 2.4).

2.3. Edit

Clicking on "Edit" opens the editing menu – Copy/Cut/Paste, Select All, etc. (Figure 6). Each function has a corresponding keyboard shortcut that is displayed next to it, e.g. <CTRL+F> (or <cmd #+F> on Mac) activates the Find function, which allows you to find specific words in the text.

File	Settings	Edit Tags Check	Analyse Searc	h Exercises Help
1	<file< th=""><th>Select All</th><th>Ctrl+A</th><th>_UCLEE.txt" tagset="uclee-en-2.0.tag"></th></file<>	Select All	Ctrl+A	_UCLEE.txt" tagset="uclee-en-2.0.tag">
2	For o	Undo	Ctrl+Z	ve always struggled to improve <fm corr="their</th></tr><tr><th></th><th>knowle</th><th>Redo</th><th>Ctrl+Shift+Z</th><th>FM> of the universe as a whole.</th></tr><tr><th>3</th><th>Let u:</th><th></th><th></th><th>strip cartoons: when Hergé <GVT</th></tr><tr><th></th><td>corr=</td><td>Сору</td><td>Ctrl+C</td><td>> Tintin's first steps on the moon, who can</td></tr><tr><th></th><td>predi</td><td>Cut</td><td>Ctrl+X</td><td>ike experience will actually take place some</td></tr><tr><th></th><th>years</th><th>Paste</th><th>Ctrl+V</th><th></th></tr><tr><th>4</th><th>These</th><th>Find</th><th>Ctrl+F</th><th>te that creative imagination can be closely</th></tr><tr><th></th><th>linke</th><th>Find next</th><th>Ctrl+G</th><th>tific">scientifice</fm> progress - in a direct or
	indire	Find previous	Ctrl+Shift+G	
5	Write	Replace	Ctrl+Shift+F	ts and other artists have more than once <xvco< th=""></xvco<>
	corr=	Deplace	CLI CLID D	ng">contributed to break fresh <xnuc< th=""></xnuc<>
	corr=	Keplace all	Ctri+Snift+R	> in the study of science and technology.
6	More a	and more peo	ople compla	ain about the consequences of automation: if it

Figure 6: "Edit" menu functionalities

2.4. Tags (Error tagging)

This functionality allows you to insert, update or remove error tags in your texts. Proceed as follows to tag detected errors (Figure 7):

- 1. Highlight the erroneous word/word string in the learner text;
- 2. Insert the correction(s) in the box situated at the bottom of the window;
- 3. Select the tag that applies to the error. This can be done in one of two ways:
 - a. Go to the 'Tag' menu at the top of the interface and select the appropriate tag or
 - b. Click on the 'Add' button at the bottom right of the dialogue box.

The error taxonomy is described in the Louvain Error Tagging Manual, which can be found by clicking on "Help" in the menu bar (cf. <u>Section 2.9</u>).



Figure 7: Highlighting erroneous words/strings of words in the text

In the example above (Figure 7), the concord error *was* is corrected to *were* and tagged as GVN (Grammar Verb Number). The tag assigned by the user appears on the screen with the markup illustrated in Figure 8.



Figure 8: Error code format

In some situations you may need to insert a tag to mark a missing word. To do this, place the cursor at the omission point, enter the missing word(s) in the Correction field and insert the appropriate tag (cf. first sentence in Figure 9).

Conversely, to mark a word or words as superfluous, select it/them, enter 0 (backslash followed by 0) in the Correction field and insert the appropriate tag (cf. second sentence in Figure 9).

S UCLEE	DUCLEE – 🗆 🗙					
File Settings Edit Tags Check Analyse Search Exercises Help						
<pre>1 <file name="" tagset="uclee-en-2.0.tag"></file></pre>						
$\stackrel{3}{}$ We might need some potatoes <ql corr="for">\0</ql> tonight. 4						
5 I love to drive <lcc corr="\0">the</lcc> cars ! 6						
7						
	_					
Correction		Add				
	Uş	odate				
	Re	move				

Figure 9: Error tagging missing and superfluous words

If you only want to tag an error without providing a correction, you should select the appropriate tag and leave the correction box empty. This will ensure that the attribute corr="" is automatically inserted beside the error tag (see Figure 10).

Note: If you don't want to add any corrections at all in the texts you are error-tagging, select the 'no correction' option in Settings (see <u>Section 2.2.2</u>).

2	UCLEE	_	\times
File	Settings Edit Tags Check Analyse Search Exercises Help		
1	<file name="" tagset="uclee-en-2.0.tag"></file>		
2	There <gvn corr="">was</gvn> people everywhere.		
3			

Figure 10: Error tag with no correction

2.5. Check

This menu contains two functionalities: "Check all" and "Check only" (Figure 11). The "Check all" option gives a broad overview of the corrected error(s); it can be used to check whether you have been consistent in your error tagging system throughout a given text. "Check only" allows you to review a limited number of tags by selecting them from a scroll-down list.

1	UCLEE	_		\times
File	Settings Edit Tags Check Analyse Search Exercises Help			
1	<file all="" check="" ctrl+enter="" name="" tagset="uclee-en-2.0.tag"></file>			
2	Since the mos Checkonly of progress <gvaux corr="can">i:</gvaux>	s to<	/GVAU	JX>
	be found in science and technology, many people tend to blame the	ese a	reas	of
	knowledge for the loss of imagination and its correlated dreams.			
3				

Figure 11: "Check" menu

2.5.1. "Check all"

You can review your corrections at any time using the "Check all" function in the "Check" menu (Figure 12). When you click on "Check all", all the tags used and corrections made are displayed in colour-coded boxes (one colour per error category). To return to the raw text view, deselect the "Check all" option in the "Check" menu.

1	<file name="FR1 - Copie.txt" tagset="uclee-en-1.3.tag"></file>	-
2	The urgent need for imagination LS: During \rightarrow in the last decade of the	
	twentieth century, GPP: it $\rightarrow 0$ is LS: easy \rightarrow perfect to exemplify how	
	science, industrialisation and technology have pervaded our everyday life.	
3	Some will claim that the increasing number of factories has spoiled the	
	GNC: landscape's IS: attraction \rightarrow beauty \rightarrow beauty of the landscape	
	Others LS: assert -> believe that imagination and dreams have definitely	
	vanished from our modern society.	
4	In order to avoid any narrow-mindedness, it would be more appropriate to	
	tackle the problem of technology versus imagination within a larger	
	framework.	
5	For obvious reasons, men have always struggled to improve their knowledge	
	of the universe as a whole. Since the most striking evidence of progress	
	GVAUX: is to -> can be found in science and technology, many people tend	
	to blame these areas of knowledge for the loss of imagination and $FS:$ it	
	socrrelated \rightarrow its correlated dreams.	
6	However, Z: it very often occurs that imagination -> imagination often	
	operates as a kind of catalyst for major scientific or technological	-
Corr	rection	
	New Tag	
	Lindata Tan	
	Opdate lag	
	Domous Tog	
	Remove rag	

Figure 12: Error tagging in the "Check all" display

To change the error tag on a given word/string of words (for example, in Figure 13 Figure 13 the user wants to replace the GVM tag with GVAUX), click on the coloured box you want to update and then on the "Update tag" button. You can now select a new error tag.



