universitätbonn

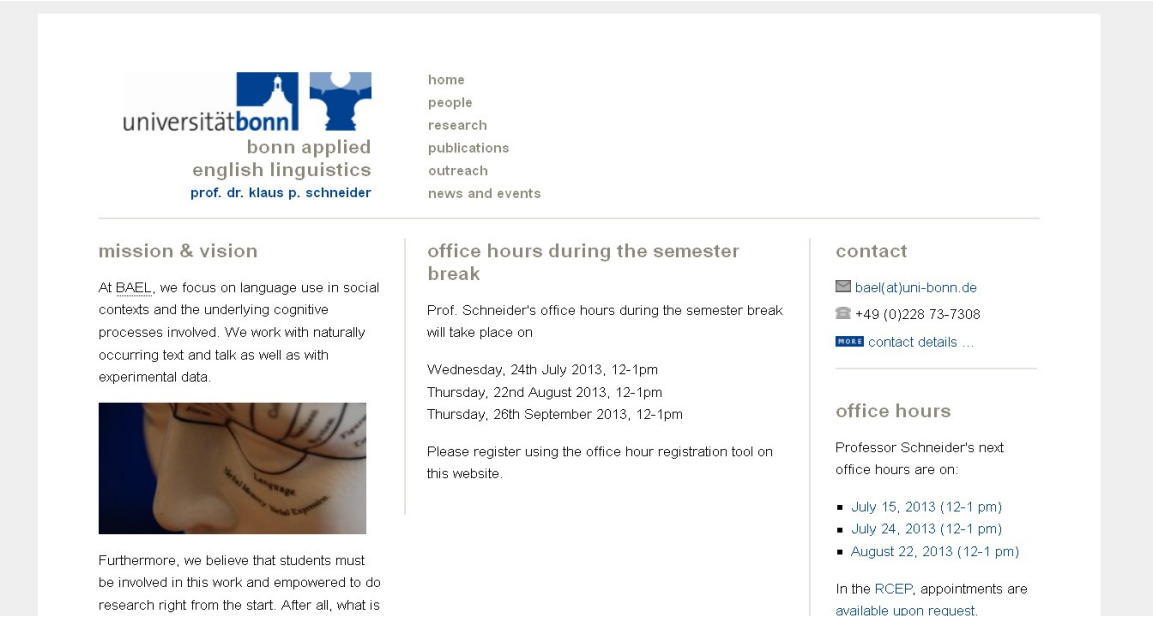
Praxiseinblicke in die linguistische Korpusanalyse und die Erwartungen / Möglichkeiten bibliothekarischer Unterstützung

Susanne Strubel-Burgdorf,

Institut für Anglistik, Amerikanistik und Keltologie

Bonn Applied English Linguistics (BAEL), Uni Bonn

Bonn Applied English Linguistics (BAEL)



universitätbonn

bonn applied
english linguistics
prof. dr. klaus p. schneider

home
people
research
publications
outreach
news and events

mission & vision

At BAEL, we focus on language use in social contexts and the underlying cognitive processes involved. We work with naturally occurring text and talk as well as with experimental data.

Furthermore, we believe that students must be involved in this work and empowered to do research right from the start. After all, what is

office hours during the semester break

Prof. Schneider's office hours during the semester break will take place on

Wednesday, 24th July 2013, 12-1pm
Thursday, 22nd August 2013, 12-1pm
Thursday, 26th September 2013, 12-1pm

Please register using the office hour registration tool on this website.

contact

bael(at)uni-bonn.de
+49 (0)228 73-7308
MORE contact details ...

office hours

Professor Schneider's next office hours are on:

- July 15, 2013 (12-1 pm)
- July 24, 2013 (12-1 pm)
- August 22, 2013 (12-1 pm)

