**Annotation of learner corpus data**

**Master Language**

ACTIVITY 2. Syntactic parsing

1. Open the [Hugging Face spaCy visualizer](https://huggingface.co/spaces/spacy/pipeline-visualizer).
2. Copy and paste the first sentence of the cleaned version of BRFF1065.txt into the window under ‘Text to analyze’.
3. For now, make sure all of the following options are unselected:

* Split sentences
* Collapse punct
* Collapse phrases
* Compact mode

1. Take a look at the visualization of the dependency parse.

*Note: you can right click the image to download it to your computer if you want to see it better.*

1. Click on the options ‘Compact mode’ and ‘Collapse punct’ to see what changes.
2. Look for the ‘amod’ and ‘dobj’ dependency relations.

***Question 1. What type of dependency relationships are marked by ‘amod’ and ‘dobj’?***

Hint: See [here](https://v2.spacy.io/api/annotation) a full list of the dependency tags used by spaCy.

1. Click on the option ‘collapse phrases’ to see what changes.
2. Unselect all of the options.
3. Scroll down to find the ‘Token attributes’ window. Note: You can expand the window by clicking on the arrows in the top right corner.

*Note:* You need to hover with your mouse over the table for the arrows to appear.

1. Click on the column ‘dep\_’ to sort by dependency type.

***Question 2. How many amod dependencies are there in this sentence?***

***Question 3. What adjective modifies the word ‘ideas?***

***Question 4. What is the object of the verb ‘debate?***

1. Now open the [LINDAT UDPipe tool.](https://lindat.mff.cuni.cz/services/udpipe/) Select the most recent UD model (UD 2.15) trained on English (english-atis-ud-2.15241121). Make sure both ‘Tag and lemmatize’ as well as ‘Parse’ are checked.
2. Upload the cleaned version of BRFF1065.txt under ‘Input File’ and click on ‘Process Input’.
3. Click on ‘Save Output File’ to download the parsed text to your computer. Note that the file has the extension ‘.conllu’ but it can still be opened normally in Excel. In order to do so, you must manually change the extension to ‘.xls’ (.xlsx will not work).
4. After you have changed the extension, open the file in Excel and count the number of adjectival modifiers (‘amod’) by typing the following formula into a blank cell: =COUNTIF(H5:H520, "amod").

***Question 5. How many adjectival modifiers are there in the entire text?***

*Bonus*

1. Find a way to count the number of direct object relations in the text (e.g., ‘debate + ideas’).