

# lexpan – Lexical Pattern Analyzer

---

*Bedienungsanleitung – verfasst von Annelen Brunner*

## Inhalt

Einführung .....	2
Erster Programmstart (und Updates).....	2
Programmüberblick.....	3
Bestandteile des Programms.....	3
Grundlegende Funktionsweise.....	3
Hauptfenster .....	3
Archive.....	5
Anlegen eines Archivs.....	5
Stores.....	5
Stores konvertieren.....	5
Details zum Store.....	6
Hinweise zu Import-Formaten .....	6
COSMAS II.....	6
Sketch Engine .....	8
Einfaches Textformat .....	9
Suchanfragesyntax .....	10
Überblicksdateien.....	11
Details zu Suchanfragen .....	11
KWIC-Liste .....	11
Füllertabellen.....	11
Kookkurrenzlisten.....	12
Annotationen.....	12
Tags.....	12
Formatierungsregeln für Tags .....	12
Sortieren und Filtern nach Tags .....	13
Kommentare.....	13
Sortierungsnummern .....	13
Export .....	14
KWICs (de)aktivieren .....	17

Fehler- und Warnmeldungen .....	17
Warnmeldung bei der Konvertierung von KWIC-Listen .....	17
Überzählige < > -Zeichen .....	18
Unvollständige KWIC-Zeilen .....	18
Neu konvertierter Store hat 0 Zeilen .....	18
Programm meldet, es laufe bereits, ist aber nicht sichtbar .....	19
Archiv kann nicht geladen werden.....	19
Tag, Kommentar oder Sortierungsnummer wurde nicht gespeichert .....	19

## Einführung

Das Programm *lexpan* dient der explorativen Untersuchung von Festigkeit und Varianz von syntagmatischen Mustern. Es hat folgende Hauptfunktionalitäten:

- Definieren von Suchanfragen über KWIC-Zeilen; Bündeln und Darstellen der KWIC-Zeilen auf Basis dieser Suchanfragen
- Auszählen von Lückenfüllern für variable Suchanfrage-Elemente
- Annotieren von Lückenfüllern mit Tags und Kommentaren und Sortierung auf Basis der Tags
- Einlesen von Kookkurrenzlisten; Annotieren der Kookkurrenzcluster mit Tags und Kommentaren und Sortierung auf Basis der Tags

## Erster Programmstart (und Updates)

*lexpan* ist ein Java-Programm und betriebssystemunabhängig, läuft also unter Windows, Linux und MacOS, wenn eine aktuelle Java-Version installiert ist. Es wird in Form einer Archiv-Datei (.zip-Datei) ausgeliefert.

Schritte zur ersten Ausführung von *lexpan*:

- Falls nötig: Aktuelle Version des Java Runtime Environments herunterladen und installieren (<https://www.java.com/de/download/>)
- Aktuelle .zip-Datei mit dem Namen *lexpan\_Jahr\_Monat\_Tag.zip* (z.B. *lexpan-2016-06-13.zip*) auf der eigenen Festplatte speichern.
- WICHTIG! Die .zip-Datei entpacken (Unter Windows: Rechtsklick auf die Datei und "alle extrahieren..." wählen).
- Dadurch entsteht ein neuer Ordner mit den Unterverzeichnissen „lib“ und „conf“ und den Dateien „lexpan.jar“ und „lexpan\_Bedienungsanleitung.pdf“ und „README.txt“.
- Dieser *lexpan*-Programmordner kann an einer beliebigen Stelle auf der Festplatte liegen. Die ursprüngliche zip-Datei kann jetzt gelöscht werden. Das Programm *lexpan* wird mit Doppelklick auf „lexpan.jar“ gestartet.

Wenn eine neue Version von *lexpan* verfügbar ist, wird diese auf die gleiche Weise heruntergeladen und entpackt. Der alte *lexpan*-Ordner kann einfach durch den neuen ersetzt bzw. gelöscht werden.

# Programmüberblick

## Bestandteile des Programms

Im *lexpan*-Programmordner befinden sich folgende Elemente:

- **lexpan.jar**: Durch Doppelklicken auf diese Datei wird das Programm gestartet
- **lib**: Ordner mit Zusatzbibliotheken (*muss vom Benutzer nie geöffnet werden*)
- **conf**: Ordner mit diversen Hilfsdateien (*muss vom Benutzer nie geöffnet werden*)

Folgende Dateien werden beim ersten Ausführen des Programms automatisch angelegt:

- **config.properties**: Konfigurationsdatei (speichert Benutzereinstellungen, wie z.B. Exportmodus und das letzte gewählte Archivverzeichnis; fehlt diese Datei, wird sie mit Default-Einstellungen neu angelegt)
- **log.txt**: Log-Datei (speichert Warnungen und Fehlermeldungen des Programms für Debugging-Zwecke)

## Grundlegende Funktionsweise

*lexpan* arbeitet immer auf Grundlage von Daten, die aus einem Korpusanalysesystem exportiert wurden. Dies sind entweder Belege in der Form von KWIC-Zeilen (Keyword in Context) oder Kookkurrenzlisten. Wenn diese Daten in *lexpan* eingelesen sind, werden sie [Stores](#) genannt. Es ist nicht möglich, vollständige Texte oder gar ein komplettes Korpus in *lexpan* einzulesen.

Die Analysen in *lexpan* sind in [Archiven](#) organisiert. Einem Archiv können beliebig viele Stores zugeordnet werden. Technisch bedeutet dies, dass diese Stores in einer gemeinsamen Datenbank verwaltet und im gleichen Unterordner abgelegt werden.

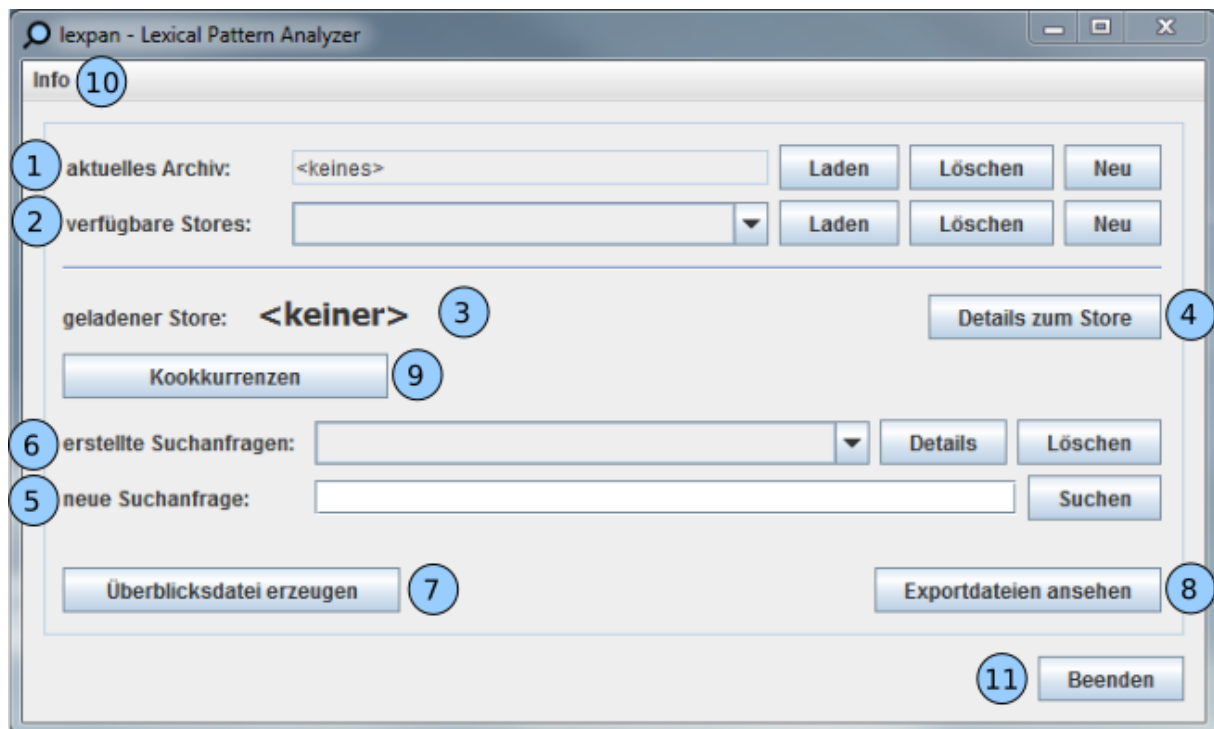
Stores, die aus KWIC-Zeilen bestehen, können mit Hilfe von Suchanfragen durchsucht werden, die auf regulären Ausdrücken beruhen (siehe [Suchanfragensyntax](#), [Details zu Suchanfragen](#)). Wenn die Suchanfragen lexikalisch nicht fest bestimmte Elemente enthalten (sogenannte „Lücken“), können [Füllertabellen](#) erstellt werden, die die Frequenzen der lexikalischen Lückenfüller auflisten.