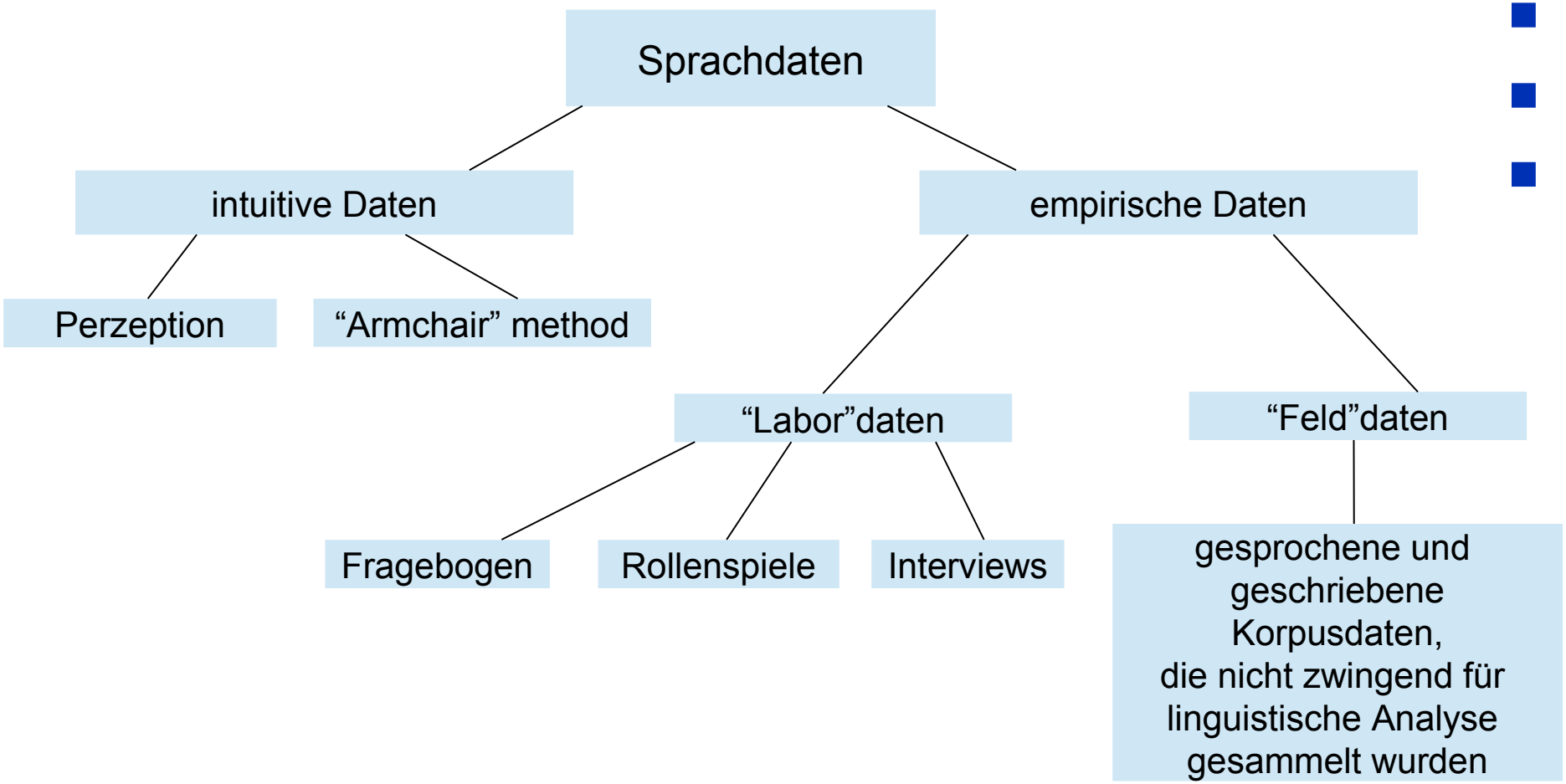
In the RCEP, appointments are available upon request.

linguistics.uni-bonn.de

Facebook

Twitter

Sprachdaten – unsere Leidenschaft



Ein Augenmerk auf Korpora

1) Was ist ein Korpus und wie wird es genutzt?

2) Praxisbeispiele (aus der BAEL Perspektive)

3) Wie würde der “perfekte Korpuspool” aussehen?

1) Was ist ein Korpus und wie wird es genutzt?

- (elektronische) Sammlung von (vielen wortreichen) Texten, die mit Software zur Korpusanalyse (oder mit quantitativen Textanalysetools) unter linguistischen Gesichtspunkten bearbeitet werden kann

→ *Key word in Context (KWIC), Konkordanzen, Frequenzberechnung, ...*

1) *Kriterien für eine Typologie englischsprachiger Korpora*

Dimension	Korpustypen		
Größe	*first generation*	vs.	*second generation*
	standard	vs.	*mega-corpora*
Inhalt	*synchronic*	vs.	*diachronic*
	spoken	vs.	*written*
	general	vs.	*specialized*
	native	vs.	*second language / learner*
	monolingual	vs.	*bilingual / multilingual*
Design	*sample-text*	vs.	*full-text*
	plain-text	vs.	*annotated*
	static	vs.	*dynamic*
	corpora	vs.	*archives / data collections*

(siehe auch Mukherjee 2009: 41)

1) *Korpora: typologisiert und Verfügbarkeit (eine Auswahl!)*

Korpus	Größe	Inhalt	Design	Verfügbarkeit
Brown 1961 **Frown** frühe 1990er	*first generation,* standard, ~ 1 Million Wörter	synchron, amerikanisches Englisch geschriebene Texte (Sammlung verschiedener Genres)	*plain-text* und *annotated, sample-text, static*	ICAME Kollektion *(International Computer Archive of Modern and Medieval English, Kauf-CD)*
LOB 1961 **FLOB** frühe 1990er	*first generation,* standard, ~ 1 Million Wörter	synchron, britisches Englisch geschriebene Texte (Sammlung verschiedener Genres)	*plain-text* und *annotated, sample-text, static*	ICAME Kollektion
COCA 1990-2012	*second generation*, mega-corpus, ~ 450 Millionen Wörter	synchron/diachron, amerikanisches Englisch, gesprochene (überwiegend Nachrichtensender) und geschriebene Texte	*annotated, full text, dynamic*	http://corpus.byu.edu/coca/
MICASE 2001-2002	"first generation", ~ 2 Millionen Wörter	synchron, gesprochenes (amerikanisches) Englisch in akademischem Kontext	*annotated, full text, static*	http://micase.elicorpora.info/
VOICE 2001-2007	"first generation", ~ 1 Million Wörter	synchron, gesprochenes ELF in verschiedenen Textsorten	*annotated (POS tagging), full text, static*	http://www.univie.ac.at/voice/page/index.php

universitätbonn

1) *meistgenutzte Analysesoftware (quantitative Analyse)*

- WordSmith

(für .txt Dateien; mit auf ICAME CD; neuere Lizenzversionen aus dem Internet, siehe: http://lexically.net/LexicalAnalysisSoftware/)

