# Manual til LANCHART-søgemaskinen (dgcsssearch.ku.dk)

Philip Diderichsen, september 2020

# 1. Kort om LANCHART-korpusset og søgemaskinen

LANCHART-centret, i daglig tale Sprogforandringscentret, har som sit helt grundlæggende mål at undersøge hvordan det danske talesprog forandrer sig. Det gør vi ved hjælp af et af verdens største korpusser over talesprog: LANCHART-korpusset. Korpusset er er indsamlet ad flere omgange med optagelser af de samme informanter med flere årtier imellem og opdateres løbende med nye optagelser, transskriptioner og opmærkning. Centret har en omfattende it-infrastruktur der understøtter disse aktiviteter, herunder et korpussøgeværktøj der i daglig tale går under navnet Søgemaskinen.

# 2. Adgang til søgemaskinen

Du kan få adgang til søgemaskinen via internettet hvis du er tilknyttet Københavns Universitet og har et KU-id (abc123). Det kræver at du underskriver en fortrolighedserklæring og bliver registreret som bruger via identity.ku.dk.

Se den særskilte vejledning "Anmod om adgang til dgcsssearch.ku.dk".

# 3. Søgning

Søgning foregår ved at specificere et søgesæt og nogle søgekriterier. Herefter kan man få vist søgeresultatet i forskellige formater. Man søger fra søgemaskinens forside, se Figur 1.



Figur 1. Søgemaskinens forside. Man søger ved at indtaste søgetermer i det orange felt.

## 3.1. Filer

Under Filer kan man specificere hvilke dele af LANCHART-korpusset man vil søge i. Korpusset består af en stor mængde udskrevne samtaler der tilhører de forskellige delprojekter som Sprogforandringscentret beskæftiger sig med. Hver samtale udgør en selvstændig fil i korpusset. Læs mere om delprojekterne på https://dgcss.hum.ku.dk/forskning/undersoegelsesomraader

Filer  Alle Sæt 70'erne rediger Enkelt amager3gI-09-AIH	Marker feltet "Alle" hvis du vil søge i alle Centrets udskrevne samtaler.
Filer Alle Sæt ✓ 70'erne 80'erne 90'erne alle filer fra København alle filer fra København alle filer fra Næstved	Hvis søgningen skal foretages i en specifik del af Centrets korpus, skal feltet "Sæt" markeres, og herefter vælges det ønskede delkorpus i drop down-listen
Filer Alle Sæt 70'erne rediger © Enkelt ✓ amager3gl-09-AIH amager3gl-09-FWA amager3gl-09-HCM	Det er også muligt at lave en søgning i en specifik fil. Her markeres feltet "Enkelt", og den ønskede fil markeres.

Det er desuden muligt at oprette sit eget sæt af filer hvis de ønskede sæt ikke passer til ens søgning. Klik på "Rediger" lige under feltet "Sæt". Herefter åbner en ny side.

