

‘Sketch’ i INESS

Veiledning i bruk av søketemplater

Helge Dyvik, Paul Meurer

Søketemplater, som f.eks. **V-arg1(@V)**, er uttrykk som tillater en bruker å søke etter eksempler i trebanken NorGramBank uten å beskjeftige seg direkte med søkespråket. Det nevnte templatet finner alle handlere (argument 1) til et oppgitt verb; brukeren behøver bare å spesifisere verbet (@V).

Denne veiledningen beskriver mulighetene så langt som de hittil er implementert. Planen er å videreutvikle grensesnittet og å definere nye søketemplater basert på leksikografenes og andres behov.




Innhold

1. Pålogging	2
2. Valg av trebanker	3
3. Lisenser	7
4. Veien til ‘Sketch’-siden	7
Valg av deltrebanker å søke i	8
5. Søk på ‘Sketch’-siden	9
(a) Bruk templatene enkeltvis	12
Se beskrivelser av templatene	12
Gi parametrene verdier	14
Kjør søket: Run query	14
Kopier fra trefflisten	16
Opsjonen ‘Group by’	18
Snarvei via ‘Query history’	19
Forfatter og tittel: Sorter etter metadata	20
(b) Søk med flere templatene samtidig	22
6. Flere muligheter utenfor ‘Sketch’	28
Sidene ‘Sentence Overview’ og ‘Sentence’	28
7. Advarsel	32
8. Appendiks: Templatoversikt	33

1. Pålogging

Startsiden for INESS finnes her: <http://clarino.uib.no/iness/>

Den ser ut som i Figur 1:

	<p>Sign in: Local Clarin SPF eduGAIN OpenIdP</p> <h2>Welcome to INESS</h2>
<ul style="list-style-type: none"> Main Page Knowledge center The project Documentation FAQ Publications Links Resources 	<p>INESS is the Norwegian <i>Infrastructure for the Exploration of Syntax and Semantics</i>. This infrastructure provides access to <i>treebanks</i>, which are databases of syntactically and semantically annotated sentences.</p> <p>INESS offers an open, interactive, language independent platform for <i>building, accessing, searching and visualizing</i> treebanks. All its functionality can be used online through a web browser.</p> <p>INESS hosts treebanks for many languages, including treebanks which have been created by others. The INESS project is also building its own large treebank for Norwegian, NorGramBank, obtained by parsing automatically with an LFG grammar. Parts of NorGramBank have been efficiently manually disambiguated with the LFG Parsebanker, while the rest is stochastically disambiguated.</p>
<p>Treebanks</p> <ul style="list-style-type: none"> Treebank Selection Parallel Sentences 	<p>The XLE-Web interface can be used for interactive parsing and disambiguation with LFG grammars for various languages.</p>
<p>Tools</p> <ul style="list-style-type: none"> XLE-Web Grammars 	<p>INESS is part of the CLARINO Bergen Centre.</p> <h3>How to cite INESS</h3>
	<p>If you use any INESS services in your research, for instance, to search or annotate treebanks, we request an acknowledgement to INESS, mentioning the webpage http://clarino.uib.no/iness and the following reference in your publications:</p> <p>Victoria Rosén, Koenraad De Smedt, Paul Meurer, and Helge Dyvik. An open infrastructure for advanced treebanking. In Jan Hajič, Koenraad De Smedt, Marko Tadić, and António Branco (eds.) <i>META-RESEARCH Workshop on Advanced Treebanking at LREC2012</i>, pages 22–29, Istanbul, Turkey, May 2012.</p> <div style="text-align: center;">   </div>


Figur 1: INESS' startside

Søking forutsetter konto på clarino.uib.no/iness. Uten akademisk tilknytning har man kanskje ikke tilgang til Feide-innlogging. Med Feide-mulighet velger man 'Clarin SPF' øverst på siden. Uten Feide-mulighet bør man skaffe seg en Clarin-konto, som kan bestilles under <https://user.clarin.eu/user/register>.

Med Feide- innlogging har man straks tilgang til NorGram-trebankene. Det har også andre som er med i eduGain. Etter førstegangsinnlogging på annen måte kan Paul Meurer (paul.meurer@uib.no) gi tilgang til NorGram-trebankene. Man må da vente med å gå videre til Paul har gitt beskjed om at rettighetene er tildelt.

2. Valg av trebanker

Gå til 'Treebank selection' i menyen til venstre. Dette åpner denne siden (Figur 2):



Treebank Selection

[Sign in: Clarin SPF](#) | [eduGAIN](#) | [OpenIdP](#)

Main Page

Knowledge center

The project

Documentation

FAQ

Publications

Links

Resources

Treebanks

[Treebank Selection](#)

Parallel Sentences

Tools

XLE-Web

Grammars

Select a set of treebanks to work with. ?

Languages: [All](#) · [Afrikaans](#) (3) · [Akkadian](#) (2) · [Amharic](#) (2) · [Ancient Greek \(to 1453\)](#) (17) · [Arabic](#) (13) · [Armenian](#) (2) · [Assyrian Neo-Aramaic](#) (1) · [Bambara](#) (2) · [Basque](#) (8) · [Belarusian](#) (3) · [Bhojpuri](#) (1) · [Breton](#) (2) · [Bulgarian](#) (9) · [Buriat](#) (3) · [Catalan](#) (6) · [Chinese](#) (17) · [Church Slavic](#) (10) · [Classical Armenian](#) (1) · [Coptic](#) (4) · [Croatian](#) (9) · [Czech](#) (26) · [Danish](#) (10) · [Dutch](#) (14) · [English](#) (39) · [Erzya](#) (2) · [Estonian](#) (10) · [Faroese](#) (3) · [Finnish](#) (21) · [French](#) (26) · [Galician](#) (11) · [Georgian](#) (9) · [German](#) (26) · [Gothic](#) (8) · [Hebrew](#) (8) · [Hindi](#) (10) · [Hungarian](#) (12) · [Icelandic](#) (3) · [Indonesian](#) (12) · [Irish](#) (8) · [Italian](#) (21) · [Japanese](#) (12) · [Karelian](#) (1) · [Kazakh](#) (6) · [Komi](#) (4) · [Komi-Permyak](#) (1) · [Korean](#) (7) · [Latin](#) (25) · [Latvian](#) (7) · [Lithuanian](#) (4) · [Livvi](#) (1) · [Maltese](#) (2) · [Marathi](#) (3) · [Mbyá Guaraní](#) (2) · [Modern Greek \(1453-\)](#) (9) · [Moksha](#) (1) · [Nigerian Pidgin](#) (2) · [Northern Kurdish](#) (3) · [Northern Sami](#) (28) · [Norwegian](#) (5) · [Norwegian Bokmål](#) (55) · [Norwegian Nynorsk](#) (17) · [Old English \(ca. 450-1100\)](#) (5) · [Old French \(842-ca. 1400\)](#) (3) · [Old Norse](#) (4) · [Old Russian](#) (20) · [Persian](#) (8) · [Polish](#) (34) · [Portuguese](#) (22) · [Romanian](#) (11) · [Russian](#) (20) · [Sanskrit](#) (4) · [Scottish Gaelic](#) (1) · [Serbian](#) (3) · [Skolt Sami](#) (1) · [Slovak](#) (5) · [Slovenian](#) (14) · [Spanish](#) (17) · [Swedish](#) (19) · [Swedish Sign Language](#) (4) · [Swiss German](#) (1) · [Tagalog](#) (2) · [Tamil](#) (8) · [Telugu](#) (3) · [Thai](#) (2) · [Turkish](#) (11) · [Uighur](#) (5) · [Ukrainian](#) (5) · [Upper Sorbian](#) (3) · [Urdu](#) (6) · [Vietnamese](#) (5) · [Warlpiri](#) (2) · [Welsh](#) (1) · [Wolof](#) (4) · [Yoruba](#) (2) · [Yue Chinese](#) (3)

Treebank Collections: [All](#) · [Acquis](#) (7) · [Alpino](#) (1) · [BulTreeBank](#) (1) · [CLARIN-PL](#) (5) · [DELPH-IN](#) (2) · [GEGO](#) (4) · [GeoGram](#) (4) · [HunGram](#) (4) · [ISWOC](#) (9) · [JOS](#) (1) · [Menotec](#) (4) · [Mercurius](#) (1) · [NAOB](#) (14) · [NDT](#) (4) · [NorGram](#) (55) · [NorGramBank](#) (31) · [POLFIE](#) (23) · [PROIEL](#) (10) · [PaHC](#) (2) · [ParGram](#) (12) · [ParTMA](#) (14) · [Sami-open](#) (15) · [Sami-restricted](#) (7) · [Sofie](#) (9) · [TOROT](#) (22) · [TiGer](#) (4) · [Universal Dependencies 1.1](#) (19) · [Universal Dependencies 1.2](#) (36) · [Universal Dependencies 1.3](#) (53) · [Universal Dependencies 1.4](#) (63) · [Universal Dependencies 2.0](#) (63) · [Universal Dependencies 2.1](#) (103) · [Universal Dependencies 2.3](#) (130) · [Universal Dependencies 2.5](#) (157) · [WolGram](#) (3) · [XPar](#) (2)

Treebank Types: [All](#) · [Ifg](#) (117) · [constituency](#) (19) · [constituency-alpino](#) (1) · [dependency](#) (45) · [dependency-cg](#) (662) · [dependency-tuebadz](#) (1) · [hpsg](#) (2)

Figur 2: Siden 'Treebank Selection'

Trebanker (eller sett av trebanker) kan velges med utgangspunkt i språk ('Languages'), navngitte trebanker ('Treebank Collections'), eller trebanktype ('Treebank Types'). For vårt leksikografiske formål er det NorGramBank som er av interesse. Klikk da på 'NorGramBank' under 'Treebank Collections' (tallet 31 angir hvor mange deltrebanker NorGramBank består av). Siden får da dette utseendet (Figur 3):

iness

Treebank Selection

Sign in: [Clarín SPF](#) | [eduGAIN](#) | [OpenIdP](#)

Main Page

Knowledge center

The project

Documentation

FAQ

Publications

Links

Resources

Treebanks

Treebank Selection

Parallel Sentences

Tools

XLE-Web Grammars

Select a set of treebanks to work with. ?

Languages: [All](#) · [Afrikaans](#) (0/3) · [Akkadian](#) (0/2) · [Amharic](#) (0/2) · [Ancient Greek \(to 1453\)](#) (0/17) · [Arabic](#) (0/13) · [Armenian](#) (0/2) · [Assyrian Neo-Aramaic](#) (0/1) · [Bambara](#) (0/2) · [Basque](#) (0/8) · [Belarusian](#) (0/3) · [Bhojpuri](#) (0/1) · [Breton](#) (0/2) · [Bulgarian](#) (0/9) · [Buriat](#) (0/3) · [Catalan](#) (0/6) · [Chinese](#) (0/17) · [Church Slavonic](#) (0/10) · [Classical Armenian](#) (0/1) · [Coptic](#) (0/4) · [Croatian](#) (0/9) · [Czech](#) (0/26) · [Danish](#) (0/10) · [Dutch](#) (0/14) · [English](#) (0/39) · [Erzya](#) (0/2) · [Estonian](#) (0/10) · [Faroese](#) (0/3) · [Finnish](#) (0/21) · [French](#) (0/26) · [Galician](#) (0/11) · [Georgian](#) (0/9) · [German](#) (0/26) · [Gothic](#) (0/8) · [Hebrew](#) (0/8) · [Hindi](#) (0/10) · [Hungarian](#) (0/12) · [Icelandic](#) (0/3) · [Indonesian](#) (0/12) · [Irish](#) (0/8) · [Italian](#) (0/21) · [Japanese](#) (0/12) · [Karelian](#) (0/1) · [Kazakh](#) (0/6) · [Komi](#) (0/4) · [Komi-Permyak](#) (0/1) · [Korean](#) (0/7) · [Latin](#) (0/25) · [Latvian](#) (0/7) · [Lithuanian](#) (0/4) · [Livvi](#) (0/1) · [Maltese](#) (0/2) · [Marathi](#) (0/3) · [Mbyá Guaraní](#) (0/2) · [Modern Greek \(1453-\)](#) (0/9) · [Moksha](#) (0/1) · [Nigerian Pidgin](#) (0/2) · [Northern Kurdish](#) (0/3) · [Northern Sami](#) (0/28) · [Norwegian](#) (5) · [Norwegian Bokmål](#) (21/55) · [Norwegian Nynorsk](#) (5/17) · [Old English \(ca. 450-1100\)](#) (0/5) · [Old French \(842-ca. 1400\)](#) (0/3) · [Old Norse](#) (0/4) · [Old Russian](#) (0/20) · [Persian](#) (0/8) · [Polish](#) (0/34) · [Portuguese](#) (0/22) · [Romanian](#) (0/11) · [Russian](#) (0/20) · [Sanskrit](#) (0/4) · [Scottish Gaelic](#) (0/1) · [Serbian](#) (0/3) · [Skolt Sami](#) (0/1) · [Slovak](#) (0/5) · [Slovenian](#) (0/14) · [Spanish](#) (0/17) · [Swedish](#) (0/19) · [Swedish Sign Language](#) (0/4) · [Swiss German](#) (0/1) · [Tagalog](#) (0/2) · [Tamil](#) (0/8) · [Telugu](#) (0/3) · [Thai](#) (0/2) · [Turkish](#) (0/11) · [Uighur](#) (0/5) · [Ukrainian](#) (0/5) · [Upper Sorbian](#) (0/3) · [Urdu](#) (0/6) · [Vietnamese](#) (0/5) · [Warlpiri](#) (0/2) · [Welsh](#) (0/1) · [Wolof](#) (0/4) · [Yoruba](#) (0/2) · [Yue Chinese](#) (0/3)

Treebank Collections: [All](#) · [Acquis](#) (7) · [Alpino](#) (1) · [BulTreeBank](#) (1) · [CLARIN-PL](#) (5) · [DELPH-IN](#) (2) · [GEGO](#) (4) · [GeoGram](#) (4) · [HunGram](#) (4) · [ISWOC](#) (9) · [JOS](#) (1) · [Menotec](#) (4) · [Mercurius](#) (1) · [NAOB](#) (14) · [NDT](#) (4) · [NorGram](#) (55) · [NorGramBank](#) (31) · [POLFIE](#) (23) · [PROIEL](#) (10) · [PaHC](#) (2) · [ParGram](#) (12) · [ParTMA](#) (14) · [Sami-open](#) (15) · [Sami-restricted](#) (7) · [Sofie](#) (9) · [TOROT](#) (22) · [TiGer](#) (4) · [Universal Dependencies 1.1](#) (19) · [Universal Dependencies 1.2](#) (36) · [Universal Dependencies 1.3](#) (53) · [Universal Dependencies 1.4](#) (63) · [Universal Dependencies 2.0](#) (63) · [Universal Dependencies 2.1](#) (103) · [Universal Dependencies 2.3](#) (130) · [Universal Dependencies 2.5](#) (157) · [WolGram](#) (3) · [XPar](#) (2)

Treebank Types: [All](#) · [Ifg](#) (31/117) · [constituency](#) (0/19) · [constituency-alpino](#) (0/1) · [dependency](#) (0/45) · [dependency-cg](#) (0/662) · [dependency-tuebadz](#) (0/1) · [hpsg](#) (0/2)

Figur 3: Siden ‘Treebank Selection’ etter valg av ‘NorGramBank’

Under ‘Languages’ ser vi nå at ‘Norwegian’, ‘Norwegian Bokmål’ og ‘Norwegian Nynorsk’ er fremhevet; det angir språkene i deltrebankene. ‘Norwegian’ er deltrebanker som både omfatter bokmål og nynorsk. I NorGramBank gjelder det stortingsforhandlingene. Hvis f.eks. nynorsk ikke skal være med, klikk på ‘Norwegian Bokmål’ og ‘Norwegian’. Man får da med stortingsforhandlingene, men nynorske eksempler der kan filtreres fra i søkeuttrykkene.