- AntConc

(für .txt, .html, .htm, .xml, .ant; Freeware aus dem Netz, siehe: http://www.antlab.sci.waseda.ac.jp/antconc_index.html)

universitätbonn

1) *Analysesoftware (qualitative Analyse)*

- verschiedene Programme zur Qualitativen Text-/Datenanalyse (aus der Soziologie und Politologie)

zum Beispiel **MaxQDA**

(http://www.maxqda.de/)

2) Praxisbeispiele
(aus der BAEL Perspektive)

© Pawel Sickinger

universitätbonn

2) *Forschungsfragen für korpusbasierte Studien*

- diskursive Funktionen von zB *question tags, discourse markers, quotatives,* etc.
- Sprechakte, wie zB *apologies, requests* und *compliments*
- grammatische Strukturen in Textsorten oder Varietäten, wie zB *Passivkonstruktionen* in gesprochenen und geschriebenen Texten, im jamaikanischen und amerikanischen oder indischen Englisch, ...
- Diskurs- und Konversationsanalyse, wie sich Gespräche entwickeln, zB *academic office hours, small talk, doctor-patient interaction, ...*

universitätbonn

2) *Themenauswahl bei Hausarbeiten, Abschlussarbeiten und eigener Forschung*

Hausarbeiten und Abschlussarbeiten von BA und MA Studierenden

→ “The get-passive in Jamaican and American English”, “The structure of if-clauses in British English and Learner English”, “Requests in Canadian and New Zealand English”, “The use of 'wonderful' and 'marvelous' in different genres”, ...

Eigene Forschungsarbeit

→ *compliment sequences in conversation, development of diminutives in American English, figurative language use in (B)ELF context, evaluations and assessments in news language, …*

2) *qualitative Textanalyse (SBCSAE 011)*

Transkript einer Unterhaltung

375.79 378.43	DORIS:	Your shirt and beads= .. are most becoming.
378.43 379.61	ANGELA:	.. (H)= Well you know,
379.61 382.27		I think they weigh about a quarter of an ounce.
382.27 384.31		(H) [And that's] the right kind of beads,
382.61 382.90	DORIS:	[Unhunh].
384.31 385.12		.. for summer.
385.12 387.41	ANGELA:	.. (H) A little girl the other day,
387.41 389.90		... remarked about them,
[illegible] 390.34		and,
390.34 392.88		(H) ... almost on impulse,
392.88 394.79		... but I didn't do it.
394.79 396.39		... I thought I'd pull em off,
396.39 398.48		just give em to her if she liked em that [well].
398.21 399.80	DORIS:	[@][2@@2][3@3][4@@@4]
398.50 398.89	SAM:	[2@@2][3@3]
398.48 398.90	ANGELA:	[2You know2].
398.90 400.03		[4(H) But then I4] thought,
400.03 402.66		(H) ... <Q these are the only summer beads I ha=ve Q>.
402.66 403.15	DORIS:	Right.

universitätbonn

2) *qualitative Textanalyse (SBCSAE 011)*

Analyse des Textes
mit eigenen Kodierungen
und Kategorisierungen

C:\Dokumente und Einstellungen\Sue\Eigene Dateien\MAXqdaCodings\Aktuelle Arbeitsdateien\Version_20111130.mx3 - MAXQDA 2007

Projekt Bearbeiten Text Codes Memos Variablen Analyse Visual Tools Fenster MAXDictio ?

Liste der Texte

Texte	17617
SBCSAE Part 1	5703
sbc0001ActualBlacksmithing	248
SBC002Lambada	917
SBC003ConceptualPesticides	1685
SBC0004RagingBureaucracy	402
SBC011ThisRetirementBit	577
SBC013AppeaseTheMonster	1874

tusp	24
tuoth	0
rere	0
difficult	15
_A2_ PRON / DET_Head VERB (INT) ADJ	0
_A3_ (DET) NOUN VERB (INT) ADJ	0
Ø NOUN VERB (INT) ADJ	19
tusp	8
tuoth	0
rere	0
difficult	3
ART NOUN VERB (INT) ADJ	10
tusp	3
tuoth	0
rere	0
difficult	2
POSS NOUN VERB (INT) ADJ	7
tusp	2
tuoth	0
non-adressee	1
addressee	2
rere	0
difficult	2
DET NOUN VERB (INT) ADJ	9
tusp	2
tuoth	0

Text: SBCSAE Part 1\SBC011ThisRetirementBit

273	375.79 378.43	DORIS:	Your shirt and beads= .. are most becoming.
274	378.43 379.61	ANGELA:	.. (H)= Well you know,
275	379.61 382.27		I think they weigh about a quarter of an ounce.
276	382.27 384.31		(H) [And that's] the right kind of beads,
277	382.61 382.90	DORIS:	[Unhunh].
278	384.31 385.12		.. for summer.
279	385.12 387.41	ANGELA:	.. (H) A little girl the other day,
280	387.41 389.90		... remarked about them,
281	389.90 390.34		and,
282	390.34 392.88		(H) ... almost on impulse,
283	392.88 394.79		... but I didn't do it.
284	394.79 396.39		... I thought I'd pull em off,
285	396.39 398.48		just give em to her if she liked em that [well].
286	398.21 399.80	DORIS:	[@][2@@2][3@3][4@@@4]
287	398.50 398.89	SAM:	[2@@2][3@3]
288	398.48 398.90	ANGELA:	[2You know2].