Die Füllertabellen können mit diversen [Annotationen](#) versehen werden, die auch benutzt werden können, um das lexikalische Material gezielt zu filtern. Bei Stores, die aus [Kookkurrenzlisten](#) bestehen, können die Kookkurrenzcluster nach dem gleichen Prinzip annotiert und gefiltert werden.

*lexpan* erlaubt zudem den [Export](#) von unterschiedlichen Tabellen im XML oder HTML-Format.

## Hauptfenster

So sieht das Hauptfenster von *lexpan* aus, das beim Start sichtbar wird:



- (1) Im Textfeld steht der Pfad des aktuell geladenen Archivs. Mit „Laden“ kann man einen Archiv-Ordner von der Festplatte auswählen, mit „Löschen“ einen Archiv-Ordner entfernen und mit „Neu“ einen neuen Archiv-Ordner erstellen (siehe Abschnitt [Archive](#)).
- (2) Das Textfeld enthält eine Drop-Down-Liste der aktuell verfügbaren Stores. Mit „Laden“ kann der ausgewählte Store geladen werden, mit „Löschen“ wird der ausgewählte Store gelöscht und mit „Neu“ wird ein neues Fenster geöffnet, in dem eine KWIC-Sammlung oder Kookkurrenzliste für die Arbeit in *lexpan* konvertiert werden kann (siehe Abschnitt [Stores](#)).
- (3) Hier sieht man, welcher Store aktuell geladen ist. Alle weiteren Aktionen (Suchanfrage definieren, Annotieren usw.) beziehen sich immer auf diesen Store und sind nicht durchführbar, wenn kein Store geladen ist.
- (4) Dieser Button öffnet ein neues Fenster mit Informationen zum aktuell geladenen Store; hier kann auch der Name des Stores verändert werden (siehe Abschnitt [Details zum Store](#)).
- (5) Im Textfeld kann eine neue Suchanfrage angegeben werden (siehe Abschnitt [Suchanfragesyntax](#)), die dann mit Klick auf „Suchen“ im aktuellen Store gesucht wird.
- (6) Das Textfeld enthält eine Drop-Down-Liste der bereits erstellen Suchanfragen. Mit „Details“ können die Ergebnisse genauer untersucht werden (siehe Abschnitt [Details zu Suchanfragen](#)), mit „Löschen“ wird die gewählte Suchanfrage gelöscht.
- (7) Hier kann man ein neues Fenster zum Erzeugen einer Überblicksdatei aufrufen (siehe Abschnitt [Überblicksdateien](#)).
- (8) Mit „Exportdateien ansehen“ wird der Export-Ordner des aktuellen Archivs geöffnet. Dort findet man die bereits erstellten Exporte (siehe Abschnitt [Export](#)).
- (9) Wenn eine Kookkurrenzlisten-Datei eingelesen wurde, kann man hier auf die Kookkurrenzcluster zugreifen (siehe [Kookkurrenzlisten](#)).
- (10) Am dieser Stelle können Informationen zu *lexpan* abgerufen werden.
- (11) Ein Klick auf „Beenden“ schließt das Programm.

## Archive

*lexpan*-Analysen sind in Archiven organisiert, denen [Stores](#) – KWIC-Sammlungen oder Kookkurrenzlisten - zugeordnet werden. Jedes Archiv ist völlig unabhängig und liegt in einem eigenen Ordner auf der lokalen Festplatte.

Ein neu angelegter Archivordner enthält zwei Unterordner:

- **db**: Enthält die Datenbank, in der die Analysen gespeichert sind (*darf nicht verändert werden und muss von Benutzer normalerweise nicht geöffnet werden*)
- **print**: Hier werden alle Export-Dateien von Daten aus diesem Archiv abgelegt, außer der Benutzer verändert den Exportpfad (siehe [Export](#)).

Die Struktur und Benennung der Unterordner darf nicht verändert werden, es ist aber möglich, zusätzliche Dateien (z.B. die Originale von eingelesenen KWIC-Listen oder Kookkurrenzlisten) im Archivordner abzulegen. Diese Dateien werden von *lexpan* ignoriert.

## Anlegen eines Archivs

- Im Hauptfenster auf „Neu“ klicken.
- Unter „Name“ einen Namen für das Archiv angeben, welcher keine Umlaute, Sonderzeichen oder Leerzeichen enthalten darf. Damit die *lexpan*-Ordner besser erkennbar sind, wird der Name des Archivordners bei Erstellung automatisch mit dem Präfix *lexpan\_* versehen. Wenn man also als Archivname „AufKnopfdruck2“ angibt, wird ein Ordner namens „lexpan\_AufKnopfdruck2“ erzeugt.
- Unter „Verzeichnispfad“ kann ein beliebiger Ort auf der Festplatte gewählt werden. Dort wird der neue Archivordner angelegt.

Hinweise:

- Es ist nicht empfehlenswert, Archive im *lexpan*-Programmordner anzulegen. Dies ist zwar technisch kein Problem, aber bei einem Programmupdate ist es am einfachsten, den gesamten *lexpan*- Programmordner zu ersetzen. Wenn dieser Archivverzeichnisse enthält, besteht die Gefahr, dass diese versehentlich überschrieben werden.
- Fertig angelegte Archive können über das Dateiverwaltungssystem (z.B. Windows Explorer) frei verschoben und umbenannt werden, allerdings darf dies niemals geschehen, während das Archiv gerade in *lexpan* geladen ist.

## Stores

*lexpan* arbeitet auf Grundlage von Daten, die aus einem Korpusanalysesystem exportiert wurden. Dies können entweder Belege in Form von KWIC-Zeilen oder Kookkurrenzlisten sein. Die Daten werden für die Arbeit in *lexpan* konvertiert und dann ‚Stores‘ genannt. Es ist nicht möglich, vollständige Texte oder gar ein komplettes Korpus in *lexpan* einzulesen.

Der Zugriff auf Stores und das Konvertieren von neuen Stores ist nur möglich, wenn ein [Archiv](#) geladen ist, in das die Stores dann gespeichert werden.

## Stores konvertieren

- In der Zeile „verfügbare Stores“ den Button „Neu“ anklicken; ein Dialogfenster öffnet sich
- Unter „Name“ einen Namen für den Store eingeben. Es gibt keine Beschränkung der zu verwendenden Zeichen, der Name darf aber nicht länger als 50 Zeichen sein.
- Unter „Typ“ den Typ auswählen, der zum Format der Datei passt, die eingelesen werden soll.

- Unter „Datei auswählen“ die Datei mit den zu konvertierenden Daten wählen. Ihr Format muss einem der von *lexpan* unterstützten Formate entsprechen.

*lexpan* wurde ursprünglich für die Arbeit mit COSMAS II, dem Korpusanalysetool des Instituts für deutsche Sprache, entwickelt. Es unterstützt jedoch auch die Arbeit mit Daten aus SketchEngine, sowie mit einem einfachen Textformat. Unter [Hinweise zu Import-Formaten](#) findet man genauere Informationen, was bei der Erstellung der zu konvertierenden Daten zu beachten ist.

