

Frå bokmål til nynorsk

Evaluering av maskinomsetjingsverktøy
hausten 2025

1 Innleiing

Språkrådet fører tilsyn med bruken av dei to norske skriftspråka i staten. Språklova set krav om veksling mellom bokmål og nynorsk (§ 13) og krav om parallell bruk av skriftspråka (§ 14). I tillegg krev lova at statsorgan og fylkeskommunar skal ha den skrivekompetansen som trengst for å stetta krava (§ 16).

Ei av oppgåvene Språkrådet har i denne samanhengen, er å gje råd om utvikling, innkjøp og bruk av språkteknologiske produkt. Denne evalueringa er eit ledd i den rådgjevinga.

Dei fleste offentlege verksemder som er omfatta av krava i språklova, nyttar eitt eller fleire digitale verktøy til å omsetja tekstar frå bokmål til nynorsk. I tillegg engasjerer dei språkkonsulentar som brukar slike verktøy. Men det er vanskeleg å finna objektiv eller uavhengig informasjon om kvaliteten på omsetjingane ein får ved hjelp av desse verktøya.

Målet med denne undersøkinga er å fylla noko av kunnskapsholet, slik at dei som har språkansvar i offentlege organ, får eit betre grunnlag for å avgjera

- om organet bør kjøpa inn / bruka eit omsetjingsverktøy
- kva verktøy som høver best
- kva som er forsvarleg bruk av verktøya
- korleis omsetjingane kan kvalitetskontrollerast

I tillegg kan undersøkinga gje utviklarane eit grunnlag for å forbetra verktøya.

Dei seks verktøya som er granska, er nærare omtalte i neste kapittel. Tre språkkonsulentar har evaluert omsetjingane. Dei er omtalte som annotatørar nedanfor.

2 Verktøya som vart testa

Maskinomsetjingsverktøya på marknaden er av to slag:

- verktøy baserte på regelsett
- verktøy baserte på maskinlæring / kunstig intelligens

I samband med [tilsynet for 2024](#) spurde Språkrådet statsorgana kva verktøy dei brukte til å omsetja frå bokmål til nynorsk. Til undersøkinga valde vi ut dei fem mest brukte verktøya pluss eit nyutvikla verktøy som kombinerer regelbasert og KI-basert omsetjing. Dette er dei seks verktøya, med våre forkortingar i parentes etter namnet:

- Apertium (AP), eit regelbasert gratisprogram med open kjeldekode
- Nynorskoversetteren (NNO), eit produkt frå Totaltekst basert på Apertium
- Nynorskroboten (NR), eit produkt frå NTB basert på Apertium
- Nynorskroboten med KI-støtte (NRKI), lansert våren 2025
- Copilot (CO), ein KI-basert samtalerobot i Microsoft 365-pakka, som er allment utbreidd i staten
- ChatGPT (CH), ein KI-basert samtalerobot som finst både i gratisversjon og betalingsversjon

I alle Apertium-programma kan ein fastsetja ei subnorm, det vil seia eit sett med rettskrivingsvariantar som verktøyet held seg konsekvent til. Det er svært vanskeleg eller uråd å få Copilot og ChatGPT til å halda seg til ei subnorm ved hjelp av instruksar («promptar»).

I Nynorskoversetteren og Nynorskroboten kan ein òg i samråd med leverandøren få lagt inn termar som er spesifikke for fagfeltet eller statsorganet. Denne funksjonen vart ikkje brukt i undersøkinga.

Sjølve omsetjingane vart gjorde hausten 2025, i denne rekkjefølgja: Apertium: 28.10., 13.11., Nynorskroboten og Nynorskoversetteren: 31.10. og 13.11., Copilot: 11.11. og ChatGPT: 13.11. Copilot hadde funksjonen GPT-5.0 slått på, og ChatGPT var innstilt på 5.1 auto.

3 Norminnstillingar og instruksar

I undersøkingar som denne kan rettskrivingspreferansar hos annotatørane (dei som evaluerer omsetjingane) gje skeivt resultat. For å motverka ei slik slagside og minimera støyen frå ulike subnormer bør verktøya vera nokolunde likt innstilte. Vi valde ein etter måten tradisjonell nynorsk. Ein føremon med det valet er at avstanden mellom kjeldespråket og målspråket blir større, slik at det kjem betre fram kor godt verktøya meistarar breidda i målspråket.