Liste der Codings

SBCSAE Part 1
SBC011ThisRetirementBit
[273-273
ADJ
_A3_ (DET) NOUN VERB (INT)
POSS NOUN VERB (INT) ADJ

SBCSAE Part 1
SBC011ThisRetirementBit
[274-276
ADJ
_A3_ (DET) NOUN VERB (INT)
POSS NOUN VERB (INT) ADJ
tuoth
addressee

375.79 378.43 DORIS: Your shirt and beads= .. are most becoming.

378.43 379.61 ANGELA: .. (H)= Well you know,
379.61 382.27 I think they weigh about a quarter of an ounce.
382.27 384.31 (H) [And that's] the right kind of beads,

1 9 2 0 Oder-Kombination (Standard Analyse)

2) *quantitative Korpusanalyse (COCA)*

KEYWORD IN CONTEXT DISPLAY

Help / Information / Contact

12	1999	SPOK	Fox_OReilly	A	B	C	" Hey , I 'm like cold " or "	I	'm	like	hungry	,	"	No , you 're not like cold .
13	2012	SPOK	CNN_Burnett	A	B	C	, though , as I was watching on video .	I	'm	like	maybe	the	president	needs to spend some time with some m
14	2008	FIC	Analog	A	B	C	response when she went off on a tangent . "	I	'm	like	one	of	the	girls in my house , " she said
15	1990	SPOK	ABC_20/20	A	B	C	full of nothing but success to talk to you .	I	'm	like	one	of	you	because I failed , I 've been embarrassed
16	2006	SPOK	CNN_King	A	B	C	son , Chance has the same thing . BARR :	I	'm	like	one-stop	crazy	,	KING : Like you do n't shake hands
17	1996	SPOK	Ind_Springer	A	B	C	love ? KENNY : Personality , looks . SHERRY :	I	'm	like	some	trophy	,	That 's how he -- that 's how
18	2007	SPOK	ABC_Nightline	A	B	C	I 'm just like a regular person . Never think	I	'm	like	somebody	else	,	I 'm just the same as other people
19	2004	SPOK	CNN_King	A	B	C	they go , " Hi , Frank , " and	I	'm	like	sorry	,	it	's Sharon KING You might say Hillary Clinton
20	1995	NEWS	Chicago	A	B	C	assured us was his spotless Los Angeles home . "	I	'm	like	that	on	the	road as well . When I check into
21	2001	SPOK	CNN_King	A	B	C	or Julia Roberts . I wish -- I hope that	I	'm	like	that	when	I	'm grandmother RIVERS Let me tell you
22	1997	NEWS	SanFranChron	A	B	C	that new things are going to happen and I guess	I	'm	like	that	,	I	've had a really good year and I
23	2005	FIC	SouthernRev	A	B	C	to go to the store , " I say .	I	'm	like	that	,	I	can lie on the spot if I have
24	2006	FIC	KenyonRev	A	B	C	whiskey coming back up . " Yeah , baby .	I	'm	like	the	J.	R.	Ewing of the northern territory . Rip up
[illegible]	[illegible]	[illegible]	[illegible]	A	B	C	tub and throw parties . It 's my place .	I	'm	like	the	love	child	. Of this spot . So was Einar
26	1992	FIC	SouthernRev	A	B	C	. You fool . " # " But here ,	I	'm	like	the	only	person	that died . Is n't that so ?
[illegible]	[illegible]	[illegible]	[illegible]	A	B	C	Milosevic gives in ? MR-DOLE : Well , no ,	I	'm	like	the	other	two	here . We should have not taken the
28	1995	FIC	Bk:UntilEndTime	A	B	C	last night you got here much later . " "	I	'm	like	the	wind	,	" I fluttered my left arm by way
[illegible]	[illegible]	[illegible]	[illegible]	A	B	C	did n't get what I should 've gotten . But	I	'm	like	this	,	you	know . I hang in this type of
30	2001	SPOK	Ind_Oprah	A	B	C	's name ... Mr-GRAMMER : Yeah . WINFREY : ...	I	'm	like	you	is	a	dreamin ' man . Mr-GRAMMER : I 've
[illegible]	[illegible]	[illegible]	Analog	A	B	C	" " Shut up ! You do n't know what	I	'm	like	!	You	're	only saying this so I 'll think it
32	2006	MAG	Cosmopolitan	A	B	C	Why are you talking to that girl ? ' and	I	'm	like	"	She	's	got a nice voice . ' ' "
[illegible]	[illegible]	[illegible]	[illegible]	A	B	C	be cool . ' And then they kept coming .	I	'm	like	'	This	is	going to be the toughest thing I 've
34	2009	SPOK	CNN_Cooper	A	B	C	And as soon as he made the big U-turn ,	I	'm	like	,	cool	,	we 're headed back home . I mean
35	2003	FIC	LiteraryRev	A	B	C	'm all hot inside , hot , real hot .	I	'm	like	,	Hey	,	money . Suck my dick . Then I
36	2009	SPOK	ABC_GMA	A	B	C	daughters in that commercial and we 're boys ? And	I	'm	like	,	I	do	n't know . Because , you know ,
37	1997	FIC	Commentary	A	B	C	" Shit , yeah-I bin studyin . Oweez . Since	I	'm	like	,	I	don	know-three years old ! . . . Karate
38	2011	SPOK	CNN_Morgan	A	B	C	? What are you up to ? You know ,	I	'm	like	,	I	'm	16 . He 's like , try this
39	1997	MAG	Jet	A	B	C	of the store with her breasts hanging out , and	I	'm	like	,	Listen	,	those are really nice breasts , but this
40	2000	MAG	Entertainment	A	B	C	's lobby a few minutes after their arrival . "	I	'm	like	,	No	,	we 're not . " It just sounds
41	2002	MAG	Essence	A	B	C	straight hair . They have this mad Coke-bottle shape and	I	'm	like	,	Oh	,	I want to look like that . I
42	1999	SPOK	NPR_ATC	A	B	C	if I have a bad day , you know ,	I	'm	like	,	Oh	,	whatever , ' you know ? How am
43	2010	SPOK	ABC_20/20	A	B	C	the Navy and he kidnapped my kids . " And	I	'm	like	,	oh	,	really ? CHRIS-CUOMO-1-ABC# (Off-camera) She
44	2008	SPOK	ABC_GMA	A	B	C	's not a normal world , and so sometimes when	I	'm	like	,	oh	,	I want to go to the mall ,
45	2005	SPOK	Ind_Oprah	A	B	C	Ms-RAY : Took the mirror down . VAN-LOKEREN : And	I	'm	like	,	OK	,	Oprah Winfrey is coming to my house .

