Guide to the LANCHART search engine (dgcsssearch.ku.dk)

Philip Diderichsen, september 2020

1. Short introduction to the LANCHART corpus

The LANCHART Centre (LANCHART is short for *language change in real time*) is a research centre at the University of Copenhagen focusing on Danish spoken language change. The core resource of the centre is the LANCHART corpus, a world-class corpus of sociolinguistic interviews. The corpus has been built up through several rounds of recordings of the same informants across several decades. New recordings, transcriptions, and linguistic markup from various research projects are being added on a regular basis. The corpus is searchable through a web interface, which is described in this document.

2. Access to the search engine

The search engine can be accessed online if you are a registered associate of the University of Copenhagen and have an official ID (KU-ID, of the form abc123). It is further required that you sign a non-disclosure agreement, and that you are registered as a user of the search engine through identity.ku.dk.

See the separate guide "Access to dgcsssearch.ku.dk".

3. Search

Searches are performed by specifying a subset of the corpus and a set of search criteria. The results can be displayed in various formats. Searches are performed from the search engine front page, see Figure 1.



Figure 1. Front page of the search engine. The search term is entered in the orange field.

3.1. Files

Under "Filer" ('files'), one may specify which subset of the LANCHART corpus one wishes to search in. The corpus consists of a sizeable collection of transcriptions of the interviews and conversations belonging to the various projects of the LANCHART Centre. Each conversation corresponds to a single file in the corpus. Read more about the individual projects here: https://dgcss.hum.ku.dk/forskning/undersoegelsesomraader.

| Filer Alle Sæt 70'erne rediger Enkelt amager3gI-09-AIH | Select the option "Alle" ('all') if you wish to search all transcriptions. |
|---|--|
| Filer Alle Sæt ✓ 70'erne 80'erne 90'erne alle filer fra København alle filer fra København alle filer fra Næstved | To search in a specific part of the corpus, choose the option "Sæt" ('set'). A specific, named subset of files can then be chosen from the drop-down menu. |
| Filer Alle Sæt 70'erne rediger ● Enkelt ✓ amager3gI-09-AIH amager3gI-09-FWA amager3gI-09-HCM | It is also possible to search in a single file. Choose the option "Enkelt" ('single'), and select the desired file. |

It is even possible to create a new subset of files if the existing sets are not sufficient. Click "Rediger" ('edit') just below the "Sæt" ('set') field. This will send you to the page "Rediger interviewsæt" ('edit interview set').

| | | | e://daceee.com | h ku dk/soo | rch/interviewco | at html?inter | viewcet- | 70 | | | ~ | lu\ | Ē | | -0 | = |
|--|---|-------|--|--------------|-----------------|---------------|--|--|--|---|---|-----|---|---|----|---|
| | w C | | os://ugcsssearc | n.ku.uk/sear | ch/interviewse | e.numinnuer | viewset= | /0 | | S | พ | | Ē | ۲ | | = |
| Ree | | r int | | et hedder: | rsæt | <u>t</u> | | | | | | | | | | |
| alle filer fra Næ alle filer fra Odc alle filer fra Vinc Alle prioriterede alt i AmDa_febr alt i GanDa_febr alt i CanDa_feb alt i CanDa_febr | stved jer jerup filler uar 2019 ruar 2019 ruar 2019 ebruar 2019 ebruar 2019 | | Tilgænge | lige inte | erviews | • | Va | algte inter | rviews valgt | | | | | | | |
| amda_iClavEr amda_iClavEr amda_iLlavEr AmDa-kbi - 09- AmDa-kbi - 09- AmDa-kbi - 09- ArgDa - 09-11-2 ArgDa SLE 201 Argda_0_7 Argda_0_8 Argda_0_9 Argda_0_9 | ubkorpus 11-2017 - 09-11-2017 2017 - 7 | | amager3gl-09-AlH amager3gl-09-FW/ amager3gl-09-HCN amager3gl-09-ISH amager3gl-09-MMI amager3gl-09-MSN amager3gl-09-SSN amager3gl-09-SSN amager3gl-09-ZIN | | | | vin vin vin vin vin vin vin vin vin | Jerup2gi-78-AEL Jerup2gi-78-AEL+P(Jerup2gi-78-AFH Jerup2gi-78-AFH+B/ Jerup2gi-78-ALH Jerup2gi-78-ALH+IN Jerup2gi-78-BAB Jerup2gi-78-BAB Jerup2gi-78-FAN | gu+rdn+sse Ab+jbg+ldr Mn+lhs+njm | E | | | | | | |
| ArgDa-korpus_ ArgDa-korpus_ ArgDa/guni Argdakomplet Bysoc 0gl, 2gl, Bysoc1gl Bysoc1gl Bysoc1ny Bysoc2ny | 280516 3 gl rsoc2-ny | | amager3gl-10-AAS amager3gl-10-ABC amager3gl-10-FTM amager3gl-10-JCB amager3gl-10-JIL amager3gl-10-KHL amager3gl-10-KHL | | | | vin | lerup2gl-78-GPD lerup2gl-78-GPD+K lerup2gl-78-IGU lerup2gl-78-IGU+JN lerup2gl-78-IMN lerup2gl-78-IMO lerup2gl-78-IMO+JF lerup2gl-78-IMO+JF | (NY+KRS+LIK- MO+LET+MJJ FA+KES+SCH | 1 | | | | | | |
| Bysoc2ny Bysoc3/Clarin is Bysoc3/Clarin is Bysoc3gl CanDa - 09-11- Clarin Clarin3 gruppes | ruppe nterview 2017 samtaler | | amager3gi-10-MCF amager3gi-10-MTE amager3gi-10-MW, amager3gi-10-NAI amager3gi-10-NAS amager3gi-10-NAS | 5 5 5 | | | vin vin vin vin vin vin | derup2gl-78-JBG derup2gl-78-JBG derup2gl-78-JFA derup2gl-78-JMO derup2gl-78-KES derup2gl-78-KNY | an an iter view | | | | | | | |