I dei Apertium-baserte verktøya kan ein velja både standardinnstillingar (meir og mindre bokmålsnær nynorsk) og detaljinnstillingar. Standardinnstillingane er noko fagleg problematiske, mellom anna fordi dei inneheld mekaniske kombinasjonar som er uvanlege i menneskeomsett tekst. Fininnstillingsmenyen er på si side svært finmaska, med lange alfabetiske lister over val av nokså ulike slag. Det å velja tenlege innstillingar krev ein filologisk kompetanse som ein ikkje kan venta at statstilsette skal ha. Men så lenge verktøya identifiserer orda rett, fungerer innstillingane i både Nynorskroboten og Nynorskoversetteren som ho skal. Det same kan ein ikkje seia om dei KI-baserte verktøya.

Samtalerobotane (Copilot og ChatGPT) fekk ein instruks med oppsummering av vala i Apertium. Det viste seg at ingen av dei var i stand til å hugsa instruksen særleg lenge om gongen. Dette har truleg ikkje påverka resultatet av denne evalueringa nemnande, sidan vi ikkje har undersøkt konsekvens i rettskrivingsval. Begge robotane graviterer mot gjennomsnittsnynorsk som ikkje verkar påfallande på setningsnivå. Men det er viktig å understreka at samtalerobotane har sloppe mest ustraffa unna instruksbrotet, og at dei regelbaserte verktøya dermed har fått minimal utteljing for det største fortrinnet sitt: konsekvent formval.

Konsekvens i formval kan bli emne for gransking i framtida.

4 Tekstane

Vi valde å prøva verktøya på utdrag frå autentiske tekstar som er relevante for offentlig forvaltning. Tekstmaterialet var likt delt mellom desse tre sjangrane:

- lovtekst
- NOU-tekst
- tekst frå planar, rettleiingar o.l.

Så vidt vi veit, har ingen av tekstane vorte omsette til nynorsk før.

Omsetjingsprogram blir sjeldan brukte til å omsetja lovtekst og NOU-tekst, men desse sjangrane er likevel svært relevante for ei undersøking som denne. Lovtekst gjennom-syrer alt forvaltningsspråk og byr på nett dei same grammatiske utfordringane som annan statleg tekst, berre i rikare mon. NOU-tekstar er på si side ofte meir lettlesne enn typisk forvaltningstekst, men dei speglar den faglege breidda i forvaltninga godt.

Vi lét verktøya omsetja heile tekstutdrag, slik at dei kunne gjera val baserte på konteksten etter evne. Deretter delte vi kjeldeteksten og omsetjingane opp i parallelle enkelt-setningar ved hjelp av [Bertalign](#).

Eit problem med å bruka autentiske tekstar er at dei gjerne inneheld feil og uklåre formuleringar. Vi retta dei fleste opplagde teiknsetjingsfeila, sletta nokre unødvendige innskot og gjorde nokre andre små justeringar, men dei fleste problema med kjeldeteksten vart ikkje løyste. Annotatørane vart difor bedne om å ikkje straffa omsetjinga for problem som kjeldeteksten må lastast for.

5 Undersøkinga

5.1 Val av metode

Det er vanleg å evaluera modellar med såkalla referansemåling (eller benchmarking). Ytinga til modellane blir testa automatisk mot ein manuelt utarbeidd fasit/standard. Eit døme er NTB si eiga testing av [Nynorskroboten mot ChatGPT](#) ved hjelp av målemetodane [Meteor](#) og [Bleu](#). Ei utfordring med referansemåling er normvariasjonen i norsk. Det er vanskeleg å laga ein menneskeleg fasit som skal ta omsyn til heile formverket. Vi valde difor å ikkje bruka denne metoden.

Ein mindre moderne måte å vurdere tekst på er manuell feilteljing, jf. [Språkrådet si undersøking av skriveferdigheitene til samtalerobotar i 2025](#). Vi forkasta denne metoden òg på grunn av vanskar med å laga ein fasit.

Vi valde difor ein tredje metode: manuell testing utan fasit, basert på den spontane individuelle vurderinga til eit utval av språkekspertar. Denne metoden har sine lyte (sjå 5.3), og han gjev ikkje noko absolutt mål på omsetjingskvalitet. Men når metoden er brukt på eit stort nok materiale, skal han gje eit nokså godt grunnlag for rangering av dei undersøkte verktøya.

5.2 Opplegg og gjennomføring

Til annotatørar valde vi tre profesjonelle språkkonsulentar som har greidd Språkrådet si kompetanseprøve i nynorsk med særleg godt resultat.