Når disse valgene er gjort, kan man scrolle lenger ned på siden for å få se deltrebankene (Figur 4):

Click on a treebank name below to proceed. All selected treebanks will be available for viewing and searching. | [Show treebank descriptions](#)

Selected	Name	Collection	Type	Sentences	Words	Indexed	License	Downloads
all none				8 890 640	124 591 351			
	Norwegian Bokmål (nob)			7 670 731	100 505 846			
<input checked="" type="checkbox"/>	nob-avis	NorGram, NorGramBank	lfg	246 397	3 157 550	yes	CC-BY (Accepted)	no
<input checked="" type="checkbox"/>	nob-child	NorGram, NorGramBank	lfg	389 556	4 110 969	yes	CLARIN_ACA (Accepted)	no
<input checked="" type="checkbox"/>	nob-fn	NorGram, NorGramBank	lfg	489 341	8 321 494	yes	CLARIN_RES- DEP (Accepted)	no
<input checked="" type="checkbox"/>	nob-lbk-av	NorGram, NorGramBank	lfg	1 336	18 971	yes	CLARIN_ACA- NC (Accepted)	no
<input checked="" type="checkbox"/>	nob-lbk-sa	NorGram, NorGramBank	lfg	189 137	2 911 173	yes	CLARIN_ACA- NC (Accepted)	no
<input checked="" type="checkbox"/>	nob-lbk-tv	NorGram, NorGramBank	lfg	18 043	127 844	yes	CLARIN_ACA (Accepted)	no
<input checked="" type="checkbox"/>	nob-naob	NAOB, NorGram, NorGramBank	lfg	705 756	10 332 528	yes	(Accepted)	no
<input checked="" type="checkbox"/>	nob-naob_1	NAOB, NorGram, NorGramBank	lfg	627 665	9 314 943	yes	(Accepted)	no
<input checked="" type="checkbox"/>	nob-naob_2	NAOB, NorGram, NorGramBank	lfg	999 909	15 414 736	yes	(Accepted)	no
<input checked="" type="checkbox"/>	nob-ndt-lfg	NDT, NorGram, NorGramBank	lfg	20 045	276 942	yes	CC-BY (Accepted)	no
<input checked="" type="checkbox"/>	nob-novel	NorGram, NorGramBank	lfg	271 366	3 111 321	yes	CLARIN_ACA (Accepted)	no
<input checked="" type="checkbox"/>	nob-novel_1	NorGram, NorGramBank	lfg	406 281	4 370 000	yes	CLARIN_ACA (Accepted)	no
		NorGramBank					(Accepted)	
<input checked="" type="checkbox"/>	nob-novel_2	NorGram, NorGramBank	lfg	498 130	5 529 358	yes	CLARIN_ACA (Accepted)	no
<input checked="" type="checkbox"/>	nob-novel_3	NorGram, NorGramBank	lfg	441 721	4 639 538	yes	CLARIN_ACA (Accepted)	no
<input checked="" type="checkbox"/>	nob-novel_4	NorGram, NorGramBank	lfg	467 551	5 184 128	yes	CLARIN_ACA (Accepted)	no
<input checked="" type="checkbox"/>	nob-novel_5	NorGram, NorGramBank	lfg	443 891	4 817 445	yes	CLARIN_ACA (Accepted)	no
<input checked="" type="checkbox"/>	nob-novel_6	NorGram, NorGramBank	lfg	395 700	5 121 559	yes	CLARIN_ACA (Accepted)	no
<input checked="" type="checkbox"/>	nob-novel_7	NorGram, NorGramBank	lfg	221 444	3 230 447	yes	CLARIN_ACA (Accepted)	no
<input checked="" type="checkbox"/>	nob-novel_8	NAOB, NorGram, NorGramBank	lfg	570 541	7 201 101	yes	CLARIN_ACA (Accepted)	no
<input checked="" type="checkbox"/>	nob-novel_9	NAOB, NorGram, NorGramBank	lfg	265 770	3 298 575	yes	CLARIN_ACA (Accepted)	no
<input checked="" type="checkbox"/>	nob-sofie (aligned)	NorGram, NorGramBank	lfg	1 151	15 224	yes	unspecified (Accepted)	no
	Norwegian (nor)			1 219 909	24 085 505			
<input checked="" type="checkbox"/>	nor-stortinget	NorGram, NorGramBank	lfg	227 699	4 477 383	yes	NLOD (Accepted)	no
<input checked="" type="checkbox"/>	nor-stortinget_1	NorGram, NorGramBank	lfg	257 129	5 090 162	yes	NLOD (Accepted)	no
<input checked="" type="checkbox"/>	nor-stortinget_2	NorGram, NorGramBank	lfg	236 006	4 648 667	yes	NLOD (Accepted)	no
<input checked="" type="checkbox"/>	nor-stortinget_3	NorGram, NorGramBank	lfg	252 845	4 924 530	yes	NLOD (Accepted)	no
<input checked="" type="checkbox"/>	nor-stortinget_4	NorGram, NorGramBank	lfg	246 230	4 944 763	yes	NLOD (Accepted)	no

Figur 4: Nedre del av siden 'Treebank Selection' etter valg av bokmålsdelene av NorGramBank

Mer omfattende beskrivelser av den enkelte deltrebank vises ved valg av 'Show treebank descriptions' øverst på siden. En kortversjon finnes på denne websiden (ikke INESS-side):

<https://norgramtall.w.uib.no/deltrebankene-i-norgrambank/>

3. Lisenser

Det er 38 trebanker som vises (september 2020), og lisensene må aksepteres for dem alle før man kan søke i dem. (Dette kan bare gjøres etter innlogging og ev. tildeling av tilgang.) I Figur 4 står det 'Accepted' under alle lisensene; det vil det ikke gjøre ved første gangs innlogging. Klikk på én etter én i kolonnen som begynner med '**nob-avis**' i Figur 4. og aksepter lisensene. ('nob-novel'-lisensen gjelder for nob-novel_1, ... nob-novel_5, og tilsvarende for 'nor- stortinget'). **Dette gjøres kun én gang.** Ved senere klikk på trebank-navnet åpner man deltrebanken direkte. (Last siden inn på nytt når lisensene er akseptert, hvis du straks skal åpne en trebank.)

4. Veien til 'Sketch'-siden

Etter at tildeling av rettigheter og aksept av lisenser er på plass, klikk på ett (hvilket som helst) av trebank-navnene i oversikten på 'Treebank Selection'-siden (Figur 4), f.eks. nob-novel_1. Du kommer til 'Sentence Overview'-siden, der setningene i første dokument i den valgte deltrebanken er å se, sammen med en rekke valgmuligheter som det ikke er nødvendig å gå inn på her – vi skal videre til 'Sketch'-siden. Den velges fra menyen til venstre på siden:

Treebanks

Treebank Selection

Treebank Details

Sentence Overview

Sentence

Query

Sketch

Parallel Sentences

Metadata

Statistics

Disc. Statistics

Comments

Query expression

2441 sentences chosen. | [Add to custom treebank](#)

Previous

Next

Rows per page: 50

Hide: -

-

↑

↓

Include: -

↑

↓

#	Id	Solutions	Sentence
1	1	*1	Bekjennelsen : roman
2	2	2	I
3	3	6	Selvbedrageren
4	4	10+	For refleksjonen gjør oss alle feige;
5	5	408p	beslutsomhetens friske morgenfarve vil s skjene ut, og miste navn av handling...
6	6	4	Hysj!
7	7	1	Den skinnne Onhalla!

Figur 5: Litt av ‘Sentence Overview’-siden, med menyen for å gå videre til ‘Sketch’-siden

Ved valg av ‘Sketch’ i menyen dukker siden opp (Figur 6):

Signed in as *Helge Julius Jakhelln Dyvik (helge)*. [Sign out](#) | [Edit is off](#)

Sketch

Search in:

nob-avis, nob-child, nob-fn, nob-lbk-av, nob-lbk-sa, nob-lbk-tv, nob-naob, nob-naob_1, nob-naob_2, nob-ndt-lfg, nob-novel, nob-novel_1, nob-novel_2, nob-novel_3, nob-novel_4, nob-novel_5, nob-novel_6, nob-novel_7, nob-novel_8, nob-novel_9, nob-sofie, nor-stortinget, nor-stortinget_1, nor-stortinget_2, nor-stortinget_3, nor-stortinget_4

max #:

| fragments:

☐ none ☐ only

| fully disamb.:

☐ none ☐ only

disambiguated:

☐ none ☐ only

| unambiguous:

☐ none ☐ only

[Select query templates ...](#) | [Select sketch ...](#) | [Query history ...](#)

Figur 6: ‘Sketch’-siden

Valg av deltrebanker å søke i

Øverst på ‘Sketch’-siden vises alle deltrebankene som det skal søkes i (i utgangspunktet de som ble valgt på Treebank Selection-siden). Ved å klikke på ‘Search in’ kan man velge noen bort. Da dukker det opp en meny (Figur 7):

Search in: max #: disambiguation

Select query

Select treebanks to search in:

<input checked="" type="checkbox"/> nob-avis	<input checked="" type="checkbox"/> nob-sofie
<input checked="" type="checkbox"/> nob-child	<input checked="" type="checkbox"/> nor-stortinget
<input checked="" type="checkbox"/> nob-fn	<input checked="" type="checkbox"/> nor-stortinget_1
<input checked="" type="checkbox"/> nob-lbk-av	<input checked="" type="checkbox"/> nor-stortinget_2
<input checked="" type="checkbox"/> nob-lbk-sa	<input checked="" type="checkbox"/> nor-stortinget_3
<input checked="" type="checkbox"/> nob-lbk-tv	<input checked="" type="checkbox"/> nor-stortinget_4
<input checked="" type="checkbox"/> nob-naob	
<input checked="" type="checkbox"/> nob-naob_1	
<input checked="" type="checkbox"/> nob-naob_2	
<input checked="" type="checkbox"/> nob-ndt-lfg	
<input checked="" type="checkbox"/> nob-novel	
<input checked="" type="checkbox"/> nob-novel_1	
<input checked="" type="checkbox"/> nob-novel_2	
<input checked="" type="checkbox"/> nob-novel_3	
<input checked="" type="checkbox"/> nob-novel_4	
<input checked="" type="checkbox"/> nob-novel_5	
<input checked="" type="checkbox"/> nob-novel_6	
<input checked="" type="checkbox"/> nob-novel_7	
<input checked="" type="checkbox"/> nob-novel_8	
<input checked="" type="checkbox"/> nob-novel_9	

Select all Select none

Apply Cancel

amb.: only

Figur 7: Meny for valg av deltrebanker å søke i

Her kan man fjerne eller legge til avkryssninger ved den enkelte deltrebank. **NB! Ikke glem å klikke på 'Apply' når valgene er gjort!**

Øverst på siden (Figur 6) kan man også bl.a. velge om det skal søkes i fragment-analyser eller ikke. Fragment-analyser er analyser der parseren ikke har funnet en fullstendig analyse av setningen og isteden har returnert analyser av de maksimale setningsfragmentene som den har funnet analyser av. I syntaktisk avanserte søk (f.eks. søk

etter argumenttrammene til et verb) er det tilrådelig å velge ‘fragments’ = ‘none’ øverst på siden, og dermed bare søke i komplette analyser. Det reduserer antallet feilanalyser i søkeresultatet.

5. Søk på ‘Sketch’-siden

Pr. mars 2020 er det definert 54 søketemplater, de fleste av antatt leksikografisk interesse; dette inventaret blir utvidet og modifisert etter hvert som behovene melder seg. Søketemplatene finner man under lenken ‘Select query templates...’ oppe på ‘Sketch’-siden; se Figur 6. Menyen i Figur 8 nedenfor dukker opp; Figur 8 viser den første av tre sider. Det finnes også en oversikt over templatene, med beskrivelser, som en appendiks til slutt i denne veiledningen.

Templatnavnene er prefigert med den ordklassen de angår (ADJ-, ADV-, N-, V-, eller WD- når de ikke er begrenset til ordklasse). Et templatnavn er forsynt med én eller flere parametre, som står for informasjon som brukeren skal oppgi. @ADJ vil da være et spesifikt adjektiv, osv. Til høyre for templatnavnene står korte beskrivelser. Ved å klikke på dem får man mer detaljert dokumentasjon av templatet. Hvis man f.eks. klikker på beskrivelsen av templatet **N-defmascofem(@N)** i Figur 8, dukker resten av beskrivelsen opp i menyen (se Figur 9 nedenfor).

Query templates

Page 1 of 3 [Previous](#) [Next](#) | Go to page: [Go](#) | ☒ public ☐ private

Filter by name: | by description: | by template:

Click on a name to choose a template. (55 stored templates shown.) | ☐ Show template expansion

[Select marked templates](#) [Store marked templates as sketch ...](#)

Name	Description
<input type="checkbox"/> * ADJ-coord(@ADJ)	Adjectives coordinated with an adjective
<input type="checkbox"/> * ADJ-degreeadvs(@ADJ)	Degree adverbs modifying an adjective
<input type="checkbox"/> * ADJ-modifies(@ADJ)	Nouns modified by an adjective
<input type="checkbox"/> * ADJ-modnominadj(@ADJ)	Adjectives modifying a nominal head adjective
<input type="checkbox"/> * ADJ-suff(@SUFF)	Adjectives derived with a suffix
<input type="checkbox"/> * ADV-degmodifies(@ADV)	Adjectives modified by a degree adverb
<input type="checkbox"/> * ADV-types(@ADV)	The types of an adverb
<input type="checkbox"/> * AT-partofcadv(@lang)	at as part of adverbial complementizers
<input type="checkbox"/> * AT-selectedprep(@lang)	at-clauses governed by selected prepositions
<input type="checkbox"/> * AT-semgrep(@lang)	at-clauses governed by semantic prepositions
<input type="checkbox"/> * AT-verbwithandwithout(@verb)	Complement clauses of a verb with and without at
<input type="checkbox"/> * IDIOM-v-obj-p(@V,@OBJ,@P)	Verb-object-preposition idioms
<input type="checkbox"/> * IDIOM-v-obj(@V,@OBJ)	Verb-object idioms
<input type="checkbox"/> * N-adjmod(@N)	The adjectives modifying a noun
<input type="checkbox"/> * N-argofpreps(@N)	Prepositions of which a noun is an object
<input type="checkbox"/> * N-argofverbs(@N)	Verbs of which a noun is argument 1 or argument 2
<input type="checkbox"/> * N-argofverbs-witharg2(@N)	Verbs of which a noun is argument 1 (plus the argument 2 of the verb, if any) or of which it is argument 2
<input type="checkbox"/> * NAT-nationalnoun(@NAT-N)	Gender and definiteness of language-denoting nouns
<input type="checkbox"/> * N-defmascofem(@N)	Feminine vs. masculine inflection of a noun
<input type="checkbox"/> * N-forms(@N)	The forms of a noun

Figur 8: Øverste del av menyen over søketemplater

Feminine vs. masculine inflection of a noun

Sorts all occurrences in the definite singular of the noun @N according to inflection as masculine (+) or feminine (-). Search is limited to Bokmål texts.

