


DIE WICHTIGSTEN KORPUSLINGUISTISCHEN BEGRIFFE

Begriff	Erklärung	Beispiel								
Korpusmanager	etwa wie Suchmaschine im Internet	<i>CQP, (Corpus Query Processor), Cosmas, Bonito...</i>								
Suchfeld	Abfragefenster									
Suchanfrage:										
Word	Wortform: beliebige Buchstaben/ Graphemhäufung	<i>saß, Ach, Ring, usw,</i>								
Interval	Wortteil, Affix (Vorsilbe/Nachsilbe)	Suche nach: <i>~saß, blöd~, ~ring~, ~erl</i>								
Wortverbindung	mehrere Wörter	<i>an der Quelle, taube Nuss</i>								
Lemma	Grundform eines Lexems; Lexikoneintrag (Nominativ Singular, bzw. Infinitiv Präsens)	<i>Apfel (zu Apfels, Äpfel, Äpfeln); hören (zu höre, höret, hörend, gehört...)</i>								
Tag	Morphologische Kategorie (falls das Korpus getaggt ist); in jedem Korpus andere Kategorien (je nach Sprache) und anderes System (Manual notwendig!)	Suche nach allen Interjektionen: <i>ITJ</i> ^(InterCorp) Suche nach allen Imperativen: <i>VAIMP</i> ^(InterCorp)								
Suchoperatoren	Wörter oder Zeichen für kompliziertere Abfragen: z.B. beim suchen des Verbs <i>einigen</i> ; nicht das Adjektiv <i>einig(en)</i> (Manual notwendig!)	<i>nicht, oder, ...</i> &,								
Platzhalter	Zeichen zum Suchen von gleichzeitig mehreren Formen, die einen gemeinsamen Teil haben – z.B. alles Wörter, die auf <i>~erl</i> enden. (Manual notwendig!)	<i>.*erl</i> ^(InterCorp) <i>++erl</i> ^(DeReKo)								
Token	laufende „Wort“form, alle vom Computer analysierten sprachliche Zeichen („ <i>Wort ist ein Zeichen</i> “ de Saussure)	<i>oder so was , ich sitze an der Quelle . “ (11 tokens)</i>								
Konkordanz/Konkordanzzeile	Derjenige Teil des Textes, in dem das Gesuchte vorkommt	<table border="1" data-bbox="1279 1225 2018 1315"> <tr><td>A98</td><td>Auslöser war eine einseitige Sprachenpolitik und Unterdrückung der</td></tr> <tr><td>E98</td><td>etwa durch die einseitige Sprachenpolitik und die Arbeitsplatz-</td></tr> <tr><td>N98</td><td>Bedeutung einer systematischen Sprachenpolitik und die Förderung der</td></tr> <tr><td>P93</td><td>sein, gilt es umzudenken, was Sprachenpolitik und -unterricht</td></tr> </table>	A98	Auslöser war eine einseitige Sprachenpolitik und Unterdrückung der	E98	etwa durch die einseitige Sprachenpolitik und die Arbeitsplatz-	N98	Bedeutung einer systematischen Sprachenpolitik und die Förderung der	P93	sein, gilt es umzudenken, was Sprachenpolitik und -unterricht
A98	Auslöser war eine einseitige Sprachenpolitik und Unterdrückung der									
E98	etwa durch die einseitige Sprachenpolitik und die Arbeitsplatz-									
N98	Bedeutung einer systematischen Sprachenpolitik und die Förderung der									
P93	sein, gilt es umzudenken, was Sprachenpolitik und -unterricht									
KWIC	„key word in context“: das Gesuchte in seinem (natürlichen) Umfeld	<table border="1" data-bbox="1279 1337 2018 1426"> <tr><td>A98</td><td>Auslöser war eine einseitigeSprachenpolitikund Unterdrückung der</td></tr> <tr><td>E98</td><td>etwa durch die einseitigSprachenpolitikund die Arbeitsplatz-</td></tr> <tr><td>N98</td><td>Bedeutung einer systematischeSprachenpolitikund die Förderung der</td></tr> <tr><td>P93</td><td>sein, gilt es umzudenken, wasSprachenpolitikund -unterricht</td></tr> </table>	A98	Auslöser war eine einseitige Sprachenpolitik und Unterdrückung der	E98	etwa durch die einseitig Sprachenpolitik und die Arbeitsplatz-	N98	Bedeutung einer systematische Sprachenpolitik und die Förderung der	P93	sein, gilt es umzudenken, was Sprachenpolitik und -unterricht
A98	Auslöser war eine einseitige Sprachenpolitik und Unterdrückung der									
E98	etwa durch die einseitig Sprachenpolitik und die Arbeitsplatz-									
N98	Bedeutung einer systematische Sprachenpolitik und die Förderung der									
P93	sein, gilt es umzudenken, was Sprachenpolitik und -unterricht									

NJ KORP Korpusová lingvistika (Káňa)

Beleg/Treffer	Ergebnis einer Korpussuche, belegt die (Nicht-)Existenz des Gesuchten im Korpus. (Interpretation notwendig!)		
Parallelen	Einsprachige Teile in einem Parallelkorpus	<i>de – en - it</i>	
Segment	Entsprechende Passage in einer anderen Sprache. Entsprechung, die durch Alignment entstanden ist.	» <i>Wir sind nämlich nicht blöd !</i> «	" <i>'Cos we 're not stupid !</i> "
Alignment	Anordnung der sich entsprechenden Textpassagen im Parallelkorpus.		
Kookkurrenz/ Kollokation	semantische oder grammatikalische Verbindungen	die kantonale Sprachenpolitik Tagung [...] Sprachenpolitik in Mittel einseitige Sprachenpolitik und ... der europäischen Sprachenpolitik Sprachenpolitik [klar] definiert Sprachenpolitik des Kantons Sprachenpolitik in Europa	
Frequenz/ Häufigkeit	Absolute oder relative Häufigkeit eines Wortes, Lexems etc. im Korpus (statistische Angabe)		
Mi-Score/ T-Score	„mutual information“(statistische Angabe)		
LLR	„log-likelihood-ratio“ (statistische Angabe)		
corpus-based/ korpusbasiert	verifiziert bisherige Hypothesen über Sprache(n)	Ist das Wort <i>Schmankerl</i> tatsächlich nur ein Austriazismus?	
corpus-driven/ korpusgesteuert	betrachtet Korpus als „Ausgangspunkt“ für die Beschreibung der Sprache	Was kommt rechts vom Wort <i>Schmankerl</i> vor?	