Aktuell unterstützt *lexpan* folgende Formate, hier sortiert nach Quelle und Datentyp:

- Quelle: [COSMAS II](#)
  - KWIC-Liste
  - Kookkurrenzliste ohne KWICs zu Kookkurrenzclustern
  - Kookkurrenzliste mit KWICs zu Kookkurrenzclustern (*Achtung! Konvertierung kann lange dauern.*)
- Quelle: [SketchEngine](#)
  - KWIC-Liste
  - Kookkurrenzliste ohne KWICs zu Kookkurrenzclustern
- Einfaches Textformat (auch erzeugt von [CQPWeb](#))
  - KWIC-Liste

## Details zum Store

Wenn ein Store geladen ist, lässt sich mit Klick auf „Details zum Store“ ein Dialogfenster öffnen, das zwei Funktionen hat:

- Umbenennung des Stores: Die Angabe „Name“ kann an dieser Stelle frei verändert werden.
- Anzeige von Metainformationen zum Store: Die meisten Informationen werden automatisch aus den konvertierten Korpusdaten extrahiert (Datum, Korpus, Suchanfrage, Weitere Infos). Je nach Format der Korpus-Datei sind die Informationen unterschiedlich detailliert. Diese Angaben können nicht verändert werden.

## Hinweise zu Import-Formaten

Im Folgenden wird erklärt, was zu beachten ist, wenn man Daten aus einem Korpusanalyzesystem exportiert, um sie dann in *lexpan* zu importieren. Dabei wird vorausgesetzt, dass der Benutzer mit der Bedienung der Korpusanalyzesysteme selbst vertraut ist. Informationen dazu:

- COSMAS II: <http://www.ids-mannheim.de/cosmas2/uebersicht.html>
- Sketch Engine: <https://www.sketchengine.co.uk/user-guide/>

## COSMAS II

### KWIC-Listen

Die folgenden Einstellungen müssen bei Export von KWIC-Daten aus COSMAS II unbedingt beachtet werden:

- Dateiformat muss auf TXT (bei COSMAS II win) bzw. ASCII (bei COSMAS II web) gestellt sein, niemals auf RTF.
- KWIC-Export muss aktiviert sein, denn *lexpan* berücksichtigt nur KWIC-Exporte, keine Volltextexporte

Zu Problemen beim Konvertieren von COSMAS-II-KWIC-Listen siehe auch [Warnmeldung bei der Konvertierung von KWIC-Listen](#).

### Kookkurrenzlisten

Prinzipiell besteht die Möglichkeit, COSMAS-II-Kookkurrenzlisten mit und ohne KWICs in *lexpan* einzulesen. Ohne KWICs nimmt sowohl der Export aus COSMAS II als auch das Einlesen in *lexpan* deutlich weniger Zeit in Anspruch, allerdings muss man bei der Auswertung dieser Kookkurrenzlisten dann auf den sprachlichen Kontext der Kookkurrenzpartner verzichten.

Bei der Durchführung wie auch beim Export der Kookkurrenzanalyse in COSMAS II müssen einige Einstellungen vorgenommen werden, damit die Ergebnisse korrekt in *lexpan* konvertiert werden können. Die rot umrandeten Einstellungen sind obligatorisch, die restlichen können auch verändert werden.

Parametereinstellungen für die Durchführung der Kookkurrenzanalyse (rot = obligatorisch):

<b>Kontext</b>	5	Wörter links	<input checked="" type="checkbox"/> höchstens 1 Satz
	5	Wörter rechts	
<b>Analyse</b>	Granularität:	Grob	<input checked="" type="checkbox"/> Autofokus
	Zuverlässigkeit:	Analytisch	<input type="checkbox"/> Funktionswörter ignorieren
	Clusterzuordnung:	Mehrfach	<input type="checkbox"/> Lemmatisierung verwenden
<b>Ausgabe</b>		<input checked="" type="checkbox"/> L.R-Wert anzeigen	
		<input checked="" type="checkbox"/> Nummerierung des Hauptkollokators	
<div>Zurücksetzen</div> <div>StartenHilfe</div>			

Einstellungen beim Export der Kookkurrenzanalyse aus COSMAS II (rot= obligatorisch). Wenn man an der blau umrandeten Stelle einen Haken setzt, sollte die COSMAS-II-Datei in *lexpan* mit dem Typ „Kookkurrenzliste COSMAS II (mit KWICs)“ eingelesen werden, wenn man ihn entfernt, mit dem Typ Kookkurrenzliste COSMAS II (ohne KWICs).

**Allgemeines**

Dateiname:  Schriftgröße:

Dateiformat:  Orientierung: ☐ Querformat ☐ Hochformat

☒ Ergebnisansicht

☒ KWIC

vorher:  Wörter  nachher:

☐ Volltext

vorher:  Sätze  nachher:

Annotationen: ☐ einblenden ☒ ausblenden

**gemeinsame Optionen für KWIC und Volltext**

Kontext umschließt: ☐ gesamten Treffer ☒ 1. Wort des Treffers

Sortierung:

☐ Nur ausgewählte Treffer

☐ Einschränken auf die ersten:  Treffer

**Quellennachweis**

☒ exportieren ☐ nachher ☐ fett

☐ Expansionslisten

☐ Korpusansicht

Hinweis (nur relevant für „Kookkurrenzliste mit KWICs“):

Der KWIC-Export ist in COSMAS II auf 100.000 Zeilen limitiert. Da bei der Kookkurrenzanalyse mit Mehrfach-Zuordnung KWIC-Zeilen mehrfach aufgelistet werden, kann diese Grenze auch dann überschritten werden, wenn die Gesamttrefferzahl der Suchanfrage unter 100.000 liegt. In diesem Fall kann die Kookkurrenzliste zwar problemlos in *lexpan* eingelesen werden, allerdings fehlen für Kookkurrenzcluster der unteren Ränge die KWICs, da diese nicht aus COSMAS II exportiert wurden. Um dies zu verhindern, ist es ratsam, bei Suchanfragen mit hoher Trefferzahl mit der COSMAS-II-Zufallsauswahl in zu arbeiten, bevor die Kookkurrenzanalyse durchgeführt wird. Erfahrungen haben gezeigt, dass das Problem bei einer Beschränkung der Trefferzahl auf etwa 10.000 vermieden werden kann.

### Sketch Engine

Beim Export aus Sketch Engine muss immer das Format XML gewählt werden (sowohl für KWIC-Exporte als auch für den Export von Kookkurrenzlisten).

**Hinweis:** SketchEngine hat zwischen April und Juni 2018 sein XML-Exportformat deutlich geändert. Deswegen gibt es aktuell in *lexpan* zwei unterschiedliche Einlesefilter:

- „KWIC-Liste – SketchEngine“ bzw. „Kookkurrenz-Liste ohne KWICS – SketchEngine“: Diese Filter funktionieren mit Dateien, die nach dem **13.06.2018** aus SketchEngine exportiert wurden.
- „KWIC-Liste – SketchEngine (alt)“ bzw. „Kookkurrenz-Liste ohne KWICS – SketchEngine (alt)“: funktionieren mit Exporten älteren Datums, sicher vor April 2018. Für Dateien, die in der Zeit



zwischen 3.4. und 13.06.18 exportiert wurden, können wir leider nicht garantieren, dass sie mit *lexpan* verwendet können, da SketchEngine in diesem Zeitraum teilweise nicht valides XML produziert hat.

### Kookkurrenzlisten

Beim Export von Kookkurrenzlisten (in Sketch Engine „Collocations“), kann ein beliebiges Kookkurrenzmaß gewählt werden. Es wird empfohlen, nur jeweils ein Kookkurrenzmaß pro Tabelle zu wählen (Sketch Engine erlaubt mehrere), da *lexpan* nur eine Kookkurrenzspalte pro Kookkurrenzliste verarbeiten kann.