Figur 9: Utvidet beskrivelse av et templat i menyen.

Med templatene kan man søke på to ulike måter:

(a) Bruk templatener enkeltvis

Ved å klikke på templatnavnet til venstre (Figur 8) velges dette templatet umiddelbart for søk. Menyen lukkes og templatet kommer til syne, sammen med kortforklaringen og tekstbokser der man kan gi parametrene konkrete verdier. Ved valg av templatet **N-defmascorfem(@N)**, får vi da bildet i Figur 10:

Select query templates ... | Query history ...

Template: * N-defmascorfem(@N)

Description: Feminine vs. masculine inflection of a noun

Parameters:

@N:

Figur 10: 'Sketch'-siden etter valg av et templat

Se beskrivelser av templatene

Igjen kan man få se hele beskrivelsen ved trykk på kortbeskrivelsen **Feminine vs. masculine inflection of a noun** på siden:

[Select query templates ...](#) | [Query history ...](#)

Template: * N-defmascorfem(@N)

Description: Feminine vs. masculine inflection of a noun

Sorts all occurrences in the definite singular of the noun @N according to inflection as masculine (+) or feminine (-). Search is limited to Bokmål texts.

Parameters:

@N:

Run query

Reset

Figur 11: 'Sketch'-siden med utvidet templatbeskrivelse

Det er også mulig å inspisere selve søkeuttrykket, hvis man skulle være interessert i det. Da klikker man på templatnavnet **N-defmascorfem(@N)**:

[Select query templates ...](#) | [Query history ...](#)

Template: * N-defmascorfem(@N)

Expansion: #f_ >PRED #p:'(@N)((*|\&).*)?' & #f_ >NUM 'sg' & #f_ >DEF-MORPH '\+' & #f_ >(GEND MASC) #masc & #f_ >(GEND NEUT) '-' & #f_ >(NTYPE NSYN) 'common' :: _language='nob'

Description: Feminine vs. masculine inflection of a noun

Sorts all occurrences in the definite singular of the noun @N according to inflection as masculine (+) or feminine (-). Search is limited to Bokmål texts.

Parameters:

@N:

Run query

Reset

Figur 12: 'Sketch'-siden med søkeuttrykket vist ('Expansion')

De utvidede beskrivelsene fjernes igjen ved nye klikk.

Gi parametrene verdier

Neste trinn er å gi søketemplatets parameter en verdi. La oss si at vi ønsker å studere bruken av maskulin vs. feminin bøyning i bestemt form entall av substantivene *kvinne*, *dame* og *kone*. Man kan gi parametrene alternative verdier ved å skille dem med loddrett strek: **kvinne|dame|kone**. Vi fyller da ut slik:

[Select query templates ...](#) | [Query history ...](#)

Template: * N-defmascofem(@N)

Description: Feminine vs. masculine inflection of a noun

Parameters:

@N:

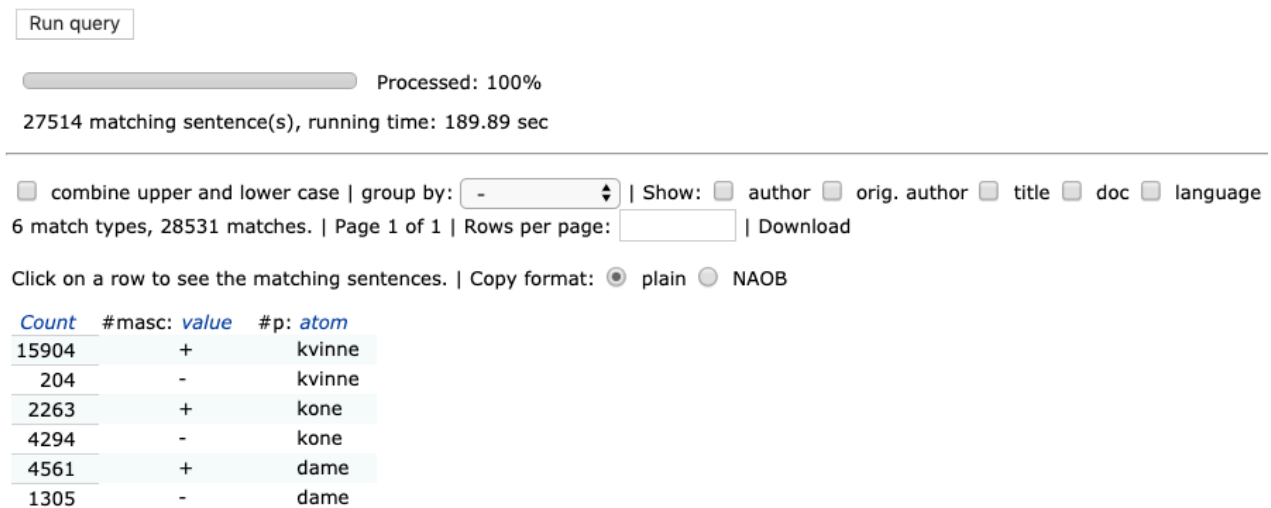
Run query

Reset

Figur 13: Parameterverdi angitt for et templat

Kjør søket: Run query

Når man har gjort det, klikker man på 'Run query'. Søket evalueres, fremgangen vises med prosenttall, og når søket er fullført (eller man utålmodig har klikket på 'stopp'), vises trefftypene i en tabell, som i Figur 14:



Figur 14: Søkeresultat: tabell med frekvenser for maskulin (+) og feminin (-) bøyning av det enkelte substantiv

Tabellen viser én trefftype per linje. 'match types' forteller hvor mange slike linjer det er. En trefftype er da kombinasjonen av verdiene for ulike valgte attributter i søkeuttrykket, f.eks. ordformer, lemmaformer eller grammatiske trekk. Dette er nærmere forklart i templatbeskrivelsene; i dette tilfellet er det altså maskulinum-verdi (+ eller -) pluss lemmaform.

Klikker man på en rad i tabellen i Figur 14, vises alle de tilhørende setningene, sammen med trebank og dokument de hører til, f.eks. for linje 2 med femininumformer av 'kvinne' (Figur 15):

Page 1 of 11 Previous Next Go to page: <input type="text"/> Go Download					
Click on a row to go to the sentence. Mouse over a row to see the structures.					
Treebank	Document	Trans.	Id	Sentence	
nob-novel_9	oai:bibsys.no:biblio...	no	5663	– Å herregud, sier kvinna og gjømmes ansiktet i armene:	Copy
nob-novel_9	oai:bibsys.no:biblio...	no	6617	– Temperatursjokk, sier kvinna og bærer inn bakverket, - jeg hadde nesten glemt hvor bister det kan være i Trondhjem på denne tida.	Copy
nob-novel_9	oai:bibsys.no:biblio...	no	6625	Kvinna nikker.	Copy
nob-novel_9	oai:bibsys.no:biblio...	no	6636	Kvinna smiler.	Copy
nob-novel_9	oai:bibsys.no:biblio...	no	6643	Kvinna glatter ut en ikke-eksisterende skrukk på duken.	Copy
nob-novel_9	oai:bibsys.no:biblio...	no	7136	Nei da, hun skulle ikke si noe til den kvinna som het Anne og som hadde en mann som het Kjell.	Copy
nob-novel_9	oai:bibsys.no:biblio...	no	7295	Kvinna med faneskrekke og gå-i-takt-fobi.	Copy
nob-novel_9	oai:bibsys.no:biblio...	no	7341	Det var bare det at den kvinna som satt i den andre stolen, ikke var en omreisende predikant.	Copy
nob-novel_9	oai:bibsys.no:biblio...	no	8906	Hun snakker så lavt at Annekin nesten ikke oppfatter, kvinna snakker vel mest til seg sjøl.	Copy
nob-novel_9	oai:bibsys.no:biblio...	no	8909	Kvinna rister på hodet, mumler.	Copy
nob-novel_9	oai:bibsys.no:biblio...	no	8915	– Tilgi dem, sier kvinna stille.	Copy
nob-novel_9	oai:bibsys.no:biblio...	no	9237	Anne-kin ser måpende på kvinna foran seg.	Copy

Figur 15: Toppen av trefflisten for linjen '205 - kvinne'

Kopier fra trefflisten

Med knappen 'Copy' kan setningen kopieres inn i bufferen og limes inn i et annet dokument. Her kan man velge mellom to formater, 'plain' og 'NAOB, slik det fremgår etter 'Copy format' i figur 14. Ved valg av 'plain' kopieres bare setningen, f.eks. første setning i Figur 15:

– Å herregud, sier kvinna og gjømmer ansiktet i armene:

Ved valg av NAOB inkluderes metadata i det formatet som brukes av NAOB-prosjektet:

```
<sitatledd><sitat>– Å herregud, sier kvinna og gjømmer ansiktet i
armene:</sitat><kilde><forf>Kim Småge</forf> <verk>En kjernesunn
død</verk> <ref>232</ref> <urn>https://urn.nb.no/URN:NBN:no-
nb_digibok_2010042708020</urn></kilde></sitatledd>
```

Man kan også laste ned alle setningene med ‘Download’ øverst. Da inkluderes metadata slik de vises i vinduet (toppen av listen):

```
nob-novel_9 7700704      – Å herregud, sier kvinna og gjømmer ansiktet i armene:
nob-novel_9 7702168      – Temperatursjokk, sier kvinna og bærer inn bakverket, - jeg
hadde nesten glemt hvor bistert det kan være i Trondhjem på denne tida.
nob-novel_9 7702176      Kvinna nikker.
nob-novel_9 7702187      Kvinna smiler.
nob-novel_9 7702194      Kvinna glatter ut en ikke-eksisterende skrukk på duken.
nob-novel_9 7702687      Nei da, hun skulle ikke si noe til den kvinna som het Anne og
som hadde en mann som het Kjell.
nob-novel_9 7702846      Kvinna med faneskrekk og gå-i-takt-fobi.
nob-novel_9 7702892      Det var bare det at den kvinna som satt i den andre stolen,
ikke var en omreisende predikant.
nob-novel_9 7704457      Hun snakker så lavt at Annekin nesten ikke oppfatter, kvinna
snakker vel mest til seg sjøl.
nob-novel_9 7704460      Kvinna rister på hodet, mumler.
nob-novel_9 7704466      – Tilgi dem, sier kvinna stille.
nob-novel_9 7704788      Anne-kin ser måpende på kvinna foran seg.
```

Når man «muser over» en setning, vises til høyre en forenklet versjon av setningens LFG-analyse i trebanken. Klikker man på setningen, vises setningens fulle analyse i et eget ‘Sentence’-vindu, der man også kan se den i kontekst (se mer om dette vinduet i avsnitt 6 på s. 27).

Hvis man «muser over» informasjonen under ‘Document’ (Figur 15), vises metadata med informasjon om forfatter, tittel, utgivelsesår og

markering av om teksten er en oversettelse eller ikke (Figur 16):

<i>Treebank</i>	<i>Document</i>	<i>Trans.</i>	<i>Id</i>	<i>Sentence</i>
nob-novel_9	oai:bibsys.no:biblio...	no	5663	– Å herregud, s
nob-novel_9	oai:bibsys.no:biblio:982280904			– iktet i arme
nob-novel_9	oai:bibsys.no:biblio:982280904			emperatursj
nob-novel_9	oai:bibsys.no:biblio:982280904			bakverket,
nob-novel_9	oai:bibsys.no:biblio:982280904			or bistert det
nob-novel_9	oai:bibsys.no:biblio:982280904			ne tida.
nob-novel_9	oai:bibsys.no:biblio:982280904			nna nikker.
nob-novel_9	oai:bibsys.no:biblio...	no	6636	Kvinna smiler.

Figur 16: Metadata vises ved «musing over» dokument-informasjonen

Tabellen (Figur 14) lar seg manipulere på ulike måter. Den kan sorteres etter antall treff ved å klikke på 'Count', og alfabetisk etter de viste attributtene, ved å klikke på dem i tittelraden. Det er en avkrysningsboks ('combine upper and lower case') for å slå sammen verdier med store og små bokstaver; da vises alle verdiene med små bokstaver.

Opsjonen 'Group by'

Verdiene i tabellen kan også grupperes etter et valgt attributt (**group by**). Det kan være nyttig hvis tabellen f.eks. har samme substantiv i ulike linjer (som i Figur 14), og man gjerne vil ha bare én linje pr. substantiv, f.eks. fordi man vil laste tabellen ned (Via 'Download') og kopiere den in i Excel el.l. for videre behandling. I tittelraden vises da alle verdiene dette attributtet antar i treffene, og antall forekomster står i kolonnene under. Denne funksjonen er bare meningsfull dersom attributtet antar få forskjellige verdier i søkeresultatet.

Hvis man vil gjøre dette med resultatet i Figur 14, kan man velge 'group by masc value', som i Figur 17:

☐ combine upper and lower case | group by: ✓ -
 doc ☐ language ☐ treebank ☐ size
 6 match types, 28531 matches. | Page 1 of 1

Click on a row to see the matching sentences. | Copy format: ☒

Count	#masc: value	#p: atom
15904	+	kvinne
204	-	kvinne
2263	+	kone
4294	-	kone
4561	+	dame
1305	-	dame

Figur 17: Menyvalg for 'group by'

Resultatet blir da slik (Figur 18):

☐ combine upper and lower case | group by: | Show: ☐ author
 doc ☐ language ☐ treebank ☐ size
 3 match types, 28531 matches. | Page 1 of 1 | Rows per page: | Down

Click on a row to see the matching sentences. | Copy format: ☒ plain ☐ NAOB

Count	#masc: value: + -		#p: atom
16108	15904	204	kvinne
6557	2263	4294	kone
5866	4561	1305	dame

Figur 18: Resultatet av 'group by masc value'

Snarvei via 'Query history'

Via lenken **Query history** kan man gå tilbake til tidligere templat-søk og kjøre dem igjen med valget **Run query**. Ikke bare søket med parameterverdi, men også søkeresultater lagres i en cache som varer innenfor sesjonen (en brukersesjon varer i 16 timer før man må logge

inn på nytt). Det innebærer at valg av et tidligere søk under **Query history**, når dette søket er gjort i samme sesjon, etterfulgt av **Run query**, vil vise de lagrede søkeresultatene øyeblikkelig.

Forfatter og tittel: Sorter etter metadata

Det er også mulig å sortere treffene etter metadata som f.eks. forfatter og verktittel (i den grad tekstene har oppgitte forfattere – det mangler i stortingsforhandlingene og noen avisartikler.)