KWIC Suche nach "I'm like": 1845 mal in 450 Millionen Wörtern → *Konkordanz,* ungefiltert, sortiert

2) *quantitative Korpusanalyse (COCA)*

MILLION WORDS, 1990-2012

-- START --

CLICK ON BARS FOR CONTEXT

CLICK ON COLUMN HEADINGS FOR FREQUENCY IN SUB-SECTION

COMPARE

SECTION	ALL	SPOKEN	FICTION	MAGAZINE	NEWSPAPER	ACADEMIC	1990-1994	1995-1999	2000-2004	2005-2009	2010-2012
FREQ	1843	1041	230	313	233	26	180	332	361	584	386
PER MIL	3.97	10.89	2.54	3.28	2.54	0.29	1.73	3.21	3.51	5.72	7.44

SEE ALL SUB-SECTIONS AT ONCE

SECTION
2000-2004

TOKENS
361

SIZE
102,940,701

PER MILLION
3.51

Help / information / contact

Frequenzvergleich über Genres und Zeit (ungefiltert)

2) *quantitative Korpusanalyse (COCA)*

weiter eingrenzende
Suche nach “I'm like,”
(Suchbefehl: I 'm like ,)

KEYWORD IN CONTEXT DISPLAY

Help / information / contact

11	2003	FIC	LiteraryRev	A	B	C	'm all hot inside , hot ,	real	hot	.	I 'm like ,	Hey	,	money	. Suck my dick . Then I
12	2009	SPOK	ABC_GMA	A	B	C	daughters in that commercial and we	're	boys	? And	I 'm like ,	I	do	n't know	. Because , you know ,
13	2005	NEWS	Denver	A	B	C	with the back of a hand . # "	Sometimes			I 'm like ,	I	do	n't want	to hear it , and other
14	2011	SPOK	CNN_Showbiz	A	B	C	like , dress and suddenly I	hear	my	name .	I 'm like ,	I	do	n't know	what to do . TURNER: So
15	2012	SPOK	ABC_GMA	A	B	C	could n't find it anywhere on	the	internet	. And	I 'm like ,	I	guess	they	, they 've totally scrubbed it from
16	2012	SPOK	CBS_ThisMorning	A	B	C	I -- I never knew what	it	was	, but	I 'm like ,	I	might	as	well get me some beer . TIM-GOLDMAN:
[illegible]	[illegible]	[illegible]	[illegible]	A	B	C	, What do you want that	tag	for	? '	I 'm like ,	It	's	got	your name on it , you know
18	2008	SPOK	NBC_Dateline	A	B	C	, Well , we need to	drive	up	there .	I 'm like ,	Listen	,	the	bottom line is , is she 's
[illegible]	[illegible]	[illegible]	[illegible]	A	B	C	. PRICE : It was like	a	gut	kick .	I 'm like ,	Mike	Jacobs	?	Where did that name come up on
20	2007	SPOK	NBC_Dateline	A	B	C	when they got the news .	(Carl-McLawrence-dr)	TRACEY	:	I 'm like ,	Missing	?	What	do you mean , missing ? How
[illegible]	[illegible]	[illegible]	[illegible]	A	B	C	, Are these magically blood vessels	or	muscles	? '	I 'm like ,	No	,	there	's more than that going on .
22	2011	SPOK	NBC_Dateline	A	B	C	in the morning .	(Photo-of-Emerson-a	Ms-KARRENBAUER:	And	I 'm like ,	oh	,	Ella	's not here . It 's --
23	2007	SPOK	NPR_TalkNation	A	B	C	friends that are supposed to be	traveling	this	way and	I 'm like ,	oh	,	great	. No one 's going to see
24	2011	SPOK	NBC_Dateline	A	B	C	. (Clothes ;	-suitcase-	RENEE	: (Voiceover)	I 'm like ,	Oh	,	OK	. You know ? (Door-of-bedroom) RENEE :
25	2010	MAG	Redbook	A	B	C	I turned into one giant eye	roll	, and	now	I 'm like ,	Oh	,	yeah	. That 's that Rodney that I
26	1998	SPOK	NPR_Saturday	A	B	C	from a perfect score on the	SAT	. "	And	I 'm like ,	oh	...	And	, " Yeah , I 'm the
27	2005	MAG	Esquire	A	B	C	beautiful angel . And I 'm	looking	at	it and	I 'm like ,	Oh	...	my	... God . And then , like
28	1995	SPOK	Ind_Geraldo	A	B	C	my chin and the other one	is	down	here .	I 'm like ,	OK	,	Enough	is enough . ' RIVERA : My
29	2006	SPOK	CBS_Morning	A	B	C	night , as they say ,	you	know	? And	I 'm like ,	Ooh	,	I	like this . ' (Photos-of-Nat) HUGHES :
30	2010	SPOK	NBC_Today	A	B	C	out in the rain , I	'm	freaking	out .	I 'm like ,	Please	,	somebody	. I 'd put a bag on
31	2006	MAG	Cosmopolitan	A	B	C	Why are you talking to that	girl	? '	and	I 'm like "	She	's	got	a nice voice . ' ' "
32	2002	SPOK	NPR_Saturday	A	B	C	lot of the players in the	draft	were	foreign .	I 'm like ,	So	?	They	're good players . I want to