	褼 Rediger intervi	ewsæt	× +											
÷) → C'	۵ (	0 🔒 I	https://dgcsssear	ch. <b>ku.dk</b> /se	arch/interviev	vset.html?i	interviewse	t=70	©	7 ☆	) I	1\ [	•	$\equiv$
Re	dige	r ir	nterv	iew	/sæ	et								
90'erne alle filer fra Køt alle filer fra Køt alle filer fra Næ alle filer fra Od alle filer fra Od alle filer fra Vin Alle prioriterede	benhavn benhavn stved der derup a filer a inkl. Odder 2 ekstra		Interviewsæt	tet hedder:	70'erne									
alt i AmDa_febr alt i ArgDa_febr alt i CanDa_febr alt i CoAmDa_f Amager 3gl	ruar 2019 ruar 2019 oruar 2019 jebruar 2019		Tilgænge	elige in	terviews	;	,	Valgte interv 91 interviews val	iews <sup>Igt</sup>					
amda_canica_r amda_CLaVE- amda_LLaVE- amda_LLaVE- amda_skit=00- AmDa-kt-dana AmDa-kt-dana AmDa-kt-dana AmDa-kt-dana AmDa-kt-dana AmDa-kt-dana AmDa-kt-dana AmDa-kt- Bysoc Cyl-2gl Bysoc	svembor 2017 subkorpus 11-2017 - 09-11-2017 200516 3 gl ssoc2-ny aruppe ntarview 2017		amager3g1-09-FW amager3g1-09-FW amager3g1-09-FW amager3g1-09-FK amager3g1-09-FK amager3g1-09-FK amager3g1-09-FK amager3g1-09-FK amager3g1-09-FK amager3g1-09-FK amager3g1-10-FK amager3g1-10-FK amager3g1-10-FK amager3g1-10-FK amager3g1-10-FK amager3g1-10-FK amager3g1-10-FK amager3g1-10-FK amager3g1-10-FK amager3g1-10-FK amager3g1-10-FK amager3g1-10-FK amager3g1-10-FK amager3g1-10-FK amager3g1-10-FK amager3g1-10-FK	4 (A) (A) (A) (A) (A) (A) (A) (A) (A) (A)			>	winderug2g1-78-AEL winderug2g1-78-AEL winderug2g1-78-AEL PGU- winderug2g1-78-AEL Minderug2g1-78-ALH Minderug2g1-78-ALH Minderug2g1-78-BEJ winderug2g1-78-BAB winderug2g1-78-BAB winderug2g1-78-GPD winderug2g1-78-GPD- Minderug2g1-78-GPD- Minderug2g1-78-MN winderug2g1-78-MN winderug2g1-78-MN winderug2g1-78-MN winderug2g1-78-MN winderug2g1-78-MN winderug2g1-78-MN winderug2g1-78-MN winderug2g1-78-MN winderug2g1-78-MN winderug2g1-78-MN winderug2g1-78-MN winderug2g1-78-MN winderug2g1-78-MN	+RDN+SSE +JBG+LDR +LHS+NJM +KRS+LIK-1 +LET+MJJ KES+SCH interview					

Figur 2. Siden "Rediger interviewsæt". Her kan kan definere sit eget sæt af interviews at søge i.



Figur 3. Bunden af siden "Rediger interviewsæt" hvor nye interviewsæt oprettes ved at klikke på "Opret nyt sæt".

På siden Rediger Interviewsæt klikkes der på funktionen "Opret nyt sæt" i bunden af listen over interviewsæt til venstre. Herefter vælges et navn til det nye sæt, og de ønskede filer i sættet markeres i den venstre spalte. Flyt herefter filerne over i højre spalte ved hjælp af pilene mellem spalterne. Til sidst gemmes ændringerne, og sættet er oprettet.

## 3.2. Informanter

Under Informanter på søgemaskinens forside kan man specificere hvilke samtaledeltagere i den udvalgte del af korpusset man er interesseret i.

Informanter  Alle Sæt Alle prioriterede inkl. Odder 2 ekstra rediger Enkelt E	Hvis man ønsker at medtage data fra samtlige samtaledeltagere der indgår i det udvalgte filsæt, skal feltet "Alle" markeres. Dette medtager såvel
	prioriterede informanter, interviewere og alle andre
	samtaledeltagere – fx en datter til en informant, der

	7
	blot deltager ved at sige "hej" eller et spædbarn der
	pludrer.
Informanter	Det er også muligt at udvælge bestemte talere i det
Alle Sæt  Alle prioritierede inkl. Odder 2 ekstra	pågældende filsæt – eksempelvis prioriterede
Enkelt ArgDa-de fem versets informanter_010616 BySoc 1 og 3	informanter, interviewere, bestemte aldersgrupper eller
Danske Stemmer-Interviewere Emil22112019 amda	deltagere med et bestemt køn. Feltet "Sæt" markeres,
	og den ønskede talersammensætning vælges i drop
	down-listen. Prioriterede talere er talere der er udvalgt
	af forskningsmæssige grunde på baggrund af
	forskellige sociolingvistiske variable – se
	https://dgcss.hum.ku.dk/forskning/undersoegelsesomra
	ader.
Informanter	Det er ligeledes muligt at vælge en enkelt informant i
Alle Sæt Alle prioriterede inkl. Odder 2 ekstra	det pågældende filsæt. Dette gøres ved at markere
Enkelt AIH	feltet "Enkelt" og skrive informantkoden for den
	pågældende taler.

Det er muligt at sammensætte sit eget talersæt hvis den ønskede sammensætning ikke allerede findes i drop down-listen. Her skal man følge den samme procedure som bruges til redigering af selve filsættet – se ovenfor.