Vi tar igjen utgangspunkt i søket etter maskuline og feminine substantivformer, vist i Figur 14 ovenfor. I øverste linje av treffene i Figur 14 finner vi disse mulighetene for å krysse av for metadata:

Show: ☐ author ☐ orig. author ☐ title ☐ doc ☐ language ☐ treebank ☐ size

Ved å krysse av for ett eller flere av disse oppnår man å fordele treffene på henholdsvis forfatter, originalens forfatter (ved oversettelser – da står oversetteren som 'author'), verktittel, dokument, språk (f.eks. 'nob' for bokmål og 'nno' for nynorsk), deltrebank i NorGramBank, og størrelse på deltrebanken målt i antall setninger. Hvis vi krysser av ved 'author' og dermed fordeler treffene på forfatter, blir toppen av trefflisten (sortert slik at '-' og 'kvinne' kommer først gjennom klikk på kolonneoverskriftene – klikk på 'Count' først) som i Figur 19:

<i>Count</i>	#masc: <i>value</i>	#p: <i>atom</i>	globals: <i>author</i>
77	-	kvinne	Småge, Kim
25	-	kvinne	Pedersen, Bente
21	-	kvinne	Wassmo, Herbjørg
9	-	kvinne	Børli, Hans
8	-	kvinne	Hoel, Tore Elias
7	-	kvinne	Gjessing, Gutorm; Norsk Arkeologisk Selskap
6	-	kvinne	Osland, Erna
5	-	kvinne	Nome, John
4	-	kvinne	
4	-	kvinne	Børli, Hans; Lindrupsen, Tor
4	-	kvinne	Haalke, Magnhild
4	-	kvinne	Nordby, Marit
3	-	kvinne	Alfsen, Merete
2	-	kvinne	Fønhus, Mikkjel
2	-	kvinne	Kittelsen, Erling
2	-	kvinne	Lange-Nielsen, Sissel
2	-	kvinne	Markusson, Andreas
2	-	kvinne	Skard, Åse Gruda
1	-	kvinne	Beyer, Edvard
1	-	kvinne	Bratteli, Randi
1	-	kvinne	Dahl, Willy
1	-	kvinne	Engelstad, Vibeke
1	-	kvinne	Gundersen, Gunnar Bull
1	-	kvinne	Haavind, Hanne
1	-	kvinne	Haff, Bergljot Hobæk
1	-	kvinne	Halvorsen, Ida; Finstad, Liv; Høigård, Cecilie
1	-	kvinne	Harsson, Margit
1	-	kvinne	Johannesen, Georg
1	-	kvinne	Mjøen, Sonja
1	-	kvinne	Nissen, Ingjald
1	-	kvinne	Skagen, Fredrik
1	-	kvinne	Spilde, Ingrid
1	-	kvinne	Stokke, Bernhard; Vang Nyman, Ingrid
1	-	kvinne	Thorn, Finn
1	-	kvinne	Vold, Jan Erik

Figur 19: Toppen av trefflisten med forfattere inkludert

Det fremgår at størsteparten av forekomstene av 'kvinna' skriver seg fra tre forfattere.

(b) Søk med flere templatener samtidig

Alternativt kan man velge flere templatener for samtidig evaluering, ved å krysse av for de ønskete templatene i templatemenyen (se Figur 8), f.eks. alle som omhandler verb. De markerte templatene velges for evaluering når man klikker på 'Select marked templates' øverst i menyen. Som før lukkes menyen, og de valgte templatene vises, sammen med tekstbokser for alle parametrene til de valgte templatene. Parametre som er felles for flere templatener (f.eks. @N) listes bare én gang, slik at man velger samme verdi for samme parameter i alle templatene. Når man så klikker på 'Run queries' evalueres alle templatene etter hverandre, og til slutt vises resultatene i en liste, med hvert templat pluss dets parameterverdier på én linje, inkludert informasjon om antall treffsetninger. Ved å klikke på templatene i resultatlisten kan man inspisere trefftypene, slik som under (a).

Eksempel: Vi vil definere en «pakke» der vi for et gitt verb vil ha: en liste over dets handlere eller 'logiske subjekter' (**V-arg1(@V)**), en liste over dens 'patienter' eller 'logiske objekter' (**V-arg2(@V)**), en oversikt over dets argumentrammer (**V-argframes(@)**) og en oversikt over s-passiv-bruk (**V-passive-s(@V)**). Vi krysser da av for disse i den menyen som er vist ovenfor i Figur 8, som i Figur 20:

<input checked="" type="checkbox"/>	* V-arg1(@V)	Agents of a verb
<input checked="" type="checkbox"/>	* V-arg2(@V)	Patients of a verb
<input checked="" type="checkbox"/>	* V-argframes(@V)	Argument frames of a verb
<input type="checkbox"/>	* V-forms(@V,@FORM)	The a verbo forms of a verb
<input type="checkbox"/>	* V-maleorfemale(@V)	Male vs. female agents of a verb
<input type="checkbox"/>	* V-morphpassforms-nno(@V)	Forms of the morphological passive of a verb, Nynorsk
<input checked="" type="checkbox"/>	* V-morphpassforms-nob(@V)	Forms of the morphological passive of a verb, Bokmål

Figur 20: Valg av flere templatere på én gang

Så velges 'Select marked templates' oppe i menyen, og resultatet blir slik (Figur 21):

Select query templates ... | Query history ...

Template: * V-arg1(@V)

Description: Agents of a verb

Template: * V-arg2(@V)

Description: Patients of a verb

Template: * V-argframes(@V)

Description: Argument frames of a verb

Template: * V-morphpassforms-nob(@V)

Description: Forms of the morphological passive of a verb, Bokmål

Parameters:

@V:

Run queries

Figur 21: 'Sketch'-siden med en «pakke» templatere for samtidig evaluering

Hvis vi gir @V verdien 'anse' og klikker på 'Run queries', blir resultatet slik (Figur 22):

Parameters:
@V:

Run queries

Processed: 100%

2984 matching sentence(s), running time: 68.68 sec

Templates

* **V-arg1()**: @V=anse
653 sentences

* **V-arg2()**: @V=anse
1260 sentences

* **V-argframes()**: @V=anse
652 sentences

* **V-morphpassforms-nob()**: @V=anse
419 sentences

Results
☐ combine upper and lower case | group by:

188 match types, 658 matches. | Page 1 of 4 [Next](#) | Go to page: [Go](#) | Rows per page: | [Download](#)

Click on a row to see the matching sentences.

Count	#a: atom	#arg1: value
19	anse	jeg
14	anse	vi
9	anse	man
8	anse	pro
6	anse	de
6	anse	han
6	anse	regjering
5	anse	statsråd
4	anse	den
4	anse	hun
3	anse	forsker
3	anse	komité
2	anse	flertall
2	anse	folk
2	anse	Fremskrittspartiet
2	anse	medlem
2	anse	myndighet

Figur 22: Resultatet (toppen av listen) av søk med fire templatere med verbet anse som parameter

Til venstre i Figur 22 ser vi hvor mange setningstreff hvert templat fikk. Resultatene kan så inspiseres ved at man klikker på det aktuelle templatet i denne oversikten. I Figur 22 vises resultatet for templatet **V-arg1()** til høyre. '188 match types' betyr at det er funnet 188 ulike predikater (substantiver, pronomen, evt. verb) som opptrer som agens for 'anse'. I Figur 22 er de sortert etter frekvens ved klikk på 'Count'. I kolonnen #a finner vi 'anse'-predikatene, som også omfatter eventuelle flerordsuttrykk med selekterte preposisjoner o.l. (ikke vist i figuren). 'pro' betyr ikke-uttrykt subjekt, som i ikke-kontrollerte infinitiver, f.eks. *To vanlige måter å anse begrepet på, er ...*, eller subjekter med bare determinativ (**Flere** anser at gratis heroinutdeling har bidratt til dette). For å inspisere resultatene for de andre templatene i pakken, klikker man på dem, f.eks. **V-argframes()**, og får resultatet i Figur 23 nedenfor. Trefflisten sorterer mellom de ulike valensrammene på grunnlag av hvilke argumenter (1, 2 eller 3) som er uttrykt i en setning, og hva som uttrykker dem: substantiv (common), egennavn (proper), pronomen (pronoun), infinitivsuttrykk (inf) osv. – se beskrivelsen av templatet på 'Sketch'-siden eller i Appendiks nedenfor.

Templates

* **V-arg1()**: @V=anse
653 sentences

* **V-arg2()**: @V=anse
1260 sentences

* **V-argframes()**: @V=anse
652 sentences

* **V-morphpassforms-nob()**: @V=anse
419 sentences

Results

☐ combine upper and lower case | group by: -
60 match types, 659 matches. | Page 1 of 2 [Next](#) | Rows per page:
 | [Download](#)

Click on a row to see the matching sentences.

Count	#a: atom	#arg1: value	#arg2: value	#arg3: value
56	anse		common	inf
50	anse	pronoun	fin	
41	anse	common	fin	
18	anse		pronoun	inf
13	anse		sup	
9	anse	proper	fin	
4	anse	pronoun	common	inf
4	anse		proper	inf
4	anse		fin	
3	anse	pronoun	pronoun	inf
2	anse		inf	inf
2	anse	common	pronoun	inf
1	anse		pronoun	fin
1	anse	common	common	inf
35	anse*for		common	inf
24	anse*for	pronoun	common	inf
22	anse*for	pronoun	pronoun	inf
15	anse*for		pronoun	inf
7	anse*for	common	common	inf
6	anse*for		common	common
6	anse*for	proper	common	inf
4	anse*for	pronoun	common	common
3	anse*for	common	common	common
3	anse*for		proper	inf
3	anse*for	proper	pronoun	inf
2	anse*for	pronoun	pronoun	common
2	anse*for	common	pronoun	inf
2	anse*for		pronoun	common
1	anse*for	proper	proper	common
1	anse*for	proper	common	common

Figur 23: Øvre del av resultatet for templatet **V-argframes()** og verbet anse

I øverste linje i Figur 23 (med 56 eksempler) er kolonnen #arg1 tom. Det vil typisk bety en agensløs passiv setning. #arg2 i den samme linjen er et fellesnavn (common), og #arg3 et infinitivsuttrykk (inf). Ved klikk på linjen får vi se eksemplene; Figur 24 viser de fire øverste:

Count	#a: atom	#arg1: value	#arg2: value	#arg3: value
56	anse		common	inf

Page 1 of 3 [Previous](#) [Next](#) | Go to page: [Go](#) | [Download](#)

Click on a row to go to the sentence. Mouse over a row to see the structures.

Treebank	Document	Id	Sentence	
nor-stortinget_2	s111110	500	Dette medfører også at slik bistand er ansett å falle utenfor bestemmelsens rekkevidde.	Copy
nor-stortinget_2	s111116	432	En del av inntektene anses å være knyttet til konjunkturer og er derfor midlertidige.	Copy
nor-stortinget_2	s111220	438	«Konvensjonen kan ikke anses å stille krav om særrettigheter for urbefolkningene.»	Copy
nor-stortinget_2	s120207	970	I lovforslaget er det likevel innarbeidet enkelte sikkerhetsventiler som vil kunne aktualiseres dersom søkermassen anses å utgjøre en utfordring for militære interesser.	Copy

Figur 24: Toppen av eksempellisten for rammen $\text{arg1}=\text{NULL}$, $\text{arg2}=\text{common}$ og $\text{arg3}=\text{inf}$

Ved klikk på det siste templatet, **V-morphpassforms-nob()**, får vi dette resultatet frem:

Templates

* **V-arg1()**: @V=anse

653 sentences

* **V-arg2()**: @V=anse

1260 sentences

* **V-argframes()**: @V=anse

652 sentences

* **V-morphpassforms-nob()**: @V=anse

419 sentences

Results

☐ combine upper and lower case | group by:

-

6 match types, 420 matches. | Page 1 of 1 | Rows per page: | [Download](#)

Click on a row to see the matching sentences.

Count	#a: <i>atom</i>	#form: <i>atom</i>
88	anse	pres
22	anse	inf
37	anse*for	pres
10	anse*for	inf
180	anse*som	pres
83	anse*som	inf

Figur 25: Resultater for templatet **V-morphpassforms-nob()**

I Figur 25 ser vi frekvensene for s-passiv i infinitiv og presens for tre predikater med verbet *anse*. For å se de 83 eksemplene på passiv

infinitiv av *anse som*, klikker man på nederste linje. Toppen av resultatet vises i Figur 26:

83 anse*som inf

Page 1 of 5 Previous Next | Go to page: Go | Download

Click on a row to go to the sentence. Mouse over a row to see the structures.

Treebank	Document	Id	Sentence	
nor-stortinget_2	s110414	2805	Det er vanskelig å se at dette kan anses som en likebehandling av aktørene.	Copy
nor-stortinget_2	s110511	642	Begrepet «kraftverk» er også nærmere presisert i skatteloven med tilhørende forskrift, herunder hva som skattemessig skal anses som ett eller flere kraftverk.	Copy
nor-stortinget_2	s110520	646	Man nærmest antar at EU vil komme med sanksjoner langt utover det som må anses som forholdsmessige reaksjoner.	Copy
nor-stortinget_2	s111019	474	Skattyter kan fritas for tilleggsskatt når hans eller hennes forhold må anses som unnskyldelig på grunn av sykdom, alderdom, uerfarenhet eller annen årsak.	Copy
nor-stortinget_2	s111110	1727	Høyesterett kom imidlertid til at en øvelse over fire dager ikke kunne anses som en arbeidsprosess i lovens forstand.	Copy

Figur 26: Toppen av eksemplene på passiv infinitiv av *anse som*

6. Flere muligheter utenfor ‘Sketch’

‘Sketch’-siden med sine templatere er ment som et hjelpemiddel for dem som ikke vil gå inn i selve søkeuttrykkene eller trenger de mer varierte mulighetene som søkesystemet ellers gir. Her skal vi likevel nevne noen ytterligere muligheter utenfor ‘Sketch’-siden, ifall behov for dem skulle oppstå.

Sidene ‘Sentence Overview’ og ‘Sentence’

Ovenfor har vi sett at eksempelsetningene kan inspiseres i menyer ved klikk på trefflinjene på ‘Sketch’-siden. Det ble også nevnt at kontekst og analyse for setningene kan bringes frem i et eget ‘Sentence’-vindu ved klikk på en setning. Menyen med setninger må da hentes frem igjen på nytt hvis man så går tilbake igjen til ‘Sketch’-siden. Hvis man ønsker en mer stabil oversikt over treffsetningene i et ‘Sentence Overview’ i menyen til venstre – for eksempel etter søket

vist i Figur 14, gjentatt i Figur 27:

Sentence Overview
Sentence Query
Sketch
Parallel Sentences
Metadata

Statistics
Disc. Statistics
Comments
Administration

Tools
XLE-Web

@N: kvinne|dame|kone

Run query

Processed: 100%

27517 matching sentence(s), running time: 265.81 sec

☐ combine upper and lower case | group by: - |

6 match types, 28533 matches. | Page 1 of 1 | Rows per page:

Click on a row to see the matching sentences. | Copy format:

Count	#masc:	value	#p: atom
15901	+	kvinne	
205	-	kvinne	
2264	+	kone	
4294	-	kone	
4564	+	dame	
1305	-	dame	

Figur 27: Søket fra Figur 14 med menyvalg til venstre

Ved klikk på 'Sentence Overview' kommer man til 'Sentence Overview'-siden, der samtlige treffsetninger fra søket listes opp, ikke lenger sortert etter kategoriene i søkeuttrykket (Figur 28):

13492 matching sentence(s), running time: 88.98 sec

Add to custom treebank

Previous Next Rows per page: 50 Hide: - Show: - Download mode: Sentence Exclude:

- Include: -

#	Id	Treebank	Document	Solutions	Sentence
1	58	nob-lbk-av	Politiet må gi f...	82	Hun gjør ikke mye gale ting, kvinnen fra Bergen.
2	14	nob-lbk-av	Jernkvinnen som ...	4	Legg merke til denne kvinnen.
3	141	nob-avis		12	Plasserer seg ved siden av kona i sofaen og prøver å beskrive en ekstrem situasjon.
4	37	nob-avis		18	Men han gikk så videre med damen.
5	3	nob-avis		*10+p	Den norskmarokkanske kvinnen hadde satt seg i moskeen noen få hundre meter fra Regjeringskvartalet da hun hørte et kraftig smell.
6	49	nob-avis		40	Å ringe kona kostet 25 kroner minuttet.
7	112	nob-avis		18	Den argentinske kvinnen legger ikke skjul på at hun er både stolt og fornøyd.
8	88	nob-avis		4p	Den ugifte, barnløse kvinnen blir skildrer som svært dyktig og som et ordensmenneske med enorm arbeidskapasitet.
9	43	nob-avis		12	Alle visste hvem kvinnen med de tykke, runde brilleglassene og den mørke parykken var.
10	42	nob-avis		8004p	Ifølge kjønnsforskeren er han i mindre grad opptatt av kvinnen som omsorgsperson, og han skriver ikke så mye om oppdragelse og omsorg i det såkalte manifestet.
11	230	nob-avis		256+p	En av båttheltene bor ved Tyrifjorden, og han og kona hørte skudd, måtte de ut på fjorden, sa han i retten.
12	75	nob-avis		992p	Flere av danikvinnen Laikalias fingre er kuttet av nedenfor det første leddet, i noe som kan virke som en grotesk skikk som nå er forbudt.
13	74	nob-avis		170+	Når vi kommer til toppen får kvinnen sin fortjente betaling og jeg overtar sekken.
14	72	nob-avis		32	Men den lille kvinnen orker mer og vil gjerne ta hånd om ryggsekken min mens vi sammen jobber oss oppover i dalsiden.
15	70	nob-avis		4	Uten frykt for spanskrøret har kvinnen Laikalia bedre tid enn de unge skoleelevene.
16	195	nob-avis		30	Det var for eller mot kvinnens frie valg.
17	58	nob-avis		58p	Men hun har snakket mye om kvinnens soleklare rett til å bestemme over egne fødsler.
18	63	nob-avis		2	Damen bak disken ler.
19	21	nob-avis		9+	Noen timer før hadde han og kona planlagt blåskjellmiddag.
20	37	nob-avis		88	Kona, Hanne har hjemmekontor og følger halvveis med i samtalen.
21	3	nob-avis		569	Kvinnen bak essayet, Anne-Marie Slaughter, jobbet tidligere som direktør for politisk planlegging i det amerikanske Utenriksdepartementet.
22	91	nob-avis		186p	Den sprudlende, engasjerte kvinnen er 31 år i år.
23	33	nob-avis		4992p	Men kanskje kunne kvinnen som ble best i sitt kull hatt jobben hun drømte om dersom arbeidslivet var tilpasset dagens moderne samfunn? spør Slaughter.
24	198	nob-avis		12	At damen på Kiwi-butikken, når hun har glemt nøkkelen til postboksen, tilbyr seg å hente posten hennes.
25	18	nob-avis		6	Kvinnen på utsiden var designantropolog og psykolog Anna Kirah.
26	115	nob-avis		2204p	Konen til Mood var en av dem som tilfeldig hadde fri fra sin jobb i regjeringskvartalet 22. juli i fjor sommer.

Figur 28: Nedre del av 'Sentence Overview'-siden med samtlige 13492 treffsetninger over flere sider

På denne siden kan setningene sorteres etter ulike kriterier, f.eks. lengde (klikk på 'Show' og velg 'Word Count', og deretter på overskriften 'Words' i den kolonnen som da dukker opp – med flere klikk sorteres den enten ovenfra eller nedenfra).

Ved klikk på en setning i 'Sentence Overview' (Figur 29) kommer man til 'Sentence'-vinduet.

569	569	Kvinnen bak essayet, Anne-Marie Slaughter, jobbet tidligere som direktør for politisk planlegging i det amerikanske Utenriksdepartementet.
186p	186	<u>Den sprudlende, engasjerte kvinnen er 31 år i år.</u>
4992p	4992	Men kanskje kunne kvinnen som ble best i sitt kull hatt jobben hun drømte om dersom arbeidslivet var tilpasset dagens moderne samfunn? spør Slaughter.

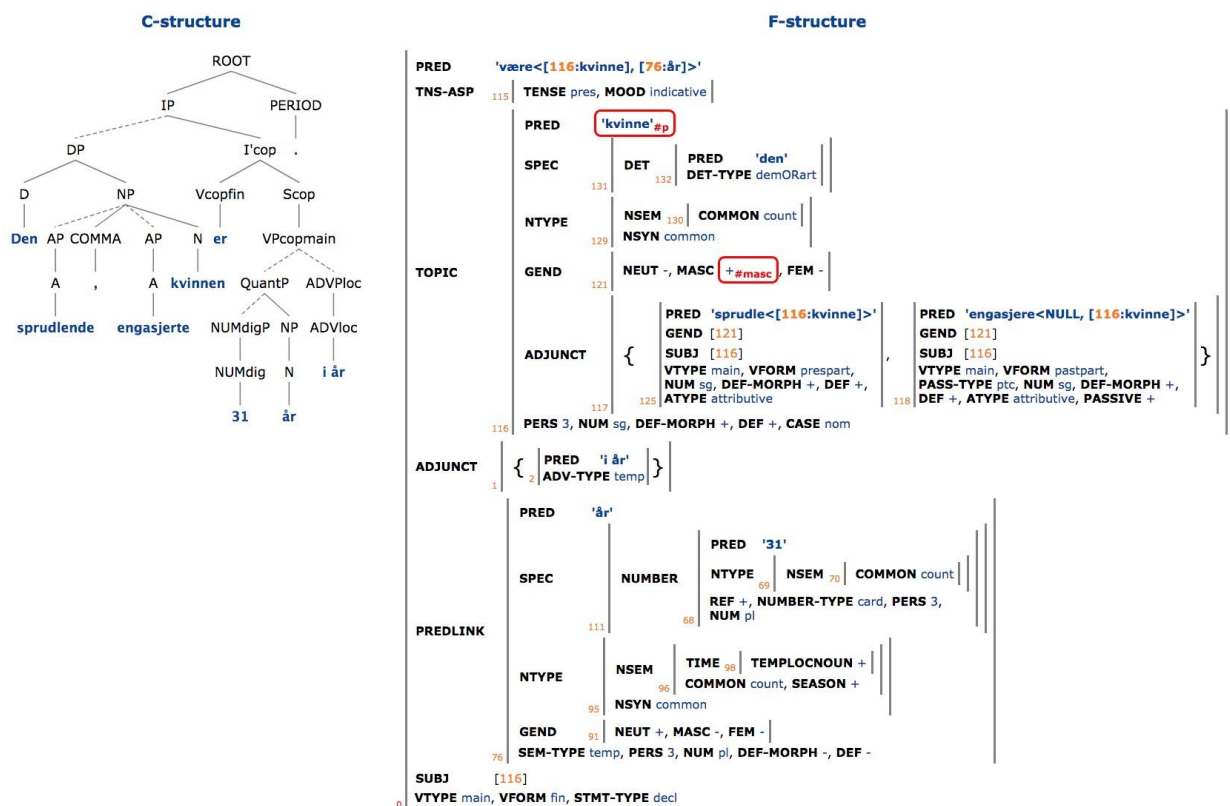
Figur 29: Klikk på en setning i 'Sentence Overview'

I 'Sentence'-vinduet kan man blant annet se setningen i kontekst (Figur 30) og analysen av setningen (Figur 31). I analysen er de egenskapene ved setningen som førte til at den ble funnet av søket, markert med rødt.

Sentence #91: Neste skritt blir å overta regjeringen, og til slutt presidentembetet! – Hvem skulle eventuelt lede en slik ko
Den sprudlende, engasjerte kvinnen er 31 år i år. Hun er fraskilt og bor i en sentrumsbydel i Kairo sammen m
 fra Banha i provinsen Qalyubia nord for Kairo.

(Pruned; 186 solutions, 0.440 CPU seconds, 17.480MB max mem, 463 subtrees unified; Grammar date: May 25, 2016 17:22;

Figur 30: Litt av setningens kontekst på 'Sentence'-siden



Figur 31: Setningens analyse på 'Sentence'-siden

Ved et nytt klikk på 'Sentence Overview' i menyen til venstre kommer man tilbake til setningsoversikten.

7. Advarsel

Der vil være et varierende innslag av feilanalyser i treffene, avhengig av fenomen. Treffene må derfor inspiseres før frekvenser brukes i andre sammenhenger.

Spørsmål og kommentarer kan sendes til

`iness@uib.no`

eller direkte til

`paul.meurer@uib.no`

og/eller

`helge.dyvik@uib.no`

Appendiks: Templatoversikt

Søketemplatene nedenfor er gruppert etter de ordklassene de angår, og forsynt med de beskrivelsene som er å finne på skjermen.

Innhold

1. [Adjektivtemplater](#)
2. [Adverbtemplater](#)
3. [at-setninger](#)
4. [Idiomer](#)
5. [Substantivtemplater](#)
6. [Preposisjonstemplater](#)
7. [Pronomentemplater](#)
8. [Syntakstemplater](#)
9. [Verbtemplater](#)
10. [Templater for flere ordklasser](#)

1. Adjektivtemplater

Til innholdsfortegnelsen

ADJ-attrib-or-nominal(@ADJ)

Function of an adjective as attributive or nominal

Sorts, with frequencies, the occurrences of the adjective **@ADJ** according to its function as either attributive to a noun or as an NP head (strictly speaking, as attributive to an empty NP head expressing the predicate 'pro'). In the output, the occurrences with an empty NP head (i.e. 'nominal adjective') are marked 'pro', and the rest are unmarked in the same column.

If **@ADJ** is a participle it is advisable to enter the participle form and the infinitive form as alternatives (e.g. **forurettet|forurette**), since some (or even all) adjective occurrences of participles may be derived from verbal lexical entries.

ADJ-coord(@ADJ)

Adjectives coordinated with an adjective

Lists, with frequencies, 'og'-coordinations of two adjectives of which one is the adjective **@ADJ**. While **@ADJ** should be the lemma form, the listed results are the occurring inflectional forms of the adjectives. Thus, the **@ADJ god** will yield i.a. **bedre og bedre**.

Det interessante ved koordinerte adjektiver er at slike koordinasjoner ofte har idiom-preg (*stor og sterk, syk og elendig*) og derfor kan være interessante å kartlegge.

ADJ-degreeadvs(@ADJ)

Degree adverbs modifying an adjective

Lists, with frequencies, the adverbs modifying the adjective @ADJ.

ADJ-suff(@SUFF)

Adjectives derived with a suffix

Lists, with frequencies, the lemma forms of all adjectives derived with the suffix @SUFF (e.g., **løs**) – or more precisely, all adjectives whose lemma forms end with the sequence @SUFF (which doesn't have to be a suffix grammatically). Search is limited to Bokmål texts.

ADJ-modifies(@ADJ)

Nouns modified by an adjective

Lists, with frequencies, all nouns modified by the adjective @ADJ, attributively or predicatively.

ADJ-modnominadj(@ADJ)

Adjectives modifying a nominal head adjective

Lists, with frequencies, the lemma forms of all adjectives modifying the adjective @ADJ when @ADJ functions as the head of a nominal phrase (or more technically: when @ADJ is the second adjective in an NP with no N head). @ADJ should be specified in lemma form, e.g. **gammel** in order to find **de pleietrengende gamle** and **de pleietrengende eldre**. In the case of past participles like **ansatt** this means the infinitive form **ansette**. However, some participle forms, like **etterlatt**, are listed in the

lexicon as adjectives, in which case the lemma form is **etterlatt**. Hence both possibilities should be tried if one of them yields no results. In the case of present participles like **medvirkende**, the participle form should always be specified.

2. Adverbtemplatet

Til innholdsfortegnelsen

ADV-degmodifies(@ADV)

Adjectives modified by a degree adverb

Lists, with frequencies, all adjectives modified by the adverb @ADV as a degree adverb (ADVdeg). @ADV should be given in its actual text form, e.g. participle form in the case of participles functioning as degree adverbs.

ADV-types(@ADV)

The types of an adverb

Lists, with frequencies, the adverb types to which @ADV belongs. The types are described by means of subscripts to the category name ADV and are distinguished on the basis of the syntactic distribution of @ADV, which tends to be correlated with semantic functions. The types are given below.

The form specified in the @ADV field is the form appearing in the text. This means that, e.g., **frem** and **fram** are separate entries, as are **sent** and **seint**. Thus, it is recommendable to specify alternative forms, e.g. **frem|fram**. Inflected forms of verbs occurring as adverbs must be specified in those inflected forms – e.g. **hoppende** when occurring as a degree adverb (ADVdeg) in **hoppende glad**. In the case of adjectives in -t, like **sterkt** in **sterkt forsinket**, they should also be given in the inflected form when occurring as degree adverbs. However, this template does not cover adverbial use as sentence adverbials (**Han kjørte sterkt**) since they are not actually analyzed as adverbs in this position.

nexus position = before the non-finite verb (if any) of a main clause: Jeg har **ofte** kjørt bil;

verb phrase position = after the non-finite verb (if any) of a main clause: Jeg har kjørt bil **forsiktig**

ADV: manner adverb, modifying the verbal action, typically occurring in verb phrase position, not in nexus position: *Han kan kjøre **fremover**.*

ADVatt: attitude adverb (the speaker's attitude to the statement), typically occurring in nexus position immediately after the finite verb (or post-verbal subject if any): *Vi har **dessverre** ikke sett resultater ennå; Derfor har vi **dessverre** ikke sett resultater ennå.*

ADVcmnt: commitment adverb (the speaker's commitment to the statement), typically occurring in nexus position after the ADVatt position: *Det må dessverre **selvfølgelig** være riktig.*

ADVdeg: degree adverb, typically modifying an adjective: *Dette har pågått over **veldig** lang tid.*

ADVdegcmp: comparison degree adverb, the first adverb in a 'jo ... desto' construction: ***Jo** tidligere det skjer desto bedre.*

ADVdegdst: comparison degree adverb, the second adverb in a 'jo ... desto' construction: *Jo tidligere det skjer **desto** bedre.*

ADVdegint: interrogative degree adverb, questioning the degree of an adjective or quantifier in a wh-interrogative: ***Hvor** mange kommer?*

ADVdegloc: locative/temporal degree adverb, typically modifying a locative or temporal adverbial phrase: ***Først** da er vi ferdige.*

ADVdegnum: numeral degree adverb, typically modifying a numerical expression: *Det var slik i **over** 50 år.*

ADVdegpost: postposed degree adverb, placed after the adjective which it modifies: *Det er ikke et godt **nok** tilbud.*

ADVdegqnt: quantifier degree adverb, modifying a quantifier: *Vi samarbeider på **veldig** mange områder.*

ADVfoc: focus adverb, preposed to a phrase and bringing it into focus: *Vibør gjøre **nettopp** det.*

ADVfocpost: focus adverb, postposed after a phrase and bringing it into focus: *Dette handler om prioritering **også**.*

ADVint: interrogative adverb, questioning an adverbial relation in a wh-

interrogative: **Hvordan** blir hverdagen nå?