Wählt man dennoch mehrere Kookkurrenzmaße, so werden nur die Werte des Maßes übernommen, das im Sketch-Engine-Export an letzter Stelle steht.

### Einfaches Textformat

Als weiteres Eingabeformat erlaubt *lexpan* zudem eine KWIC-Liste in einem einfachen Textformat.

Dieses Format kann von CQPweb, der Weboberfläche der Corpus Workbench, (<http://cw.b.sourceforge.net/cqpweb.php>) erzeugt werden, indem man dort „Download with typical settings for copy-paste into Word, Excel etc.“ wählt.

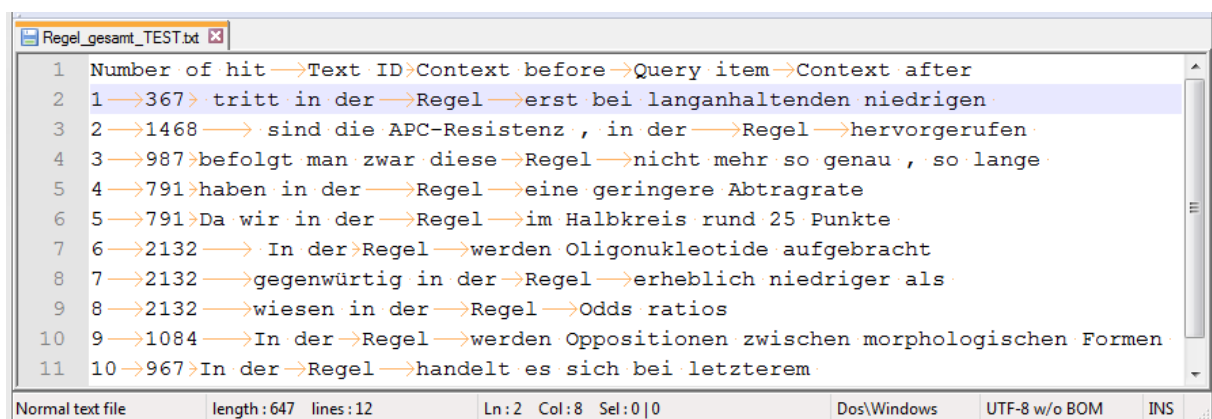
Der besondere Vorteil ist jedoch, dass dieses Format so schlicht ist, dass der Benutzer es mit relativ geringem Aufwand selbst erzeugen kann. KWIC-Listen, die aus einem Korpusanalysesystem stammen, das von *lexpan* nicht unterstützt wird, können vom Benutzer in dieses Format konvertiert werden und über diesen Umweg in *lexpan* eingelesen werden.

Die Eingabedatei ist eine einfache Textdatei in der Kodierung UTF-8, die eine Tabelle repräsentiert. Jede Zeile ist durch Tabstopp-Zeichen in 5 Felder geteilt. Die erste Zeile stellt die Kopfzeile der Tabelle dar und benennt die fünf Spalten:

- Number of Hit: Eindeutige ID für den KWIC
- Text ID: Eindeutige ID für den Text, aus dem der KWIC stammt
- Context before: Kontext vor dem Trefferelement
- Query Item: Trefferelement (kann auch mehrere Wörter enthalten)
- Context after: Kontext nach dem Trefferelement

Felder dürfen leer sein, allerdings muss trotzdem die richtige Menge von Tabstopp-Zeichen in jeder Zeile vorhanden sein (genau 4 Tabstopps zur Trennung der 5 Felder).

Hier eine Beispieldatei mit KWICs zum Trefferelement „Regel“. Die orangen Pfeile markieren die Tabstopp-Zeichen.



```
Regel_gesamt_TEST.txt
1 Number of hit→Text ID→Context before→Query item→Context after
2 1→367→ tritt in der →Regel→erst bei langanhaltenden niedrigen
3 2→1468→ sind die APC-Resistenz , in der →Regel→hervorgerufen
4 3→987→befolgt man zwar diese →Regel→nicht mehr so genau , so lange
5 4→791→haben in der →Regel→eine geringere Abtragsrate
6 5→791→Da wir in der →Regel→im Halbkreis rund 25 Punkte
7 6→2132→ In der →Regel→werden Oligonukleotide aufgebracht
8 7→2132→gegenwärtig in der →Regel→erheblich niedriger als
9 8→2132→wiesen in der →Regel→Odds ratios
10 9→1084→In der →Regel→werden Oppositionen zwischen morphologischen Formen
11 10→967→In der →Regel→handelt es sich bei letzterem
```

Normal text file   length: 647   lines: 12   Ln: 2   Col: 8   Sel: 0 | 0   Dos\Windows   UTF-8 w/o BOM   INS

## Suchanfragesyntax

In dem Feld „Neue Suchanfrage“ kann ein Muster angegeben werden, das dann in den KWICs des aktuell geladenen Stores gesucht wird.

Die Syntax der Suchanfrage ist angelehnt an reguläre Ausdrücke. Jedes Element repräsentiert mindestens ein ‚Wort‘ (Zeichenkette ohne Leerzeichen). Mögliche Elemente sind:

- konkrete Zeichenketten
- abstrakte Platzhalter für nicht fest spezifizierte (Folgen von) Zeichenketten

Folgende Platzhalter stehen zur Verfügung:

Platzhalter, die ein Wort oder mehrere Wörter erfassen

#	erfasst ein einziges, beliebiges Wort. Das Zeichen kann mehrfach hintereinander gebraucht werden, um eine genau bestimmte Menge von Wörtern zu erfassen (# # → genau 2 beliebige Wörter)
#*	erfasst 0 bis beliebig viele aufeinanderfolgende Wörter
#+	erfasst 1 bis beliebig viele aufeinanderfolgende Wörter
N(wort)	erfasst genau ein Wort, das nicht die Oberfläche „wort“ aufweist. ("nicht wort")
wort1 wort2	erfasst entweder die Zeichenfolge „wort1“ oder die Zeichenfolge „wort2“

Platzhalter für Zeichen innerhalb eines Wortes

Wo*	Das Zeichen * steht für 0 bis beliebig viele Buchstaben. „Wo*” erfasst also entweder das Wort „Wo“ oder irgendein Wort, das mit „Wo“ beginnt („Wort“, „Woche“ ...). Es können beliebig viele * in Zeichenketten eingefügt werden.
W?rt	Das Zeichen ? steht für 0 oder genau einen Buchstaben. „W?rt“ erfasst also „Wort“ oder „Wirt“ und auch die Zeichenfolge „Wrt“, aber z.B. nicht „Wahlort“. Es können beliebig viele ? in Zeichenketten eingefügt werden.

Hinweise:

- Es ist nicht möglich, nach Satzzeichen zu suchen.
- Suchanfragen sind so definiert, dass sie keine satzübergreifenden Strukturen erfassen. D.h. konkret, es werden keine Treffer gefunden, die zwischen dem ersten und letzten Element der Suchanfrage eines der Zeichen ., ! oder ? aufweisen.
- Es ist nicht sinnvoll, die Zeichen #\* oder #+ direkt an Anfang und Ende einer Suchanfrage zu verwenden, sondern nur zwischen enger beschränkten Suchelementen (z.B. „auf #\* Grund“). Wenn man beliebige Wörter an Anfang oder Ende der Anfrage erfassen will, sollte man # verwenden (z.B. „in Gottes #“).
- Wenn eine Suchanfrage auf mehrere Stellen innerhalb des gleichen KWIC passt, so wird immer das erste Vorkommen erfasst. Zum Beispiel:

Suchmuster: die # (Zeichenfolge „die“ gefolgt von Platzhalter für genau ein Wort)  
KWIC-Zeile: er sah **die Frauen** in bunten Kleidern, **die dort** standen

Auch wenn das Suchmuster in diesem Beispiel sowohl auf *die Frauen* als auch auf *die dort* passt, so wird nur der erste (hier grün markierte) Treffer berücksichtigt. Der Platzhalter erfasst also nur das Wort *Frauen*, nicht das Wort *dort*. Dies ist vor allem relevant, wenn man [Füllertabellen](#) erstellt.

