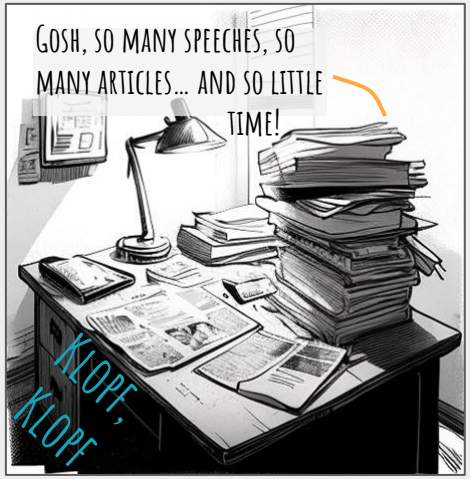
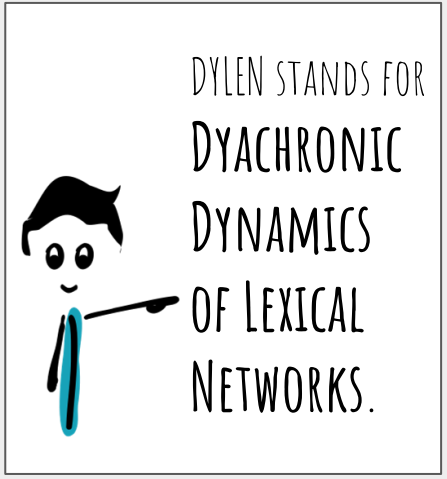
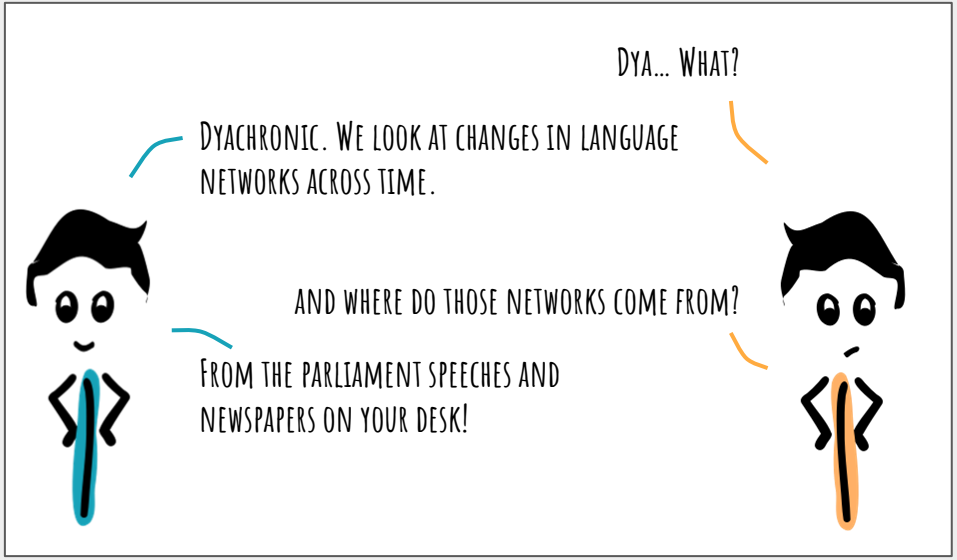

Somewhere in Austria...


Gosh, so many speeches, so many articles... and so little time!
KLOPF, KLOPF

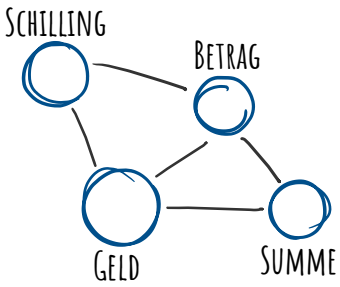

I think I have a solution for you!
Do you?
Ever heard of DyLen?


DyLen stands for
Dyachronic Dynamics of Lexical Networks.