ADVloc: locative or temporal adverb, occurring in nexus as well as verb phrase position: *Vi skal **nå** befinne oss **her**.*

ADVneg: negative adverb, modifying a sentence in its nexus position: *Det vil **ikke** skje.*

ADVpar: parenthetical adverb, occurring as a parenthetical between commas: *Per, **for eksempel**, kommer.*

ADVpre: pre-finite adverb, occurring in nexus position before the finite verb in main clauses without being in topic position (and hence seemingly violating the verb-second constraint): *Jeg **bare** sier det først.*

ADVprt: particle adverb, functioning as an unstressed discourse particle in nexus position: *Det har **nok** vært for tidlig.*

ADVqnt: quantity adverb, postposed in a nominal phrase and expressing an additional quantity: *Jeg fikk et eple **til**.*

ADVqntneg: quantifier negation adverb, attaching to a quantifier and negating it: ***Ikke** alle klarer dette.*

ADVroot: root adverb, an adverb capable of forming an independent statement, otherwise typically occurring sentence initially, separate from the clause structure: ***Ja**, det stemmer.*

ADVs: sentence adverb, an adverb typically occurring in nexus position, but also in verb phrase position, modifying the proposition:

*Han måtte **ofte** reise; Han måtte reise **ofte**.*

ADVsroot: adverb occurring before the topic position but still within the clause structure (unlike ADVroot), and hence seemingly in violation of the verb-second constraint (mostly limited to 'kanskje'): ***Kanskje** han kommer.*

ADVtmp: temporal adverb limited to certain positions not accessible to ADVloc in general, e.g. between the finite verb and the subject in main clauses: *I tillegg hadde **nylig** Stortinget en debatt om temaet.*

Search is limited to Bokmål texts.

RECOMMENDATION: search only in non-fragmented analyses ('fragments' = 'none').

3. *at*-setninger

Til innholdsfortegnelsen

AT-partofcadv(@lang)

***at* as part of adverbial complementizers**

This template should be given either **nob**, **nno** or **nob|nno** as parameter, depending on whether search should target Bokmål, Nynorsk or both.

The template lists the frequencies of the adverbial multiword complementizers **fordi at**, **tross at**, **slik at** and **sånn at**.

AT-selectedprep(@lang)

***at*-clauses governed by selected prepositions**

This template should be given either **nob**, **nno** or **nob|nno** as parameter, depending on whether search should target Bokmål, Nynorsk or both.

The template finds the examples of *at*-clauses governed by selected prepositions, i.e., prepositions selected by preceding adjectives, nouns or verbs. In the output the prepositions are listed, and also the predicates selecting them, with adjectives, nouns and verbs in separate columns. The output can be sorted by clicking the heads of the columns, either according to frequency, preposition, adjective, noun or verb.

AT-semprep(@lang)

at-clauses governed by semantic prepositions

This template should be given either **nob**, **nno** or **nob|nno** as parameter, depending on whether search should target Bokmål, Nynorsk or both.

The template finds the examples of *at*-clauses governed by semantic (regular) prepositions, e.g. **etter at**, **uten at** etc. (Unfortunately, due to shortcomings of the grammar some of the examples of **på at** belong under the template **AT-selectedprep(@lang)**, covering selected prepositions.)

AT-verbwithandwithout(@verb)

Complement clauses of a verb with and without *at*

Finds all nominal complement clauses of the verb @verb and sorts them according to the presence or absence of the complementizer *at*.

4. Idiommer

Til innholdsfortegnelsen

IDIOM-v-obj-p(@V,@OBJ,@P)

Verb-object-preposition idioms

Finds sentences with the specified combination of verb plus object noun plus preposition or adverb (whether they are idioms or not). Example: with **@V** = 'kaste', **@OBJ** = 'øye' and **@P** = 'på' we get examples like *Han kastet ofte sine øyne på henne* etc. The template both finds ordinary compositional examples and examples analyzed as idioms in NorGram.

The latter will appear in a separate column **#idiom**. It is possible to enter alternative parameter values, e.g. **kaste|få** as the value of **@V**.

IDIOM-v-obj(@V,@OBJ)

Verb-object idioms

Finds sentences with the specified combination of verb plus object noun (actually whether they are idioms or not). Example:
with **@V** = *ta*, **@OBJ** = *kveld* we get examples like *Han tok aldri kvelden* etc. The template both finds ordinary compositional examples and examples analyzed as idioms in NorGram. The latter will appear in a separate column **#idiom**. It is possible to enter alternative parameter values, e.g. **fri|kveld** (separated by a vertical bar) as the value of **@OBJ**.

5. Substantivtemplater

[Til innholdsfortegnelsen](#)

N-adjmod(@N)

The adjectives modifying a noun

Lists, with frequencies, all adjectives modifying the noun **@N**, attributively or predicatively.

N-argofpreps(@N)

Prepositions of which a noun is an object

Finds, with frequencies, the prepositions, semantic as well as selected, of which the noun @N is an object. In addition the predicate modified by or governing the preposition is listed.

N-argofverbs(@N)

Verbs of which a noun is argument 1 or argument 2

This template finds, with frequencies, all verbs and verbal expressions of which the noun @N occurs as either ARG1 or ARG2, sorting the verbs in two columns #arg1of and #arg2of. (The variables are actually called **#Aarg2of**, **#Bnoun** and **#Carg1of** in order to get the columns in an intuitive order, placing the noun @N after the verb of which it is ARG2 and before the verb of which it is ARG1.)

N-argofverbs-witharg2(@N)

Verbs of which a noun is argument 1 (plus the argument 2 of the verb, if any) or of which it is argument 2

This template finds, with frequencies, all verbs and verbal expressions of which the noun @N occurs as either ARG1 or ARG2, sorting the verbs in two columns #arg1of and #arg2of. (The variables are actually called **#Aarg2of**, **#Bnoun** and **#Carg1of** in order to get the columns in an intuitive order, placing the noun @N after the verb of which it is ARG2 and before the verb of which it is ARG1.) In addition, whenever the verb of which @N is ARG1 is transitive (or has a predicative complement), the ARG2 of this verb is listed under the variable **#Darg2**. Thus, searching with

the parameter **kvinne**, in an example like *kvinnen så passasjeren*, the verb *se* would be **#Carg1of** and *passasjer* would be **#Darg2**.

N-defmascofem(@N)

Feminine vs. masculine inflection of a noun

Sorts all occurrences in the definite singular of the noun @N according to inflection as masculine (+) or feminine (-). Search is limited to Bokmål texts.

N-forms(@N)

The forms of a noun

Gives the frequencies of the forms indef.sg. def.sg., indef.pl. and def.pl. of the noun @N.

It is possible to specify multiple nouns, which must be separated by a vertical bar, e.g. **kvinne|dame|kone** .

The form specified in the @N field is actually the predicate value (PRED) of the noun, which may be the same across alternative stem forms. Thus **grøt** will yield forms of both **grøt** and **graut**, while **graut** alone gives no hits (not being a predicate value in the lexicon). Since the user cannot know when this is the case, it is recommended to specify alternative stem forms when in doubt, and when relevant, e.g. **grøt|graut**. This will give the same result as **grøt** alone in this particular case.

N-genders(@N)

Genders of a noun

Lists, with frequencies, the genders of the noun @N in the treebank. There are three columns #masc, #fem and #neut with the values + and - listed. Only gender possibilities mentioned in the lexicon are registered. Only occurrences with syntactically or morphologically unambiguous gender are registered as either masc, fem or neut. This means only singular forms, either definite, or if indefinite, specified by a gender-inflected determiner. A fourth column, #wtotal, gives the total frequency of the noun @N, irrespective of unambiguous gender or not.

Search is limited to Bokmål texts.

N-mwe(@N)

Multi-word expressions with a noun

Lists, with frequencies, types of multi-word expressions (MWEs) with

the noun @N, typically expressions with selected prepositions.

NB! Dette er begrenset til flerordsuttrykk som er registrert som flerordsuttrykk i leksikon.

N-possof(@N)

Nouns of which a noun is a possessive

Lists, with frequencies, the nouns of which the noun @N functions as a possessive. Both s-possessives (*guttens sykkel*), 'garp' possessives (*gutten sin sykkel*) and prepositional possessives (*sykkelen til gutten*) are included.

N-possposition(@N)

Possessives before or after a noun

Sorts, with frequencies, occurrences of the noun @N when occurring with possessive determiners, according to the placement of the possessive before or after the noun. (This distribution tends to be correlated with semantic properties of the noun, as well as with style.) Examples with **egen/eigen** between possessive and noun (*hans egen sykkel*) are disregarded since they uniformly demand pre-position of the possessive.

Because of the syntactic analysis of the possessives, in which only postposed possessives are dominated by NP, the postposed cases are marked **NP** in the output, and the preposed cases are marked **POSS**.

N-SmplOrDbIDef(@N)

Simple or double definiteness with a noun

Sorts, with frequencies, the occurrences of the noun @N in determiner phrases (DPs) with *den* or *denne* according to definite or indefinite inflection of the noun. The search is limited to Bokmål texts.

N-verbderived-noun-cmpds(@VRBSTEM,@LANG)

Nouns derived from a verb stem, including compounds

Lists with frequencies the lemma forms of all ing-, ning-, nad- and else-derived nouns from the verb stem (or verb stems) @VRBSTEM, and also gives the frequency of the verb lemma @VRBLEMMA. Occurrences of the derived nouns in compounds are listed separately, in one column for each suffix. To exclude compounds, use the template **N-verbderived-noun(@VRBSTEM,@VRBLEMMA,@LANG)**

@VRBSTEM is normally like the infinitive of the verb with the possible final unstressed -e removed. It is possible to enter more than one verb stem, separated by a vertical bar: **les|spis|forstå**. If the verb stem ends in a double consonant which will be simplified before an ending beginning with a consonant (e.g., **bygg** > **bygning**), the simplified stem has to be entered as an alternative: **bygg|byg**. Other possible verb stem variants must also be taken into account, e.g. **unntak|unntag** in addition to **unnta** (to account for the nouns *unntakelse* and *unntagelse*).

@VRBLEMMA is the citation form – the infinitive – of the verb, as a basis for search for verb forms. Finite, infinitival and supine forms of the verb lemma are counted.

@LANG should be either **nob** (Bokmål), **nno** (Nynorsk) or **nob|nno** (both).

N-verbderived-noun(@VRBSTEM,@LANG)

Nouns derived from a verb stem

Lists with frequencies the lemma forms of all ing-, ning-, nad- and else-derived nouns from the verb stem (or verb stems) @VRBSTEM, and also gives the frequency of the verb lemma @VRBLEMMA.

@VRBSTEM is normally like the infinitive of the verb with the possible final unstressed -e removed. It is possible to enter more than one verb stem, separated by a vertical bar: **les|spis|forstå**. If the verb stem ends in a double consonant which will be simplified before an ending beginning with a consonant (e.g., **bygg** > **bygning**), the simplified stem has to be entered as an alternative: **bygg|byg**. Other possible verb stem variants must also be taken into account, e.g. **unntak|unntag** in addition to **unnta** (to account for the nouns *unntakelse* and *unntagelse*). Only simplex occurrences of the derived nouns are counted, not occurrences in compounds.

@VRBLEMMA is the citation form – the infinitive – of the verb, as a basis for search for verb forms. Finite, infinitival and supine forms of the verb lemma are counted.

@LANG should be either **nob** (Bokmål), **nno** (Nynorsk) or **nob|nno** (both).

6. Preposisjonstemplater

Til innholdsfortegnelsen

P-prepgovernedby(@PREP,@LANG)

Verbal and other predicates governing a preposition

Lists, with frequencies, the verbal, adverbial and nominal predicates governing the preposition **@PREP**, whether **@PREP** is a selected or a semantic preposition. If it is selected, the predicate will appear in the column **#governing2** (and the predicate name will contain **@PREP** after an asterisk), if it is semantic, the predicate name will appear in the column **#governing1**.

Only uses of **@PREP** as a preposition are included, not possible uses as adverb or particle.

@LANG limits search to Bokmål or Nynorsk – or allows both – and can take the values **nob**, **nno**, or (if both) the disjunction **nob|nno**.

It is also possible to search for more than one preposition at the same time by use of disjunction, e.g. **fra|ifra|frå|ifrå**

P-prepobjpred(@PREP)

Object predicates of a preposition

Sorts, with frequencies, occurrences of the preposition **@PREP** according to the predicate expressed by the object of the preposition. In occurrences where the preposition is semantic, they are listed under **#governed**. In occurrences where the preposition is selected, the objects are listed under **#oblth** ('oblique theta', the function of a prepositional phrase selected by a verb).

The hits are also sorted according to language (**nno** or **nob**).

P-prepobjtypes(@PREP)

Types of objects of a preposition

Sorts, with frequencies, occurrences of the preposition **@PREP** according to the object of the preposition – a common noun (**common**), a proper noun (**proper**), a pronoun (**pronoun**), a web address (**uri**), a finite clause (**fin**), an infinitive (**inf**), etc. In occurrences where the preposition is semantic, they are listed under **#governed**. In occurrences where the preposition is selected, the objects are listed under **#oblth** ('oblique theta', the function of a prepositional phrase selected by a verb).

The hits are also sorted according to language (**nno** or **nob**).

7. Pronomentemplater

Til innholdsfortegnelsen

PRON-objoroblben(@V)

Personal pronouns as OBJ-BEN or OBL-BEN (indirect object or oblique)

Finds pronouns which function as either OBJ-BEN (indirect object) or OBL-BEN (benefactive oblique, i.e., a *til*-phrase) of the verb @V, sorting them

according to function. The output lists the pronouns in a column headed **#form**, and their function expressed as the presence or absence of **til** in the column headed **#oblben**, with frequencies. More than one verb can be entered as the value of the parameter @V by separating the verbs with a vertical bar |.

Search is limited to Bokmål texts.

8. Syntakstemplater

Til innholdsfortegnelsen

SYNT-fillergap(@FILLER,@PATH,@GAPFN)

Filler-gap (long-distance dependency) constructions

Finds examples of long-distance dependencies as specified by the parameters, which allow selection of filler type (**@FILLER**), path between filler and gap (**@PATH**) and gap function (**@GAPFN**).

@FILLER can have the values **TOPIC**, **TOPIC-REL** and **FOCUS-INT**. **TOPIC** limits the search to topic constructions, i.e. clauses where the topic is the constituent in initial position (SPEC, IP) binding a gap further down. **TOPIC-REL** limits the search to relative clauses, where **TOPIC-REL** is the possibly empty anaphoric element in the relative clause binding the gap further down. **FOCUS-INT** limits the search to direct and indirect wh-interrogatives, where **FOCUS-INT** is the wh-expression binding a gap further down. In the LFG-analysis filler-gap binding is analyzed with unification, which means that the **@FILLER** and the **@GAPFN** will have the same value in the structure.

@PATH specifies the hierarchy of syntactic functions separating the filler from the gap by means of a sequence of zero, one or more function names. Relevant function names are **COMP**, **XCOMP**, **PREDLINK**, **ADJUNCT** \$ and in certain

combinations **OBJ . COMP** is a closed verbal complement, i.e. a finite subclause or an infinitive construction where the subject is not controlled from the outside. **XCOMP** is an open verbal complement, i.e. an infinitive or participle construction with a controlled subject. This covers complements of auxiliaries as well as complements of control verbs like *prøve*, *like*, etc. **PREDLINK** is a predicative complement, e.g. a complement of a copula verb like *være*. **ADJUNCT \$** is an adverbial, e.g. a prepositional phrase. (The reason why **\$** must be added is that the value of **ADJUNCT** is a set, and **\$** is an operator picking out a member of the set (actually **€**).) **OBJ** can be a member of the path in cases where an object noun has a verbal complement, as in *ambisjoner om å ...*, *lyst til å ...*, *inntrykk av at ...*, where a gap may be located, as in *Det har jeg inntrykk av at kommunene gjør*. The value of **@PATH**, then, can be sequences like **XCOMP COMP**, **COMP ADJUNCT \$** or **OBJ COMP** (where **OBJ** is necessarily followed by **COMP**). It is also possible to specify repetitions of a member of the path by means of Kleene-star ***** (zero, one or more occurrences) or Kleene-plus **+** (one or more occurrences). The element to be repeated must be enclosed in parentheses, e.g.: **(XCOMP)* OBJ COMP**. This path specification will allow one or more auxiliaries before the main verb taking the **OBJ**, as in: *Dette vil jeg ta initiativ til å videreutvikle*. The path can also be empty – i.e., filler and gap are on the same syntactic level – but since parameters cannot be completely unspecified it is then necessary to enter a single space.