Figure 13: "Update tag" option

To update a correction while keeping its tag, click on the box with the correction you want to change, click on Update at bottom right, and select the "Only update correction" option that you will see above the list of tags. (Figure 14). Insert the new correction in the correction box and click on "enter". This allows you to replace the correction you previously entered (*scientifical*) with a new one (*scientific*).

File	Settings Edit Tags Check Analyse Search Exercises Help	
1	<file name="test helen_text_UCLEE.txt" tagset="uclee-en-2.0.tag"></file>	A
2	For obvious reasons, men have always struggled to improve FM:	
	theirknowledge $ ightarrow$ their knowledge of the universe as a whole.	
3	Let us take an example from strip cartoons: when Hergé GVT: sketches	→
	sketch Tintin's first steps on the moon, who can predict that such a	dream-
	like experience will actually take place some years later?	
4	These two examples illustrate that creative imagination can be closely	У
	linked with FM: scientifice \rightarrow scientifical progress - in a direct of	or
	indirect way.	
5	Writers, painters, architects and other artists have more than once	KVCO:
	contributed to break \rightarrow contributed to breaking fresh XNUC: grounds	→
	ground in the study of science and technology.	Only update correction
6	More and more people complain about the consequences of automation:	only apaate concetion
	GVAUX: does lighten \rightarrow lightens the burden of LSADJ: fastidious \rightarrow	F 🕨
	tiresome tasks in daily life, it may also lead to a paralysis of ou	G 🕨
	creativeness \rightarrow creativity .	L 🕨
7	LCLS: Indeed, $\rightarrow 0$ in the twentieth century, fears for the future	Q 🕨
coio	mankind have predominated over the optimism about inevitable progres	W 🕨
3010	Ac	X 🕨
		Z
	Upda	te
	Remo	ve

Figure 14: "Only update correction" option

To remove a tag and the corresponding correction, select the relevant box and click on the "Remove" button at bottom right.

2.5.2. Review specific tags ("Check only")

To review only certain tags, select them in the scroll-down list in the "Check only" submenu. Multiple tags can be reviewed at the same time by holding the $\langle CTRL \rangle$ (or $\langle cmd$ $\mathbb{H} \rangle$ in Mac) key down while (de)selecting the relevant tags (Figure 15).

Check only	>
Select tags you want to highlight (keep CTRL down and drag the mouse to (un)select several tags)	Select All
GPO GPP GPR GPU	•
GVAUX GVM GVN GVNF	
GVT GVV GWC LCC	
	Close Highlight

Figure 15: Review specific tags only

Once the specific tag(s) have been selected (e.g. LS, see Figure 16), the main window will foreground the selected tag(s), while greying out the others.



Figure 16: "Check only" option after the selection of a specific tag

2.5.3. Nested error tags

To correct more than one error in the same string of words, repeat the steps detailed above, following the order of your own thought processes. For example, let us imagine that you want to correct the erroneous string of words "landscape's attraction" (see Figure 19). You first notice that "beauty" would be a better choice than "attraction" and therefore highlight the word "attraction", adding a new tag (in this case LS), and the correction "beauty" by following the steps described in Section 2.4.

Then you realise that "beauty of the landscape" would be a better formulation than "landscape's beauty". Select the same string of words (so 'landscape's beauty') (as if there were no previous correction), add the correction in the correction box and select a new tag (here GNC). This will produce the result shown in <u>Figure 17</u>, where the red tag (LS) is nested inside the orange tag (GNC).

2	The urgent need for imagination During the last decade of the twentieth
	century, it is easy to exemplify how science, industrialisation and
	technology have pervaded our everyday life.
3	Some will claim that the increasing number of factories has spoiled the
	GNC: landscape's LS: attraction \rightarrow beauty \rightarrow beauty of the landscape
	Others assert that imagination and dreams have definitely vanished from our
	modern society.

Figure 17: Nested error tags

2.6. Analyse

The built-in concordancer in the "Analyse" tool allows you to count, sort and analyse errors, together with their corrections. It also offers the possibility of drawing up an error profile for specific users or user groups. To open the concordancer, click on the "Analyse" menu (Figure 18).

Note: The "Analyse" function only works with error-tagged texts (i.e. files with the .era extension).

ا 🕑	JCLEE	_			×
File	Settings Edit Tags Check Analyse Search Exercises Help				
1	<file name="FR1.txt" tagset="uclee-en-1.3.tag"></file>				-
2	The urgent need for imagination <ls corr="in">During</ls> t	he last	dec	ade	of
	the twentieth century, <gpp corr="0">it</gpp> is <ls< td=""><td></td><td></td><td></td><td></td></ls<>				
	corr="perfect">easy to exemplify how science, industriation of the science industriatin of the science industriation of the	alisati	on a	nd	
	technology have pervaded our everyday <gnn corr="lives">lif</gnn>	e	•		
3	Some will claim that the increasing number of factories has	spoile	d th	e	
	landscape's <ls corr="beauty">attraction</ls> <gnc corr="th</td><td>e beaut</td><td>y of</td><td>the</td><td></td></tr><tr><td></td><td>landscape">landscape's attraction</gnc> . Others <ls< td=""><td></td><td></td><td></td><td></td></ls<>				
	corr="believe">assert that imagination and dreams $<\!\!\text{GVN}$				
	corr="">have definitely vanished from our modern soci	ety.			
4	In order to avoid any narrow-mindedness, it would be more a	ppropria	ate	to	
	tackle the problem of <ls corr="correction 1"><lp <="" corr="correction 1" p=""></lp></lp></lp></lp></lp></lp></lp></lp></lp></lp></lp></lp></lp></lp></lp></lp></lp></lp></lp></lp></lp></lp></lp></lp></lp></lp></lp></lp></lp></lp></lp></lp></lp></ls>	rection			
	2">technology versus imagination within a larger $% \lambda = 0 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 -$	framewo	rk.		
5	For obvious reasons, men have always struggled to improve t	heir kno	owle	dge	of
	the universe as a whole. Since the most striking evidence o	f progre	ess	<gva< td=""><td>UX</td></gva<>	UX
	corr="can">is to be found in science and technology $% \mathcal{A} = \mathcal{A} = \mathcal{A} = \mathcal{A}$, many p	peop	le	
	tend to blame these areas of knowledge for the loss of imag	ination	and	(FS)
	it socrrelated \$its correlated\$ dreams.				-
Corr	ection		Add		

Figure 18: "Analyse" menu

Clicking on "Analyse" opens a separate window with a concordance tool (see Figure 19).