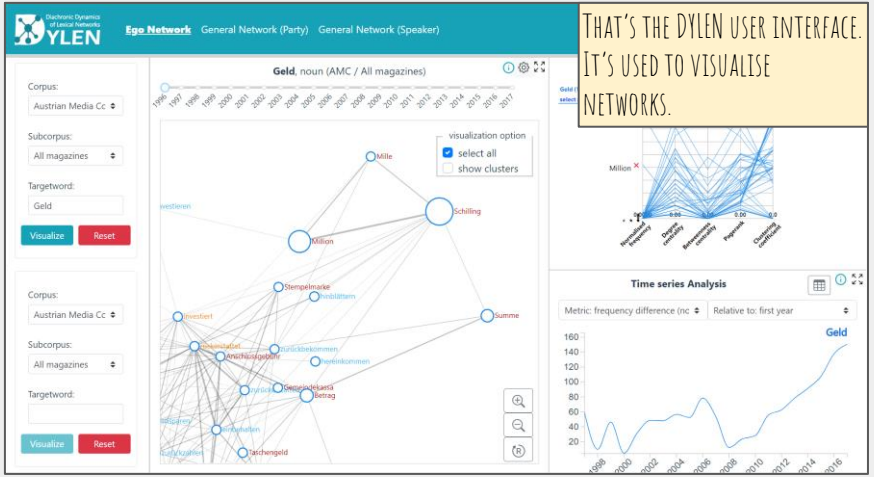

Dya... What?
Dyachronic. We look at changes in language networks across time.
And where do those networks come from?
From the parliament speeches and newspapers on your desk!


To be exact, from the Austrian Media Corpus (AMC)
And the Corpus of Austrian Parliamentary records (ParlAT)

Word embeddings help to generate networks from these texts.



Words that are connected are called **semantic neighbours**. They are also interchangable.



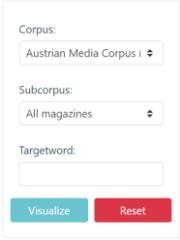
That's the DYLEN user interface. It's used to visualise networks.

Phew, that's quite a lot to take in...

No Worries, let's do it step by step.

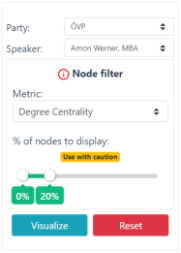
On top, you can choose between **Ego-Networks** und **General Networks**.

Ego-Networks are about **a specific word**, for example „Geld" (money). It will not be shown itself, because all words in the network are connected to it– otherwise you couldn't see anything anymore because of all the lines.



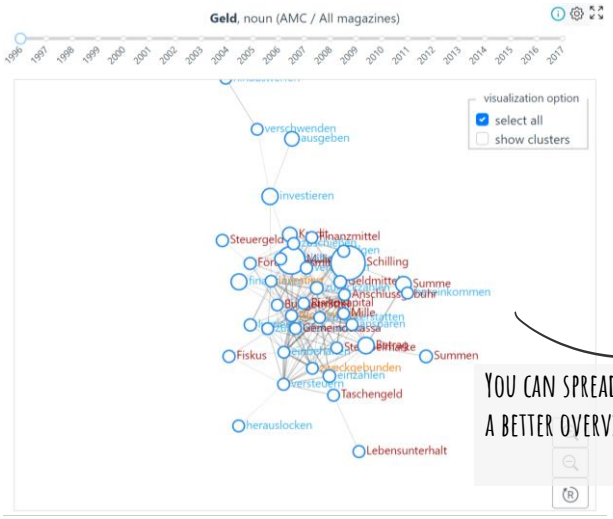
In your input form on the left, you can select the corpus (AMC or PARLAT), the Subcorpus, e.g. a specific newspaper, and type the target word.

General Networks comprise the speeches of **a specific politician** or **a whole party** and are much larger. You need to use a filter to restrict the number of nodes, before you can display the network.



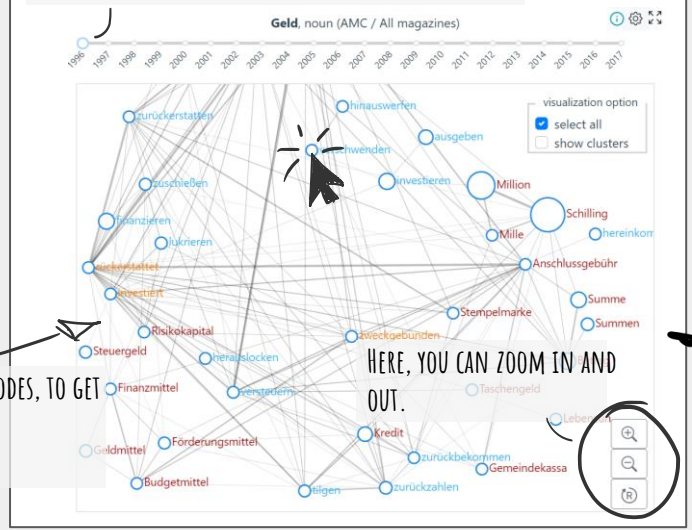
Choose the party and the politician on the left. Additional filters like in the example can show only 20% of the words with the lowest Degree Centrality.

When you click on „Visualize" it looks a bit like this:



You can spread out the nodes, to get a better overview.

On top on the timeline you can choose the year.



Here, you can zoom in and out.