33	2012	SPOK	NBC_Dateline	A	B	C	Mr-GIRARDOT: Linda Sohus is everything	in	this	case .	I 'm like ,	the	story	of	this case , whichever way it goes
34	2006	SPOK	PBS_Tavis	A	B	C	the room , and he introduces	himself	to	me .	I 'm like ,	this	guy	,	(laugh) This dude . This
35	2005	SPOK	Ind_Oprah	A	B	C	-- and I 'm like --	run	to	Lisa .	I 'm like ,	We	got	ta	call an ambulance . We might have
36	2001	MAG	Entertainment	A	B	C	guys capture the tedium of prison	life	'	and	I 'm like	We	're	not	acting . " ' says Dean Winters

2) *quantitative Korpusanalyse (COCA)*

weiter eingrenzende Suche nach “I'm like “/:” (Suchbefehl: I 'm like “|:)

KEYWORD IN CONTEXT DISPLAY

Help / information / contact

CLICK FOR MORE CONTEXT [?] SAVE LIST CHOOSE LIST --------- CREATE NEW LIST [?]

1	2001	SPOK	CNN_KingWknd	A	B	C	miss you so much . You look beautiful . "	I 'm like "	All the way , " and I 'm like "
2	2009	ACAD	EnvironHealth	A	B	C	power . It was literally a plywood shack . And	I 'm like :	Are you serious ? There are people in our country
3	2012	NEWS	NYTimes	A	B	C	flight was taking off in 45 minutes . # "	I 'm like :	Come on , man . I just got a call
4	1999	FIC	Mov:Dogma	A	B	C	" Get the hell out of here ! " And	I 'm like "	Fuck you , you bald cocksucker ! I 'll kick
5	2011	SPOK	NPR_FreshAir	A	B	C	, you know , I look at my dad .	I 'm like :	Hey , Dad , sorry I overslept . And he
6	2009	MAG	Cosmopolitan	A	B	C	your mother better . " That cracks me up .	I 'm like "	Hey , me too ! " C In what ways-if
7	2003	MAG	Entertainment	A	B	C	Planet : You 're the gazelle being chased ! "	I 'm like :	I love you . EW So how vindicated did you
8	[illegible]	[illegible]	[illegible]	A	B	C	, okay . # This is my big moment .	I 'm like :	it 's an emergency , Susan , hold my calls
9	1999	SPOK	CNN_Showbiz	A	B	C	that , and let 's try this , " and	I 'm like "	OK , but we got to go . " (BEGIN-VIDEO-CLIP-)
10	[illegible]	[illegible]	[illegible]	A	B	C	I 'm like " All the way , " and	I 'm like "	OK , next song . " ANGELIL : And we
11	2006	MAG	Cosmopolitan	A	B	C	Why are you talking to that girl ? ' and	I 'm like "	She 's got a nice voice . ' ' "
12	2006	MAG	Cosmopolitan	A	B	C	we go out , Suzie asks how she looks ,	I 'm like "	Uh , you look fine , " but I 'm
13	1997	FIC	Mov:ChasingAmy	A	B	C	how to deal with what I 'm feeling . And	I 'm like "	What the fuck is wrong with you ? " and
14	1998	SPOK	NPR_Weekly	A	B	C	of obscenities back and forth , you know . And	I 'm like "	What 's your problem , you testing me ? I
15	1993	MAG	RollingStone	A	B	C	the end of the week , ' says Pauly .	I 'm like :	Yeah , that 's great , but how long is
16	2011	SPOK	NPR_FreshAir	A	B	C	And he 's like : I 'm right . And	I 'm like :	Yeah , you 're right . People really do n't
17	2005	MAG	Cosmopolitan	A	B	C	the mirror and do n't like what I see .	I 'm like "	You could start jogging more . " I 'm never
18	1999	FIC	Commentary	A	B	C	happened ? she says : Richard dumped me . And	I 'm like :	you 've got to be messing with me big time
19	2011	MAG	MotherJones	A	B	C	you to give a speech in San Francisco . "	I 'm like :	" Whoa , start me on the ground floor .
20	1999	NEWS	Chicago	A	B	C	says . " She really likes Elliott 's drumming .	I 'm like :	' I know he 's a good drummer , but
21	2003	MAG	PsychToday	A	B	C	or ' You 're much prettier in person . '	I 'm like :	' Oh , is that a compliment ? " #
22	2008	NEWS	Chicago	A	B	C	tennis player , ' ' Sharapova said , ' and	I 'm like :	' Uhh , uhh , yeah . It 's me.
23	2012	NEWS	CSMonitor	A	B	C	So I get to Baltimore as an adult , and	I 'm like :	' Who dropped the ball here ? " # He
24	2005	MAG	Entertainment	A	B	C	, " You should work on this , " and	I 'm like "	-she huffs (inadvertently channeling Joyce Brabne

