

Suche in ANNIS



Bevor man suchen kann, muss das gewünschte Korpus in der Korpusliste ausgewählt werden (z.B. das Teilkorpus *mo* (monoethnisches Ergänzungskorpus) oder KiDKo *mu* (multiethnisches Korpus)).

ANNIS (ANNIS Corpus Search) - Chromium

ANNIS (ANNIS Corpus x)

localhost:8080/annis-gui/

About ANNIS [Help us to make ANNIS better!](#)

Please enter AQL query

Query Builder

Search More History

Welcome to ANNIS! A tutorial is available on the right side.

Corpus List Search Options

Visible: All

Name	Texts	Tokens		
mo	82	146,704		
mu	189	345,813		

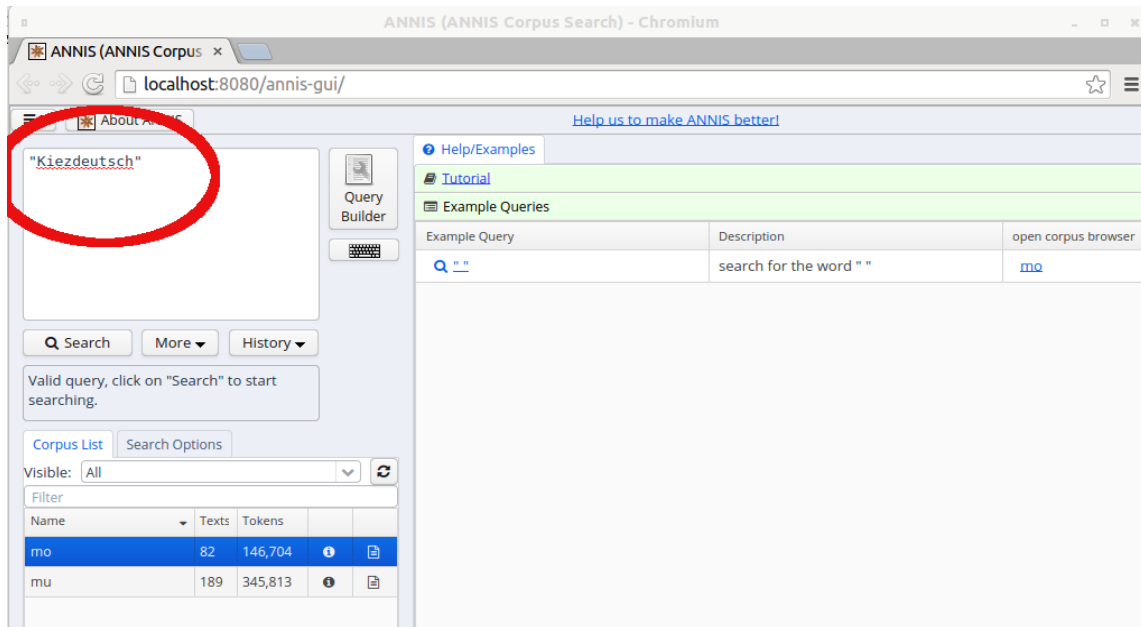
Help/Examples