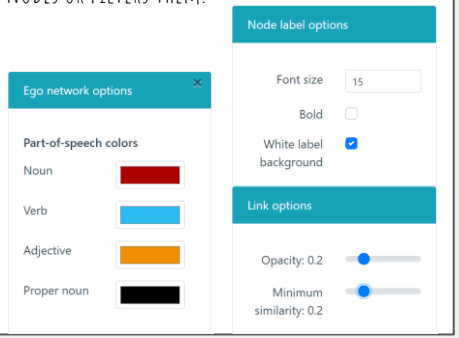
Why are the circles a different size?

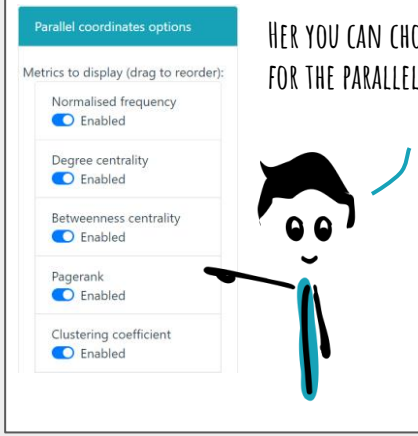
Size stands for Frequency, so how often a word is used. The bigger the circle, the more frequently it can be found in the corpus.

And what do the different colours mean?

They stand for different parts of speech. Nouns are red, verbs blue, adjectives yellow and proper names are black. But you can change that in the setting in the cogwheel on the top right.

In the sidebar that opens when you click the cogwheel, you can change the formatting of the annotations. „Link options" lightens the connections between the nodes or filters them.

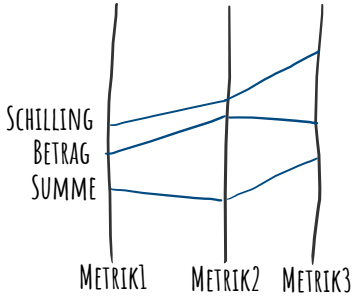




Her you can choose the metrics for the parallel coordinates.

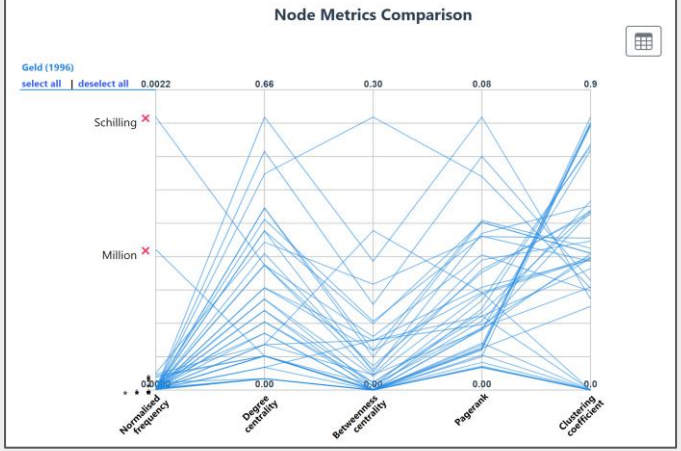
What coordinates?

Different metrics are calculated for every word in the network. Each of these Metrics is shown on an individual axis in the Parallel Coordinates Plot.



So, each line stands for a word in the network and every intersection with a vertical axis stands for the respective value in the metric..

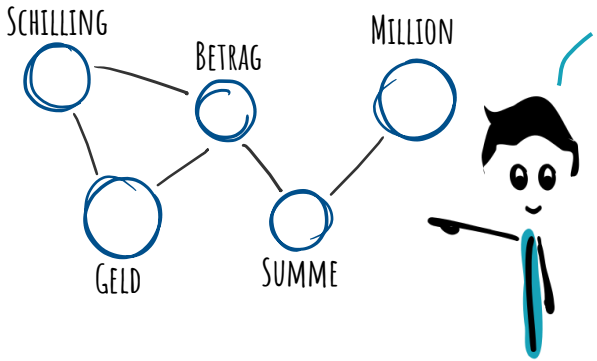
Im DYLEN-Tool findet man diese Visualisierung oben rechts:



Aber was bedeuten denn diese Metriken überhaupt?

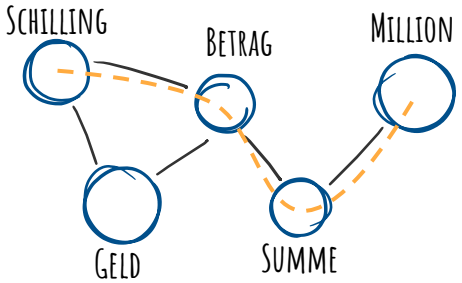
Die Frequency kennen wir ja schon. Lass uns noch einen Blick auf ein paar der anderen Metriken werfen...

Degree Centrality shows the number of connections with other words in the network.



