

COSMAS

(Corpus Storage, Maintenance and Access System)

1. Möglichkeiten von COSMAS

- Suchen und Kombinationssuchen in einem vordefinierten Korpus. Die vordefinierten Korpora sind in Anlage 1 aufgelistet.
- In der Suche können folgende Suchobjekte verwendet werden:
 - Worte im lexikalischen oder grammatischen Sinn,
 - Satzzeichen,
 - Sonderzeichen,
 - Ziffern, Ziffernfolgen,
 - Wortsegmente über Wildcards (* für Zeichenfolge, ? für Einzelzeichen),
 - Grundformen (mit flektierten Formen) über Grundformenoperator & ,
 - morphosyntaktische Annotationen¹,
 - logische Operatoren: und, oder, nicht,
 - Abstandsoperatoren: Wortabstand, Satzabstand, Absatzabstand, Zeilenabstand.

2. Benutzeroberflächen

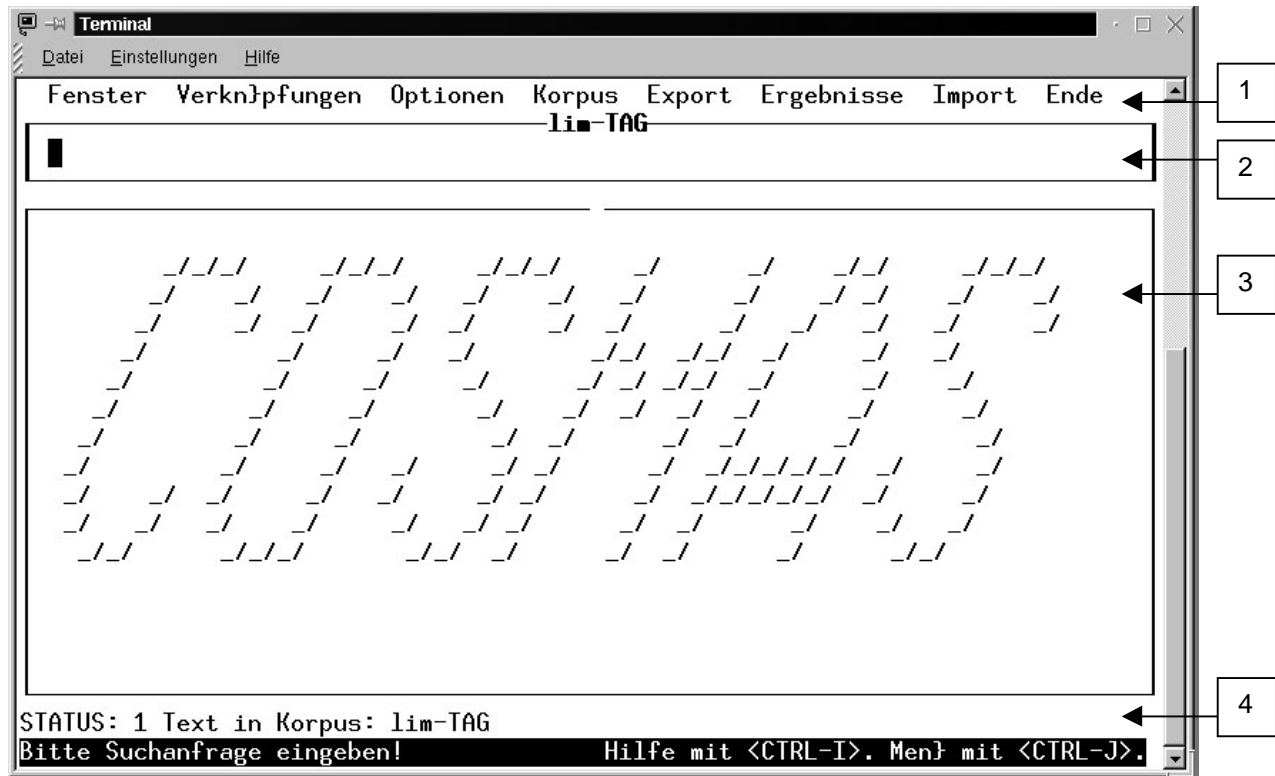
- Das IDS bietet zwei Benutzeroberflächen zum COSMAS an:
 - Die klassische Terminal-Oberfläche, die über Telnet erreichbar ist (Befehl: *telnet corpora.ids-mannheim.de*², Login-ID: *cosmas-demo*, Passwort: [Enter-Taste], anschließend Eingabe der E-Mail-Adresse, diese wird aber nicht überprüft.). Diese Testversion soll sich laut IDS selbständig nach einer halben Stunde beenden. Es ist daher damit zu rechnen, daß die Verbindung plötzlich abbricht. Die bisherigen Suchergebnisse sind danach nicht mehr vorhanden.
 - Die WWW-Schnittstelle des COSMAS ist mit jedem Browser, der Framesets erkennt³, erreichbar unter *http://www.ids-mannheim.de/~cosmas/*. Sie ist auf den ersten Blick etwas unübersichtlich, aber von der Bedienungsführung doch benutzerfreundlicher (eher „Guided Tour“). Die zahlreichen Tastaturkurzbefehle (s. 2.1.) fallen weg, und auch der gesamte Abfragevorgang wird transparenter. Die WWW-Schnittstelle befindet sich jedoch noch in der Beta-Phase (d.h. längerfristige Erprobung), so daß einige Möglichkeiten hier noch nicht gegeben sind.