UCLI	EE - Concord	and the second	1		
File					
	Files		22		Search
File Nar	ne	#Tokens	#Tags	Tag	
C:\Use	rs\Adrien\Documents\Travaux\ROGE 2 MA\CECL\UCLEE\UCLEE - fichiers\FR2 - Copie.era	908	93 ^	Tag	
C:\Use	rs\Adrien\Documents\Travaux\ROGE 2 MA\CECL\UCLEE\UCLEE - fichiers\FR3 - Copie.era	777	32	Error	Match outside tags
C:\Use	rs\Adrien\Documents\Travaux\ROGE 2 MA\CECL\UCLEE\UCLEE - fichiers\FR4 - Copie.era	690	35	Correction	
C:\Use	rs\Adrien\Documents\Travaux\ROGE 2 MA\CECL\UCLEE\UCLEE - fichiers\FR5 - Copie.era	739	50		
C:\Use	rs\Adrien\Documents\Travaux\ROGE 2 MA\CECL\UCLEE\UCLEE - fichiers\FR6 - Copie.era	699	62	Sort by	
C:\Use	rs\Adrien\Documents\Travaux\ROGE 2 MA\CECL\UCLEE\UCLEE - fichiers\FR7 - Copie.era	841	56		
C:\Use	rs\Adrien\Documents\Travaux\ROGE 2 MA\CECL\UCLEE\UCLEE - fichiers\FR8.era	752	98		Find & Sort
C:\Use	rs\Adrien\Documents\Travaux\ROGE 2 MA\CECL\UCLEE\UCLEE - fichiers\FR9.era	911	37 🗸		- Ind d boit
Sei	arch Results Frequencies				
_					
Expor	tto V Tokens in environment 10 V				
Tag	Result				
QC	, then people in Europe would have been more involved $! \rightarrow$. O	course , this tend	s to be " a bit		A
LS	tends to be " a bit " exaggerated; this far-fetched reflexion \rightarrow the	ught came to me	during the Gulf	War . Life in	
Z	bit " exaggerated; this far-fetched thought came to me during this time of war $ ightarrow$ th	e Gulf War . Life ir	our Western s	society did not chang	je very
GA	not change very significantly in January 1990 , except that $0 \rightarrow$ the 1	elevision , the rad	io and newspa	pers bombarded us v	with Saddam
GA	significantly in January 1990 , except that the television , $0 \rightarrow$ the	adio and newspap	ers bombarde	d us with Saddam , E	Bush and
LS	, except that the television , the radio and newspapers " <code>bombed</code> " \rightarrow <code>bom</code>	nbarded us with Sa	addam , Bush a	and the allied raids .	
GDI	radio and newspapers bombarded us with Saddam , Bush and some \rightarrow th	allied raids . The	only field whic	h underwent some c	hanges
z	s with Saddam , Bush and the allied raids . The only activity which had to modify its way of living was journalism $ ightarrow$ The second se	e only field which	underwent son	ne changes is that of	f journalism . This is what I felt at that time ,
FS	that period of conflict . The dictatorial power of " $tyran \rightarrow tyra$	nt Saddam " cons	tantly affected	people 's lives befo	re the
LS	. The dictatorial power of " tyrant Saddam " constantly infected \rightarrow affe	cted people 's liv	es before the v	var broke out ,	
GNN	of " tyrant Saddam " constantly affected people 's life \rightarrow lives	before the war bi	oke out , and	it further intensified	
LS	" constantly affected people 's lives before the war occurred \rightarrow broken broken broken and the second seco	e out , and it furt	ner intensified	afterwards . Furtherr	nore , both
LS	s lives before the war broke out , and it still \rightarrow furth	er intensified after	wards . Furthei	rmore , both Saddan	n and God impose
	494 result(;)			

Figure 19: Overview of the UCLEE concordance tool

"Files" (red box, top left) displays the names of the .era files that have been uploaded, with their respective numbers of tokens and tags.

"Search" (green box, top right) allows you to launch a linguistic query by entering a tag, an error and/or a correction and to filter the query results. It is composed of four fields (Tag, Error, Correction, Sort by), three of which are *matchers* (Tag, Error, Correction). *Matchers* are currently not case-sensitive. Queries are performed by entering a search target into one matcher (or more) and clicking on "Find & Sort". Queries in the Error field can be extended to non-tagged instances by checking "Include correct uses". An empty *matcher* field means that the field will accept all results. If all three *matcher* fields are empty, the concordancer will retrieve all error tag instances. The wildcard symbol (*) can be used at the beginning or end of a word to specify that the word can be followed or preceded by anything (e.g. *the* * *of*). A *matcher* can contain several words or tags separated by spaces (e.g. GVAUX GVT). The "Sort by" field defines the order in which the search results are reported in the concordancer. There are three possible sorting types (tag, error, correction), which can be combined.

It is possible to combine multiple search criteria. These are a few examples:

- you can search for all cases of missing personal pronouns by entering "\0" in the Error box and "GPP" in the Tag box;
- the wildcard (*), representing a sequence of zero or more characters, may be used to group multiple Tags, Errors or Corrections. For example, querying "un*" in the Error box will yield all instances of error-tagged text that includes a word beginning with "un";
- you can search for all tags denoting grammatical verb errors by querying "GV*" in the Tag box.

The concordance lines (Figure 19, blue circle) are sorted by order of appearance in the selected file(s). Concordance lines can be exported to PDF or CSV^3 formats by using the "Export to" option situated on the left-hand side of the window (see Figure 20). You can adapt the length of the concordance lines displayed on the screen (ranging from 10 to 50 tokens) by selecting the desired number in the "Tokens in environment" scroll-down menu ("Search Results" tab).

³ The CSV format is better suited to subsequent data processing in Microsoft Excel.

VCLEE - Concord			Sec.	-	
File					
Files					Search
File Name	#Tokens	#Tags		Tag	
C:\Users\Adrien\Documents\Travaux\ROGE 2 MA\CECL\UCLEE\UCLEE - fichiers\FR2 - Copie.era	908	93	*	Tay	
C:\Users\Adrien\Documents\Travaux\ROGE 2 MA\CECL\UCLEE\UCLEE - fichiers\FR3 - Copie.era	777	32		Error	Match outside tags
C:\Users\Adrien\Documents\Travaux\ROGE 2 MA\CECL\UCLEE\UCLEE - fichiers\FR4 - Copie.era	690	35		Correction	
C:\Users\Adrien\Documents\Travaux\ROGE 2 MA\CECL\UCLEE\UCLEE - fichiers\FR5 - Copie.era	739	50			
C:\Users\Adrien\Documents\Travaux\ROGE 2 MA\CECL\UCLEE\UCLEE - fichiers\FR6 - Copie.era	699	62		Sort by	
C:\Users\Adrien\Documents\Travaux\ROGE 2 MA\CECL\UCLEE\UCLEE - fichiers\FR7 - Copie.era	841	56			
C:\Users\Adrien\Documents\Travaux\ROGE 2 MA\CECL\UCLEE\UCLEE - fichiers\FR8.era	752	98			Find & Sort
C:\Users\Adrien\Documents\Travaux\ROGE 2 MA\CECL\UCLEE\UCLEE - fichiers\FR9.era	911	37	-		This door
Search Results Frequencies					
export to • Tokens in environment. 10 •					
QC CSV , then people in Europe would have been more involved ! → . Of cour	se , this tends	s to be "a	bit		
LS PDF tends to be " a bit " exaggerated; this far-fetched reflexion → thought	came to me d	luring the	Gulf W	'ar . Life in	
Z bit " exaggerated; this far-fetched thought came to me during this time of war → the Gul	War . Life in	our Weste	ern soc	ciety did not ch	ange very
GA not change very significantly in January 1990 , except that 0 → the televis	ion , the radi	o and new	spaper	rs bombarded i	us with Saddam
GA significantly in January 1990 , except that the television , $0 \rightarrow$ the radio	and newspape	ers bomba	rded u	is with Saddam	, Bush and
LS , except that the television , the radio and newspapers " bombed " → bombard	ed us with Sa	ddam , Bu	ish and	d the allied raid	s.
GDI radio and newspapers bombarded us with Saddam , Bush and some → the allie	d raids . The	only field v	which u	underwent som	e changes
Z s with Saddam , Bush and the allied raids . The only activity which had to modify its way of living was journalism → The only	y field which u	underwent	some	changes is tha	t of journalism . This is what I felt at that time ,
FS that period of conflict . The dictatorial power of " tyran → tyrant Sa	ddam " const	antly affec	ted pe	ople 's lives b	efore the
LS . The dictatorial power of " tyrant Saddam " constantly infected → affected	people 's live	s before ti	he war	broke out ,	
GNN of " tyrant Saddam " constantly affected people 's life → lives befo	re the war bro	oke out , a	and it fi	urther intensifi	ed
LS " constantly affected people 's lives before the war occurred → broke ou	, and it furth	er intensifi	ied aft	erwards . Furth	nermore , both
LS s lives before the war broke out , and it still further inte	nsified afterv	vards . Fur	rtherm	ore , both Sado	dam and God impose 👻
494 result(s)					