## 3.3. Søgekriterier

Her skriver man det man ønsker at søge efter. Centrets korpus består af en lang række TextGrids – et for hver samtale – der er inddelt i mange lag (tiers), som hver især er knyttet til en bestemt arbejdsgruppe – fx udskrivning, grammatik, fonetik m.v. Der kan søges i alle tiers. Man skal dog være opmærksom på at det ikke er alle samtaler i LANCHART-korpusset der er blevet behandlet af samtlige analysegrupper. Alle filer i korpusset indeholder dog et ortografitier for hver taler, som indeholder den ortografiske udskrift af det den pågældende taler siger i løbet af samtalen.

Korpusset der ligger til grund for søgemaskinen, er opbygget sådan at alle indholdselementer i de enkelte lag optræder i deres eget tidsafgrænsede interval – fx optræder der kun ét ord i hvert interval i de forskellige ortografitiers. Søgemaskinen fungerer på den måde at søgeresultater hentes fra alle de intervaller der imødekommer søgekriterierne. For at få de korrekte søgeresultater frem er det derfor vigtigt at ens søgekriterier ikke kræver indholdet fra mere end ét interval ad gangen.

Søgekriterier tilføj søgekriterie fjern ✓ ortografi Almarkering Aktivitetstype Aktivitetstype Aktivitetstype	Først skal man specificere hvilket tier man vil foretage søgningen i – drop down-listen viser alle søgetiers, uanset om de findes i det valgte datasæt eller ej.
	Derefter skrives den søgestreng som man vil søge på – feltet er rødt hvis det ikke er udfyldt, da der ikke må optræde tomme søgekriterier.
Søgekriterier         fjern ortografi         Søg Der må ikke være tomme søgekriterier         Søg defter gentagelser i interview         Gå til statistiksiden.	Når søgestrengen er skrevet, er det vigtigt at specificere om strengen udgør hele intervallet – dvs. al data i det pågældende interval – eller om den blot er en del af intervallet. De forskellige muligheder kan ses i drop down- listen.

Det er muligt at søge på flere søgestrenge på én gang – dvs. en såkaldt kontekstsøgning.

Søgekriterier	
tilføj søgekriterie	
fjern ortografi 💿	Hele ordet
fjern ortografi 🕤 hvad	Hele ordet 🧿 + / - 0 - 5 - 4 - 3 - 2 - 1 🗹 0 1 2 3 4 5 + / -
Sag Der må ikke være tomme søgekriterier	

Figur 4. Kontekstsøgning. Klik på "tilføj søgekriterie" for at tilføje et søgefelt. Husk at sætte flueben i den rigtige position i forhold til den øverste søgestreng så du ikke fx kommer til at søge på to forskellige ord i samme position.

Klik på funktionen "tilføj søgekriterie" - så vil der dukke endnu en søgestrengslinje op. Denne linje udfyldes på samme måde som den foregående søgestreng, men her skal man desuden specificere om den tilføjede søgestreng skal komme i intervallet før (-5 til -1), efter (+1 til +5) eller samtidigt med den første (0). Det er både muligt at vælge forskellige søgetiers i kontekstsøgningen og at anvende det samme søgetier i begge søgestrenge. Ved hjælp af regulære udtryk er det muligt at foretage søgninger hvor man fx kan søge efter tomme intervaller eller en mængde af former (tekststrenge) på en gang. Søgestrengen markeres som et regulært udtryk ved hjælp af menuen til højre for søgestrengen. Eksempler på regulære udtryk:

Tabel 1. Regulære udtryk. Husk at vælge "regulært udtryk" i menuen til højre for søgefeltet hvis du bruger regulære udtryk.

Symbol	Søgestreng	Forklaring
.*	.*	Finder alle intervaller.
.+	.+	Finder alle intervaller der ikke er tomme.
	man du	Finder intervaller der indeholder sekvensen <i>man</i> eller <i>du</i> .
( )	l(i æ)gge	Finder intervaller der indeholder sekvensen <i>ligge</i> eller <i>lægge</i> .
^	^G	Finder intervaller der starter med <i>G</i> .
\$	ik\$	Finder intervaller der slutter med <i>ik</i> .
token.*	sur.*	Finder intervaller der indeholder sekvensen <i>sur</i> evt. efterfulgt af flere tegn.
		Fx <i>sure, glasuren, armbåndsuret</i> etc.

