**Probs / fragen**

- flesch-kincait bei nachberechnung teilweise anderes ergebnis auch wenn silbenzahl korrekt

- Silbenzählen funzt nicht (Python module Pyphen)

--> Zusätzliche Spalte für Silben gewünscht?

--> wenn nach dem satz ein leerzeichen = +1 Silbe

- Gunning Fox bläst auf 100 Wörter auf, wenn weniger abgebildet

- ARI klappt ganz gut ==> zeichen ohne leerzeichen

**FRE deutsch einfügen**

- teilsätze noch offen **🡪 satzzeichen zählen, punkt, komma, semikola, doppelpunkte**

- wie wortfrequency bei ganzen sätzen/texten? --> **SUM(wortfrequenzen)/ANZ(Wörter)?**

**- Nomen/Verben/Adjektive**: Tagger muss installiert und trainiert werden da deutsch == **NLTK**?   
--> Anwendung auf workstations oder einem server? 🡪 tree-tagger! 🡪 zusätzlich trainieren.. geht auch ‚untrainiert‘ = Standardform kommt mit lexicon!!! in zellen zerlegen als zwischenschritt 🡪 daraus mittelwert aus wortfrequenzen 🡪 mittlere wortfrequenz für CONTENTWÖRTER

Treetagger = stuttgarter tagset?

- txt einlesen in excel optional im skript oder **extraskript**?

- speichern in fu.box? 🡪 In github arbeiten!!!

TODO

Github

Skript txt 🡪 Excel

Case Sensitiv? ‚Leben‘ im SUBTLEX. ‚leben‘ nicht drin

SUBTLEX hat Verkehrsunfall, Verkehrsunfalls, Verkehrsunfällen

SUBTLEX hat nicht Verkehrsunfälle

🡪 Genaues Wort vergleichen oder Lemma?

GitHUB Jana: janafu

Doppelte Worte im Korpus Mergen! 🡪 doppelte gibt es nicht!