¹ Nur in speziell morphosyntaktisch annotierten Korpora!

² Einzugebende Befehle werden in kursiv, Eingabebegriffe in [] gesetzt.

³ z.B. Netscape Σ4.0 oder Internet Explorer Σ4.0

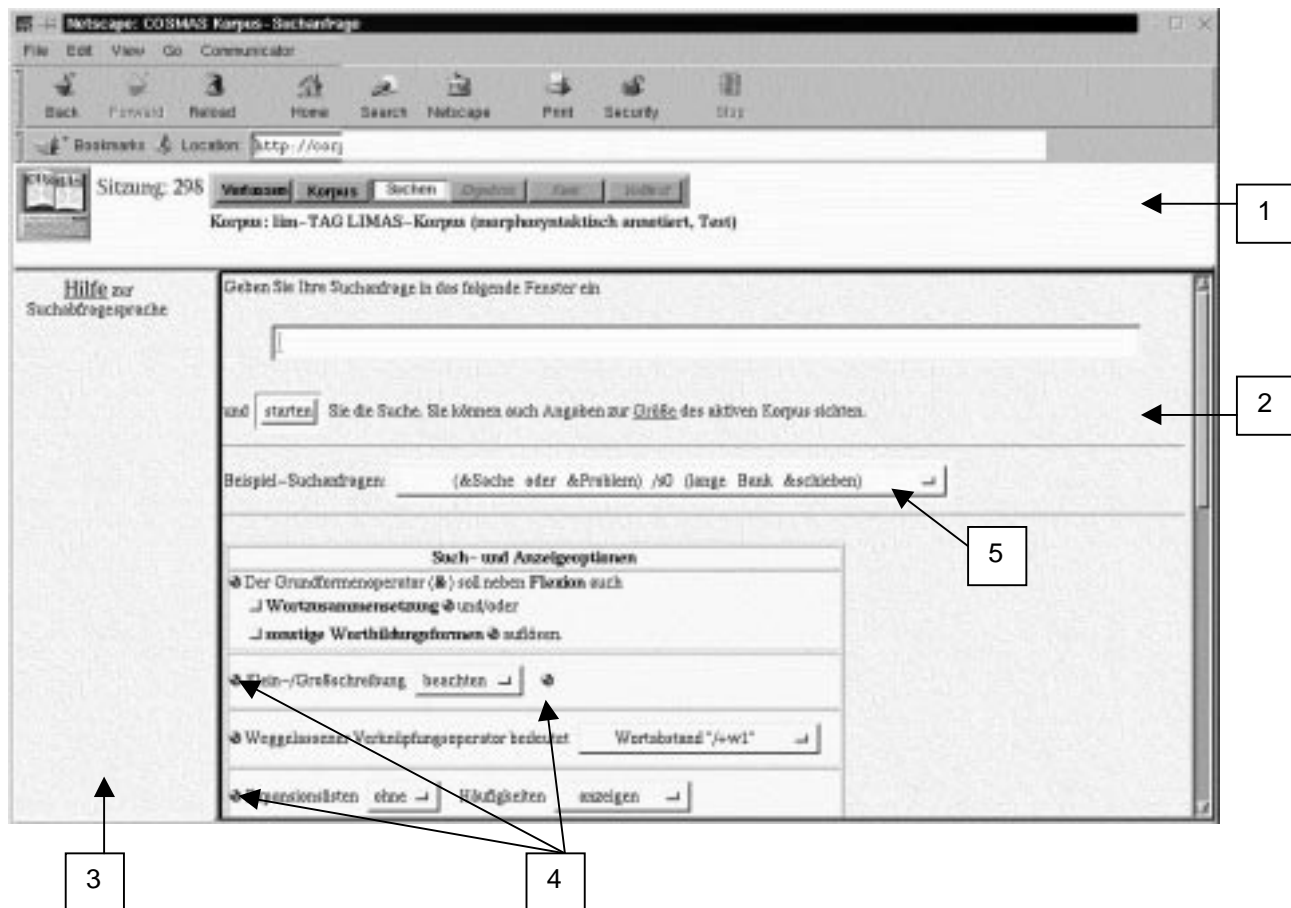
2.1 Terminal-Oberfläche über Telnet



- Das Terminal-Fenster gliedert sich in 4 Abschnitte: Befehlsleiste (1), Anfrageleiste (2), Ausgabefenster (3) und Statusleiste (4). Der wichtigste Befehl zum Wechsel in die Befehlsleiste (1) ist [Strg J], bzw. [Ctrl J]. Hiermit wechselt man zwischen Sucheingabe in (2) und Befehls-/Optionseingabe in (1). Die Menüpunkte unter (1) sind unterteilt in:
 - Fenster: Suchanfrage, Wortformenliste, Grundformenliste, morphosyntaktische Annotationen, Zwischenergebnisse, Ergebnisse, frühere Suchen.
 - Verknüpfung: und, oder, nicht, Wortabstand, Zeilenabstand, Satzabstand, Absatzabstand.