Figure 20: Exporting the concordance lines

The "Frequencies" tab (Figure 21) displays the following results:

1. "Absolute frequencies"

This option indicates the raw number of errors sorted by tag type. For example, the first file on the list (indicated in the first column of the window) contains 93 error tags (see second column). Among these 93 error tags there are four FS errors (fourth column), six GA errors (fifth column), one GADJN error (sixth column), etc. If the grid is read vertically, the results found at the very bottom of the page indicate that if all the sub-corpora are put together, the FM error tag represents 1% of all the errors found in the entire corpus, with a total of three occurrences. Similarly, the FS tag represents 7% of all the error tags in the corpus, with 35 occurrences.

Note: For all error types to appear, the search boxes need to be empty.

UCLEE - Concord		100.000	Contract			Sec. 1	-		Read of the	-		10	74		1000		-	-		- 0	23
File																					
					Files												Searc	:h			
File Name										#Tok	kens	#Tags			T						
C:\Users\Adrien\Documents\Trav	aux\ROGE	2 MA\CEC	L\UCLEE\U	CLEE - fich	niers\FR2 - 0	opie.era				908		93	-		tag						
C:\Users\Adrien\Documents\Trav	aux\ROGE	2 MA\CEC	L\UCLEE\U	CLEE - fich	iers\FR3 - 0	opie.era				777		32		E	rroc			-)	Aatch outs	ide tags	
C:\Users\Adrien\Documents\Trav	aux\ROGE	2 MA\CEC	L\UCLEE\L	CLEE - fich	niers\FR4 - 0	opie.era				690		35		Correc	bon						
C:\Users\Adrien\Documents\Trav	aux\ROGE	2 MA\CEC	L\UCLEE\L	CLEE - fich	iers\FR5 - 0	Copie.era				739		50									
C:\Users\Adrien\Documents\Trav	aux\ROGE	2 MA\CEC	L\UCLEE\U	CLEE - fich	iers\FR6 - 0	opie.era				699		62		Sort	t by						
C:\Users\Adrien\Documents\Trav	aux\ROGE	2 MA\CEC	LUCLEEN	CLEE - fich	niers\FR7 - 0	opie.era				841		56									
C:\Users\Adrien\Documents\Trav	aux\ROGE	2 MA\CEC	L\UCLEE\U	CLEE - fich	iers\FR8.er	8				752		98	10				Find 8.9	2nit			
C:\Users\Adrien\Documents\Trav	aux\ROGE	2 MA\CEC	L\UCLEE\U	CLEE - fich	iers\FR9.er	a				911		37					PILLURS				
Search Results Freque	ocies	N																			
Export to + Absolute frequ	uencies	% by teo	% by file	Per n tok	ens n= (1	• (00															
	-					Terreto	1		Letter.	1.000	1.000				1	1	1 and		12.00		
File	Total	FM	P5	GA	GADJN	GADVO	ept	GDO	GNC	GNN	GPF	6	PΡ	GPR	GPU	GVAUX	GVM	GVN	GVNP	GVI	G
C:\Users\Adrien\Documents\Tra	93		4	0	1		1			8		1		2	2		1			4	1
C:\Users\Adrien\Documents\Tra	32		8	1		1										2					
C:\Users\Adrien\Documents\Tra	35		8													1				2	
C:\Users\Adrien\Documents\Tra	50	1	4			2			1	3		1						1	1		1
C:\Users\Adrien\Documents\Tra	62		3	4		1		1	1	2					3	2				2	
C:\Users\Adrien\Documents\Tra	56			3		1				3									1	1	
C:\Users\Adrien\Documents\Tra	98	2	5	4		1			1	2	1				5	1		1		1	2
C:\Users\Adrien\Documents\Tra	. 37		1							1		1			2	1				3	
C:\Users\Adrien\Documents\Tra	31		2	1					1	2		1				3			1	1	
4		11111	-				1440	1227		12.23	1443				1222	14400	1221	1922	1210	1225	,
% out of total number of tags	100%	1%	7%	4%	0%	1%	0%	0%	1%	4%	0%	19		0%	2%	2%	0%	0%	1%	3%	1
Total	494	3	35	19	1	6	1	1	4	21	1	4		2	12	10	1	2	3	14	4

Figure 21: Absolute frequencies in the concordance tool

2. <u>"% by tag"</u>

The results displayed in this sub-header should be read vertically. They indicate the relative distribution of individual error tags across corpus files (with 100% representing the whole corpus). For example, <u>Figure 22</u> indicates that 33.3% of the FM error tags are found in the fourth file, while the remaining 66.7% are in the seventh file.