@GAPFN specifies the syntactic function of the gap bound by the filler, and can have values such as **SUBJ**, **OBJ**, **OBJ-BEN**, **OBL-TH** and **PREDLINK**. **OBJ-BEN** is indirect object. **OBL-TH** ('oblique-theta') denotes the function of objects of selected prepositions (examples like *Hva tenker du på?*). **PREDLINK** is predicative complement.

NB! The automatic analysis of filler-gap constructions, in particular those with longer paths, is complex and error-prone. **It is therefore strongly recommended only to search in non-fragmented analyses (check 'fragments' = 'none' before searching)**. Even so, the output should be checked. Automatic correct separation between direct and indirect objects (**OBJ** and **OBJ-BEN**) in filler-gap constructions is especially difficult.

Example: The following choice of parameters:

@FILLER: TOPIC
 @PATH: COMP COMP
 @GAPFN: OBJ

finds 141 examples, among them the following:

1. *Den form for samhandling tror jeg det er veldig bra at vi nå gjennomfører.*
2. *Det mener jeg det er viktig at en greier i politikken.*
3. *Det er det ikke alle kommuner som ser at dei greier, så derfor er det få kommunesamanslåingar.*
4. *Det tror jeg faktisk det er mulig å forklare.*
5. *Men det synes jeg faktisk det er nødvendig at statsråden klargjør.*

In 1. the TOPIC *den form for samhandling* functions as OBJ of *gjennomfører*, and in between there are two COMP boundaries, one after *tror jeg* and one after *veldig bra*. Both COMPs are finite clauses. In 4. the TOPIC *det* is OBJ of *forklare*, and the COMP boundaries are after *faktisk* and *mulig*. The latter COMP is infinitival, but has a non-controlled SUBJ with arbitrary reference, and is therefore not an XCOMP.

SYNT-fillergap-vform(@FILLER, @PATH, @GAPFN, @LASTVFORM)

Filler-gap (long-distance dependency) constructions with specified verb form

This template is an extended and modified version of **SYNT-fillergap(@FILLER, @PATH, @GAPFN)**, whose definition is copied in below. The modifications are intended to reduce the amount of erroneous examples in the output, and also to enable limiting the search according to the last verb form before the gap. The last-mentioned modification involves the added parameter **@LASTVFORM**:

@LASTVFORM specifies the form of the last verb governing the gap in the construction. This is desirable because a long-distance @PATH ending in **COMP** will include both finite clauses and non-nontrolled infinitives, which both have the function COMP. Examples: *Dette er det viktig at du*

forstår, Dette er det viktig å forstå. **XCOMP** as end of path can also have more than one verb form – both infinitive and supine. **@LASTVFORM** can have the values:

fin (finite verb)

inf (infinitival)

sup (supine, i.e. uninflected past participle)

If more than one of these are to be included, they can be entered separated by a vertical bar: **fin|inf|sup**.

A further modification reduces the number of erroneous analyses: Fillers which can be either adverbials or verb arguments are excluded. Example: **det året** in a sentence like *Det året tror jeg han nesten ikke spiste*. Such examples are frequently syntactically ambiguous between two readings (*det året* as a temporal adverbial or as object of *spiste*), and as TOPICS they are nearly always adverbials, whereas the parser often chooses the argument reading. Hence excluding them avoids many undesirable hits.

Here is the description of SYNT-fillergap(@FILLER, @PATH, @GAP):

Finds examples of long-distance dependencies as specified by the parameters, which allow selection of filler type (**@FILLER**), path between filler and gap (**@PATH**) and gap function (**@GAPFN**).

@FILLER can have the values **TOPIC**, **TOPIC-REL** and **FOCUS-INT**. **TOPIC** limits the search to topic constructions, i.e. clauses where the topic is the constituent in initial position (SPEC, IP) binding a gap further down. **TOPIC-REL** limits the search to relative clauses, where **TOPIC-REL** is the possibly empty anaphoric element in the relative clause binding the gap further down. **FOCUS-INT** limits the search to direct and indirect wh-interrogatives, where **FOCUS-INT** is the wh-expression binding a gap further down. In the LFG-analysis filler-gap binding is analyzed with unification, which means that the **@FILLER** and the **@GAPFN** will have the same value in the structure.

@PATH specifies the hierarchy of syntactic functions separating the filler from the gap by means of a sequence of zero, one or more function names. Relevant function names are **COMP**, **XCOMP**, **PREDLINK**, **ADJUNCT** \$ and in certain

combinations **OBJ. COMP** is a closed verbal complement, i.e. a finite subclause or an infinitive construction where the subject is not controlled from the outside. **XCOMP** is an open verbal complement, i.e. an infinitive or participle construction with a controlled subject. This covers complements of auxiliaries as well as complements of control verbs like *prøve*, *like*, etc. **PREDLINK** is a predicative complement, e.g. a complement of a copula verb like *være*. **ADJUNCT \$** is an adverbial, e.g. a prepositional phrase. (The reason why **\$** must be added is that the value of **ADJUNCT** is a set, and **\$** is an operator picking out a member of the set (actually **€**).) **OBJ** can be a member of the path in cases where an object noun has a verbal complement, as in *ambisjoner om å ..., lyst til å ..., inntrykk av at ...*, where a gap may be located, as in *Det har jeg inntrykk av at kommunene gjør*. The value of **@PATH**, then, can be sequences like **XCOMP COMP**, **COMP ADJUNCT \$** or **OBJ COMP** (where **OBJ** is necessarily followed by **COMP**). It is also possible to specify repetitions of a member of the path by means of Kleene-star ***** (zero, one or more occurrences) or Kleene-plus **+** (one or more occurrences). The element to be repeated must be enclosed in parentheses, e.g.: **(XCOMP)* OBJ COMP**. This path specification will allow one or more auxiliaries before the main verb taking the **OBJ**, as in: *Dette vil jeg ta initiativ til å videreutvikle*. The path can also be empty – i.e., filler and gap are on the same syntactic level – but since parameters cannot be completely unspecified it is then necessary to enter a single space.

@GAPFN specifies the syntactic function of the gap bound by the filler, and can have values such as **SUBJ**, **OBJ**, **OBJ-BEN**, **OBL-TH** and **PREDLINK**. **OBJ-BEN** is indirect object. **OBL-TH** ('oblique-theta') denotes the function of objects of selected prepositions (examples like *Hva tenker du på?*). **PREDLINK** is predicative complement.

NB! The automatic analysis of filler-gap constructions, in particular those with longer paths, is complex and error-prone. **It is therefore strongly recommended only to search in non-fragmented analyses (check 'fragments' = 'none' before searching)**. Even so, the output should be checked. Automatic correct separation between direct and indirect objects (**OBJ** and **OBJ-BEN**) in filler-gap constructions is especially difficult.

Example: The following choice of parameters:

@FILLER: TOPIC
 @PATH: COMP COMP
 @GAPFN: OBJ
 @LASTVFORM: fin

finds the following examples, among many others:

1. *Den form for samhandling tror jeg det er veldig bra at vi nå gjennomfører.*
2. *Det mener jeg det er viktig at en greier i politikken.*
3. *Det er det ikke alle kommunar som ser at dei greier, så derfor er det få kommunesamanslåingar.*
4. *Men det synes jeg faktisk det er nødvendig at statsråden klargjør.*

In 1. the TOPIC *den form for samhandling* functions as OBJ of *gjennomfører*, and in between there are two COMP boundaries, one after *tror jeg* and one after *veldig bra*. Both COMPs are finite clauses.

SYNT-possposition-adjunct(@ADJTYPE)

Possessives before or after a noun sorted by adjunct types in the NP

Sorts, with frequencies, occurrences of NPs with possessive determiners in addition to an adjunct of type **@ADJTYPE**, according to the placement of the possessive before or after the noun.

Because of the syntactic analysis of the possessives, in which only postposed possessives are dominated by NP, the postposed cases are marked **NP** in the output, and the preposed cases are marked **POSS**.

The values of the parameter **@ADJTYPE** may be:

ATYPE = adjective

PTYPE = prepositional phrase (This option is error prone since PPs belonging to the sentence may have been wrongly analysed as belonging to the NP)

TOPIC-REL = relative clause (finite or infinitival) (This option is also somewhat error prone)

It is possible to search for several of these at once by separating the alternatives with a vertical bar |, **adding a space before and after the vertical bar**:

ATYPE | PTYPE

Search is limited to Bokmål texts.

SYNT-possposition-det(@DETERM)

Possessives before or after a noun sorted by determiner types in the NP

Sorts, with frequencies, occurrences of NPs with possessive determiners in addition to a determiner of type **@DETERM**, according to the placement of the possessive before or after the noun.

Because of the syntactic analysis of the possessives, in which only postposed possessives are dominated by NP, the postposed cases are marked **NP** in the output, and the preposed cases are marked **POSS**.

The values of the parameter **@DETERM** may be:

QUANT = quantifiers like **alle**, **hele**, **noen** etc.

AQUANT = quantifiers like **mange** (which can be inflected for degree, e.g. **flere**, **fleste**) etc.

DET = demonstratives like **den**, **denne** etc.

ADET = 'adjective-like' determiners like **annen**, **egen** etc.

It is possible to search for several of these at once by separating the alternatives with a vertical bar |, **adding a space before and after the vertical bar**:

QUANT | AQUANT

Search is limited to Bokmål texts.

SYNT-possposition-fn(@FN)

Possessives before or after a noun sorted by NP function

Sorts, with frequencies, occurrences of NPs with possessive determiners filling the syntactic function **@FN**, according to the placement of the possessive before or after the noun.

Because of the syntactic analysis of the possessives, in which only postposed possessives are dominated by NP, the postposed cases are marked **NP** in the output, and the preposed cases are marked **POSS**.

The values of the parameter @FN may be:

SUBJ = subject

OBJ = object (of a verb or a preposition)

OBL-TH = object of a selected preposition

OBJ-BEN = indirect object

PREDLINK = predicative complement

It is possible to search for several of these at once by separating the alternatives with a vertical bar |, **adding a space before and after the vertical bar**:

OBJ | OBJ-BEN

Search is limited to Bokmål texts.

9. Verbtemplater

Til innholdsfortegnelsen

V-allprepsrtsadv(@V)

Prepositions and adverbs governed by a verb

Lists, with frequencies, all prepositions, particles and adverbs, non-selected as well as selected, that are governed by the verb @V. Adverbs are restricted to manner adverbs, including adverbs of direction (**inn**, **ut** etc.). The intention of the template is to gather all such governed items irrespective of the grammar's possible analysis of them as selected, in order to provide material for the unbiased manual identification of selected items and other collocations. To search specifically for items described as selected by the grammar and lexicon, the template **V-mwe(@V)** can be used.

Search with this template may be slow for frequent verbs. In order to search for combinations of a specific verb and specific prepositions, particles or adverbs, the templates

V-selorsemprep(@V,@PREP) (prepositions) or

V-prtoradv(@V,@PRT) (particles or adverbs) should be used.

It should be borne in mind that PP attachment is perhaps the most notorious source of syntactic ambiguity there is. Hence there will inevitably be a number of analyses where a prepositional phrase has been wrongly attached to the verb @V.

V-arg1(@V)

Agents of a verb

Lists, with frequencies, the predicate values of all nouns, verbs etc. functioning as the first argument (the 'actor') of the verb @V, i.e., as

subject in the active, oblique agent in the passive, or in other constructions.

V-arg2(@V)

Patients of a verb

Lists, with frequencies, the predicate values of all nouns, verbs etc. functioning as the second argument (the ‘patient’) of the verb @V, i.e., as object in the active, subject in the passive, or in other constructions.

V-argframes-incl-adj-pro(@V)

Argument frames of a verb, including predicative adjectives and *pro*-arguments

Lists, with frequencies, all argument frames (valency frames) of the verb @V by means of columns headed ‘arg1’, ‘arg2’ and ‘arg3’. The frames are further sorted within the columns according to whether the argument is:

- a common noun (**common**)
- a proper noun (**proper**)
- a pronoun (**pronoun**)
- an infinitival (**inf**)
- a supine (**sup**)
- a past participle (**pastpart**)
- a present participle (**prespart**)
- a finite clause (**fin**)
- a predicative adjective (**predicative**)
- a web address (**uri**)
- or a *pro* argument (**pro**), which includes unexpressed subjects of non-controlled infinitives, ‘gap’ arguments in relative clauses, and arguments expressed by nominal phrases without a head noun, like *noen*, *mange*, *de gamle*, etc.

Verbal expressions with @V, i.e. @V with selected prepositions, with selected particles, or in idioms, are listed as separate predicates in the

column '#a'. The search is limited to Bokmål texts.

*The search may take several minutes with frequent verbs. The templates **V-argframes-incl-adj(@V)**, not including pro arguments, and **V-argframes(@V)**, neither including pro arguments nor adjective arguments, take less time.*

RECOMMENDATION: search only in non-fragmented analyses ('fragments' = 'none').

V-argframes-incl-adj(@V)

Argument frames of a verb, including predicative adjectives

Lists, with frequencies, all argument frames (valency frames) of the verb @V by means of columns headed 'arg1', 'arg2' and 'arg3'. The frames are further sorted within the columns according to whether the argument is:

- a common noun (**common**)
- a proper noun (**proper**)
- a pronoun (**pronoun**)
- an infinitival (**inf**)
- a supine (**sup**)
- a past participle (**pastpart**)
- a present participle (**prespart**)
- a finite clause (**fin**)
- a predicative adjective (**predicative**)
- or a web address (**uri**).

Verbal expressions with @V, i.e. @V with selected prepositions, with selected particles, or in idioms, are listed as separate predicates in the column '#a'. The search is limited to Bokmål texts.

*The search may take several minutes with frequent verbs. The template **V-argframes(@V)**, not including predicative adjectives, takes less time.*

RECOMMENDATION: search only in non-fragmented analyses ('fragments' = 'none').

V-argframes(@V)

Argument frames of a verb

Lists, with frequencies, all argument frames (valency frames) of the verb @V by means of columns headed 'arg1', 'arg2' and 'arg3'. The frames are further sorted within the columns according to whether the argument is:

a common noun (**common**) a
 proper noun (**proper**)
 a pronoun (**pronoun**) an
 infinitival (**inf**)
 a supine (**sup**)
 a past participle (**pastpart**)
 a present participle (**prespart**) a
 finite clause (**fin**)
 or a web address (**uri**).

Verbal expressions with @V, i.e. @V with selected prepositions, with selected particles, or in idioms, are listed as separate predicates in the column '#a'. The search is limited to Bokmål texts.

The search may take several minutes with frequent verbs.

RECOMMENDATION: search only in non-fragmented analyses ('fragments' = 'none').

V-attr-or-pred-ptc(@V)

Attributive or predicative/main verb function of a participle

Sorts occurrences of the past participle or supine form of the verb @V (**to be entered in its infinitive form**) according to whether the participle has attributive or predicative/main verb function – in the latter case after the verbs **være** or **bli**. Both singular and plural examples are included. The

attributive examples are tagged **attributive** under #type, the predicative or main verb examples are tagged **main** under #type. Search is limited to Bokmål texts.