0.811

3) Wie würde der “perfekte Korpuspool” aussehen?

zum Beispiel ungefähr so
...
nur mit Rechnern

© pixelio.de,
gefunden auf: http://www.uzh.ch/news/articles/2010/kann-ich-mich-schlau-trainieren-/lernen460.jpg

universitätbonn

3) *Ausstattung und Verfügbarkeit*

- mehrere Arbeitsplätze, die (Studierenden wie Forschenden) zu Bibliotheksöffnungszeiten zur Verfügung stehen

→ Computer (online, keine Speichermedien nutzbar, Drucker angeschlossen)

→ Informationsmanual zu Korpusauswahl (ggf. elektronische *read me* Datei auf Desktop und Korpusliteratur in greifbarer Nähe)

→ online/offline Auflistung von Korpora, Informationen zu Korpora

→ Analysesoftware (quantitativ wie qualitativ)

universitätbonn

3) *Kooperationen und Schulungen / Kurse*

- Schulungen mit Lehrenden / Bibliotheksangestellten und Studierenden (Einführung in die Nutzung)
- Sprachwissenschaftskurse vor Ort (eventuell auch ***fächerübergreifend,*** andere Geisteswissenschaften, IT Studiengänge, ...)
- Fortbildungskurse für Lehrer (“corpus literacy”, *Data Driven Learning,* Erstellen von Unterrichtsmaterialien)
- ***fächerübergreifendes*** Forschen und Möglichkeit der Verfügbarmachung originärer Datensammlungen der Uni (evtl. Datenbanksysteme)
- ...

universitätbonn

3) *last but not least im Traumwunschkonzert*

- die Möglichkeit für Studierende wie Lehrende, eigene Kopora zu sammeln und zu verarbeiten (d.h. Aufnahmegeräte, Räumlichkeiten, Transkriptionssoftware, POS-Tagger etc.)

- in kleinstem Rahmen momentan im RCEP von Bonn Applied English Linguistics, aber nur zwei Computer, zwei Räume, veraltete Aufnahmegeräte, geringe Personalressourcen, ….

universitätbonn

3) *ein bereits existierender "Mini-Korpuspool" bei BAEL im Research Centre for Empirical Pragmatics*

universitätbonn
bonn applied
english linguistics
prof. dr. klaus p. schneider

home
people
research
publications
outreach
news and events

current projects
variational pragmatics
further projects
small talk
rcep

corpora available

corpora available at the rcep

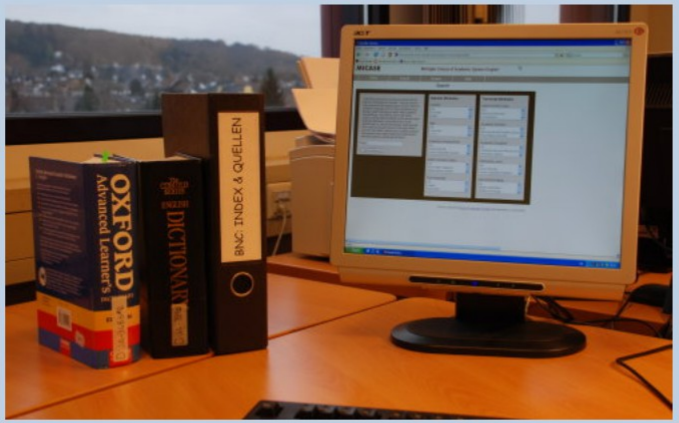
Corpora are large electronic collections of language data. Many corpora not only consist of written data such as newspaper articles but also contain samples of spoken language or even transcripts of conversations - which is one of the reasons why the use of corpora has become increasingly attractive for pragmatic research.

at the rcep, students have access to the following corpora:

Corpus name	Size (in words)	Spoken/ written?	Variety sampled	Tagged?	Ava at RCI
ACE – Australian Corpus of English	1 million	both	Australian	no	y
more information					
ANC ? American National Corpus	100 million (aim)	both	American	yes	y
more information					
BROWN Corpus	1 million	written	American	yes	y
more information					