UCLEE - Concord																		-	- 0 -	×
					Files											Searc	h			
File Name										#To	kens	#Tags		-						
C:\Users\Adrien\Documents\Trav	aux\ROGE	2 MA\CECL	UCLEE\UC	LEE - fichie	ers\FR2 - C	opie.era				908		93 -		Tag						
C:\Users\Adrien\Documents\Trav	aux\ROGE	2 MA\CECL	UCLEE\UC	LEE - fichie	ers\FR3 - C	opie.era				777		32		Error			101	latch outsi	de tags	
C:\Users\Adrien\Documents\Trav	aux\ROGE	2 MAICECLI	UCLEENUC	LEE - fichia	ers\FR4 - C	opie.era				690		35	Com							
C:\Users\Adrien\Documents\Trav	aux\ROGE	2 MAILCECL	UCLEE\UC	LEE - fichie	ers\FR5 - C	opie.era				739		50	CONT	neoun						
C:\Users\Adrien\Documents\Trav	aux\ROGE	2 MAICECL	UCLEE\UC	LEE - fichie	ers\FR6 - C	opie.era				699		62	54	ort by						
C:\Users\Adrien\Documents\Trav	aux\ROGE	2 MA\CECL	UCLEE/UC	LEE - fichie	ers\FR7 - C	opie.era				841		56								
C:\Users\Adrien\Documents\Trav	aux\ROGE	2 MA\CECL	UCLEE/UC	LEE - fichie	ers\FR8.era	2				752		98								
C:\Users\Adrien\Documents\Trav	aux\ROGE	2 MA\CECL	UCLEE\UC	LEE - fichie	ers\FR9.era					911		37 🗸				Find & S	ort			
Search Results Freque	incies	_																		
Export to + Absolute freq	uencies	% by tag	by file	Per n toke	ns n= (1	• (00														
File	Total		PS	GA	GADJN	GADVO	GDI	GDO	GNC	GNN	GPF	GPP	GPR	GPU	GVAUX	GVM	GVN	GVNF	GVT	
C:\Users\Adrien\Documents\Tra.	. 93		11.4%	31.6%	100%		100%			38.1%		25%	100%	16.7%		100%			28.6%	1
C:\Users\Adrien\Documents\Tra.	32		22.9%	5.3%		16.7%									20%					
C:\Users\Adrien\Documents\Tra.	35		22.9%												10%				14.3%	
C:\Users\Adrien\Documents\Tra	. 50	33.3%	11.4%			33.3%			25%	14.3%		25%					50%	33.3%		
C:\Users\Adrien\Documents\Tra.	62		8.6%	21.1%		16.7%		100%	25%	9.5%				25%	20%				14.3%	
C:\Users\Adrien\Documents\Tra.	56			15.8%		16.7%				14.3%								33.3%	7.1%	
C:\Users\Adrien\Documents\Tra.	. 98	66.7%	14.3%	21.1%		16.7%			25%	9.5%	100%			41.7%	10%		50%		7.1%	4
C:\Users\Adrien\Documents\Tra.	. 37		2.9%							4.8%		25%		16.7%	10%				21.4%	
C:\Users\Adrien\Documents\Tra.	. 31		5.7%	5.3%					25%	9.5%		25%			30%			33.3%	7,1%	
4																				,
Total %		100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	
% out of total number of tags	100%	1%	7%	4%	0%	1%	0%	0%	1%	4%	0%	1%	0%	2%	2%	0%	0%	1%	3%	3
Total	40.4	3	35	10	1	6	1		4	21	+	4	2	12	10	1	2		1.4	

Figure 22: Detailed percentages of errors sorted by tag

3. <u>"% by file"</u>

The results displayed in the option "% by file" (Figure 23) should be read horizontally. They indicate the relative distribution of error tags in a single corpus file (as opposed to the previous "% by tag" option). In this case, 100% represents the total number of errors in an individual file (and not the entire corpus).

Contra Contracto			Contract	-				ι.						-		100		100			×
File					Files													2			
					FIDES												Searc	9			
File Name										#To	kens	#Tage			Tag						
C:\Users\Adrien\Documents\Trav	aux/ROGE 2	MAYCECL	AUCLEEVUC	LEE - fichi	ers\FR2 - C	opie.era				908		93	ΰÎ.								
C:\Users\Adrien\Documents\Trav	aux\ROGE 2	MAYCECL	\UCLEE\UC	LEE - fichi	ers\FR3 - C	opie.era				777		32			Errar				Match outsi	de tegs	
C:\Users\Adrien\Documents\Trav	Bux\ROGE 2	MA\CECL	/UCLEE/UC	LEE - fichi	ers\FR4 - C	opie.era				690		35		Corre	ction						
C:\Users\Adrien\Documents\Trav	aux\ROGE 2	MAICECL	/UCLEE/UC	LEE + fichi	ers\FR5 - C	opie.era				739		50			460						
C:\Users\Adrien\Documents\Trav	aux/ROGE 2	MAVCECL	/UCLEE/UC	LEE + fichi	ers\FR6 + C	opie.era				699		62		50	te ny						
C:\Users\Adrien\Documents\Trav	aux\ROGE 2	MAVCECL	/UCLEE/UC	LEE - fichi	ers\FR7 - C	opie.era				841		56									
C:\Users\Adrien\Documents\Trav	www.ROGE 2	MAVCECL	\UCLEE\UC	LEE + fichi	ers\FR8_ers	r				752		98	10				Find & S	no			
C:\Users\Adrien\Documents\Trav	aux/ROGE 2	MA\CECL	/UCLEE/UC	LEE - fichi	ers\/R9.eri					911		37									
Search Results Freque	incies																				
Export to = Absolute freq	uencies % I	by tep	% by file	Per n toke	ns n= 🖸	+ (00															
File	Total	FH 1	15	GA	GADIN	GADVO	GDI	600	GNC	GNN	GPF	GP	p.	GPR	GPU	GVAUX	GVN	GVN	GVNP	GVT	G
C:\Users\Adrien\Documents\Tra.	93 (100%)		4.3%	6.5%	1.1%		3.1%			8.6%		1.1	56	2.2%	2.2%		1.1%			4.3%	1
C:\Users\Adrien\Documents\Tra.	32 (100%)		25%	3.1%		3.1%										6.3%					
C:\Users\Adrien\Documents\Tra	35 (100%)		22.9%																	5.7%	
																2.9%					1
C:\Users\Adrien\Documents\Tra.	50 (100%)	2%	8%			4%			2%	6%		2%	2			2.9%		2%	2%		2
C:\Users\Adrien\Documents\Tra. C:\Users\Adrien\Documents\Tra	50 (100%) 62 (100%)	2%	8% 4.8%	6.5%		4% 1.6%		1.6%	2% 1.6%	6% 3.2%		2%	R.		4,8%	2.9%		2%	2%	3.2%	2
C:\Users\Adrien\Documents\Tra. C:\Users\Adrien\Documents\Tra. C:\Users\Adrien\Documents\Tra.	50 (100%) 62 (100%) 56 (100%)	2%	0% 4.8%	6.5% 5.4%		4% 1.6% 1.8%		1.6%	2% 1.6%	6% 3.2% 5.4%		2%	i.		4,8%	3.2%		2%	2%	3.2%	2
C:\Users\Adrien\Documents\Tra C:\Users\Adrien\Documents\Tra C:\Users\Adrien\Documents\Tra C:\Users\Adrien\Documents\Tra	50 (100%) 62 (100%) 56 (100%) 98 (100%)	2%	8% 4.8% 5.1%	6.5% 5.4% 4.1%		4% 1.6% 1.8% 1%		1.6%	2% 1.6%	6% 3.2% 5.4% 2%	1%	2%	6		4.8% 5.1%	2.9%		2%	2% 1.8%	3.2%	2 74
C: (Users\Adrien\Documents\Tra C:\Users\Adrien\Documents\Tra C:\Users\Adrien\Documents\Tra C:\Users\Adrien\Documents\Tra C:\Users\Adrien\Documents\Tra	50 (100%) 62 (100%) 56 (100%) 98 (100%) 37 (100%)	2%	8% 4.8% 5.1% 2.7%	6.5% 5.4% 4.1%		4% 1.6% 1.8% 1%		1.6%	2% 1.6% 1%	6% 3.2% 5.4% 2% 2.7%	2%	2%	-		4.8% 5.1% 5.4%	2.9% 3.2% 1% 2.7%		2%	2%	3.2% 1.8% 1% 8.1%	N (4:
C: Users Vadrien/Documents/Tra C: Users Vadrien/Documents/Tra C: Users Vadrien/Documents/Tra C: Users Vadrien/Documents/Tra C: Users Vadrien/Documents/Tra C: Users Vadrien/Documents/Tra	50 (100%) 62 (100%) 56 (100%) 90 (100%) 37 (100%) 31 (100%)	2%	0% 4.8% 5.1% 2.7% 6.5%	6.5% 5.4% 4.1% 3.2%		4% 1.6% 1.8% 1%		1.6%	2% 1.6% 1% 3.2%	6% 3,2% 5,4% 2% 2.7% 6,5%	2%	2%	46 16		4.8% 5.1% 5.4%	2.9% 3.2% 1% 2.7% 9.7%		2%	2%	3.2% 1.8% 1% 0.1% 3.2%	2
C: Users VAdrien Documents (Tra C: Users VAdrien Documents) Tra C: Users VAdrien Documents (Tra C: Users VAdrien (Documents) Tra C: Users VAdrien (Documents) Tra C: Users VAdrien (Documents) Tra	50 (100%) 62 (100%) 56 (100%) 90 (100%) 37 (100%) 31 (100%)	2%	0% 4.8% 5.1% 2.7% 6.5%	6.5% 5.4% 4.1% 3.2%		4% 1.6% 1.8%		1.6%	2% 1.6% 1% 3.2%	6% 3.2% 5.4% 2% 2.7% 6.5%	2%	2%	46		4.8% 5.1% 5.4%	2.9% 3.2% 1% 2.7% 9.7%		2%	2% 1.8% 3.2%	3.2% 1.8% 1% 8.1% 3.2%	2
C: (Users)Adrien/Documents/Tra C: (Users)Adrien/Documents/Tra C: (Users)Adrien/Documents/Tra C: (Users)Adrien/Documents/Tra C: (Users)Adrien/Documents/Tra C: (Users)Adrien/Documents/Tra	50 (100%) 62 (100%) 56 (100%) 90 (100%) 37 (100%) 31 (100%)	2%	8% 4.8% 5.1% 2.7% 6.5%	6.5% 5.4% 4.1% 3.2%		4% 1.6% 1.8% 1%		1.6%	2% 1.6% 1% 3.2%	6% 3.2% 5.4% 2% 2.7% 6.5%	2%	2%	46		4.8% 5.1% 5.4%	2.9% 3.2% 1% 2.7% 9.7%		2%	2% 1.8% 3.2%	3.2% 1.8% 1% 8.1% 3.2%	2
C:/Users/Adrien/Documents/Tra C:/Users/Adrien/Documents/Tra C:/Users/Adrien/Documents/Tra C:/Users/Adrien/Documents/Tra C:/Users/Adrien/Documents/Tra 	50 (100%) 62 (100%) 56 (100%) 90 (100%) 37 (100%) 31 (100%)	2%	0% 4.8% 5.1% 2.7% 6.5%	6.5% 5.4% 4.1% 3.2%	0%	4% 1.6% 1.8% 1%	0%	1.6%	2% 1.6% 1% 3.2%	6% 3.2% 5.4% 2% 6.5% 4%	3%	2%	44	0%	4.8% 5.1% 5.4%	2.9% 3.2% 1% 2.7% 9.7%	0%	2%	2% 1.8% 3.2%	3.2% 1.8% 1% 0.1% 3.2%	2 12 12