V-forms(@V,@FORM)

The a verbo forms of a verb

Lists, with frequencies, the word forms realizing the form @FORM of the verb @V. @FORM may have the following values:

inf = infinitive;

pres = present tense;

past = past tense;

sup = supine (uninflected past participle);

It is possible to specify multiple values, which must be separated by a vertical bar, e.g. **pres|past|sup** . This is also possible in the @V field, e.g. **kaste|kreve|bite** .

The form specified in the @V field is actually the predicate value (PRED) of the verb, which may be the same across alternative stem forms. Thus **lyve** will yield forms of both **lyve**, **lyge** and **ljuge**, while **lyge** alone gives no hits. Since the user cannot know when this is the case, it is recommended to specify alternative stem forms when in doubt, and when relevant, e.g. **lyve|lyge|ljuge** . This will give the same result as **lyve** alone in this particular case. The search is limited to Bokmål texts.

V-maleorfemale-arg2(@V)

Patients of a verb with male vs. female agents

Sorts, with frequencies, relevant occurrences of the verb @V plus its

arg2 (the patient, i.e. the object of active and the subject of passive) according to the biological gender of arg1 (the agent, i.e., subject of active, oblique agent of passive, head of present participle (*lesende kvinne*), etc.) Male agents are marked with + in the column **#masc**, female agents are unmarked. The template **V-maleorfemale(@V)** below does the same, but disregards the patient.

Male agents are delimited by 4556 male names, the pronoun *han*, and the nouns *mann, herre, gutt, gut, far, bror, onkel, fetter, farfar, morfar, bestefar*. Female agents are delimited by 3582 female names, the pronoun *hun* or *ho*, and the nouns *kvinne, dame, frue, jente, mor, søster, tante, kusine, farmor, mormor, bestemor*.

Poenget med dette templatet, sammenlignet med det neste, er at det i tillegg til å vise hvilke verb som særlig har mannlige eller kvinnelige handlere, også viser eventuelle forskjeller mellom hva slags objekter (eller arg2) verbene har med mannlige vs. kvinnelige handlere. En ting er hvor ofte kvinner vs. menn 'slår', en annen ting er *hva de typisk slår*.

V-maleorfemale(@V)

Male vs. female agents of a verb

Sorts, with frequencies, relevant occurrences of the verb @V according to the biological gender of arg1 (the agent, i.e., subject of active, oblique agent of passive, head of present participle (*lesende kvinne*), etc.) Male agents are marked with + in the column **#masc**, female agents are unmarked.

Male agents are delimited by 4556 male names, the pronoun *han*, and the nouns *mann, herre, gutt, gut, far, bror, onkel, fetter, farfar, morfar, bestefar*. Female agents are delimited by 3582 female names, the pronoun *hun* or *ho*, and the nouns *kvinne, dame, frue, jente, mor, søster, tante, kusine, farmor, mormor, bestemor*.

V-morphpassforms-nno(@V)

Forms of the morphological passive of a verb, Nynorsk

Sorts, with frequencies, the occurrences of the morphological passive (*st*-passive) of the verb @V according to the form of the verb: **pres**, **past** or **inf**. The template only searches Nynorsk texts, i.e., the nno treebanks and the Nynorsk parts of the nor treebanks. For Bokmål, use **V-morphpassforms-nob(@V)**

V-morphpassforms-nob(@V)

Forms of the morphological passive of a verb, Bokmål

Sorts, with frequencies, the occurrences of the morphological passive (*s*-passive) of the verb @V according to the form of the verb: **pres**, **past** or **inf**. The template only searches Bokmål texts, i.e., the nob treebanks and the Bokmål parts of the nor treebanks. For Nynorsk, use **V-morphpassforms-nno(@V)**

V-mwe(@V)

Multi-word expressions with a verb

Lists, with frequencies, types of multi-word expressions (MWEs) with the verb @V, typically expressions with selected prepositions or selected particles, or verb phrase idioms with selected content words.

V-objshift(@V)

Shifted and non-shifted object pronoun *det* with negation

Sorts sentences with the verb @V, a negative adverb, and the pronoun *det* as direct object, according to the placement of the pronoun before (shifted) or after (non-shifted) the negative adverb.

V-passtypes-nno(@V)

Morphological and periphrastic passives of a verb, Nynorsk

Sorts, with frequencies, the occurrences of morphological (-st) and periphrastic (*bli* and *verte*) passive forms of the verb @V into the categories **morph** and **periph**. This template only searches Nynorsk texts, i.e. the nno-treebanks and the Nynorsk parts of the nor-treebanks. For Bokmål use **V-passtypes-nob(@V)**.

V-passtypes-nob(@V)

Morphological and periphrastic passives of a verb, Bokmål

Sorts, with frequencies, the occurrences of morphological (-s) and periphrastic (*bli* and *være*) passive forms of the verb @V into the categories **morph** and **periph**. This template only searches Bokmål texts, i.e. the nob-treebanks and the Bokmål parts of the nor-treebanks. For Nynorsk use **V-passtypes-nno(@V)**.

V-perf-pres-past(@V)

Present, past and present perfect forms of a verb

Finds, with frequencies, the present, past and present perfect forms of a verb @V. The perfect forms are limited to the auxiliary 'ha'. It is possible to enter sets of verbs as parameter, separated by a vertical bar |, e.g. *spise|drikke'. Search is limited to Bokmål texts.

V-prepobj(@V,@P)

Objects of a preposition governed by a verb

Finds, with frequencies, examples where the verb @V governs an adverbial (non-selected) prepositional phrase with the (semantic) preposition @P, sorted by the object of @P.

V-prtoradv(@V,@PRT)

Particle or adverb governed by a verb

Lists, with frequencies, instances where the particle or adverb @PRT is governed by the verb @V. This template only handles particles like **ut** in *kaste ut* or **på** in *gå på*, and adverbs like **rundt** in *løpe rundt*, **not** selected prepositions like **på** in *tenke på ferien* or semantic prepositions like **på** in *sitte på bussen*. For the latter, see the template **V-selorsemprep(@V,@PREP)**.

The main purpose of this template is to get around the problem that multiword expressions (mwes) like verb-particle constructions such as '**kaste*ut**' have not always been recognized consistently in parsing, some of the cases receiving compositional analyses as. verb + adverb. This template gathers all examples with the verb @V and the particle or adverb @PRT, allowing the user to evaluate to what extent compositional analyses are motivated. In the output the mwe-analyses are listed in a separate column **#mwe**.

There are also cases where the parser lexicon does not list a certain combination as an mwe at all; then this template allows the user to inspect the frequencies of the combinations and evaluate whether they should be recognized as mwes.

In order to get a listing of all such combinations for a given verb, including selected prepositions, that are governed by a given verb, without limiting it

to a specific particle or adverb, the template **V-allprepsrtsadvs(@V)** should be used.

V-ptc-or-sup-pred(@V)

Participles of a verb inflected or non-inflected for plural agreement

Sorts, with frequencies, perfect participles of the verb @V (**which must be entered in its infinitive form**) when occurring after **være** or **bli** and a plural subject, according to whether the participle is inflected for plural agreement. The inflected examples are tagged **pastpart**, while the uninflected examples are tagged **sup** (for ‘supine’). The search is limited to Bokmål; Nynorsk participles are analyzed differently.

In the treebank the inflected examples (**pastpart**) are analyzed as adjectives and the governing verbs (**være** or **bli**) are analyzed as copula verbs. (Example: *Vi er skilte*.) The predicate forms in the output are then simply **være** and **bli**. The uninflected examples (**sup**) are usually analyzed as verbal complements to a passive auxiliary. (Example: *Vi er skilt*.) The predicate forms in the output are then **være-pass** and **bli-pass**. But then there are examples with uninflected participle (**sup**) which are nevertheless analyzed as adjectives after a copula (**være** or **bli + sup**). These are examples where the participle is analyzed as an adjective for syntactic reasons, in spite of its lack of agreement. The syntactic reason may be a degree adverb – example: *Leppene var lett skilt*.

V-selorsemprep(@V,@PREP)

Selected or semantic preposition governed by a verb

Lists, with frequencies, instances where the selected or semantic preposition @PREP is governed by the verb @V. This template only handles selected prepositions like **på** in *tenke på ferien* and semantic prepositions like **på** in *sitte på bussen*, **not** particles like **ut** in *kaste ut* or **på** in *gå på*, or adverbs like **rundt** in *løpe rundt*. For the latter, see the template **V-prtoradv(@V,@PRT)**.

The main purpose of this template is to get around the problem that multiword expressions (mwes) like constructions with selected prepositions such as '**tenke på**' have not always been recognized consistently in parsing, some of the cases receiving compositional analyses as. verb + semantic prepositional phrase. This template gathers all examples with the verb **@V** and the selected or semantic preposition **@PREP**, allowing the user to evaluate to what extent compositional analyses are motivated. In the output the mwe-analyses are listed in a separate column **#mwe**.

There are also cases where the parser lexicon does not list a certain combination as an mwe at all; then this template allows the user to inspect the frequencies of the combinations and evaluate whether they should be recognized as mwes.

In order to get a listing of all such combinations for a given verb, including particles and manner adverbs, that are governed by a given verb, without limiting it to a specific preposition, the template **V-allprepsprtsadvs(@V)** should be used.

V-separableverb(@V, @P)

Separable verb with particle

Finds combinations of the verb **@V** and the particle, adverb or preposition **@P**, where **@P** is either prefixed to **@V** (as in **frastjele**) or occurs as a particle, adverb or preposition after **@V** (as in **stjele fra**). The hits are sorted by four properties listed in four columns:

#lemma (with values like **stjele**, **stjele*fra** and **frastjele**)

#vform (with values like **inf**, **fin**, **sup** and **pastpart**)

#y (the particle, adverb or preposition whenever separate from the verb, with values like **fra**)

language (with values like **nob** (Bokmål) and **nno** (Nynorsk))

Not all complex predicates with the verb **@V** are included. For example, with **@V = sette** and **@P = til**, predicates with other elements than **til** are not included, such as **sette#liv*til** for the idiom *sette livet til*, since they are

not assumed to be relevant for comparison with **tilsette**. The only exception is reflexives. Thus, with @V = **føye** and @P = **til** we find not only the predicates **føye**, **føye*til** and **tilføye**, but also **føye*seg** whenever occurring with **til**.

V-WD-maleorfemale(@V, @WD)

male vs. female agents of a verb plus its complement

Sorts, with frequencies, relevant occurrences of the verb @V with its given arg2 @WD (i.e. the object of active, the subject of passive or a verbal or predicative complement) according to the biological gender of arg1 (the agent, i.e., subject of active, oblique agent of passive, head of present participle (*lesende kvinne*), etc.) Male agents are marked with + in the column **#masc**, female agents are unmarked. The template **V-maleorfemale(@V)** does the same, but disregards the patient. The template **V-maleorfemale-arg2(@V)** in addition lists all the patients of the verb. The present template differs in allowing the specification of a **given** arg2, e.g. @V = *lese* and @WD = *avis*, @V = *være* and @WD = *lydig* or @WD = *direktør*, etc.)

Male agents are delimited by 4556 male names, the pronoun *han*, and the nouns *mann*, *herre*, *gutt*, *gut*, *far*, *bror*, *onkel*, *fetter*, *farfar*, *morfar*, *bestefar*. Female agents are delimited by 3582 female names, the pronoun *hun* or *ho*, and the nouns *kvinne*, *dame*, *frue*, *jente*, *mor*, *søster*, *tante*, *kusine*, *farmor*, *mormor*, *bestemor*.

10. Templater for flere ordklasser

[Til innholdsfortegnelsen](#)

WD-fstel(@WD)

Productive compounds with a word as first element

Lists, with frequencies, all productive (non-lexicalized) compounds with @WD as a non-last element. The compounds found are those which have been analyzed by the compound analyzer, which does not cover compound verbs and adjectives as well as it does compound nouns. @WD should be the lemma (more precisely: the PRED form) of the element.

WD-lemmas&pos-of-form(@WD)

Lemmas and parts of speech of a word form

Lists, with frequencies, the lemmas of which @WD is a form together with the parts of speech of the lemmas.

WD-mwe-containing-wd(@WD)

Multi-word expressions containing a word

Lists with frequencies all multi-word expressions (mwes) which contain the word @WD. The search is limited to mwes that are listed as variable mwes (e.g. inflected or with possible intervening words) in the lexicon – i.e. not invariable mwes listed as ‘words with spaces’. To find the latter type of mwes, use the template **WD-mwe-with-spaces(@WD)**. Bokmål and Nynorsk examples are distinguished by the tags ‘nob’ and ‘nno’.

WD-mwe-with-spaces(@WD)

Invariable multi-word expressions containing a word

Lists, with frequencies, all invariable multi-word expressions (analyzed as ‘words with internal spaces’) which contain the word @WD as a string between spaces (or at the beginning or end). Example: the expression *på*

kryss og *tvers* will be found both with @WD = **kryss** and with @WD = **tvers** (and, with wider results, with @WD = **på** or @WD = **og**).

WD-posoflemma(@WD)

Parts of speech of a lemma form

Lists, with frequencies, the parts of speech ('pos') of which @WD is a lemma form. The listed categories are actually more fine-grained than parts of speech, since what is listed are the pre-terminal categories, i.e., the nodes immediately dominating the words in the c-structure tree. Thus we don't only get **V** for verb, but also e.g. **Vfin** for finite verb, **Vcopfin** for finite copula verb, etc. Generally, the capitalized parts of the categories ('V', 'N' etc.) indicate the simple parts of speech. A clickable overview of the pre-terminal nodes, with explanations, can be found here:

[Preterminals](#)

NB! Examples should be inspected, since homonymy (which is the issue here) often creates difficulties for parsing when more than one syntactic analysis is possible. Hence there will inevitably be a number of wrong analyses. Choose **fragments = none** before searching to reduce the number of errors.

WD-pred-selecting-wd-pos(@WD)

Predicates selecting a word in multiword expressions

Lists, with frequencies, the predicates of multiword expressions in which the word @WD occurs as a selected element, together with the part of speech of @WD. This includes both idioms, in which @WD is a 'semantically heavy' word with its own PRED (e.g. a noun (N) like 'øye' in 'få#øye*på'), and expressions with particles (PRT, e.g. 'på' in 'holde*på') and selected prepositions (Psel-n, Psel-v, e.g. 'på' in 'tenke*på'). It is possible to enter more than one alternative word by separating the alternatives with a vertical bar '|', e.g. **øye|leir**.

WD-prodcmpnds(@WD)

Productive compounds with a word as either element

Lists, with frequencies, productive compounds where the word @WD is one of the elements. Compounds with @WD as a first element are listed under **#fstEl**, compounds with @WD as the last element are listed under **#scndEl**.

@WD may be a noun, a verb or an adjective, but compound analysis has been most consistently carried out for nouns. Hence the results for other parts of speech may be more partial.

WD-wordinit(@WD)

Word forms beginning with a sequence

Lists, with frequencies, all word forms beginning with the letter sequence @WD. No analysis as compound is presupposed. Analyzed compounds are found by the template **WD-fstel**.

WD-wordsequence(@WD1,@WD2) Sequence of two word forms

Gives the frequency of the unanalyzed word form sequence @WD1 + @WD2.

De to siste WD-templatene, som bare søker etter bokstavsekvenser, tillater søk etter enkelte ting som ikke er dekket av grammatikken, f.eks. avvikende genusbruk (*et appelsin*).