Tutorial

Example Queries

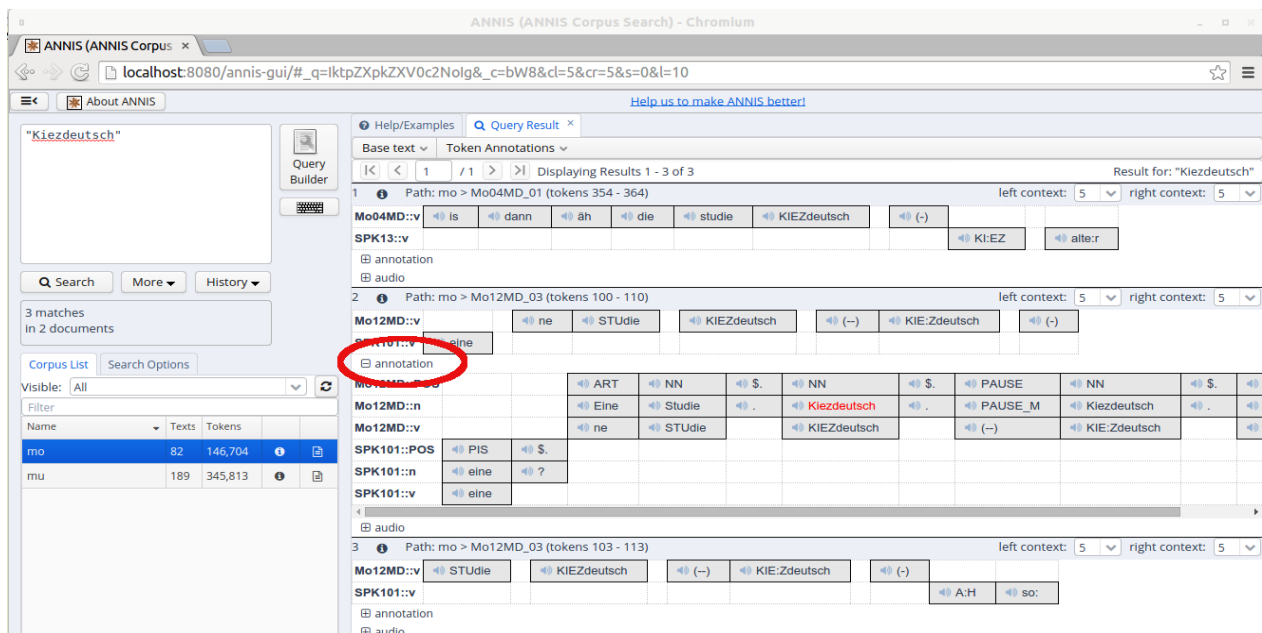
Example Query	Description
Q " "	search for the word " "
Q " "	search for the word " "

Danach kann in der Suchmaske ein Suchausdruck eingegeben werden.



Durch Anklicken des *Search*-Buttons wird die Suche gestartet.

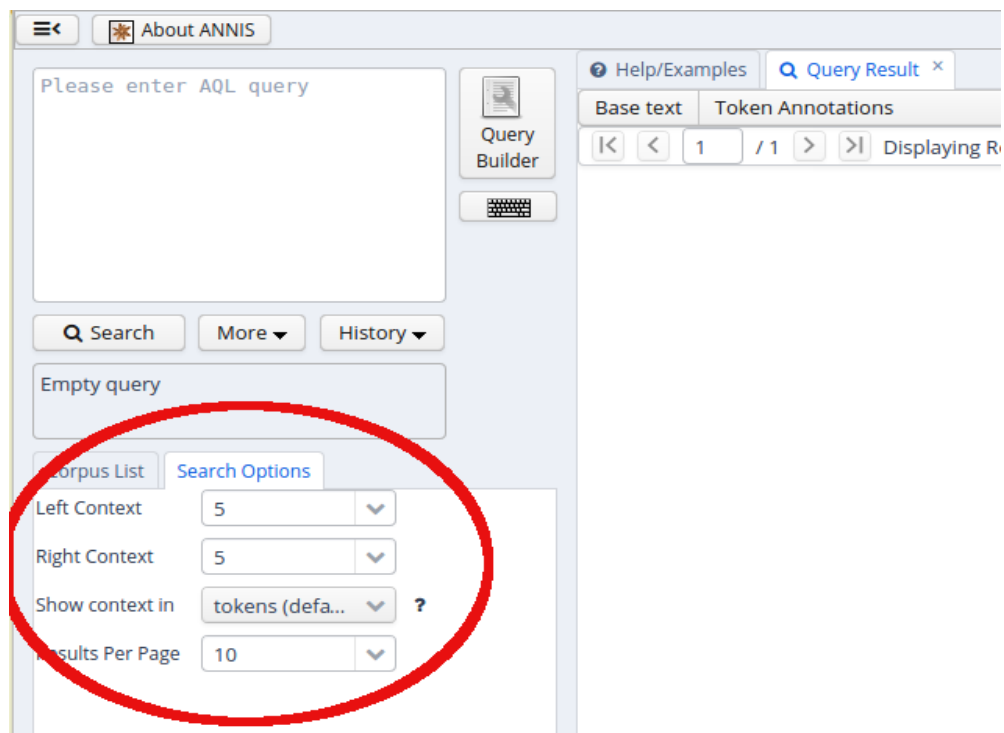
Als Ergebnis wird zuerst nur die Transkriptionsebene angezeigt. Durch einen Klick auf *annotation* werden zusätzlich auch die Normalisierungsebene und die Wortartenauszeichnungen visualisiert.



Die zugehörige Audiodatei kann durch Anklicken der kleinen Lautsprechersymbole abgespielt werden.