## Überblicksdateien

Überblicksdateien zeigen, welche KWIC-Zeilen von bestimmten Suchanfragen erfasst wurden.

- Durch Klick auf „Überblicksdatei erzeugen“ wird ein neues Fenster geöffnet.
- Jetzt können von den Suchanfragen, die für diesen Store definiert sind, beliebig viele ausgewählt werden. Die Exportdatei zeigt dann die KWICs für jede der ausgewählten Suchanfragen.
- Durch Klick auf „Export“ wird der Export angestoßen (siehe hierzu Abschnitt [Export](#)). Wenn beim Export „Restmenge ebenfalls exportieren“ ausgewählt ist, werden auch die KWICs ausgedruckt, die von keiner der ausgewählten Suchanfragen erfasst wurden. Dies ist nützlich, um zu prüfen welche KWIC-Zeilen noch nicht durch Suchanfragen abgedeckt sind. In solchen Fällen kann auch sinnvoll sein, beim Export „Nur die Restmenge exportieren“ auszuwählen. Dann werden die KWICs, welche von den ausgewählten Suchanfragen erfasst sind, gar nicht mehr ausgedruckt und die Überblicksdatei wird entsprechend kleiner.

Hinweise:

- Vorsicht: Bei großen Stores kann die Überblicksdatei sehr groß werden.
- Die Überblicksdateien werden mit Namen benannt, die den Zeitpunkt ihrer Erstellung abbilden. Jede Überblicksdatei hat damit einen eindeutigen Namen.

## Details zu Suchanfragen

Über den Button „Details“ kann man die Suchanfrage, die in der Drop-Down-Liste daneben ausgewählt ist, genauer analysieren und annotieren.

### KWIC-Liste

Der Klick auf den Button „Details“ öffnet zunächst die KWIC-Liste. Diese zeigt alle KWIC-Zeilen, die von der Suchanfrage erfasst wurden. Die einzelnen Felder können durch Klick auf die Spaltennamen sortiert werden.

### Füllertabellen

Unterhalb der KWIC-Ansicht werden alle Felder angezeigt, die von Platzhaltern definiert sind und also mit unterschiedlichen lexikalischen Elementen gefüllt sein können, die sogenannten ‚Lücken‘. Man kann diese Felder auswählen und dann durch einen Klick auf den Button „Zeige Lückenfüller-Häufigkeiten“ eine Tabelle erstellen lassen, die zeigt, welche lexikalischen Elemente (‚Füller‘) mit welcher Häufigkeit an der Lücken-Position auftreten.

Ist nur ein Feld ausgewählt, werden die unterschiedlichen Füller dieses Feldes gezählt.

Sind zwei oder mehr Felder ausgewählt, werden die Füller in Kombination gezählt. D.h. es werden nur die Zeilen zusammengezählt, in denen *alle* gewählten Felder gleich besetzt sind.

Hinweise:

- Beim Zusammenfassen der Lückenfüller werden Satz- und Sonderzeichen ignoriert. „Haus“ und „Haus,“ würden also zusammengezählt.
- Zu jeder Zeile einer Lückenfüller-Tabelle kann man durch Klick auf die Spalte „KWICs“ die zugeordneten KWIC-Zeilen ansehen.

## Kookkurrenzlisten

Wenn eine Kookkurrenzliste als Store eingelesen wurde, kann man über den Button „Kookkurrenzen“ die Kookkurrenzcluster (und die ihnen zugeordneten KWICs) ansehen, taggen und kommentieren. Die [Annotations-Funktionen](#) funktionieren bei Cluster- und Füllertabellen analog.

Hinweise:

- *Achtung:* Das Löschen eines Stores mit dem Typ „Kookkurrenzliste mit KWICs – COSMAS II“ kann **mehrere Minuten dauern**. *lexpan* reagiert in dieser Zeit nicht auf andere Eingaben. Bitte das Programm trotzdem nicht über den Taskmanager o. ä. beenden, da sonst die Datenstrukturen im geladenen Archiv beschädigt werden könnten.
- Die Tabellenspalte „Abweichung Frequenz“ ist nur für den Store-Type „Kookkurrenzliste mit KWICs – COSMAS II“ relevant und kann bei anderen Store-Typen ignoriert werden. Bei diesem Store-Typ fasst *lexpan* aus Effizienzgründen oberflächengleiche KWICs zusammen. Dies kann dazu führen, dass es für einzelne Cluster Abweichungen gibt zwischen der KWIC-Menge, die COSMAS II angibt, und der KWIC-Menge, die in *lexpan* angezeigt werden kann. Die Tabelle zeigt in diesen Fällen in der Spalte „Abweichung Frequenz“ an, wie viele KWICs weggefallen sind. Da es sich bei den wegfallenden KWICs per Definition immer um Duplikate von noch vorhandenen KWICs handelt, können dabei keine einzigartigen Oberflächenmuster verloren gehen. Allerdings ist die KWIC-Gesamtmenge nicht komplett identisch mit der in COSMAS II gefundenen.

## Annotationen

Es ist möglich, Füller- und Clustertabellen manuell mit verschiedenen Annotationen zu versehen. Hierzu werden die Felder in den Tabellenspalten „Tag“, „Kommentar“ oder „SortNr.“ angeklickt und editiert. Der neue Eintrag ist automatisch gespeichert, sobald das editierte Feld verlassen wurde, indem auf eine andere Stelle in der Tabelle geklickt wird.

## Tags

Tags können nicht nur die Tabelle mit neuen Informationen anreichern, sondern auch neue Darstellungs- und Filterfunktionen ermöglichen.

### Formatierungsregeln für Tags

- Tags werden durch Leerzeichen voneinander abgegrenzt („adj belebt“ wären also zwei unabhängige Tags, „adj“ und „belebt“).
- Bei Tag-Namen wird zwischen Groß und Kleinschreibung unterschieden („ADJ“ und „adj“ werden also als unterschiedliche Tags interpretiert).

- Tag-Namen können beliebig lang sein, allerdings darf die Länge aller Tags in einer Zeile (einschließlich Leerzeichen) 500 Zeichen nicht überschreiten.

### Sortieren und Filtern nach Tags

Sowohl das Fenster für Füllertabellen als auch das für Clustertabellen hat zwei Buttons „Alle Tag-Gruppen zeigen“ und „Tag-Kombination zeigen“. Diese erlauben es, die Tabelle neu zu sortieren und/oder zu filtern.

#### Alle Tag-Gruppen zeigen

Ein Klick auf diesen Button erzeugt für jeden definierten Tag eine Tabelle mit den Zeilen, denen er zugeordnet ist. Zeilen mit mehreren Tags tauchen mehrfach auf. Zeilen, denen gar kein Tag zugewiesen ist, werden in der Tabelle „Restmenge“ angezeigt.

#### Tag-Kombination zeigen

Diese Funktion ermöglicht es, mit Hilfe der vergebenen Tags komplexere Gruppierungen von Tabellenzeilen zu erzeugen.

In dem Textfeld neben dem Button wird zunächst eine Tag-Kombination eingegeben. Dann wird mit Klick auf den Button selbst eine Tabelle mit den Zeilen erzeugt, deren Tagmuster auf diese Kombination passt. In der Tabelle „Restmenge“ werden alle Zeilen angezeigt, die übrig bleiben.

Für Tag-Kombinationen stehen folgende Syntaxelemente zur Verfügung (ein kurzer Überblick kann auch über den Button „Hilfe“ aufgerufen werden):

OR(tag1 tag2)	Liefert alle Zeilen, die entweder tag1 oder tag2 enthalten. (Es handelt sich hier nicht um eine ausschließende Disjunktion, d.h. eine Zeile kann auch beide Tags gleichzeitig enthalten.)
AND(tag1 tag2)	Liefert alle Zeilen, die sowohl tag1 als auch tag2 enthalten.
NOT(tag1)	Liefert alle Zeilen, die tag1 nicht enthalten.
AND(tag1)	Liefert alle Zeilen, die tag1 enthalten.