Figure 2. The page "Rediger interviewsæt" ('edit interview set'). On this page one can select a subset of transcriptions to search in.



Figure 3. The bottom of the page "Rediger interviewsæt", where new interview sets are created by clicking "Opret nyt sæt" ('create new set').

On the page "Rediger interviewsæt" ('edit interview set'), scroll to the bottom, and click the button "Opret nyt sæt" ('create new set'). Choose a (new, nonexistent) name for the interview set, and enter it in the field "Interviewsættet hedder" ('name of the interview set'). The desired files can then be selected in the left column and moved to the right column using the right arrow button between the columns. Save the changes to create the new interview set.

3.2. Informants

Under "Informanter" ('informants') on the front page, you can specify which informants in the corpus you are interested in. Only results from the selected informants will then appear in the results.

| Informanter Alle Sæt Alle prioriterede inkl. Odder 2 ekstra rediger Enkelt | If you want to search in data from all informants, select the option "Alle" ('all'). This will include everyone, even incidental informants like people only appearing with a single 'hi' in passing, babies babbeling, and the like. |
|---|---|
| Alle Sæt Alle prioriterede talere ArgDa-de fem nyeste informanter_010616 BySos 1 og 3 Danske Stemmer-interviewere Emil22112019 amda | It is also possible to select named subsets of informants. For instance, all interviewers, all informants in a certain age group, or all informants of a certain gender. The option "Sæt" ('set') is selected, and the desired subset is selected in the drop-down menu. "Prioriterede" ('prioritized') informants are informants systematically singled out according to various sociolinguistic variables - see https://dgcss.hum.ku.dk/forskning/undersoegelsesomra ader. |
| Informanter Alle Sæt Alle rediger Enkelt AlH | It is also possible to choose single informants. Select the option "Enkelt" ('single'), and enter the informant code for the relevant informant. |

It is even possible to create a custom subset of informants. The procedure is very similar to creating a custom interview set - see above.

3.3. Search criteria

Under "Søgekriterier" ('search criteria'), the query can be specified. The corpus is based on a collection of files (Praat TextGrids), each associated with several annotation tiers (with phonetic annotations, grammatical annotations, etc.). All tiers are available for search, although not all files contain annotations in all tiers. Each file contains at least the principal annotation tier, i.e. an orthographic transcription of the speech of the informant in question.

A basic search is performed as follows.

| Søgekriterie fjern vorografi Admarkering Admarkering Alvinetaring | Start by specifying which tier the search should query. (The drop-down menu shows a complete list of all tiers, whether the individual tiers contain annotations for the current subcorpus or not). Next, enter the search term in the orange search field. The field is orange as long as it is empty; empty search criteria are not allowed. |
|---|--|
| Søgekriterie ijen ortografi Søg Der må ikke være tomme søgekriterier Søg efter gentagelser i interview Gå til statistiksiden. | When the search term has been entered, pick the appropriate category of the search term: Hele ordet ('the whole word'): Matches the whole word. Del af ordet ('part of the word'): Matches that contain the search term anywhere. Ikke en del af ordet ('not a part of the word'): The inverse of the above. Starten/Slutningen af ordet ('the beginning/end of the word'): Matches that start/end with the search term. Regulært udtryk: Regular expression match. |

The tiers contain one time interval per content element. For instance, each orthographic tier contains a single word for each interval. It is thus not possible to search for strings of several elements (several words, for instance) from one and the same search field. In order to do this, an additional search field has to be added for each additional search term. This context search is accomplished as follows.



Figure 4. Context search. Click "tilføj søgekriterie" ('add search criterion') to display an additional search field. Don't forget to check the correct position relative to the topmost search term. Otherwise you might end up trying to search for two different words in the same position, which is a logical impossibility.

Click "tilføj søgekriterie" ('add search criterion'). This displays an additional search field. The criterion line is filled in as described above. In addition, the position relative to the primary search term must be specified. Select 0 to search in the same position as the primary search term (logically, this has to be in a different tier). Select -1 to search in the position to the left of the primary search term, +2 to search in the position two words to the right, etc. As long as the position is not 0 (or more generally, not equal to the position of any other search field), the search can be performed in the same tier as the primary search term.