 - Optionen: Einstellen (Anzeigeoptionen, Häufigkeiten, Zufallsauswahl, Lemmatisierung, Suchmodalitäten), Zeichensätze (Tastatur, Bildschirm, Import- und Exportdateien).
 - Korpus: Vordefiniert, Laden, Sichern, Texte anwählen, Texte abwählen, Gruppe anwählen, Gruppe abwählen, Alles abwählen, Suchergebnis >> Korpus, Suchergebnis >-Korpus, Zufallsauswahl, Korpusstatistik.
 - Export: Definitionen, Typ, Ziel, Aktueller Text, Gesamt, Markiert, Wortliste.
 - Ergebnisse: Textstatistik, Belege alphabetisch, Definitionen, Wörter alphabetisch, Frequenz sortieren, Kollokationsanalyse, Analyseparameter.
 - Import: Suchanfrage, Wortliste, Importdatei löschen, Importdatei zeigen.
 - Ende: Ende, Tasten, Version.

- Die Wichtigsten Befehle:
 - Fenster --> Suchabfrage: Wechseln aus dem Menü zur Suchabfrage.
 - Fenster -->frühere Suchen: Verschnellert das Arbeiten.
 - Korpus --> Laden
 - Ende --> Ende
 - [Esc]: Einen Schritt im Abfragevorgang (s. 3.) zurück.