De forskellige symboler kan godt anvendes sammen i det pågældende regulære udtryk. Vær opmærksom på at man skal anvende tegnene ^ og \$ til at specificere start og slutning af et søgeord hvis man ikke ønsker at få fx affikser med i søgningen – hvis man eksempelvis ønsker at søge efter forekomster med enten ligge eller lægge, men ikke ønsker at medtage indlægge eller lægger, skrives det regulære udtryk på følgende måde: ^l(i|æ)gge\$. ^ og \$ kan også anvendes flere gange i det regulære udtryk. Hvis man ønsker forekomster med en forskel mellem ligget og lagt, kan det regulære udtryk skrives således: ^ligget\$|^lagt\$.

# 4. Resultater

Når søgningen er færdig, bliver der åbnet et nyt vindue med søgeresultatet øverst.



Figur 5. Den oversigt over søgeresultater der præsenteres når søgningen er færdig.

Der er nu flere måder at behandle sit søgeresultat på: Eksporter til CSV, vis som slices, vis (og eksporter) som tabel og kollokationer.

## 4.1. Eksporter til CSV

Man kan vælge at eksportere resultatet direkte ind i Excel. Klik på linket "Eksportér til CSVfil" – herefter kan du enten vælge at åbne filen eller gemme den.



Figur 6. Prompt der dukker op når man har valgt "Eksportér til CSV-fil".

CSV-filen viser samtlige intervaller i korpusset der matcher søgekriterierne. Hvert match vises i sin egen række sammen med samtale- og taleroplysninger samt indholdet på det pågældende sted i samtalen – også i alle andre tiers end det man har søgt i. Formatet er derfor velegnet til videre statistisk behandling i forbindelse med kvantitative undersøgelser.

## 4.2. Vis som slices

Hvis man klikker på linket "Vis som slices", kommer hver enkelt match frem i en konkordans.

• • •	≷ Søgeresultat		×	Søgeres	sultater	×	< +									
$\leftrightarrow$ > G	۵	🖸 🔒 ht	tps://dgcs	ssearch.k	u.dk/sear	ch/showslice	es.html				(	פ בי	liily	1	۰ 🍯	≡
Søgeresultate	erne <b>1 - 15</b> ud	af <b>472</b> for	" <b>altså</b> " i	ortograf	i i Clarin	13 gruppesa	<i>mtaler,</i> f	or alle inf	ormanter				red	igér se	<u>igekrite</u>	<u>ier</u>
1 2 3	4 5 6	7	8 9	10 11	12 1	13 14 15	i 16	17 18	19 20	21 22	23	24 25	26 27	28 2	9 30	31
clarin3	gl-10-A	EB+	AOE	HBF	G+l	LML		273 47	12 Afenil kel							274 466
ortografi		oi			iea	tror	alteå	ah	z Aispiriya	gengæld	at					274.400
-		ej			Jeg	uor	ansa		u	gengæiu	aı					
- tilføj tier -	0															
🤇 - tilføj tier -	<b>)</b>	← forrige													følger	de ⇒
clarin3 AEB (338	∎ gl-10-A .342)	← forrige EB+ <sup>337.694</sup>	AOE	E+BF	G+l	LML		338.34	12 Afspil lyd						følger	de ⇒ 339.32
- tilfej ter - Clarin3( AEB (338 ortografi	□ gl-10-A .342)	← forrige EB+ 337.694 været	AOE så	<b>+BF</b>	G+L	LML	altså	338.34 han	12 Afspil lyd har	været	Så	sådan			følger	de ⇒ 339.32
- tilfej tier -	₽ gl-10-A .342)	← forrige EB+, <sup>337.694</sup> været	AOE så	E+BF	G+L	_ML medierne	altså	338.34 han	12 Afspil lyd har	været	så	sådan			følger	de ⇒ 339.32
- tillej tier -	₿ gl-10-A .342) ₿	← forrige EB+ 337.694 været ← forrige	AOE så	<b>+BF</b>	G+l	_ML medierne	altså	338.34	12 Afspil lyd har	været	sâ	sådan			følger	de ⇒ 339.32 de ⇒
clarin3 AEB (338 ortografi clarin3 AEB (342	gl-10-A .342) gl-10-A .505)	⇐ forrige EB+ 337 694 været ⇐ forrige EB+ 341.259	AOE <sup>sà</sup>	E+BF <sup>meget</sup> E+BF	G+l G+l	LML <sup>medierne</sup>	altså	338.34 han 342.505	i2 Afspil lyd har i Afspil lyd	været	Sâ	sådan			følger følger	de ⇒ 339.32 de ⇒ 43.3825
clarin3 AEB (338 ortografi ortografi clarin3 AEB (342 ortografi ortografi clarin3	gl-10-A .342) gl-10-A .505)	← forrige EB+ 337.694 været ← forrige EB+ 341.259 fulgt	AOE sà AOE det	E+BF <sup>meget</sup> E+BF så	G+L G+L	LML medierne LML	altså	338.34 han 342.505	i2 Afspil lyd har Afspil lyd alle	været har	så jo	sådan haft			følger følger	de ⇒ 339.32 de ⇒ 43.3825