By using regular expressions, it is possible to perform searches for empty intervals or a set of different strings using a single search term. Don't forget to declare the search term a regular expression using the drop-down menu. Examples of regular expressions are listed in Table 1.

| Symbol | Search term | Explanation |
|--------|-------------|---|
| .* | .* | Zero or more characters. Finds all intervals, including empty intervals. |
| .+ | .+ | One or more characters. Finds all non-empty intervals. |
| | man du | Logical OR. Finds intervals containing <i>man</i> or <i>du</i> . |
| () | l(i æ)gge | Logical OR within string. Finds intervals containing <i>ligge</i> or <i>lægge</i> . |
| ^ | ^G | Beginning of string. Finds intervals beginning with <i>G</i> . |
| \$ | ik\$ | End of string. Finds intervals ending on <i>ik</i> . |

| Symbol | Search term | Explanation |
|---------|-------------|--|
| token.* | sur.* | Zero or more optional characters within string. Finds intervals containing <i>sur</i> followed by zero or more characters, e.g. <i>sure, glasuren, armbåndsuret</i> etc. |

Table 1. Regular expressions. Don't forget to choose "regulært udtryk" in the menu to the right of the search field in order to search using regular expressions.

The various regular expression symbols can be used together in a regular expression. Note that the symbols ^ and \$ must be used to specify matches at the beginning and/or end or words. For instance, if the exact words *ligge* and *lægge* are intended, and not *indlægge*, *lægger* etc., the following regular expression can be used: ^l(i|æ)gge\$. The symbols ^ and \$ can even be used multiple times in the same regular expression. To match the exact forms *ligget* and *lagt*, for instance, the following regular expression can be used: ^ligget\$|^lagt\$.

4. Results

When the search is completed, a result overview is shown.



Figure 5. Result overview presented when a search has been performed.

From here, several views of the search results are available: "Eksporter til CSV" ('Export to CSV'), "vis som slices" ('view as slices'), "vis som tabel" ('view as table' (also supports CSV export)), "kollokationer" ('collocation').

4.1. Export to CSV

The result can be exported directly to Excel. Click "Eksportér til CSV-fil" ('export to CSV file'). This will open a prompt to open or save the file.



Figure 6. The prompt that appears when "Eksporter til CSV-fil" ('export to CSV file') is clicked.

The CSV file shows each matching interval in the corpus. Each match is shown in a separate row along with file and speaker information as well as the content of the other tiers in the given time interval.

This format is well suited for further quantitative processing and statistic analysis.

4.2. View as slices (concordance view)

The link "Vis som slices" ('view as slices') reveals a concordance view of the result.

| ••• | 💐 Søgeresultat | | × | 減 Søgeres | sultater | × | (+ | | | | | | | | | | | |
|----------------------|-----------------------|-------------------|------------|-----------|---------------------|-------------|-------------------|--------------|--------------|---------|--------|-------|----|-------|--------|-------------|---------|-------|
| (←) → C ⁴ | ۵ (| 🛈 🔒 ht | tps://dgcs | ssearch.k | u.dk/searc | h/showslice | s.html | | | | •••• (| פ לב | | 111/ | • | ٢ | • | ≡ |
| Søgeresultate | erne 1 - 15 ud | af 472 foi | "altså" i | ortograf | i i <i>Clarin</i> 3 | gruppesa | <i>mtaler,</i> fo | or alle info | ormanter | | | | | redig | jér si | <u>øgek</u> | riterie | r |
| 1 2 3 32 | 4 5 6 | 7 | 8 9 | 10 11 | 12 13 | 14 15 | 16 | 17 18 | 19 20 | 21 22 | 23 | 24 25 | 26 | 27 2 | 8 ; | 29 | 30 | 31 |
| clarin3 | gl-10-A | EB+ | AOE | +BF | G+L | .ML | | 070 47 | 2 Afault had | | | | | | | | 07 | 4 466 |
| ortografi | -12) | 272.4644 ei | | | iea | tror | altså | 2/3.4/ | til | gengæld | at | | | | | | 21 | 4.400 |
| - tilføj tier - | • | C) | | | 109 | | unsu | | | gengæid | u | | | | | | | |
| | | ← forrige | | | | | | | | | | | | | | fø | lgende | ⇒ |
| clarin3 | gl-10-A | EB+ | AOE | E+BF | G+L | .ML | | | | | | | | | | | | |
| AEB (338. | 342) | 337.694 | | | | | | 338.34 | 2 Afspil lyd | | | | | | | | 3 | 39.32 |
| ortografi | | været | så | meget | i | medierne | altså | han | har | været | så | sådan | | | | | | |
| - tittøj tier - | V | | | | | | | | | | | | | | | | | |
| | | ← forrige | | | | | | | | | | | | | | fø | lgende | ⇒ ⇒ |
| clarin3 | gl-10-A 505) | EB+ | AOE | E+BF | G+L | .ML | | 342.505 | Afspil lyd | | | | | | | | 343 | 3825 |
| AED (342. | | fulat | det | så | tæt | | altså | | alle | har | jo | haft | | | | | | |
| ortografi | | luigi | uoi | | | | | | | | | | | | | | | |
| ortografi | | luigt | uor | | | | | | | | | | | | | | | |

Figure 7. Concordance view. Shown when "Vis som slices" ('view as slices') was chosen.