2.2 WWW-Oberfläche



- Die Adresse: <http://www.corpora.ids-mannheim.de/~cosmas/>
- Die WWW-Oberfläche ist in 3 Fenster unterteilt:
 - Statusfenster oben (1),
 - Ein-/Ausgabefenster rechts (2),
 - Informationsfenster links (3).
- Hilfe gibt es über die roten Knöpfe (4) → einfach die Maus darüber halten.
- In einem Drop-Down-Fenster (5) sind Beispiele für Suchabfragen eingetragen. Diese werden nicht automatisch übernommen, helfen aber bei der Orientierung und bei syntaktisch falschen Suchbefehlen.
- Das Menü der Statusanzeige (1) entspricht dem Abfragevorgang, siehe auch 3.: Verlassen, Korpus, Suchen, Ergebnis, Kwic, Volltext.

3. Abfragevorgang (Nur frei zugängliche Versionen):

	WWW	Terminal
1	<p>Im Browser Adresse eingeben: <i>http://www.ids-mannheim.de/~cosmas/</i></p> <p>Im mittleren Fenster den Start-Knopf drücken.</p>	<p>Im Terminal-Fenster eingeben: <i>telnet corpora.ids-mannheim.de</i> login:> <i>cosmas-demo</i> password > [Enter] E-Mail-Adresse: [Beliebige E-Mail-Adresse]</p>
2	<p>Auswahl des Korpus mit der Maus und Aktivieren-Knopf drücken.</p>	<p>Auswahl des Korpus mit den Cursor-Tasten und Enter.</p>
3	<p>Suchbefehl eingeben (s. 4). Suchoptionen im Ein-/Ausgabefenster auswählen:</p> <ul style="list-style-type: none"> - Grundformenoperator soll auch Komposita und Flexionsformen erkennen, - Groß- und Kleinschreibung beachten/nicht beachten, - Automatischer Verknüpfungsoperator, - Expansionslisten der Wildcard-Ergebnisse anzeigen - Maximale Anzeigemenge - Anzeigemenge in der Ansicht. <p>Im unteren Teil des Fensters erscheint hier auch die Statistik des Korpus. Suchbefehl ausführen mit Suchen-Knopf</p>	<p>Suchbefehl eingeben (s. 4). Die Suchoptionen können unter dem Menüpunkt [Optionen] → [Einstellen] verändert werden. Zu den Möglichkeiten siehe linke Spalte.</p> <p>Suchbefehl mit [Enter] bestätigen.</p>
4	<p>Je nach Einstellung der Expansionslisten: Anzeige der Flexionsformen und Komposita mit Auswahlmöglichkeit</p>	<p>Anzeige der Flexionsformen und Komposita mit Auswahlmöglichkeit. Auswahl über + oder -.</p>
5	<p>KWIC-Anzeige mit erweiterter Auswahl. Option: Volltext-Kontextgröße mit Satz-, Wort- und Absatzumgebung.</p>	<p>KWIC-Anzeige mit Auswahlmöglichkeit für die einzelnen Beispiele. Auswahl über + oder -.</p>
6	<p>Anzeige des Volltexts mit Quellenangabe.</p>	<p>Anzeige des Volltexts mit Quellenangabe. Dessen Größe wird von den Einstellungen im Menü [Optionen] → [Einstellen] bestimmt.</p>

4. Zur Abfrage-Syntax

4.1. Grundformen mit Grundformenoperator:

&suchen → *suchte, gesucht, suchend, such, suchen, ...*

&arbeiten → *arbeitest, arbeitete, gearbeitet, ...*

&Donau → *Donaudampfschiff(f)ahrtskapitänsmützen.....*

&Stift → *Bleistift, Buntstift, Stiftehalter,...*

4.2. Groß-/Kleinschreibung: Ignorierungsoperator:

(Falls nicht bereits in [Optionen] aktiviert)

\$sucht → *sucht (V), Sucht (N)*

4.3. Platzhalteoperatoren:

*we*en* → *weggehen, werfen, wegen*

we?en → *wegen, weben,...*

we??en → *werfen, wecken, wessen,...*

**we?en* → *Bewegen, weswegen, derentwegen,...*

we????????????????en → *wettervorhersehen*

4.4. Abstandsoperatoren:

/w_n: maximale Wortabstände

Tag /w2 offenen → *Tag der offenen Tür, einen offenen Tag,...*

Tag /+w2 offenen → *Tag der offenen Tür*

Tag /-w2 offenen → *einen offenen gelungenen Tag*

/s_n → Satzabstände

/a_n → Absatzabstände

/z_n → Zeilenabstände des Originaltexts.

für */w* ist *n=0* möglich → darauffolgendes Wort.

für */s /a /z* ist auch *n=0* möglich → Ergebnisse im selben Satz/Absatz/Zeile.