Figure 23: Detailed percentages of errors sorted by file

4. <u>"Per n tokens" (relative frequencies)</u>

This sub-header automatically calculates relative (or relative/normalized) frequencies to ensure comparability between texts of varying length (Figure 24). Following a 'rule of three', these results correspond to the division of absolute frequencies by the total number of words in the text(s), which are then multiplied by a normalisation base (e.g. relative frequency per 100 words, per 500 words, per 1,000 words).

File																				
				F	iles											Search				
File Name										#Tok	ens #	Tags		2.10						
C:\Users\Adrien\Documents\Travi	aux\ROGE	2 MA\CECL	UCLEE\UCL	EE - fichier	s\FR2 - C	opie.era				908	93	3 -		Tag						
C:\Users\Adrien\Documents\Travi	aux\ROGE	2 MA\CECL	UCLEE\UCL	EE - fichier	s\FR3 - C	opie.era				777	33	2	E	irror			E M	tch outsid	e tags	
C:\Users\Adrien\Documents\Travi	aux\ROGE	2 MA\CECL	UCLEE\UCL	EE - fichier	s\FR.4 - C	opie.era				690	35	5	Correct	tion						
C:\Users\Adrien\Documents\Travi	aux\ROGE	2 MA\CECL\	UCLEE\UCL	EE - fichier	s\FR5 - C	opie.era				739	5	0								
C:\Users\Adrien\Documents\Travi	aux\ROGE	2 MA\CECL\	UCLEE\UCL	EE - fichier	s\FR6 - C	opie.era				699	6;	2	Sor	tby						
C:\Users\Adrien\Documents\Travi	aux\ROGE	2 MA\CECL\	UCLEE\UCL	EE - fichier	s\FR7 - C	opie.era				841	5	5								
C:\Users\Adrien\Documents\Travi	aux\ROGE	2 MA\CECL\	UCLEE\UCL	.EE - fichier	s\FR.8.era	0				752	91	3				Dard # Day				
C:\Users\Adrien\Documents\Travi	aux\ROGE	2 MA\CECL\	UCLEE\UCL	EE - fichie	s\FR9.era	0)				911	31					Fille & Gu	4			
Search Results Frequen	ncies				1															
Export to + Absolute frequ	uencies 9	by tag 👋	by file P	er n token	s) n= (1	00) -														
File	Total	FM	FS	GA	~	GADYO	GDI	GDO	GNC	GNN	GPF	GPP	GPR	GPU	GVAUX	GVM	GVN	GVNF	GVT	G
C:\Users\Adrien\Documents\Tra	93		0.44052	0.6607	100		0.11013			0.88105		0.11013	0.22026	0.22026		0.11013.			0.44052	0
C:\Users\Adrien\Documents\Tra	32		1.02960.	0.12870.	500	0. 2870.									0.25740					
C:\Users\Adrien\Documents\Tra	35		1.15942	1	1000	/									0.14492				0.28985	
C:\Users\Adrien\Documents\Tra	50	0.13531	0.54127.	`	$ \rightarrow $	0.27063.			0.13531.	0.40595.		0.13531					0.13531	0.13531		0
C:\Users\Adrien\Documents\Tra	62		0.42918.	0.57224		0.14306		0.14306	0.14306.	0.28612				0.42918	0.28612				0.28612	
C:\Users\Adrien\Documents\Tra	56			0.35671.		0.11890.				0.35671.								0.11890	0.11890	
C:\Users\Adrien\Documents\Tra	98	0.26595	0.66489	0.53191		0.13297.			0.13297.	0.26595.	0.13297			0.66489	0.13297		0.13297		0.13297	. o
C:\Users\Adrien\Documents\Tra	37		0.10976							0.10976.		0.10976		0.21953	0.10976				0.32930	
C:\Users\Adrien\Documents\Tra	31		0.34542	0.17271					0.17271	0.34542		0.17271			0.51813			0.17271	0.17271	
% out of total number of tags	100%	1%	7%	4%	0%	1%	0%	0%	1%	4%	0%	1%	0%	2%	2%	0%	0%	1%	3%	1
Total	494	3	35	19	1	6	1	1	4	21	1	4	2	12	10	1	2	3	14	4

Figure 24: Relative frequencies of error tags

It should also be noted that the user can at any time export the results displayed in the various sub-headers to CSV format for later use in other programs such as Microsoft Excel or statistical packages (see Figure 25).