The view contains the following:

- At the top of the view, an overview in the form of the number of search results is still shown.
- On the left of each concordance line it is possible to choose which tiers are shown along with the transcription.
- Above each concordance line, time information for the match is shown (in seconds).
- The match is marked in blue. The context is white.
- Below each concordance line, a click on "forrige" ('previous') will expand the context to the left, and "følgende" ('next') to the right.
- The recording can be played at the current position by clicking "Afspil lys ..." above each concordance line. For information security reasons, this is however only supported for a limited amount of the data currently the interview set "Clarin3 gruppesamtaler" ('Clarin3 group conversations') and the files from the project "Danish Voices in the Americas", i.e. the files beginning with "amda" or "argda".

4.3. View (and export) table

An overview table of the matches in context is available from the link "Vis som tabel" ('view as table'). The page contains an interface to choose how much context (how many words) to the left and right of the match are shown, and which tiers are shown.

| $\bullet \bullet \bullet$ | 💐 Søgeres | sultat | | × | 📈 s | Skemavis | isning af si | øgeresult | tatet × | × 🕺 | Søgeresu | ltater | × | + | | | | | |
|---------------------------|--------------------------------|--------------|----------|---------|--------------|----------|-----------------|-----------|---------|-----|----------|--------|---|---|-----|---|------|---|-------|
| \leftrightarrow | C' 🕜 | Ø | 🔒 htt | tps://d | gcsssea | arch.ku | dk /sear | rch/sho | wgrid.h | tml | | | | | ••• | ☆ | 111\ | ۲ | ≡ |
| Tal | belegensk | abei | r | | | | | | | | | | | | | | | | |
| Slic | es før: | 5 | | 8 | | | | | | | | | | | | | | | |
| SIIC | es eπer: « resultator pr | 5 | | -1 | hetvde | ar. | | | | | | | | | | | | | |
| side | e: | Vis alle | | | betyue | | | | | | | | | | | | | | |
| Visr | ningstier 1: | ortografi | | | 0 | 3 | | | | | | | | | | | | | |
| Visr | ningstier 2: | IPA | | | • | | | | | | | | | | | | | | |
| Visr | ningstier 3: | Ikke valg | ıt . | | 0 | | | | | | | | | | | | | | |
| Visr | ningstier 4: | Ikke valg | it. | | | 3 | | | | | | | | | | | | | |
| Visi mar | fulde infor- ntoplysninger: | | | | | | | | | | | | | | | | | | |
| Vi | s Download | | | | | | | | | | | | | | | | | | |
| Tilba | age til resultatet. | <u>. Til</u> | lbage ti | l mas | <u>kinen</u> | | | | | | | | | | | | | | |
| Ingen tab | oel endnu. | | | | | | J | | | | | | | | | | | | |
| | | | | | | | | | | | | | | | | | | | |

Figure 8. Page shown after clicking "Vis som tabel" ('view as table'). Here, you can specify the parameters of the result table.

When these parameters have been set, the table can either be viewed on the page by clicking "Vis" ('view') or be downloaded as a CSV file by clicking "Download".

In Figure 9, the results are viewed on the page.