4.5. Logische Operatoren:

x und y → alle Texte, in denen beide Begriffe vorkommen.

x oder y → alle Texte, in denen entweder *x* oder *y* vorkommen

x nicht y → alle Texte, in denen *x* vorkommt und *y* nicht vorkommt.

Logische Operatoren beziehen sich immer auf ein **Text!**

4.6. Besonderheiten:

4.6.1 Satzzeichen:

werden in Anführungszeichen gesetzt.

willkommen /w0 "!" → Quakenbrück heißt Sie willkommen!

4.6.2. Satzenden werden mit "<s>" markiert:

&haben /w0 "<s>" → ...weil Ihr sie gesehen habt.

4.6.3. Wörter wie *und*, *nicht* und *oder* werden in Anführungszeichen gesetzt.

bevor /w1 "nicht" → ...bevor Du nicht gestehst...

4.6.4. Wortformen, die im Text mit Sonderzeichen enden (Abkürzungen) werden ohne dieses gesucht:

etc → etc.

Prof → Prof.

4.6.5. Klammerungen bei Mehrfachsuchen:

Bei Verwendung von mehr als einem Abstandsoperator sind runde Klammern zu setzen:

(Tag /+w2 offenen) /+w1 Tür → Tag der offenen Tür

das gilt auch bei Abstands- und logischen Operatoren:

(Landschaft oder Gegend) /s0 &schauen → ...als er in die Gegend, die vor ihm liegt, schaut,...

4.6.6. Trennbare Wörter:

&herausfordern oder (&fordern /+s0 heraus) → herausfordert, forderte ihn heraus,...

5. Morphosyntaktische Annotationen (in Limas-Korpus)

Annotationen beginnen alle mit #ANA und sind daher als eigenständige Suchbegriffe suchbar.

Beispiele:

#ANA#AUX#pl → Hilfs-/Modalverben im Plural

#ANA#ADJ*sup* → alle Adjektive im Superlativ

#ANA#NOU*pl*acc* → alle Nomen im Akkusativ Plural

Hierarchische Liste der Annotationen: #ANA +

<p>#ADJ -->Adjektive</p> <ul style="list-style-type: none"> #desc #ger #pastpart #prespart #com #pos #sup #sg #pl #m #f #n #nom #acc #dat #gen 	<p>#ADP --> Adposition</p> <ul style="list-style-type: none"> #post #pre #p-acc #p-dat #p-gen #DET #acc #dat
<p>#ART -->Artikel</p> <ul style="list-style-type: none"> #def #idef #dem #int #lim #neg #pos #pl #sg #m #f #n #nom #gen #dat #acc 	<p>#ADV -->Adverbien</p> <ul style="list-style-type: none"> #com #post #sup <p>#AUX -->Hilfs-/Modalverben</p> <ul style="list-style-type: none"> #pl #sg #m #f #n #1 #2 #3 <p>#CON -->Konjunktionen</p> <ul style="list-style-type: none"> #coor #sub #c #s

#NOU -->Nomen

#com

#prp

#sg

#pl

#m

#f

#n

#nom

#gen

#dat

#acc

#NUM -->Numeralien

#ca

#fr

#or

#sg

#pl

#m

#f

#n

#nom

#gen

#dat

#acc

#PRN -->Pronomen

#dem

#int

#pers

#ref

#rel

#sg

#pl

#m

#f

#n

#nom

#gen

#dat

#acc

#1

#2

#3

#VRB -->Verben

#npas

#pas

#indic

#inf

#imper

#subj

#pas

#pres

#fut

#sg

#pl

#m

#f

#n

#1

#2

#3