Figur 7. Konkordansvisning. Vises når man har valgt "Vis som slices".

Visningen indeholder følgende:

- Øverst i vinduet kan man fortsat se antallet af søgeresultater.
- Til venstre kan man se hvilket tier der er valgt. Her er det muligt at få vist forskellige opmærkningslag samtidig ved at tilføje flere tiers til det pågældende interval.
- Over hver enkelt forekomst kan man se hvor langt inde i den pågældende samtale forekomster findes her er tiden angivet i sekunder
- Det eller de intervaller der matcher søgekriterierne, er markeret med lyseblåt
- Under hver forekomst er det desuden muligt at se det udsnit af den pågældende samtale der ligger enten lige før eller lige efter det synlige udsnit. Dette gøres ved at klikke på hhv. "forrige" eller "følgende" under den enkelte forekomst
- Det er desuden muligt at lytte til det synlige udsnit. Klik på "Afspil lyd..." over udsnittet. Af informationssikkerhedshensyn er dette dog kun muligt for en udvalgt del af materialet, i skrivende stund interviewsættet "Clarin3 gruppesamtaler" samt filerne fra projektet "Danske Stemmer" (dvs. filer der begynder med "amda" eller "argda").

## 4.3. Vis (og eksporter) som tabel

Hvis man i stedet ønsket et skematisk overblik over sine søgeresultater i den kontekst de forekommer i, klikker man på linket "Vis som tabel" på resultatsiden. Her kommer der en side frem hvor man skal vælge hvor meget kontekst (dvs. hvor mange ord) på hver side af sit resultat man vil have vist, samt hvilke tiers man gerne vil se resultater fra.