NOT kann mit AND oder OR geschachtelt werden. Damit kann man z.B. ausdrücken: „Zeige mir alle Zeilen, die mit dem Tag ‚adj‘ aber *nicht* mit dem Tag ‚belebt‘ versehen wurden.“ Die Syntax dafür lautet: AND(adj NOT(belebt)).

### Kommentare

In die Tabellenspalte „Kommentar“ kann beliebiger Text eingetragen werden. Der Kommentar darf nicht länger als 500 Zeichen sein. Eine Formatierung des Textes ist nicht möglich.

### Sortierungsnummern

In die Spalte „SortNr.“ können beliebige ganze Zahlen eingetragen werden. Diese dienen dazu, beim Export eine bestimmte Reihenfolge der Tabellenzeilen zu erzwingen (siehe Abschnitt [Export](#)).

## Export

Export aus *lexpan* bedeutet, dass die erzeugten Tabellen außerhalb des Programms im HTML oder XML-Format gespeichert werden. Dies ist nützlich, wenn diese unabhängig von *lexpan* weitergegeben, archiviert oder weiterverarbeitet werden sollen.

Per Default werden alle Exporte automatisch im „print“-Verzeichnis des aktuellen *lexpan*-Archivs abgespeichert. Je nach Ergebnistyp legt das Programm die exportierten Dateien automatisch in die Unterordner **overviewfiles** (bei Export von Überblicksdateien), **kwicfiles** (bei Export von KWICs), **fillerfiles** (bei Export von bearbeiteten oder unbearbeiteten Lückenfüllertabellen) oder **clusterfiles** (bei Export von bearbeiteten oder unbearbeiteten Clustertabellen) ab.

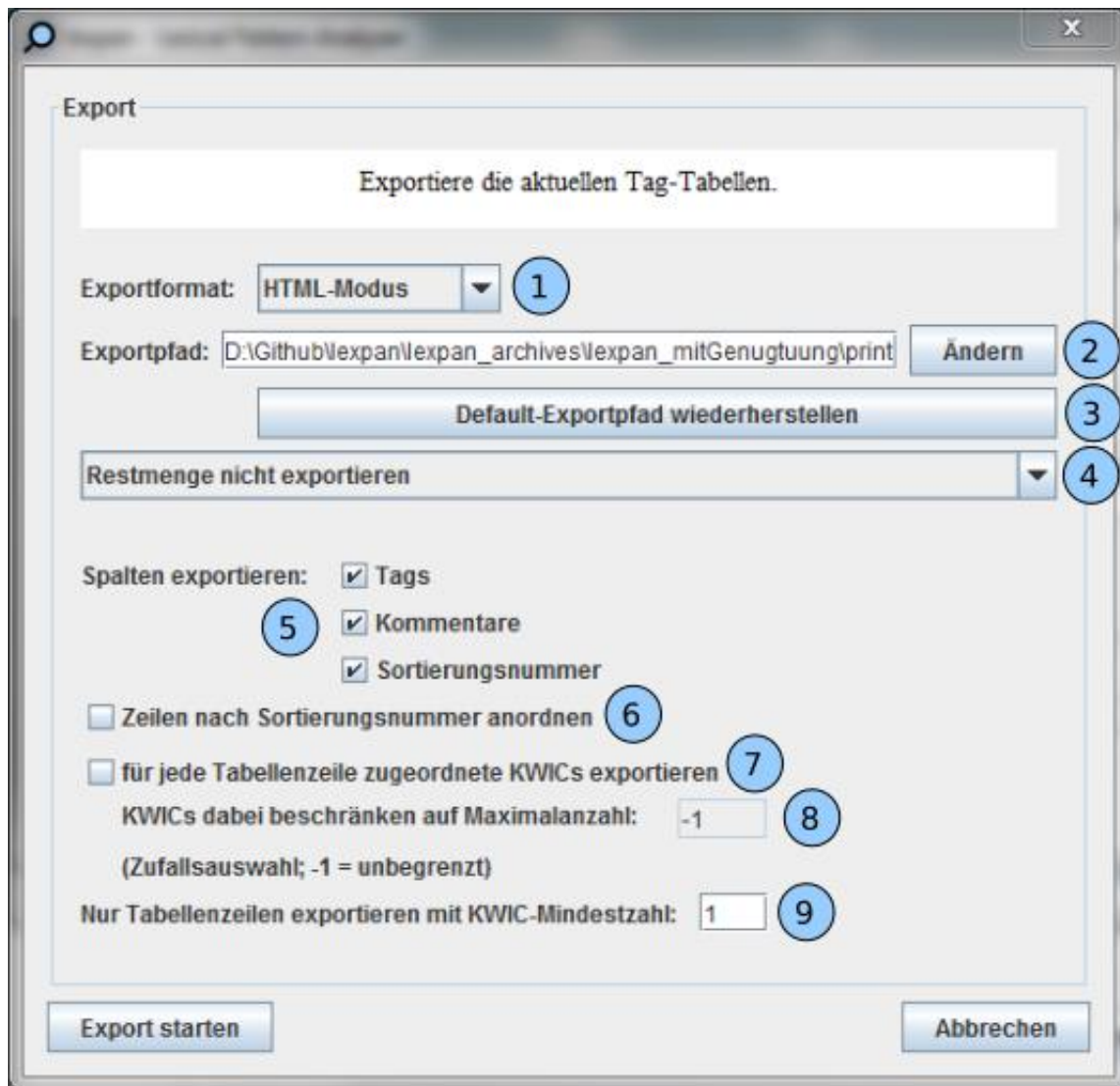
Exportmöglichkeiten gibt es an unterschiedlichen Stellen von *lexpan*. Sie werden immer durch den Button „Export“ signalisiert.

Die folgende Übersicht zeigt, welche Ergebnistypen exportiert werden können und in welchem *lexpan*-Programmfenster der Export jeweils angestoßen werden muss:

- Überblicksdatei: Fenster „Erzeuge Überblicksdatei“ (erreichbar über Button „Überblicksdatei erzeugen“); druckt eine Datei in **overviewfiles**
- Alle KWICs einer Suchanfrage und Lückenfüllertabellen für alle Lücken der gleichen Suchanfrage: Fenster „KWICs zur Suchanfrage“ (erreichbar über Button „Details“); druckt zwei Dateien, eine in **kwicfiles** und eine in **fillerfiles**
- Komplette Füllertabelle: Fenster „Füllerliste“ (erreichbar über Button „Zeige Lückenfüller-Häufigkeiten“); druckt eine Datei in **fillerfiles**
- Komplette Kookkurrenzliste: Fenster „Kookkurrenzliste“ (erreichbar über Button „Kookkurrenzen“); druckt eine Datei in **clusterfiles**
- Nach Tags sortierte Ansichten: Fenster „Tag-Überblick“ erreichbar über Button „Alle Tag-Gruppen zeigen“ oder Button „Tag-Kombination zeigen“; druckt Dateien in **fillerfiles** oder **clusterfiles**, je nachdem um welchem Tabellentyp es geht.

An allen Stellen der Programmoberfläche, an denen ein Export möglich ist, befindet sich auch der Button „Exportdateien ansehen“. Bei Klick auf diesen Button wird automatisch im Dateiverwaltungssystem des lokalen Rechners der Ort angezeigt, an dem die Exportdateien gespeichert wurden.

Wenn der Button „Export“ angeklickt wird, öffnet sich zunächst das Exportfenster, das eine Anpassung der Exporteinstellungen erlaubt. Dieses sieht, abhängig davon welcher Ergebnistyp exportiert werden soll, etwas unterschiedlich aus. Als Beispiel dient hier die Ansicht beim Export einer nach Tags sortierten Füllertabelle. Einige der beschriebenen Optionen können beim Export anderer Ergebnistypen fehlen.