UCLEE - Concord	- 68	-	_		i ince	-			-			1								I,
File																				
				F	iles											Search				
File Name										#	Tokens	#Tags								
C:\Users\Adrien\Documents\Trava	aux\ROGE 2	MA\CECL\	UCLEE\UCL	EE - fichiei	rs∖FR1 - Co	pie.era				5	79	19								
C:\Users\Adrien\Documents\Trava	aux\ROGE 2	MA\CECL\	UCLEE\UCL	.EE - fichier	rs\FR2 - Co	pie.era				91	08	93	Er	ror			— м	atch outsid	le tags	
C:\Users\Adrien\Documents\Trava	aux\ROGE 2	MA\CECL\	UCLEE\UCL	EE - fichier	rs\FR3 - Co	pie.era				7	77	32	Correct	tion						
C:\Users\Adrien\Documents\Trava	aux\ROGE 2	MA\CECL\	UCLEE\UCL	.EE - fichiei	rs∖FR4 - Co	pie.era				6	90	35								
C:\Users\Adrien\Documents\Trava	aux\ROGE 2	MA\CECL\	UCLEE\UCL	EE - fichier	rs\FR5 - Co	pie.era				7	39	50	Sort	by						
C:\Users\Adrien\Documents\Trava	aux\ROGE 2	MA\CECL\	UCLEE\UCL	EE - fichier	rs∖FR6 - Co	pie.era				6	99	62								
C:\Users\Adrien\Documents\Trava	aux\ROGE 2	MA\CECL\	UCLEE\UCL	EE - fichier	rs∖FR7 - Co	pie.era				84	41	56				Find & So	rt			
C:\Users\Adrien\Documents\Trava	aux\ROGE 2	MA\CECL\	UCLEE\UCL	EE - fichiei	rs\FR8.era					7	52	98				11110 0 00				
Frequer Frequer	ncies																			
Export to Absolute frequ	iencies %	by tag %	by file P	er n token	s n= 10	• 0														
File	Total	FM	FS	GA	GADJN	GADVO	GDI	GDO	GNC	GNN	GPF	GPP	GPR	GPU	GVAUX	GVM	GVN	GVNF	GVT	G
drien (Documents \Tra	19		1	1						2								1		
C:\Ders\Adrien\Documents\Tra	93		4	6	1		1			8		1	2	2		1			4	1
C:\Users\Adrien\Documents\Tra	32		8	1		1									2					
C:\Users\Adrien\Documents\Tra	35		8												1				2	
C:\Users\Adrien\Documents\Tra	50	1	4			2			1	3		1					1	1		1
C:\Users\Adrien\Documents\Tra	62		3	4		1		1	1	2				3	2				2	
C:\Users\Adrien\Documents\Tra	56			3		1				3								1	1	
C:\Users\Adrien\Documents\Tra	98	2	5	4		1			1	2	1			5	1		1		1	2
4		1	1	1	1		1	1												F
% out of total number of tags	100%	1%	7%	4%	0%	1%	0%	0%	1%	4%	0%	0%	0%	2%	1%	0%	0%	1%	2%	1
Total	445	3	33	19	1	6	1	1	3	20	1	2	2	10	6	1	2	3	10	4

Figure 25: Exporting frequencies

2.7. Search

The Search tab enables you to extract and highlight specific features by using Corpus Query Language⁴. For example, by entering /of|from/, all instances of the prepositions "of" and "from" are highlighted in the text (see Figure 26). The highlighted items can then be tagged using the Tags function.

File	Settings Edit Tags Check Analyse Search Exercises Help
	economic support in bad times. Furthermore we have a system where we can
	borrow a lot <mark>of</mark> "expensive" money <mark>from</mark> the state during education. When an
	immature 18-year-old boy or girl get in a situation like this, I can promise
	they dream a lot. Some people dream so much that they don't get anything
	done and pass all their exams
5	There's also another aspect with all this technological advances. People
	don't care. They use the parts <mark>of</mark> it they need like email, mobile phones and
	stuff that keep them connected to our global world 24 hours a day. Due to
	the fact that technology has come so far in it's development, people haven't
	got the slightest idea about what they actually do. In my opinion, regarding
	virtual reality and future time travels, we don't need all this crap! We
	live in a society where it is a job to make the masses addicted to unreal
	needs. The people's relationship to technology can almost be compared to the
	late Victorian era where all the workers stood by a machine and did the same
	thing over and over. We don't know what it is all about, and at the same
	time a society <mark>of</mark> moneymaking data nerds is created, they earn all the money
	to. And one more thing before I drift too far away from the topic: Isn't it
	more important to have a society where common knowledge is higher valued

Figure 26: Search tab

2.8. Exercises

This tab can be used to generate tailor-made exercises for specific users or user groups by clicking on the "Create exercise" option (see Figure 27).

⁴ More information on CQL searches can be found on the following websites <u>http://fedora.clarin-d.uni-saarland.de/teaching/Corpus Linguistics/Tutorial RegExp.html</u> or <u>https://www.sketchengine.eu/guide/regular-expressions/</u>



You can create exercises focused on particular error tags by selecting these in the scrolldown list. Multiple tags can be selected at the same time by holding the $\langle CTRL \rangle$ (or $\langle cmd$ $\Re \rangle$ in Mac) key down while (de)selecting several tags (<u>Figure 28</u>).

Create exercise	×
Select tags (keep CTRL down and drag the mouse to (un)select several tags)	Select All
FM FS FSR GA GADJCS GADJN GADJO GDD GDD GDD GDD GDD GDT	·
Cic	se Filter

Figure 28: Creating exercises based on multiple tags

When one or several tags have been selected, the software opens a new window which displays all occurrences of the selected tags in the file. The default state is that all the tags are selected; the user can then choose to keep all examples or remove unwanted instances

by unticking the box situated next to the number of the concordance line (for example, to delete instances that are too challenging or too easy for particular user groups) (Figure 29).

The Louvain Error Tagging Editor allows the user to choose between three types of gapfilling exercise. These can be saved and exported as raw text to the clipboard, as a .txt file or an html webpage. To do this, select the desired format in the scroll-down menu and specify the information to be displayed (detailed tags, simplified tags or no tags), using the buttons at bottom left, above "Format" (Figure 29).



Figure 29: Exporting tailor-made exercises

As exemplified in Figure 30, the "Detailed tags" option makes use of the full error tagset to provide detailed information about the type of error made. "Simplified tags" is designed to guide learners in the right direction but only uses the main error categories (e.g. L instead of LS). "No tags" is intended to point out the error to the learner without adding any additional information; this option allows for a variety of possible answers.

For example, in the sentence *My house has been evaluated by an expert*, the erroneous string of words displayed is *has been evaluated*. In the "Detailed tags" option the learner is provided with the LSVF tag, which indicates that the error lies in the use of a lexical verb which is a false friend. The learner therefore knows that they do not have to think

about tense, aspect, etc. Their attention is clearly directed towards a specific issue. In the "Simplified tags" option the learner is provided with the LS tag, which indicates that they need to focus on a lexical error, but is not given the "false friend" hint. Finally, in the "No tags" option the learner is presented with only the erroneous string of words, so that they have to check whether the tense, aspect, number, choice of verb, etc. are correct.