| A second se | C C C C C C C C C C C C C C C C C C C | -1 betyde | | | http: | s://dgo | essse | arc | h. ku.dk / | search/sho | wgrid.htm | l | | | | 50% | 🛛 | | IIIV | | 8 | Ξ |
|---|---|---|--|---|--|---|---------------------------------|--|--|---------------------------------------|--------------------|---|--------------------------|---------------------------|----------------------------------|-----------------------------|--|----------------------------|---------------------------------|----------------------|----|---|
| Silves fer: > Silves fer: > Silves fer: > Silves fer: > Varingstef: > Varingstef: </th <th>skaber</th> <th>-1 betyde</th> <th></th> | skaber | -1 betyde | | | | | | | | | | | | | | | | | | | | |
| Sloves feer: F Sloves feer: Very Very Very discription: | S | -1 betyde | | | | | | | | | | | | | | | | | | | | |
| Silves there 1 bolk: Yes bolk: Yes bolk: Yes device Yes Veringstaff 2: Ne device Yes Veringstaff 2: Ne devicingstaff 2: Ne <th>5 vis alle. origent lities valgt lities valgt lities</th> <th>-1 betyde</th> <th></th> | 5 vis alle. origent lities valgt lities | -1 betyde | | | | | | | | | | | | | | | | | | | | |
| Max resultative pri = 1 Verangster 2 Verangster 1 Verangster 2 Verangster 3 Verangster 4 deringster 5-10 | -1 Vis alle, entryms Nee waigt Nee Nee Nee Nee Nee Nee Nee Nee Nee Nee | -1 betyde | | | | | | | | | | | | | | | | | | | | |
| color: Vis visrangster 1: | Vis alle. orignal Rev valgt Rev valgt Rev valgt Rev valgt Rev valgt orignal Rev valgt Rev | maskinen | | | | | | | | | | | | | | | | | | | | |
| Veningster 3 Veningster 6 Veningster 6 Venin | His vargt His vargt His vargt His vargt His vargt Jis vargt Lis vargt Lis vargt Lis vargt | maskinen | | | | | | | | | | | | | | | | | | | | |
| Vieningstef 4 Neiningstef 4 Vieningstef 4 Neiningstef 4 Vieningstef 4 Neiningstef 5 Vieningstef 5 Neiningstef 5 Vieningstef 5 Neiningstef 5 Vieningstef 5 Neiningstef 5 dending 5 Neiningstef 5 dending 5 Neiningstef 5 dending 5 Neiningstef 5 dending 5 Neining 5 dending 5 Neinin | Hite valgt Hite valgt Hite valgt Hite valgt | maskinen | | | | | | | | | | | | | | | | | | | | |
| delanguida 6 | iske vargt iske vargt and let. Tilbage til m | maskinen | | | | | | | | | | | | | | | | | | | | |
| Vis Lube and physicing." | and Tibage til m | maskinen | | | | | | | | | | | | | | | | | | | | |
| Description 0 Description Rescription Control Rescription Control dendrogi - 10 AEE dendrogi - 10 AEE | et. Tilbage til m | Imaskinen | | | | | | | | | | | | | | | | | | | | |
| Boopt II resoluted. d Finite desing 1-10 AEE desing 1-10 AEE desing 1-10 AEE de | tet. Tilbage til m | maskinen | | | | | | | | | | | | | | | | | | | | |
| d First demolg-10-ABL demolg-10-ABL demolg-10-ABL | | | | | | | | | | | | | | | | | | | | | | |
| chemicagi 1-0. AEE denicagi 1-0. AEE | Filnavn | Opt | Opt.år. | Tier | Inf | XMin H | øn S. | kl. F | år. Projekt | -5 | -4 | -3 | -2 | -1 | 0 | 1 | 2 | 3 | 4 | 5 | | |
| 14. Acti Marchael Control (19) Action (| AEB+AOE+BFG+I | +LML gl | 2010 | ortografi | AEB | 273,47 | K N | A 1 | 391 CLARIN 991 CLARIN | ej | sà | meget | jeg | tror | altså | øh | til | gengæld | at så | såden | - | |
| densigi 1-0 AEI densigi 1-0 AEI | AEB+AOE+BFG+I | +LML gi | 2010 | ortografi | AEB | 342,5 | K N | A 1 | 991 CLARIN | fulgt | det | sà | tæt | -199491110 | altsá | | alle | har | jo | haft | -1 | |
| denicajo 1-0. AEE denicajo 1-0. AEE | AEB+AOE+BFG+I | +LML gl | 2010 | ortografi | AEB | 353,62 | K N | A 1 | 991 CLARIN | man | lige | pludselig | ikke | | altsá | øh | | at | at | han | | |
| Book (B) A (B) | AEB+AOE+BFG+I | +LML gl | 2010 | ortografi | AEB | 422,82 | K N | A 1 | 391 CLARIN | jeg | var | øh | jeg | synes | altsà | | det | var | ret | øh | _ | |
| denright 10-AEE (denright 10-AEE) (denright 10-A | AEB+AOE+BFG+I | +LML gi | 2010 | ortografi | AEB | 625,68 | K N | A 1 | 991 CLARIN 991 CLARIN | - | det | er | jo | øn klart | altsà | | | men | at | men | - | |
