

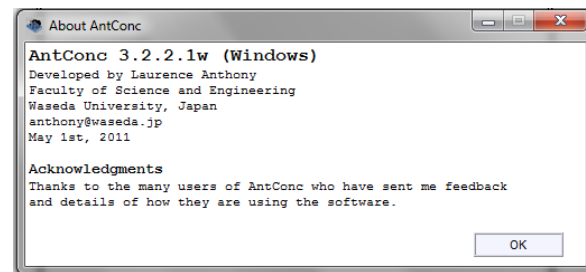
# Ko[Gloss]

**Kollaborative Konstruktionsglossare im Fachfremdsprachenlernen Deutsch,  
Estnisch, Lettisch, Litauisch**

***AntConc* – Arbeit mit digitalen Textsammlungen**

## Digitale Analyse: Sprachliche Muster recherchieren und beschreiben

Die KoGloss-Methode möchte die Möglichkeit bieten, zu einem bestimmten Thema, Sachgebiet oder Diskurs gezielt Informationen über dafür typische sprachliche Muster zu recherchieren. Dazu kann der Nutzer eigene Textsammlungen erstellen, die dann als Grundlage einer selbständigen Recherche und Analyse dienen. Um diese Textsammlungen nutzen zu können, gibt es unterschiedliche Analyse-Programme. Sie sind dafür gemacht, mittels Suchanfragen in Form von einzelnen Wörtern, Wortteilen oder Wortfolgen häufige sprachliche Muster aus den vorher zusammengetragenen Texten zu gewinnen und Informationen zu ihrer Beschreibung zu entnehmen. Solche Analyse-Softwares sind unterschiedlich komplex aufgebaut und bieten verschiedene Werkzeuge. KoGloss bedient sich des frei verfügbaren und auch für linguistische Laien leicht zugänglichen Programms *AntConc*. Das Programm wurde von Professor Laurence Anthony, dem Leiter des CELESE (Center for English Language Education in Science and Engineering) an der Universität Waseda in Tokyo entwickelt und ist besonders für kleinere Textsammlungen (bis zu 20 Mio. Textwörter) geeignet.



## Die Grundlage: Digitale Textsammlungen

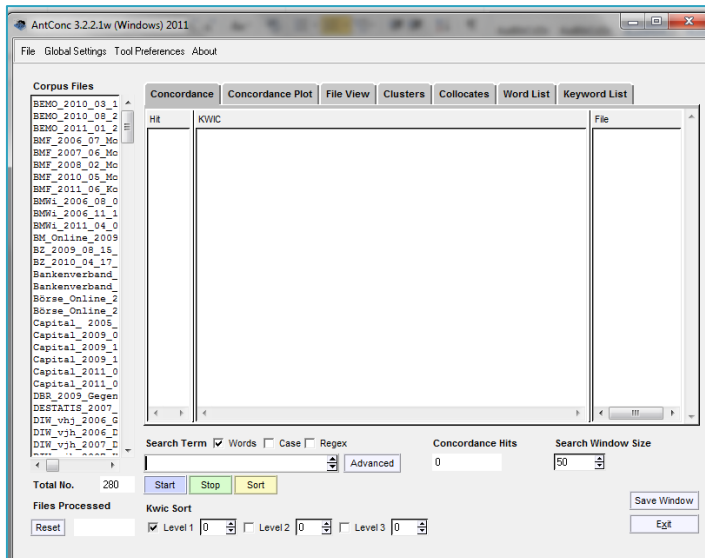
Die Methode KoGloss stützt sich auf die Arbeit mit authentischen Texten. Das heißt, die Nutzer erstellen eigenständig Sammlungen, die sich aus typischen Texten einer bestimmten Domäne oder eines Themengebietes zusammensetzen. Aufzufinden sind diese in Form von elektronischen Veröffentlichungen oder analog vorliegenden Texten, die dann vor der Analyse in digitale Form gebracht werden. Für die KoGloss -Sammlung zum Thema ‚Konjunktur und Konjunkturentwicklung‘ sieht eine Übersicht dieser Textsammlung zum Beispiel so aus:

Dateiname	Vollständige Überschrift	Länge (Wörter)	Länge (Zeichen)	Funktionale Merkmale	Textsorte	Medienkürzel	Volständiger Link
BEMO_2010_03_15_Konjunktur_Aussicht	Konjunktur. Aussicht für die deutsche Wirtschaft verdüstert sich.	1.231	8.534	Nachrichtentext	Zeitungsa rtikel	BEMO	<a href="http://www.morgenpost.de/wirtschaft/article1274587/Aussicht-fuer-deutsche-Wirtschaft-verduestert-sich.html">http://www.morgenpost.de/wirtschaft/article1274587/Aussicht-fuer-deutsche-Wirtschaft-verduestert-sich.html</a>
ZEDE_2009_10_23_Kreditloser_Aufschwung	Was hinter dem kreditlosen Aufschwung steckt	1.092	7.673	Nachrichtentext	Web-Publikation	ZEDE	<a href="http://www.zeit.de/wirtschaft/2010/studie-konjunktur-kreditentwicklung">http://www.zeit.de/wirtschaft/2010/studie-konjunktur-kreditentwicklung</a>
ZEDE_2009_03_26_Klotzen_oder_kontrollieren	Klotzen oder kontrollieren	1.906	12.340	Nachrichtentext	Zeitungsa rtikel	ZEDE	<a href="http://www.zeit.de/2009/14/G20">http://www.zeit.de/2009/14/G20</a>
WAMS_2008_05_18_Gnadenfrist_für_den_Aufschwung	Gnadenfrist für den Aufschwung	1.028	7.014	Nachrichtentext	Zeitungsa rtikel	WAMS	<a href="http://www.welt.de/wams_print/ticle2006628/Gnadenfrist_fuer">http://www.welt.de/wams_print/ticle2006628/Gnadenfrist_fuer</a>

Dieses Quellenverzeichnis zeigt alle Texte, die in der Sammlung enthalten sind und gibt u.a. Auskunft über den Fundort, die Textsorte, die Autorschaft und die Länge. Diese Informationen können zur Einordnung der späteren Rechercheergebnisse nützlich sein. Über das Kürzel kann man die Texte später den ursprünglichen Medien zuordnen.

## Das Werkzeug: Die Analysesoftware AntConc

*AntConc* ist ein viel genutztes Programm zur Analyse digitaler Textsammlungen, das einfach zu bedienen ist und das nicht installiert werden muss. Nach dem kostenlosen Download können die Nutzer die .exe-Datei aktivieren und alle Texte aus ihrer selbst erstellten Sammlung in das Programm einspeisen. Danach können Suchanfragen unterschiedlichster Art gestellt werden, die man über verschiedene Reiter erreicht:



Ganz links ist die Liste der eingespeisten Textdateien zu sehen, die hier als Kürzel erscheinen.

Über dem zentralen Fenster befinden sich die Reiter mit den unterschiedlichen Analysetools.

Unterhalb des großen Fensters lassen sich Sucheinstellungen verändern, z.B. wie groß der Wortabstand sein soll, der bei einer Suchanfrage beachtet werden soll. Möchte man zu einem bestimmten Ausdruck recherchieren, wird dieser in der Eingabezeile eingetippt.

Mit den verfügbaren Werkzeugen können die Nutzer zum Beispiel Listen mit häufig auftauchenden Wörtern und Wortmustern erstellen lassen oder recherchieren, welche Partnerwörter besonders häufig mit einem bestimmten Ausdruck zusammen auftauchen. Alle Suchergebnisse können angeklickt werden, so dass man Einsicht in den Ursprungstext erhält und Kontextinformationen entnehmen kann. Außerdem liefert das Programm Informationen über die Auftauchenshäufigkeit der gesuchten Wörter oder Muster.