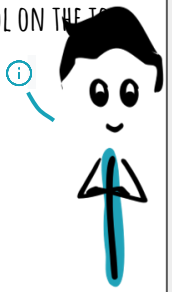
„**Million**" has the lowest Degree Centrality, because it's only connected to one other word („Summe"-sum). „**Betrag**"(amount) has three connections and therefore the highest Degree Centrality.

Betweenness Centrality defines how many other words are in between different words in the network.



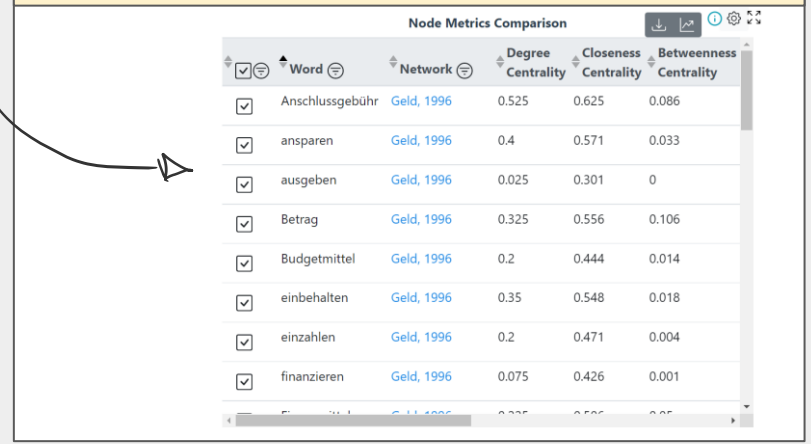
If we look for the shortest path from „Schilling" to „Million", „Betrag" and „Summe" are in **between**. The more paths pass through a node, the higher its Betweenness Centrality.

Every metric is explained in the tool in more detail in the tool. Just click on the information symbol on the t… right.

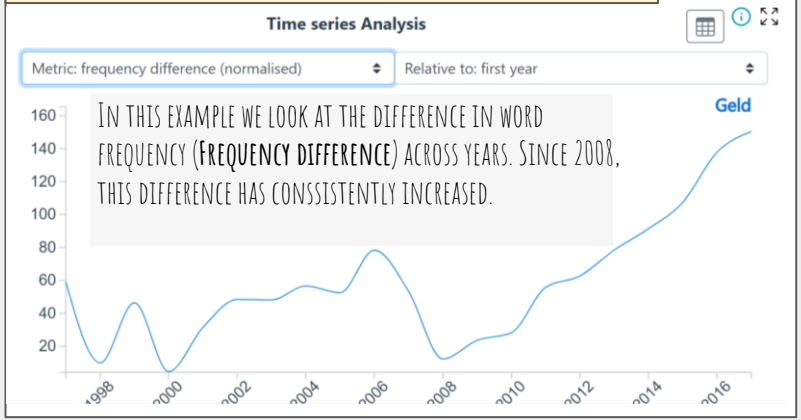


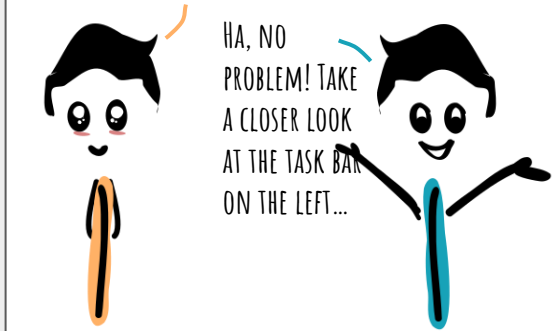
Visualisations are connected with one another. When you hover your moues over a node in the network, the word is highlighted in the parallel coordinates too – and the other way around.

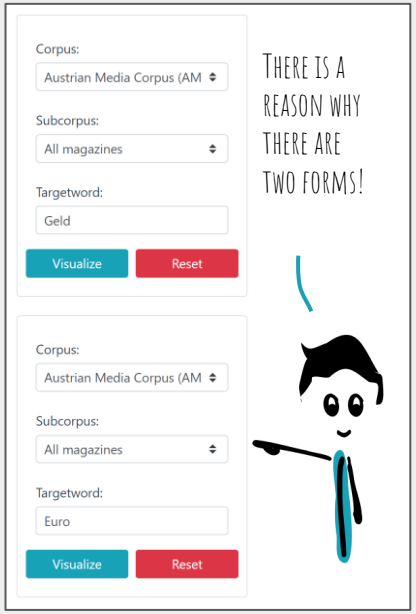
One click on the table symbol shows all words and metrics in detail. You can sort, filter and export the table.



Finally, at the bottom right, there is the Time Series Analysis. Here you can inspect the metrics of your target word.







There is a reason why there are two forms!

Both networks will be displayed one below the other.

The words from the bottom network are on the **right**. It shows the **lines in pink**.

In the line chart **the second target word** is displayed in **pink** too.