

Hypotheses

—
SpaCy's dependency abbreviations for German:

sb = subject

oa = accusative object

og = genitiv object

da = dative

nk = noun kernel element

pd = predicate

oc = clausal object
—

1. Baum: sb, oa/og/da

most probable partner word with sb?

1.1. We suspect that the word "Wald" co-occurs most frequently as a partner word of "Baum" with the dependency sb. As an example serves the following sentence: Der Baum im Wald wurde gepflanzt. → Our hypothesis is based on the fact that "Baum" and "Wald" belong to the same topic area.

most probable partner word with oa/og/da?

1.2. We hypothesize that a verb will be the most frequent partner word with the constructions oa, og and da. Probably "fällen" or "pflanzen". We settle on the word "pflanzen". → This hypothesis is again based on the fact that "Baum" and "pflanzen" belong to the same topic area.

What influence will different text sorts have?

1.3. Based on the fact that climate change is a growing topic in news articles we suspect that the word "Baum" does occur more often in news articles than in Wikipedia articles.

What difference could different time periods make?

1.4. We suspect that - as stated above - that there are more occurrences of the word "Baum" in more current News articles based on the fact that climate change is an ever growing topic in news since the Paris Climate Agreement in 2015.

1.5. We assume that the number of occurrences of "Baum" in Wikipedia articles did not grow vastly. Maybe a slight increase - which must seem a small number compared to the assumed increased number in hypothesis 1.4.

2. schön; pd, nk

most probable partner word with pd:

2.1. The most occurring predicate (pd) is “sein” in any conjugation. An example: “Der Tisch ist schön.” Whereas the dependency pd would point from “ist” to “schön”. Well there are many many predicates (verbs) that can be modified (amod) by the word “schön”.

most probable partner word with nk:

2.2. For this construction to predict a most probable partner word with “schön” is very hard since in this construction, the adjective “schön” serves as a describer or modifier of a noun. We hypothesize that the most probable partner word is “Wetter”. An exemplary sentence could be: “Morgen werden wir schönes Wetter haben.”

What influence will different text sorts have?

2.3. We will research the content between Webpages and News articles and we could imagine that “schön” occurs more frequently on webpages - for example webpages for commercial purposes including a webshop - and to market their product they might use the word “schön” as a word to describe their products in a positive way.

What difference could different time periods make?

2.4. Since societal beauty standards are changing more and more rapidly, we would expect it to become a less important factor for judging the quality of something. Thus, we would expect for the word to occur less often in more recent texts.

3. sein (old: ist); oc, ROOT

most probable partner word with oc:

3.1. We presume that the partner word as a clausal object is most probably an adjective. We think that the most probable object that would appear is the word “gut”, since it is a very common word that can occur almost context-independently.

most probable partner word with ROOT:

3.2. We assume that most words that have ‘sein’ as root verb, will be a predicate. We expect it to be an adjective, however, could not make a prediction on what the most common one could be, since ‘sein’ is such a broughtly used verb.

What influence will different text sorts have?

3.3. We hypothesize that there are approximately as many occurrences of the lemma “sein” (or any conjugation of it) in a Web corpus as in a Wikipedia corpus. → This hypothesis is based on the assumption that conjugations of “sein” are very frequently used words, regardless of the text sort.

What difference could different time periods make?

3.4. We think that there are no (significant) differences between time periods.