- (1) **Exportformat:** Hier wird angegeben, in welchem Format die Exportdateien erzeugt werden. Die Optionen sind:
  - HTML-Modus: Dieser Modus erzeugt gut lesbare HTML-Ansichten mit sortierbaren Tabellenspalten.
  - XML-Modus: Dieser Modus ist besonders geeignet, wenn man die Exporte weiterverarbeiten oder neu formatieren möchte.
- (2) **Exportpfad:** Hier wird festgelegt, wo die Exportdatei lokal gespeichert wird. Per Default werden *lexpan*-Exporte im Unterverzeichnis „print“ im aktuellen *lexpan*-Archiv-Verzeichnis abgelegt, jedoch kann über „Ändern“ auch ein beliebiger anderer Speicherort angegeben werden. Die Pfadänderung wird nach Abschluss des Exports gespeichert, solange man weiter mit dem gleichen Store arbeitet. Wird ein anderer Store geladen oder das Programm beendet, wird der Pfad automatisch auf den Default zurückgesetzt.
- (3) **Default-Exportpfad wiederherstellen:** Ein Klick auf diesen Button setzt eine Änderung des Exportpfades ebenfalls auf den Default zurück.
- (4) **Auswahlbox zur Handhabung der Restmenge:** In der Auswahlbox können verschiedene Optionen gewählt werden, wie beim Export mit der „Restmenge“ umgegangen wird. Beim Export von getaggten Füller- oder Clusterdateien bezieht sich der Begriff „Restmenge“ auf die Tabelle, welche die Zeilen der Füller- oder Clustertabelle zeigt, die keine Tags enthalten



(die sogenannte „Restmenge-Tabelle“, siehe Abschnitt [Tags](#)). Wenn es darum geht, [Überblicksdateien](#) zu exportieren, bezieht sich der Begriff „Restmenge“ auf die KWICs, welche von keinem der ausgewählten Suchmuster erfasst sind. Es kann jeweils entschieden werden, ob diese Menge zusätzlich exportiert werden sollen, nicht exportiert werden sollen oder sogar als einzige Menge exportiert werden sollen (letzteres kann bei Überblicksdateien sinnvoll sein). Diese Option ist nicht für alle Exporte verfügbar.

Die folgenden Einstellungen sind nur sichtbar, wenn Füller- oder Clustertabellen exportiert werden:

- (5) **Spalten exportieren:** Hier kann festgelegt werden, welche der editierbaren Spalten der Tabelle mitexportiert werden sollen.
- (6) **Zeilen nach Sortierungsnummer anordnen:** Wenn diese Option mit einem Haken versehen wird und Sortierungsnummern in die Tabelle eingetragen wurden, werden die Tabellenzeilen im Export nach den Sortierungsnummern geordnet ausgegeben (aufsteigend) (siehe [Sortierungsnummern](#)). Zeilen ohne Sortierungsnummern stehen am Ende der Tabelle.
- (7) **Für jede Tabellenzeile zugeordnete KWICs exportieren:** Wenn diese Option ausgewählt wird, wird nicht nur die Tabelle selbst exportiert, sondern für jede Tabellenzeile auch die KWICs, die den jeweiligen Füllern oder Kookkurrenzcluster zugeordnet sind.
- (8) **KWICs dabei beschränken auf Maximalanzahl:** Diese Option ermöglicht das Arbeiten mit einer Zufallsauswahl, um die KWIC-Mengen zu reduzieren. Wenn hier eine Zahl eingetragen wird, die größer ist als -1, aber kleiner als die KWIC-Menge, die einer Zeile zugeordnet ist, werden die KWIC-Zeilen für den Export per Zufallsauswahl auf diese Maximalmenge beschränkt. Ein Eintrag von -1 ist der Default und bedeutet ‚keine Beschränkung‘.
- (9) **Nur Tabellenzeilen exportieren mit KWIC-Mindestzahl:** Hier kann festgelegt werden, dass nur Tabellenzeilen exportiert werden, denen eine bestimmten Mindestzahl von KWICs zugeordnet ist. Der Default ist 1 (d.h. alle Zeilen werden ausgegeben) und je höher die Zahl ist, die eingetragen wird, desto mehr wird die Tabelle verkürzt. Dies ist v.a. nützlich, wenn man mit sehr großen KWIC-Mengen arbeitet, was fast immer zu einer großen Menge von Füllern mit nur einem Vorkommen und damit sehr langen Tabellen führt.

Wenn die Option „Für jede Tabellenzeile zugeordnete KWICs exportieren“ ausgewählt wird, kann der Export deutlich länger dauern und große Datenmengen können erzeugt werden. Dies hängt von zwei Faktoren ab:

- Zeilenanzahl der Tabelle: *lexpan* erzeugt für jede Zeile der Tabelle eine separate KWIC-Datei und legt diese in einem automatisch erzeugten Unterordner ab, dessen Namen von dem der Hauptdatei abgeleitet ist (z.B. Hauptdatei: *Filler\_per\_X\_6\_Field1.html*, Verzeichnis: *Filler\_per\_X\_6\_Field1\_HTML*). Mit Hilfe der Option „Nur Tabellenzeilen exportieren mit KWIC-Mindestzahl“ können die Tabellen verkürzt und damit die Zahl der zu druckenden KWIC-Dateien reduziert werden.
- Menge der KWICs, die einer Zeile zugeordnet sind: Bei hohen Trefferzahlen können die pro Zeile erzeugten KWIC-Dateien sehr groß werden. Dies kann durch die Verwendung der Zufallsauswahl („KWICs dabei beschränken...“) verhindert werden.

Exportdateien können prinzipiell frei verschoben und umbenannt werden.



Wenn es sich allerdings um Tabellen mit zugeordneten KWIC-Zeilen handelt, muss darauf geachtet werden, dass der Unterordner mit den abhängigen KWIC-Dateien immer im gleichen Verzeichnis wie die Hauptdatei liegt. Der Name des KWIC-Unterordners darf nicht verändert werden.

## KWICs (de)aktivieren

Es ist möglich, einzelne KWIC-Zeile zu deaktivieren. Dies bewirkt, dass diese Zeilen beim Export übersprungen werden, also nicht in den Export-Dateien erscheinen. Auf diese Weise kann man manuell bereinigte KWIC-Übersichten erzeugen. Der Aktiv-Status von KWICs wird auch nach Beenden von lexpan gespeichert.

ACHTUNG: Deaktivierung von KWICs wirkt sich **nicht** auf die Frequenzangaben aus, weder in lexpan selbst noch in den Exportdateien. Frequenzangaben beziehen sich immer auf alle vorhandenen KWICs, egal ob diese aktiviert sind oder nicht.

Es kann also z.B. passieren, dass in einer exportierten Tabelle angegeben ist, dass ein Füller 100 KWICs hat, aber in der verlinkten KWIC-Datei steht „90 KWICs“, weil die restlichen deaktiviert sind. Wenn in einer Exportdatei aufgrund von Deaktivierung einem Cluster oder Füller überhaupt keine KWICs mehr zugeordnet sind, so wird dies in den Export-Tabellen angezeigt („keine aktiven KWICs“).

(De)aktivieren von KWICs kann in jeder Tabellenansicht innerhalb von lexpan geschehen. Es gibt zwei Vorgehensweisen:

- Gezielter Klick auf das erste Feld (Spalte „Aktiv“) einer bestimmten KWIC-Zeile, um diese zu aktivieren oder zu deaktivieren (Haken gesetzt bedeutet der KWIC ist aktiviert)
- Verwendung der Buttons „Alle KWICs aktivieren“ und „Alle KWICs deaktivieren“. Diese Befehle beziehen sich immer auf alle KWICs der aktuell gezeigten Tabelle und verändern deren Status. ACHTUNG: Bei großen KWIC-Stores kann die Verwendung dieser Buttons deutliche Wartezeiten nach sich ziehen.