Image: Arg     Image: Arg <th>••</th> <th>🔵 🛛 💐 Søger</th> <th>esultat</th> <th></th> <th>× 🐹</th> <th>Skemavis</th> <th>sning af søger</th> <th>resultatet <math> imes</math></th> <th>🗮 Søgi</th> <th>eresultater</th> <th>×</th> <th>+</th> <th></th> <th></th> <th></th> <th></th> <th></th> <th></th>	••	🔵 🛛 💐 Søger	esultat		× 🐹	Skemavis	sning af søger	resultatet $ imes$	🗮 Søgi	eresultater	×	+						
Tabelegenskaber         Slices far:         Visinigstier 1:         Output         Visinigstier 5:         New wagt         Slices file         Vis Tuble informantoplysning:         Vis         Vis         Vis Tuble informantoplysning:         Vis Tuble informatoplysning:         Vis Tuble informatop	$\leftarrow$	) C' 🛈	Ū	🔒 https	://dgcssse	arch.ku	<b>dk</b> /search/	/showgrid.h	tml				··· 🗵 ··	\$ lii\	<b></b>	۲	<b>.</b>	≡
Slices før: 5   Slices efter: 5   Max resultater pr -1 betyder:   side: Vis alle.   Visningstier 1: organi   Visningstier 3: kke valgt   Visningstier 4: kke valgt   Visningstier 5: ikke valgt   Visningstier 7: ikke valgt   Indee informantoplysninger:   Vis Downlead   Ingen tabel endnu.	т	abelegens	kaber	•														
Slices efter: 5   Max resultater pr -1   side: Vis alle.   Visningstier 1: ortografi   Visningstier 2: IPA   Visningstier 3: likke valgt   Visningstier 4: likke valgt   Visningstier 5: ikke valgt   Vis fulde infor- mantoplysninger:   Vis Download   Ingen tabel endnu.	s	Slices før:	5	8	1													
Max resultater pr       -1       -1 betyder:         side:       Vis alle.         Visningstier 1:       ortografi         Visningstier 2:       IPA         Visningstier 3:       Ikke valgt         Visningstier 4:       Ikke valgt         Visningstier 5:       Ikke valgt         Visningstier 5:       Ikke valgt         Vis fulde infor-       Imaskinen         Vis       Download         Tilbage til maskinen       Ingen tabel endnu.	s	Slices efter:	5															
side: Vis alle. Visningstier 1: ortografi © Visningstier 2: IPA © Visningstier 3: likke valgt © Visningstier 5: likke valgt © Vis fulde infor- mantoplysninger: © Tilbage til resultatet. Tilbage til maskinen	N	Aax resultater pr	-1		-1 betvd	ler:												
Visningstier 1: ordografi   Visningstier 2: IPA   Visningstier 3: kke valgt   Visningstier 4: kke valgt   Visningstier 5: kke valgt   Visningstier 6: ikke valgt   Visningstier 7: ikke valgt   Vis Download ikke valgt   Tilbage til maskinen	S	ide:	Vis alle.															
Visningstier 2: IPA C Visningstier 3: Ikke valgt C Visningstier 4: Ikke valgt C Visningstier 5: Ikke valgt C Vis Loke valgt C Vis Download Tilbage til resultatet. Tilbage til maskinen	V	/isningstier 1:	ortografi			0												
Visningstier 3: Ikke valgt  Visningstier 4: Ikke valgt  Visningstier 5: Ikke valgt  Vis fulde infor- mantoplysninger: Vis Download Tilbage til resultatet. Tilbage til maskinen Ingen tabel endnu.	V	/isningstier 2:	( IPA															
Visningstier 4: kke valgt Visningstier 5: kke valgt Vis fulde infor- mantoplysninger: Vis Download Tilbage til resultatet. Tilbage til maskinen	V	/isningstier 3:	Ikke valg	t		0												
Visningstier 5: ikke valgt	V	/isningstier 4:	Ikke valg	t		0												
Vis tulde infor- mantoplysninger: Vis Download Tilbage til resultatet. Tilbage til maskinen	V	/isningstier 5:	Ikke valg	t		0												
Vis       Download         Tilbage til resultatet.       Tilbage til maskinen         Ingen tabel endnu.	v n	/is fulde infor- nantoplysninger:																
Tilbage til resultatet. Tilbage til maskinen Ingen tabel endnu.		Vis Download																
Ingen tabel endnu.	Ti	ilbage til resultate	et. <u>Til</u>	bage til m	naskinen													
ingen tabel endnu.																		
	Ingen	tabel endnu.																

Figur 8. Side der vises når man har valgt "Vis som tabel". Her speficiferer man hvordan man vil have sin resultattabel.

Når man har valgt sine tabelegenskaber, kan man vælge enten at se resultaterne på skærmen (klik på "Vis") eller at eksportere resultaterne til en fil (klik på "Download").

Her ses resultaterne på skærmen.