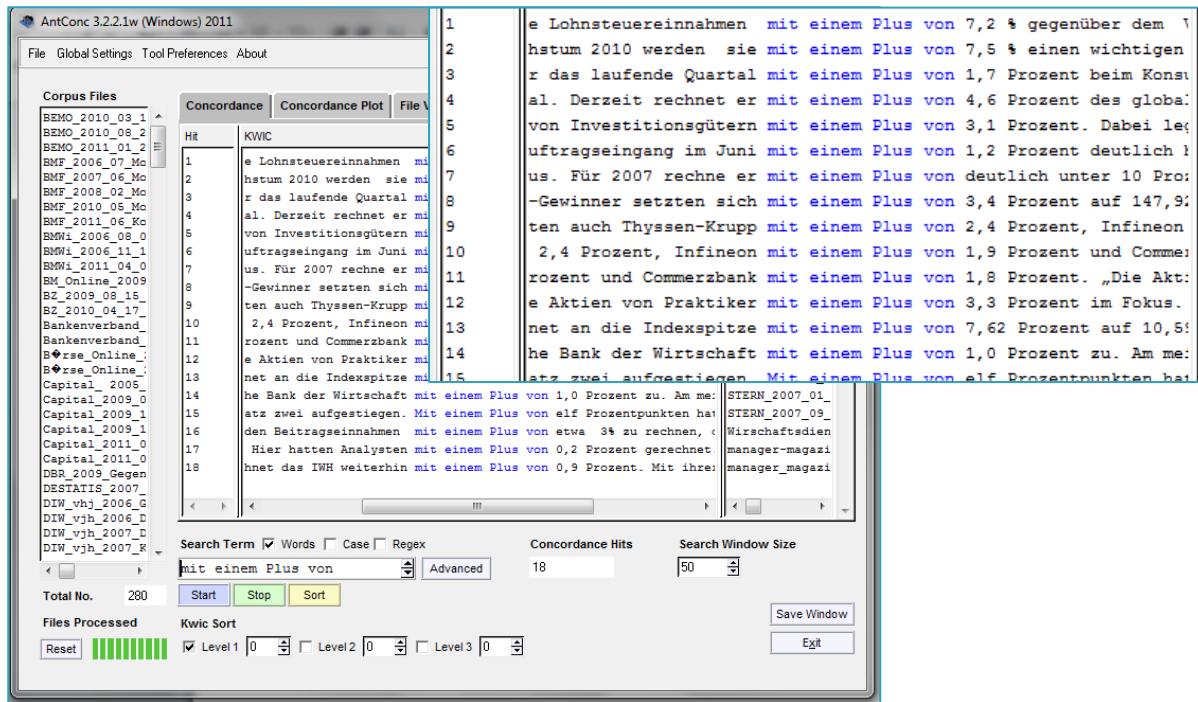
## Ein Beispiel

Bei einer sogenannten Clusteranfrage innerhalb der Textsammlung zu ‚Konjunktur und Konjunktorentwicklung‘ kann der Nutzer einsehen, welche Wortkombinationen besonders häufig in den gesammelten Texten vorkommen. In diesem Fall sind dies Muster wie *im Vergleich zum Vorjahr*, *die Zahl der Arbeitslosen* oder *mit einem Plus von*.

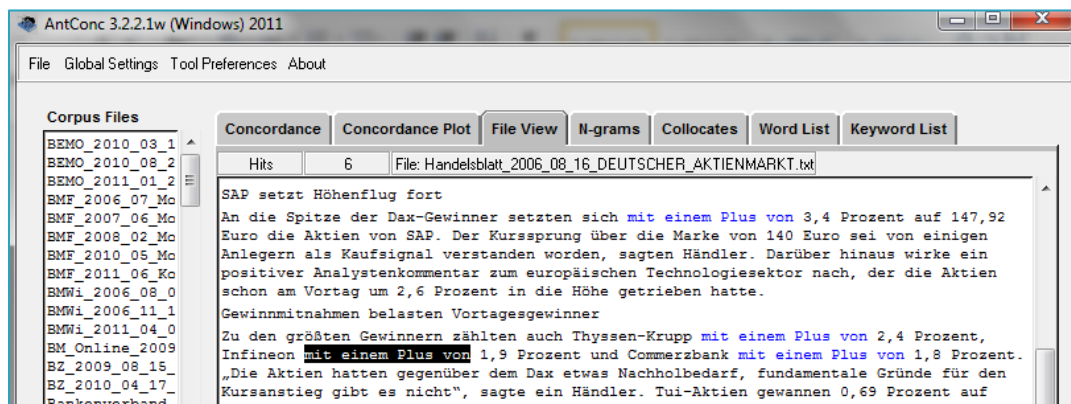
Rank	Freq	N-gram
1	63	in den vergangenen Jahren
2	33	in Höhe von Mrd
3	29	in den vergangenen Monaten
4	26	im Vergleich zum Vorjahr
5	24	in der zweiten Jahreshälfte
6	23	in den kommenden Monaten
7	21	der Europäischen Zentralbank EZB
8	19	mit Blick auf die
9	18	davon aus dass die
10	18	einem Plus von Prozent
11	18	für das kommende Jahr
12	18	in den er Jahren
13	17	in der Euro Zone
14	17	in diesem Jahr um
15	17	mit einem Plus von

Dabei macht das Programm keine Unterschiede zwischen sinnvollen Wort-Kombinationen mit eigener Bedeutung und rein statistisch auffälligen Mustern. Diese Unterscheidung kann der Nutzer dann in weiteren Analyseschritten selbst erarbeiten und mittels der gewonnenen Informationen entscheiden, welche sprachlichen Muster für den gezielten Bereich wichtig und funktional sind.

So könnte eine detaillierte Analyse zur gefundenen Wortkombination *mit einem Plus von* etwa ergeben, dass diese Wortkette kein vollständiges Muster ergibt. Es taucht regelmäßig mit weiteren Elementen zusammen auf, nämlich in der Kombination: *mit einem Plus von X Prozent*, wobei ‚X‘ eine Zahl ist, die meist unter 10 liegt und inklusive Dezimalstellen angegeben wird.



Möchte der Nutzer weitere Informationen über das Auftauchen dieses sprachlichen Musters in den verschiedenen Kontexten einsehen, so lässt sich einfach die jeweilige Zeile anklicken, in der die interessante Konstruktion angegeben ist und es erscheint der Gesamttext aus der Textsammlung. So kann man Details über den Gebrauch und die Bedeutung in einem bestimmten Kontext herausfinden. Auch die Textsorte oder das Veröffentlichungsmedium können in diesem Zusammenhang interessant sein.



Die recherchierten Informationen können anschließend in einem Glossar festgehalten werden. Hierzu nutzt KoGloss die Oberfläche der Lehr- und Lernplattform Moodle, in der alle sprachlichen Muster als Glossareinträge beschrieben werden können.