Unter der Tabelle erscheint eine Warnmeldung, wenn innerhalb dieser Tabelle KWICs deaktiviert sind. Durch Sortieren der Tabelle nach der Spalte „Aktiv“ (Klick auf die Spalte) kann man versehentlich deaktivierte KWIC-Zeilen leicht finden.

Immer wenn beim Export auch KWICs exportiert werden, erscheint ein Hinweis, dass nur aktive KWICs exportiert werden. Dies geschieht unabhängig davon, ob im aktuellen Store tatsächlich KWICs deaktiviert sind oder nicht.

## Fehler- und Warnmeldungen

### Warnmeldung bei der Konvertierung von KWIC-Listen

Bei der Konvertierung von KWIC-Stores kann eine Warnmeldung auftreten, dass die Trefferangaben in der zu konvertierenden Datei nicht mit der tatsächlich konvertierten KWIC-Menge übereinstimmen. Hier sind verschiedene Fälle zu unterscheiden:

**Sketch-Engine-Dateien**, bei denen die Treffermenge durch Zufallsauswahl reduziert wurde: In diesem Fall ist in der Sketch-Engine-Datei trotzdem die Gesamttrefferzahl – nicht die Zahl der reduzierten Treffer – angegeben, was zu der Warnmeldung führt. Die Warnung kann ignoriert werden, da im

Fälle des Arbeitens mit einer Zufallsauswahl hier ja nicht wirklich Zeilen verloren gegangen sind. Man kann das Problem aber auch beheben, indem man die Sketch-Engine-Datei in einem Text- oder XML-Editor öffnet und die Trefferzahl manuell korrigiert.

**COSMAS-II-Dateien** mit fehlerhaft formatierten Zeilen: Dateien, die aus COSMAS-II exportiert wurden, enthalten manchmal Formatierungsfehler. In diesem Fall werden die problematischen Zeilen bei der Konvertierung von *lexpan* automatisch ignoriert. Die Meldung weist auf die Abweichung in der KWIC-Menge hin, die sich dadurch ergibt. Zudem wird eine Liste der betroffenen KWIC-Zeilen unter dem Namen *missingLines.txt* im aktuellen *lexpan*-Archiv-Verzeichnis abgelegt.

Meist handelt es sich um einen der beiden folgenden Fehlertypen:

### Überzählige < > -Zeichen

In der korrekten Form wird in der COSMAS-II-Datei das Bezugswort im KWIC mit den Zeichen <B> </> umschlossen. Wenn innerhalb des KWICs weitere < oder > vorkommen, kann dieses Muster nicht mehr korrekt erkannt werden und *lexpan* ignoriert die Zeile.

Wenn man auf diese KWIC-Zeilen nicht verzichten will, kann man die COSMAS-II-Export-Datei in einem Texteditor (Empfehlung: Notepad++) öffnen, die fehlerhaften Zeilen mit Hilfe der Übersicht in *missingLines.txt* suchen, die überzähligen Zeichen (im Beispiel unten rot hervorgehoben) löschen und danach das Konvertieren wiederholen.

#### Fehlerhafte KWIC-Zeile:

A09 in Kindern spielerisch die Freude <B>an Naturwissenschaften</> wecken. Mit einem

#### Korrekte KWIC-Zeile:

A09 in Kindern spielerisch die Freude <B>an Naturwissenschaften</B> wecken. Mit einem

### Unvollständige KWIC-Zeilen

COSMAS-II-KWIC-Zeilen werden manchmal nach dem ersten Suchwort abgeschnitten, weil das darauffolgende Wort zu lang ist. Diese Zeilen sind damit so stark beschädigt, dass man sie nicht leicht reparieren kann.

Wenn man die Warnmeldung in *lexpan* vermeiden möchte und Wert darauf legt, dass die KWIC-Mengen in der eingelesenen COSMAS-II-Datei und im *lexpan*-Store identisch sind, kann man die COSMAS-II-Datei in einem Texteditor öffnen, die fehlerhaften Zeilen manuell entfernen und zusätzlich am Beginn der COSMAS-II-Datei die Menge der exportierten KWIC-Treffer entsprechend manuell korrigieren. Bei erneutem Konvertieren der so veränderten Datei sollte nun keine Warnung mehr auftreten.

#### Fehlerhafte KWIC-Zeile (zu löschen):

A09 in Kindern spielerisch die Freude <B>an

#### Zahlen, die am Beginn der Datei korrigiert werden müssen:

Ergebnis : 100.000 Treffer, zufällig reduziert aus 683.556  
KWIC (unsortiert)  
Anz. Treffer : 100.000, zufällig reduziert aus 683.556  
Anz. exportierte Zeilen: 100.000 (interne Begrenzung)  
Angezeigter Kontext : 10 Wörter links, 10 Wörter rechts  
Kontext umschließt : 1. Wort des Treffers

### Neu konvertierter Store hat 0 Zeilen

Sicherstellen, dass

- der richtige Typ unter „Typ des Stores“ ausgewählt wurde.

- die eingelesene Datei einem von *lexpan* unterstützten Format entspricht und nicht beschädigt ist.
- die eingelesene Datei nicht tatsächlich 0 KWIC-Zeilen enthält (etwa durch einen Exportfehler).

### Programm meldet, es laufe bereits, ist aber nicht sichtbar

Es ist nicht möglich, auf dem gleichen Rechner *lexpan* nochmals zu starten, während es schon läuft. Wenn diese Meldung beim Programmstart auftaucht, ohne dass das Programm sichtbar ist, liegt dies daran, dass *lexpan* bei einem früheren Aufruf zwar gestartet wurde, die Benutzeroberfläche aber nicht korrekt geladen wurde. Das Programm läuft also tatsächlich, aber unsichtbar im Hintergrund. Dies kann passieren, wenn

- versucht wurde, *lexpan* aus einem Archiv-Ordner (.zip-Ordner) heraus zu starten. Der .zip-Ordner muss unbedingt entpackt werden, bevor ein Programmstart durch Doppelklick auf die Datei *lexpan.jar* versucht wird.
- der *lexpan*-Ordner in einen Bereich kopiert wurde, an dem der ausführende Benutzer keine Schreibrechte hat (z.B. unter Windows im Ordner „Programme“, wenn der Nutzer keine Administratorenrechte hat).
- die auf dem Rechner installierte Java-Version nicht aktuell ist. Bitte Java aktualisieren: (<https://www.java.com/de/download/>)

Nachdem die oben beschriebenen Probleme behoben wurden, sollte man einen Neustart des Rechners durchführen, um das unsichtbare Programm sicher zu beenden.

### Archiv kann nicht geladen werden

Während ein Archiv geladen ist, wird sein Datenbank-Verzeichnis gesperrt. Wenn *lexpan* nicht regulär beendet wurde (z.B. wenn der Computer abgestürzt ist oder *lexpan* über den Windows-Task-Manager beendet werden musste), kann es passieren, dass diese Sperrung nicht korrekt aufgehoben wurde. Dies hat zur Folge, dass das Archiv nicht mehr geladen werden kann.

Dieser Fehler kann folgendermaßen behoben werden: Im Unterverzeichnis „db“ des betroffenen Archivs befindet sich eine Datei namens „lock“. Diese kann vom Benutzer manuell gelöscht werden. Danach sollte sich das Archiv wieder laden lassen.

### Tag, Kommentar oder Sortierungsnummer wurde nicht gespeichert

Veränderungen in den Tabellenspalten (Tags, Kommentare, Sortierungsnummern) werden automatisch gespeichert, allerdings erst, wenn die Eingabe in die Tabelle explizit beendet wurde - d.h. der Cursor darf nicht mehr in dem editierten Feld stehen. Dies ist am besten zu erreichen, indem man nach der Eingabe auf eine beliebige andere Stelle der Tabelle klickt.