	Skemavisning	af sø	øgeresultatet $ imes$	+															
€	$\rightarrow$ C' $$	0	https://dgcs	ssea	rch. <b>ku.dk</b>	/search/sho	wgrid.htm	า				50%	⊘	ជ	111		۲	<b>.</b>	≡
-																			
Stor	e fer																		
Slice	is effer: 5																		
Max	resultater pr -1 -1 betyder:																		
side	Vis alle.																		
Visn	ingstier 1: ortigraf																		
Visn	ingstier 3: Iske valgt																		
Visn	ingstier 4: Ikke valgt																		
Visn Vis f	ingstier 5: Itike valgt																		
man	toplysninger:																		
TIP	a Download Tilbace til maskinen																		
				1															
Fund	Filnavn Opt. Opt	t.år.	Tier Inf XMin Ke	n S.kl.	F.år. Projek	t -5 N ei	4	-3	-2	-1	0 A selice	1	2 51	3	4	5	_		
2	clarin3gl-10-AEB+AOE+BFG+LML gl 20	010 0	ortografi AEB 338,34 K	N/A	1991 CLARI	N været	sà	meget	jeg	medierne	altsà	han	har	været	a. Så	sådan	-		
3	clarin3gl-10-AEB+AOE+BFG+LML gl 20	010 o	ortografi AEB 342,5 K	N/A	1991 CLARI	N fulgt	det	så	tæt		altsá		alle	har	jo	haft			
4	clarin3gl-10-AEB+AOE+BFG+LML gl 20	010 0	ortografi AEB 353,62 K	N/A	1991 CLARI	N man	lige	pludselig	ikke		altsà	øh	44	at	at	han	_		
6	clarin3gl-10-AEB+AOE+BFG+LML gl 20 clarin3gl-10-AEB+AOE+BFG+LML gl 20	010 0	ortografi AEB 422,82 K	N/A	1991 CLARI 1991 CLARI	N jeg	det	var	ret	øh	altsá		OPE	var	ret	overgøn	e		
7	clarin3gl-10-AEB+AOE+BFG+LML gl 20	)10 o	ortografi AEB 625,68 K	N/A	1991 CLARI	N	det	er	jo	klart	altsá			men		men	-		
8	clarin3gl-10-AEB+AOE+BFG+LML gl 20	010 o	ortografi AEB 637,94 K	N/A	1991 CLARI	N giver	jo	ingen	mening		altsà				men				
9	clarin3gl-10-AEB+AOE+BFG+LML gl 20 clarin3gl-10-AEB+AOE+BEG+LML gl 20	010 o	ortografi AEB 656,48 K	N/A	1991 CLARI	N men	det	er	jo		altsà	det	er	jo	klart	fordi	_		
11	clarin3gl-10-AEB+AOE+BFG+LML gl 20	010 0	ortografi AEB 732,47 K	N/A	1991 CLARI	N er	jo	fra	Jehovas	Vidne	altsà	der	er	jo	ikke	noget	_		
12	clarin3gl-10-AEB+AOE+BFG+LML gl 20	)10 o	ortografi AEB 787,96 K	N/A	1991 CLARI	N sá	en	dokumentar	det	var	altsá	for	sà	noget		True			
13	clarin3gl-10-AEB+AOE+BFG+LML gl 20	010 0	ortografi AEB 811,78 K	N/A	1991 CLARI	N var		hel	hel	lille	altsà				elve		_		
15	clarin3gl-10-AEB+AOE+BFG+LML gl 20	010 0	ortografi AEB 903.26 K	N/A	1991 CLARI	N UI	fuldstændig	gour	nej	men	altså			det	er	io	-		
16	clarin3gl-10-AEB+AOE+BFG+LML gl 20	)10 o	ortografi AEB 934,81 K	N/A	1991 CLARI	N stue		f-	løft-	d-	altsà	flyttet	alle	v-	øh	møbler			
17	clarin3gl-10-AEB+AOE+BFG+LML gl 20	010 o	ortografi AEB 975,56 K	N/A	1991 CLARI	N	de	L		d-	altsá		de	er			_		
18	clarin3gl-10-AEB+AOE+BFG+LML gl 20 clarin3gl-10-AEB+AOE+BFG+LML gl 20	010 0	ortografi AEB 1029,14 K	N/A N/A	1991 CLARI 1991 CLARI	N	ja	og	nan det	er	altsá	ogsa det	var	grineren så	noget	med	_		
20	clarin3gl-10-AEB+AOE+BFG+LML gl 20	010 o	ortografi AEB 1077,91 K	N/A	1991 CLARI	N er	det	grineren		men	altsà	pointen	med	det	er	jo			
21	clarin3gl-10-AEB+AOE+BFG+LML gl 20	)10 o	ortografi AEB 1281,42 K	N/A	1991 CLARI	N er	bare	vild	med	fodbold	aitsá	lad	nu		barnet	ikke			
22	clarin3gi-10-AEB+AOE+BFG+LML gl 20 clarin3gi-10-AEB+AOE+BFG+LML nl 20	010 0	ortografi AEB 1291,84 K ortografi AEB 1312.66 K	N/A N/A	1991 CLARI 1991 CLARI	N nu N skulle	bare	køre været	lidt	ræs	altsá	det	skal	jeg	da men	ikke	_		
24	clarin3gl-10-AEB+AOE+BFG+LML gl 20	010 0	ortografi AEB 1330,84 K	N/A	1991 CLARI	N ikke		det	synes	jeg	altsà	er	for	vildt		at			
25	clarin3gl-10-AEB+AOE+BFG+LML gl 20	)10 o	ortografi AEB 1382,9 K	N/A	1991 CLARI	N det	er	slet	ikke		altsá		pà	den	ene	side			
26	clarin3gl-10-AEB+AOE+BFG+LML gl 20 clarin3gl-10-AEB+AOE+BEG+LML gl 20	10 0	ortografi AEB 1387,77 K	N/A N/A	1991 CLARI	N også N den	pænt	frygteligt	ikke		altsá		at	øh	at det	de	_		
28	clarin3gl-10-AEB+AOE+BFG+LML gl 20	10 0	ortografi AEB 1396,66 K	N/A	1991 CLARI	N fred-	øh	fredelig	gjort		altsà		der	k-	det	kunne	-		
29	clarin3gl-10-AEB+AOE+BFG+LML gl 20	010 o	ortografi AEB 1496,98 K	N/A	1991 CLARI	N	jeg	har	58-		altså	jeg	kommer	bare	hjem	med			
30	clarin3gl-10-AEB+AOE+BFG+LML gl 20 clarin3gl-10-AEB+AOE+BEG+LML gl 20	10 0	ortografi AEB 1638,9 K	N/A N/A	1991 CLARI	N vi	øh	vi	lever		altsá	vi	lever	vi	vi	lever	_		
32	clarin3gl-10-AEB+AOE+BFG+LML gl 20	010 0	ortografi AEB 1649,89 K	N/A	1991 CLARI	N at	være	bange			altsà			100	ja		-		
33	clarin3gl-10-AEB+AOE+BFG+LML gl 20	)10 o	ortografi AEB 1731,92 K	N/A	1991 CLARI	N slet	ikke	komme	nogen	ind	altsà	pà	lufthavnen	som	ikke	skulle			
34	clarin3gl-10-AEB+AOE+BFG+LML gl 20 clarin3gl-10-AEB+AOE+BEG+LML gl 20	010 0	ortografi AEB 1736,98 K	N/A N/A	1991 CLARI	N derind	alene	niende	ki-	i sårlan	altsà	medicin	niendeklasses	eà	ung	+	_		
36	clarin3gl-10-AEB+AOE+BFG+LML gl 20	010 0	ortografi AEB 1755,94 K	N/A	1991 CLARI	N pas		overhovedet	eventuelle	neose	altså	med	øh	op	i	flyet	-		
37	clarin3gl-10-AEB+AOE+BFG+LML gl 20	)10 o	ortografi AEB 1775,43 K	N/A	1991 CLARI	N sådan	at	gemme	det	øh	altsà	at	dysse	det	ned	ikke			
38	clarin3gl-10-AEB+AOE+BFG+LML gl 20	)10 o	ortografi AEB 1835,99 K	N/A	1991 CLARI	N det	var	for		vildt	altså	jeg	synes	fandeme	det	var			_