ART	NN	\$.	NN	\$.	PAUSE	NN	\$.
Eine	Studie	.	Kiezdeutsch	.	PAUSE_M	Kiezdeutsch	.
ne	STUDIE		KIEZdeutsch		(-)	KIE:Zdeutsch	

Suchoptionen: Durch Auswahl des Menus *Search Options* gelangt man zu einer Maske, in der die Einstellungen zur Ansicht der Suchergebnisse verändert werden können (z.B. rechter/linker Kontext, Anzahl der Ergebnisse pro Seite)



Annotationsebenen in KiDKo

Die aktuelle Version von KiDKo (Release 1.0) beinhaltet die folgenden Annotationsebenen:

Kürzel	Beschreibung
nv	nonverbale Ebene
v	Transkriptionsebene (verbale Ebene)
n	Normalisierungsebene
POS	Wortartenebene (part-of-speech)
tr	türkische Transkriptionsebene
trnorm	Normalisierung
trdtwwue	deutsche Übersetzung (Wort für Wort)
trdtue	deutsche Übersetzung (frei)

Bei Suchanfragen muss die jeweilige Annotationsebene mit angegeben werden, also z.B.:

v="icke" oder POS="PPER"

Die Normalisierungsebene ist die Default-Suchebene. Hier kann die Ebenenbezeichnung weggelassen werden. Die folgenden beiden Anfragen sind äquivalent:

n="Kiezdeutsch" und "Kiezdeutsch"

Beispielanfragen

Die folgenden Beispielanfragen illustrieren den Gebrauch der Suchabfragesprache in ANNIS.

Beschreibung	Anfragesyntax
<i>transkribierte Wortformen</i>	
Finde alle Instanzen der Wortform <i>ick</i> (nur unbetonte Instanzen)	v="ick"
Finde alle Instanzen der Wortform <i>ick</i> (mit oder ohne Betonung)	v="ick" v="ICK"
<i>normalisierte Wortformen</i>	
Finde alle Instanzen der (normalisierten) Wortform <i>ich</i> (nur Kleinschreibung)	"ich" or n="ich"
Finde alle Instanzen der (normaisierten) Wortform <i>sein</i>	"sein" or n="sein"
<i>Wortarten-Tags (Part-of-speech, POS)</i>	

Finde alle Instanzen eines finiten Auxiliars	POS="VAFIN"
Kombinationen von Wortformen und Wortarten	
Finde alle Instanzen der Wortform <i>sein</i> , die Auxiliare sind	"sein" & POS="VAFIN" & #1 _ _ #2
Finde alle Instanzen der Wortform <i>sein</i> , die Possessivpronomen sind	"sein" & POS="PPOSAT" & #1 _ _ #2
Sequenzen von Wortformen und/oder Wortarten	
Finde alle Instanzen von <i>lass</i> , gefolgt von <i>mal</i>	n="lass" & n="mal" & #1 . #2
Finde alle Instanzen von <i>lass=ma</i> (mit und ohne Betonung)	"lass=ma" "LASS=ma"
Finde Instanzen der Präposition <i>in</i> , gefolgt von einem Nomen (ohne Determinierer)	n="in" & POS="NN" & #1 . #2
Finde Instanzen einer beliebigen Präposition, gefolgt von einem Nomen (ohne Determiner)	POS="APPR" & POS="NN" & #1 . #2
Finde Instanzen von agglutinierten Wortformen (markiert durch =)	v=/.*=.* /
Finde Instanzen von gelängten Vokalen (markiert durch :)	v=/.*:.* /
Finde Instanzen von <i>so</i> , gefolgt von einem Nomen (z.B. <i>so Lederjacke</i>)	n="so" & POS="NN" & #1 . #2
Finde Instanzen von <i>weil</i> -V2 (<i>weil</i> , gefolgt von einem Personalpronomen, einem finiten Verb und einem Token, das <i>kein</i> Interpunktionszeichen ist)	"weil" & POS="PPER" & POS="/V.FIN/" & n!="/[,.\!?!]/ & #1 . #2 & #2 . #3 & #3 . #4
Weitere Annotationsebenen	
Suche nach beliebigem türkischen Sprachmaterial	tr=/.*/
Suche auf der nonverbalen Ebene nach <i>Lachen</i> (hier wurde ein regulärer Ausdruck benutzt, damit alle transkribierten Formen (<i>lachen, Lachen, lacht, ...</i>) gefunden werden)	nv="/[L]ach.* /