1. Detailed tags

The urgent need for imagination [LS: During]	the last decade of the twentieth century, GPP	it] is [LS: e	asy] to exemplify
how science, industrialisation and technology have pervaded our ever	ryday life.	h	
Some will claim that the increasing number of factories has spoiled the	e [GNC: landscape's [LS: attraction]		·
Others [LS: assert] that imagination and	dreams have definitely vanished from our modern so	ociety.	
0, in the twentieth century, fears for the future of mankind [LS: have	predominated] the opti	mism about inevitable progress which w	as more typical of the nineteenth century .
All these examples point out that artists (that is to say people who give	e shape to their dreams and imagination through wo	rks of art) either play the role of catalyst	s for technological progress, or that of a [LS:
railing] against the uses and abuses of sc	ence and industrialisation.		
1. Simplified tags			
The urgent need for imagination II : During	the last decade of the twentieth century. [G: it]	is II : easy	to exemplify how
science, industrialization and technology have perceded our everyday	the last decade of the twentieth century, [6, n]	is [L. easy]	to excluping now
Some will claim that the increasing number of factories has enailed th	IC: landscene's II : attraction]	h	
Out of the state o	le [O. landscape's [L. attraction]	·	
Others [L: assert] that imagination and d	reams have definitely vanished from our modern so	ciety.	
0, in the twentieth century, fears for the future of mankind [L: have p	redominated] the optim	iism about inevitable progress which wa	s more typical of the nineteenth century .
All these examples point out that artists (that is to say people who give	e shape to their dreams and imagination through we	orks of art) either play the role of catalys	ts for technological progress, or that of a [L:
railing] against the uses and abuses of sc	ience and industrialisation.		
Z. NO TAQS			
0			
The urgent need for imagination [SDuring]	the last decade of the twentieth century. [Sit]	is [Seasy]	to exemplify how science
industrialisation and technology have peruaded our everyday life	and has becaue of the mention century, [on]	in focusit	to enempiny non serence,
Some will claim that the increasing number of factories has spoiled f	e [Slandscape's [Sattraction]	1	
Others [Sassant]	are have definitely venished from our modern seei	/ ·	
Unat integritation and dre	ans have definitely vanished from our modern socia	ay.	
o, in the twentieth century, fears for the future of mankind [Shave pr	edominated] the optimis	m about inevitable progress which was	nore typical of the nineteenth century .

Figure 30: Illustration of exercise formats

2.9 Help

2.9.1. The Louvain Error Tagging Manual (2.0)

This submenu opens the Louvain Error Tagging Manual (version 2.0).

2.9.2. "UCLEE Helpfile"

This submenu opens the UCLEE user guide.

Appendix: Working with texts previously tagged with uclee-en-1.3

This manual relates exclusively to the most recent version of UCLEE, version 2.0. However, some users may wish to continue working with texts already tagged using the previous version (UCLEE 1.3) and to tag further texts within that system. This appendix describes the procedures to follow if you are in that situation.

I. To convert previously tagged texts from version 1.3 to version 2.0 display

1. Open UCLEE version 1.3 as follows:

The "Tagset selection" window that opens when you launch UCLEE provides several tagset options (Figure A 1):

UCLEE - Tagset selection
Which tagset should we use for this session?
uclee-en-2.0.tag
uclee-en-1.3.tag
uclee-en-2.0.tag
upload a new tagset
Figure A 1: Tagset selection

Select uclee-en-1.3.tag.

Click on the "Open UCLEE" button at bottom right of the window. This will take you to the software's main interface (Figure A 2).



Figure A 2: Main interface

2. Convert the file(s) from version 1.3 to version 2.0 display

Click on the "Convert files from 1.3 tagset with double dollars" option in the "File" menu (Figure A 3).

	cord Exercises Help		
New window	Ctrl+N	ag">	
Open	Ctrl+O		
Open several files (Merge)	Ctrl+Shift+O		
Save	Ctrl+S		
Save as	Ctrl+Shift+S		
Print to PDF			
Export	Ctrl+Shift+E		
Convert files from 1.3 tagset with dou	ble dollars		
Close	ΔI++F4		
Quit	00114		
prrection			New Tag
rrection			New Tag
prrection			New Tag Update Tag

Figure A 3: Convert files from 1.3. tagset with double dollars

The window shown in Figure A 4 will open.

Click on 'Select files to convert' and browse to the file(s) you wish to convert (Figure A 4).

After conversion is complete, open the corresponding .era file(s), the path to which is indicated in the 'conversion completed' message that appears if the conversion was successful (Figure A 5).

UCLEE - Convert from V1	_	×
File		
Select files to convert		
Logs		
		11

Figure A 4: Select files to convert

UCLEE - Convert from V1 -	_		×
File			
Select files to convert			
Logs			
G:\Tagged texts\Tagged texts\924952\924952XT_hs_tagged_plain G:\Tagged texts\Tagged texts\924952\924952XT_hs_tagged_plain Conversion completed!	n.txt - n.era	>	4

Figure A 5: Conversion completed

The file(s) originally tagged in 1.3 (Figure A 6) will now be displayed as shown in Figure A 7.

3	UCLEE – 🗆 X				
File	Settings Edit Tags Check Analyse Search Exercises Help SOS shelters				
65	in order to ensure adequate gender-sensitive (GNN) surrounding				
	\$surroundings\$ and				
66	psychological support to young girls;				
67	8. Calls on the Member States to set up compulsory (FS) gender specific				
	\$gender-specific\$				
68	training for (GA) the $0\$ personnel receiving unaccompanied minors in (GA) the $0\$				
69	shelters as well as for interviewers, decision makers and legal				
	representatives of				
70	unaccompanied minors, as well as for making sure that the police and the				
	legal authorities in (GA) 0 \$the\$ Member States				
71	regularly undergo (FS) gender specific \$gender-specific\$ training;				
72	2 9. Calls on the Commission and (GA) 0 \$the\$ Member States to develop and				
	implement				
73	specific measures for pregnant minors by providing day-care facilities and				
	adapting (GA)				
74	the \$0\$ shelters to their specific needs during and after the pregnancy;				
75	10. Invites the Commission, in cooperation with the EIGE, to propose	*			
Con	Add Add				
	Update				
N	Remove	10			

Figure A 6: File in 1.3 tagset with double dollars before conversion



Figure A 7: Dialog box after conversion

II. UCLEE 1.3 and 2.0 tagsets

The UCLEE 2.0 tagset is not identical to that of UCLEE 1.3: some tags have been added and others changed, so that they overlap only partially with earlier tags. For example, the LS (Lexical Single) category, which in UCLEE 1.3 comprised only two sub-categories (LS and LSF), in UCLEE 2.0 comprises ten (LSADJ, LSADV etc.). It is very important that you should be aware that the conversion process just described is a technical conversion only, which changes the way that your previous error tags are formatted and displayed, but **does not change the tag labels into those of the UCLEE 2.0 tagset**.

III. Tagging new texts with UCLEE 1.3

You may also wish to tag new texts in order to add them to those previously tagged in UCLEE 1.3 to form a single corpus. Please be aware that **the two tagsets cannot be mixed**. You must continue to work in UCLEE 1.3 until your corpus is complete. If you were to work in the two versions and then attempt to open your whole corpus for analysis, you would be unable to open it in either version.