Figur 9. Tabelvisning der kommer frem på siden når man klikker på "Vis". Man kan også vælge "Download" for at downloade tabellen som fil.

Ud for hver forekomst findes forskellige oplysninger om den enkelte taler. I position 0 findes det pågældende søgeresultat. Bemærk at der kan være flere linjer til hvert søgeresultat – en linje pr. tier.

Det er også muligt at eksportere sit resultat til en fil efter man har fået vist resultaterne på skærmen. Klik på "Download" i den lyseblå boks med tabelegenskaber. Herefter kan man vælge enten at åbne filen eller gemme den.

Hvis man vælger at eksportere resultaterne under tabelvisningen, får man den kontekst med som man har valgt under tabelegenskaberne – som også kan ses hvis man vælger at vise resultaterne på skærmen. Man får dog ikke de øvrige oplysninger fra korpusfilen med som man kan få ved at eksportere til fil direkte fra den første resultatside.

#### 4.4. Kollokationer

Hvis søgestrengen kun består af en enkelt tekststreng, er det muligt at få vist søgeresultatets kollokationer – dvs. ord som søgestrengen forekommer hyppigere sammen med end man kunne forvente ud fra deres hyppighed i datasættet.

Hvis man ønsker at se søgeresultatets kollokationer i det valgte datasæt klikker man på linket "Find kollokationer for resultatet" på resultatsiden.

Øverst på resultatsiden kan man se hvor mange ord der findes i det valgte datasæt og hvor mange forskellige ord (tokens) der findes i datasættet. Nedenunder kan man se den højre og venstre kontekst for den pågældende søgestreng.

Tabellerne viser dels kontekstordet, dels det antal gange hvor kontekstordet optræder sammen med søgestrengen, og dels Mutual Information-scoren – dvs. et indeks for hvor sandsynligt det er at finde kontekstordet sammen med søgestrengen i det valgte datasæt. Kollokationerne er sortereret efter MI-scoren.