| elenicity 1-0. AEE elenicity 1-0. AEE elenicity 1-0. AEE elenicy 1-0. AEE | AEB+AOE+BFG+I | +LML gl | 2010 | ortografi | AEB | 637,94 | K N | A 1 | 991 CLARIN | giver | jo | ingen | mening | | altsá | | | | men | | | |
| Linking House | AEB+AOE+BFG+I | +LML gl | 2010 | ortografi | AEB | 656,48 | K N | A 1 | 391 CLARIN | men | det | er | jo | | altså | det | er | jo | klart | fordi | _ | |
| clandig-10 AEE clange-10 AE | AEB+AOE+BFG+I | +LML gl | 2010 | ortografi | AEB | 732.47 | K N K N | A 1 A 1 | 991 CLARIN 991 CLARIN | er | io sa | fra | Jehovas | Vidne | altsà | der | af er | io | der | noget | - | |
| dennisty 1-0 AEE dennis | AEB+AOE+BFG+I | i+LML gl | 2010 | ortografi | AEB | 787,96 | K N | A 1 | 991 CLARIN | sà | en | dokumentar | det | var | altsá | for | sà | noget | | True | _ | |
| clam3g-10.AEB cl | AEB+AOE+BFG+I | +LML gl | 2010 | ortografi | AEB | 811,78 | K N | A 1 | 991 CLARIN | var | | hel | hel | lille | altsá | | | | elve | | | |
| clam3g110AEE | AEB+AOE+BFG+I | i+LML gl | 2010 | ortografi | AEB | 900,2 | K N K N | A 1 | 391 CLARIN 991 CLARIN | til | al fuldstandio | godt | nei | ikke | altsà | | | det | nej | nej | _ | |
| clarin3g-10-AEE | AEB+AOE+BFG+I | i+LML gl | 2010 | ortografi | AEB | 934,81 | K N | A 1 | 991 CLARIN | stue | raraananang | f- | løft- | d- | altsá | flyttet | alle | V- | øh | møbler | - | |
| claining-10-AEE | AEB+AOE+BFG+I | +LML gl | 2010 | ortografi | AEB | 975,56 | K N | A 1 | 991 CLARIN | | de | | | d- | altsá | | de | er | | | | |
| caining110AEE claning110AEE claning110AEE claning110AEE claning110AEE claning110AEE claning110AEE claning110AEE claning110AEE claning110AEE claning110AEE claning110AEE claning110AEE claning110AEE claning110AEE claning110AEE claning110AEE claning110AEE claning110AEE claning110AEE | AEB+AOE+BFG+I | +LML gl | 2010 | ortografi | AEB | 1029,14 | K N | A 1 | 391 CLARIN | | ja | | han | er | altsá | også | bare | grineren | ikke | med | _ | |
| clarin3gi-10-AEB clarin3gi-10-AEB clarin3gi-10-AEB clarin3gi-10-AEB clarin3gi-10-AEB clarin3gi-10-AEB clarin3gi-10-AEB clarin3gi-10-AEB clarin3gi-10-AEB clarin3gi-10-AEB clarin3gi-10-AEB clarin3gi-10-AEB clarin3gi-10-AEB clarin3gi-10-AEB clarin3gi-10-AEB clarin3gi-10-AEB clarin3gi-10-AEB clarin3gi-10-AEB clarin3gi-10-AEB | AEB+AOE+BFG+I | +LML gi | 2010 | ortografi | AEB | 1044,88 | K N | A 1 | 991 CLARIN 991 CLARIN | er | det | grineren | OPT | men | altsá | pointen | med | sa det | er | io | - | |
| clarin3g-10-AEE clarin3g-10-AEE clarin3g-10-AEE clarin3g-10-AEE clarin3g-10-AEE clarin3g-10-AEE clarin3g-10-AEE clarin3g-10-AEE clarin3g-10-AEE clarin3g-10-AEE clarin3g-10-AEE clarin3g-10-AEE clarin3g-10-AEE clarin3g-10-AEE clarin3g-10-AEE clarin3g-10-AEE clarin3g-10-AEE clarin3g-10-AEE | AEB+AOE+BFG+I | +LML gl | 2010 | ortografi | AEB | 1281,42 | K N | A 1 | 991 CLARIN | er | bare | vild | med | fodbold | altsá | lad | nu | | barnet | ikke | | |
| claming-10-AEE claring1-10-AEE claring1-10-AEE claring1-10-AEE claring1-10-AEE claring1-10-AEE claring1-10-AEE claring1-10-AEE claring1-10-AEE claring1-10-AEE claring1-10-AEE claring1-10-AEE claring1-10-AEE claring1-10-AEE claring1-10-AEE | AEB+AOE+BFG+I | +LML gl | 2010 | ortografi | AEB | 1291,84 | K N | A 1 | 991 CLARIN | nu | bare | køre | lidt | ræs | altsá | det | skal | jeg | da | ikke | _ | |
| clarin3gi-10-AEE clarin3gi-10-AEE clarin3gi-10-AEE clarin3gi-10-AEE clarin3gi-10-AEE clarin3gi-10-AEE clarin3gi-10-AEE clarin3gi-10-AEE clarin3gi-10-AEE clarin3gi-10-AEE clarin3gi-10-AEE | AED+AOE+BFG+I | +LML gl | 2010 | ortografi | AFB | 1312,66 | r. N. K. N. | A 1 A 1 | 991 CLARIN 991 CLARIN | skulle | nave | været | sadan | iea | altsà | agesom | sidst | men vildt | men | der | - | |
| clarin3gl-10-AEB clarin3gl-10-AEB clarin3gl-10-AEB clarin3gl-10-AEB clarin3gl-10-AEB clarin3gl-10-AEB clarin3gl-10-AEB clarin3gl-10-AEB clarin3gl-10-AEB clarin3gl-10-AEB | AEB+AOE+BFG+I | i+LML gl | 2010 | ortografi | AEB | 1382,9 | K N | A 1 | 991 CLARIN | det | er | slet | ikke | 1.48 | altsà | • | på | den | ene | side | - | |