RCEP

student workspaces



universitätbonn

was the woman , Elizabeth Wallerstein .	Thank	you	so	much	for	remembering	that . T				
esteem , which is important to me , so I	thank	you	so	much	,	and	best of luck to you				
healthy shoes . COURIC : All right . And	thank	you	so	much	.	And	by the way , you c				
ing us . Stanley Gault , Geoffrey Colvin ,	thank	you	so	much	.	We	do appreciate it . Pa				
ay Books , a division of Random House .	Thank	you	so	much	.	Prof-JAGER	: Thank you				
t YERGEN Wednesday night , PBS SNOW	Thank	you	so	much	.	Sorry	for that . As Secr				
VIEIRA : All right , Roni Cohen Sandler .	Thank	you	so	much	.	Ms-SANDLER	: Thanks				
ing to do ? BUCHANAN : Mr. Carr- Kerr ,	thank	you	very	much	for	coming-	for being				
laudia . Thank you very much , Claudia .	Thank	you	very	much	for	being	here . Mr-KA				
-VIDEOTAPE) LEMON : All right , Elaine .	Thank	you	very	much	for	that	. As you can s				
11 years old . (CROSSTALK) LEMON : So	thank	you	very	much	--	thank	you very muc				
. PETER JENNINGS : OK , Aaron Brown ,	thank	you	very	much	.	And	Barbara , I 'm st				
latelinenbc.com. HOLT: All right , Chris ,	thank	you	very	much	.	'					
answer to that question DOBBS Andrea ,	thank	you	very	much	.	Andrea	Koppel from t				
for joining us. SEC-CISNEROS : Robin ,	thank	you	very	much	.	MR-MacNeil	: What d				
take care of that for you . SIMON : OK .	Thank	you	very	much	.	SCHORR	: Are we tal				
different types of questions FITZWATER	Thank	you	very	much	.	Nice	to be here PHIL				
h us and have a good holiday . Mr. JOE :	Thank	you	very	much	SIMON	Speaking	with				
k of the camera I experienced a verbal "	Thank	you	"	from	the	model	instead of the o				
nline and again here tomorrow evening .	Thank	you	,	and	good	night	.				
er . I only wish I was there . WINFREY :	Thank	you	,	Ben	.	Ben	did not want the new				
you very much for being here . DELAY :	Thank	you	.	Brit	.	HUME	: When we return , i				

universitätbonn

Literaturhinweise (Auswahl)

Baker, Paul / Hardie, Andrew / McEnery, Tony (2006): *A Glossary of Corpus Linguistics.* Edinburgh: Edinburgh University Press.

Mukherjee, Joybrato (2009): *Anglistische Korpuslinguistik. Eine Einführung.* Berlin: Erich Schmidt Verlag.

O'Keeffe, Anne / Clancy, Brian / Adolphs, Svenja (2011): *Introducing Pragmatics in Use.* London: Routledge.

O'Keeffe, Anne / McCarthy, Michael (eds.) (2011): *The Routledge Handbook of Corpus Linguistics.* London: Routledge.

Links zu Korpuslisten (Auswahl)

BAEL: Liste von im RCEP verfügbaren Korpora (in der Regel lizensierte Korpora): http://www.linguistics.uni-bonn.de/research/rcep/corpora-available/

→ Informationen und Links zu Korpora

David Lee's Bookmarks for Corpus-Based Linguists: http://www.uow.edu.au/~dlee/CBLLinks.htm

→ Informationen und Links zu vielen verschiedenen Korpora (auch nicht englischsprachige, zT aber schon veraltete Links)

Talkbank.org: www.talkbank.org

→ dient als Webseite für einige offen zugängliche Korpora und Korpusannotationsprogramme

Links zu verschiedenen Korpora und Korpustools (Auswahl)

Korpora:

VOICE: Vienna-Oxford International Corpus of English (ELF)
http://www.univie.ac.at/voice/

MICASE:Michigan Corpus of Academic Spoken English
http://quod.lib.umich.edu/m/micase/

MICUSP: Michigan Corpus of Upper-Level Student Papers
http://search-micusp.elicorpora.info/simple/

WebCorp live: http://www.webcorp.org.uk/live/

Tools:

Wordsmith: http://www.lexically.net/wordsmith/

AntConc: http://www.antlab.sci.waseda.ac.jp/antconc_index.html

Mark Davies Korpusseiten (offen zugänglich, zT mehr Inhalt verfügbar durch individuelle Registierung)

COCA: Corpus of Contemporary American English → http://corpus.byu.edu/coca/

COHA: Corpus of Historical American English → http://corpus.byu.edu/coha/

TIME: The Time Magazine Corpus → http://corpus.byu.edu/time/

SOAP: Corpus of American Soap Operas → http://corpus2.byu.edu/soap/

GloWbE: Corpus of Global Web-Based English → http://corpus2.byu.edu/glowbe/

STRATHY: Corpus of Canadian English → http://corpus2.byu.edu/can/

BYU-BNC: British National Corpus (extract) → http://corpus.byu.edu/bnc/

Corpus of LDS General Conference Talks → http://corpus.byu.edu/gc/

andere Sprachen von Davies:

Corpus del Espanol → http://www.corpusdelespanol.org/

O Corpus do Portugues → http://www.corpusdoportugues.org/