| clamsgi-10-AEE clarin3gi-10-AEE clarin3gi-10-AEE clarin3gi-10-AEE clarin3gi-10-AEE clarin3gi-10-AEE clarin3gi-10-AEE clarin3gi-10-AEE clarin3gi-10-AEE | AEB+AOE+BFG+I | | 2010 | ortografi | AEB | 1387,77 | K N | A 1 | 991 CLARIN | også | paent | frygteligt | ikke | | altsá | | at | øh | at | de | _ | |
| clarin3gl-10-AEB clarin3gl-10-AEB clarin3gl-10-AEB clarin3gl-10-AEB clarin3gl-10-AEB clarin3gl-10-AEB clarin3gl-10-AEB clarin3gl-10-AEB | AEB+AOE+BFG+I AEB+AOE+BFG+I AEB+AOE+BFG+I | +LML gl | 2010 | ortografi | AEB | 1391,99 | K N | A 1 A 1 | 991 CLARIN 991 CLARIN | den fred- | øh | fredelig | aiort | | altsá | | det | k- | det | kunne | - | |
| clarin3gl-10-AEE clarin3gl-10-AEE clarin3gl-10-AEE clarin3gl-10-AEE clarin3gl-10-AEE clarin3gl-10-AEE clarin3gl-10-AEE | AEB+AOE+BFG+I AEB+AOE+BFG+I AEB+AOE+BFG+I AEB+AOE+BFG+I AEB+AOE+BFG+I | i+LML gl | 1 2010 | ortografi | AEB | 1496,98 | K N | A 1 | 991 CLARIN | | jeg | har | 380- | | altsà | jeg | kommer | bare | hjem | med | - | |
| clarin3gl-10-AEE clarin3gl-10-AEE clarin3gl-10-AEE clarin3gl-10-AEE clarin3gl-10-AEE | AEB+AOE+BFG+I AEB+AOE+BFG+I AEB+AOE+BFG+I AEB+AOE+BFG+I AEB+AOE+BFG+I AEB+AOE+BFG+I | +LML gl +LML gl +LML gl +LML gl | 2010 | | AEB | 1638,9 | K N | A 1 | 991 CLARIN | | øh | vi | lever | | altsà | | | vi | vi | lever | - | |
| clarin3gl-10-AEE clarin3gl-10-AEE clarin3gl-10-AEE | AEB+AOE+BFG+I AEB+AOE+BFG+I AEB+AOE+BFG+I AEB+AOE+BFG+I AEB+AOE+BFG+I AEB+AOE+BFG+I AEB+AOE+BFG+I | +LML gl +LML gl +LML gl +LML gl +LML gl | 2010 2010 2010 | ortografi | | | κ Ν. | A 1 | 391 CLARIN 991 CLARIN | vi | lever | jo | vi | ŀ | altsá | vi | lever | jo | rigtig | rigtig | - | |
| clarin3gl-10-AEE clarin3gl-10-AEE | AEB+AOE+BFG+I AEB+AOE+BFG+I AEB+AOE+BFG+I AEB+AOE+BFG+I AEB+AOE+BFG+I AEB+AOE+BFG+I AEB+AOE+BFG+I AEB+AOE+BFG+I AEB+AOE+BFG+I | +LML gl +LML gl +LML gl +LML gl +LML gl +LML gl +LML gl +LML gl | 2010 2010 2010 2010 2010 | ortografi ortografi | AEB | 1640,87 | K N | | | | ikke | komme | nogen | ind | altsá | pá | lufthavnen | som | ikke | skulle | - | |
| clarin3gl-10-AEE | AEB+AOE+BFG+I AEB+AOE+BFG+I AEB+AOE+BFG+I AEB+AOE+BFG+I AEB+AOE+BFG+I AEB+AOE+BFG+I AEB+AOE+BFG+I AEB+AOE+BFG+I AEB+AOE+BFG+I | ++LML gl +LML gl +LML gl +LML gl +LML gl +LML gl +LML gl +LML gl +LML gl | 2010 2010 2010 2010 2010 2010 2010 | ortografi ortografi ortografi ortografi | AEB AEB | 1640,87 1649,89 1731,92 | K N K N | A 1 | 991 CLARIN | slet | | | | | altab | | | | | - | | |
| claria3aL10 AFF | AEB+AOE+BFG+I AEB+AOE+BFG+I AEB+AOE+BFG+I AEB+AOE+BFG+I AEB+AOE+BFG+I AEB+AOE+BFG+I AEB+AOE+BFG+I AEB+AOE+BFG+I AEB+AOE+BFG+I AEB+AOE+BFG+I | ++LML gl ++LML gl ++LML gl ++LML gl ++LML gl ++LML gl ++LML gl ++LML gl ++LML gl ++LML gl | 2010 2010 2010 2010 2010 2010 2010 2010 | ortografi ortografi ortografi ortografi ortografi | AEB AEB AEB AEB | 1640,87 1649,89 1731,92 1736,98 | K N K N | A 1 A 1 | 991 CLARIN 991 CLARIN | slet derind | alene | niende | kl- | | anad | | niendeklasses | | ung | | _ | |
| clarin3gl-10-AEE | AEB+AOE+BFG+I AEB+AOE+BFG+I AEB+AOE+BFG+I AEB+AOE+BFG+I AEB+AOE+BFG+I AEB+AOE+BFG+I AEB+AOE+BFG+I AEB+AOE+BFG+I AEB+AOE+BFG+I AEB+AOE+BFG+I AEB+AOE+BFG+I | ++LML gl ++LML gl | 2010 2010 2010 2010 2010 2010 2010 2010 | ortografi ortografi ortografi ortografi ortografi ortografi | AEB AEB AEB AEB AEB | 1640,87 1649,89 1731,92 1736,98 1749,9 | | A 1 A 1 A 1 A 1 | 991 CLARIN 991 CLARIN 991 CLARIN | slet derind pas | alene | niende og | kl- eventuelle | i sådan | altsà | medicin | niendeklasses eller | sá | noget | fuet | _ | |
| clarin3gl-10-AEE | AEB+AOE+BFG+ AEB+AOE+BFG+t AEB+AOE+BFG+t AEB+AOE+BFG+t AEB+AOE+BFG+t AEB+AOE+BFG+t AEB+AOE+BFG+t AEB+AOE+BFG+t AEB+AOE+BFG+t AEB+AOE+BFG+t AEB+AOE+BFG+t AEB+AOE+BFG+t | I+LML gl I+LML gl | 2010 2010 2010 2010 2010 2010 2010 2010 | ortografi ortografi ortografi ortografi ortografi ortografi ortografi ortografi | AEB AEB AEB AEB AEB AEB | 1640,87 1649,89 1731,92 1736,98 1749,9 1755,94 1775,43 | | A 1 A 1 A 1 A 1 A 1 A 1 A 1 A 1 A 1 | 991 CLARIN 991 CLARIN 991 CLARIN 991 CLARIN 991 CLARIN | slet derind pas sådan | alene | niende og overhovedet gemme | kl- eventuelle det | i sådan øh | altså altså | medicin med at | niendeklasses eller øh dysse | så op det | i noget i | flyet | | |
| | AEB+AOE+BFG+ AEB+AOE+BFG+ AEB+AOE+BFG+ AEB+AOE+BFG+ AEB+AOE+BFG+ AEB+AOE+BFG+ AEB+AOE+BFG+ AEB+AOE+BFG+ AEB+AOE+BFG+ AEB+AOE+BFG+ AEB+AOE+BFG+ AEB+AOE+BFG+ AEB+AOE+BFG+ | ++LML gl i+LML gl | 2010 2010 2010 2010 2010 2010 2010 2010 | ortografi ortografi ortografi ortografi ortografi ortografi ortografi ortografi ortografi | AEB AEB AEB AEB AEB AEB AEB AEB | 1640,87 1649,89 1731,92 1736,98 1749,9 1755,94 1775,43 1835,99 | | A 1 A 1 A 1 A 1 A 1 A 1 A 1 A 1 A 1 A 1 | 991 CLARIN 991 CLARIN | slet derind pas sådan det | alene at var | niende og overhovedet gemme for | kl- eventuelle det | i sådan øh vildt | altså altså altså altså | medicin med at jeg | niendeklasses eller øh dysse synes | så op det fandeme | i noget i ned det | flyet ikke var | | |
| | AEB+AOE+BFG+I AEE+AOE+BFG+I AEB+AOE+BFG+I AEB+AOE+BFG+I AEB+AOE+BFG+I AEB+AOE+BFG+I AEB+AOE+BFG+I AEB+AOE+BFG+I AEB+AOE+BFG+I AEB+AOE+BFG+I AEB+AOE+BFG+I AEB+AOE+BFG+I | i+LML gl | 2010 2010 2010 2010 2010 2010 2010 2010 | ortografi ortografi ortografi ortografi ortografi ortografi ortografi ortografi | AEB AEB AEB AEB AEB AEB AEB AEB | 1640,87 1649,89 1731,92 1736,98 1749,9 1755,94 1775,43 1835,99 | K N K N K N K N K N | A 1 A 1 A 1 A 1 A 1 A 1 A 1 A 1 A 1 A 1 | 991 CLARIN 991 CLARIN 991 CLARIN 991 CLARIN 991 CLARIN 991 CLARIN | slet derind pas sådan det | alene at var | niende og overhovedet gemme for | kl- eventuelle det | i sådan øh vildt | altså altså altså altså | medicin med at jeg | niendeklasses eller øh dysse synes | så op det fandeme | i noget i ned det | flyet ikke var | | |
| | AEB+AOE+BFG+I AEB+AOE+BFG+I AEB+AOE+BFG+I AEB+AOE+BFG+I AEB+AOE+BFG+I AEB+AOE+BFG+I AEB+AOE+BFG+I AEB+AOE+BFG+I AEB+AOE+BFG+I AEB+AOE+BFG+I AEB+AOE+BFG+I AEB+AOE+BFG+I | ++LML gl | 2010 2010 2010 2010 2010 2010 2010 2010 | ortografi ortografi ortografi ortografi ortografi ortografi ortografi | AEB AEB AEB AEB AEB AEB AEB AEB | 1640,87 1649,89 1731,92 1736,98 1749,9 1755,94 1775,43 1835,99 | | A 1 A 1 A 1 A 1 A 1 A 1 A 1 A 1 A 1 | 991 CLARIN 991 CLARIN 991 CLARIN 991 CLARIN 991 CLARIN 991 CLARIN 991 CLARIN | slet derind pas sådan det | alene at var | niende og overhovedet gemme for | kl- eventuelle det | i sådan øh vildt | altså altså altså altså | medicin med at jeg | eller øh dysse synes | så op det fandeme | ung noget i ned det | flyet ikke var | | |

Figure 9. Table view. The table appears on the page when "Vis" ('view') is clicked. The other option is "Download", which will prompt a CSV download of the table.

In each row, speaker information is available. Position 0 contains the query match. Note that the table can contain multiple rows per search result - one extra row per additional tier selected.

The table can be downloaded as a CSV file by clicking "Download".

The table shown on this page only contains the most important corpus metadata. To obtain the full set of metadata, use "Eksportér til CSV-fil" ('export to CSV file') instead.

4.4. Collocations

If the query only consists of a single string, a collocation analysis can be performed. Collocations are calculated statistically as word pairs cooccurring more frequently than would be expected from each word's individual frequency in the corpus.

To perform a collocation analysis, click "Find kollokationer for resultatet" ('find collocations for the result').

At the top of the results page, the number of tokens in the selected subcorpus is listed along with the number of unique tokens. Below that, the right and left context of the query string are shown.

The tables show the context word, the number of times the context word occurs with the query word, and the Mutual Information score - a measure of the unexpectedness of the context word and the query word occurring together. The collocations are ordered by the MI score.