Desse annotatørane vurderte omsetjingar av enkeltsetningar presenterte i tilfeldig rekkefølge. Dei fekk sjå kjeldesetninga i den opphavlege samanhengen, slik at dei kunne vurdere om omsetjinga stemte med konteksten. Annoteringsgrensesnittet var [Label Studio](#).

Annotatørane gav kvar omsett setning karakter på skalaen 1 (dårlegast) til 5 (best) innanfor tre vurderingsdimensjonar:

- rettskriving (formelle feil)
- informasjonsoverføring
- flyt

Dei to siste dimensjonane eller parametrane svarer i prinsippet til parametrane «adequacy» og «fluency», som vart brukte i [Workshop on machine translation 2006 og 2007](#).

Vi ynskte å få fram skilnadene mellom god og dårleg nynorsk målføring uavhengig av talet på rettskrivingsfeil. Rettskriving vart difor skilt ut og plassert i ein eigen dimensjon. (Til denne dimensjonen høyrer i prinsippet inkonsekvensar innanfor rettskrivinga, men som nemnt er denne feiltypen ikkje målt, bortsett frå eventuelle påfallande inkonsekvensar innanfor setningar.)

I dimensjonen *informasjonsoverføring* skulle annotatørane vurdere om meininga var overført intakt frå bokmål til nynorsk. Informasjonsoverføring er altså om lag det same som semantisk ekvivalens.

Flyt er den minst sjølvforklarande dimensjonen eller parameteren. *Flyt* skal i teorien vera det same som *godt språk* utan omsyn til rettskrivingsfeil og omsetjingsfeil. Men dimensjonen er i praksis vanskeleg å isolera. Annotatørane må bruka mykje skjønn.

Annotatørane vart bedne om å setja skår innanfor kvar dimensjon som om feila og problema i dei andre dimensjonane var retta.

Dei vart òg bedne om å bruka skalaen fritt, men konsekvent, det vil seia å gradera problem spontant etter eige godtykke, men vurdere dei ulike setningane like strengt, etter beste evne.

Hovuddelen av instruksen til annotatørane står i vedlegg 1.

5.3 Vurdering av gjennomføringa

Ulik evaluering

Resultata, særleg resultata i kapittel 6.1, må sjåast i lys av semja mellom annotatørane. Di mindre semja er, di mindre vekt kan ein leggja på resultatet. Vi nytta semjemålet Krippendorffs alfa, som går frå -1 til 1 . Om verdien er under $0,6$ – $0,7$, kan ein ikkje leggja særleg vekt på resultatet. Det viste seg at semja om formelle feil låg på $0,51$, semja om flyt på $0,32$ og semja om informasjonsoverføringa på $0,16$.

Hovudårsaka til den låge graden av semje er truleg måleproblem. To personar som stort sett er samde om kva som er god og dårleg nynorsk, kan måla kvaliteten på heilt ulike måtar. Ein gjennomgang av materialet viser at ein av annotatørane systematisk har vurdert setningane mildare enn dei to andre, og at den strengaste annotatøren i mange tilfelle må ha straffa for same feil eller lyte i fleire dimensjonar. Kategoriseringsvanskar er truleg den viktigaste forklaringa på dei lågaste semjeverdiane. Det påverkar semjetalet kraftig om to annotatørar plasserer eit problem i ulike dimensjonar.

Om vi ser bort frå vektinga og berre fokuserer på semja om kor mange av setningane som i det heile inneheld lyte, får vi talet $\alpha = 0,38$. Det er heller ikkje eit høgt tal.

På grunn av den ulike evalueringa er gjennomsnittsskårane, særleg for kvar dimensjon for seg, mindre informative enn dei kunne ha vore. Men skårane er ikkje i noko tilfelle eit godt mål på absolutt kvalitet. Dei kan berre brukast til jamføring.

Systematisk mild evaluering

Alle annotatørane har vore svært tilbakehaldne med dei låge karakterane. Ein av dei har sjeldan trekt meir enn eitt poeng i éin kategori. Noko av forklaringa kan vera at karakteren 5 skulle dekkja både lytelause og nesten lytelause setningar. Men stikkprøver av materialet viser at mange setningar med alvorlege feil har fått høg skår, opp til 14 av 15 poeng. Ein av grunnane til det kan vera at ikkje alle lyte er oppdaga, sidan annotatørane skulle vurdera setningane mest mogleg spontant.

Men den klårt viktigaste grunnen til den høge poengsummen er at annotatørane rett og slett ikkje har utnytta dei låge karakterane på skalaen. Dei har i praksis reservert dei låge karakterane for setningar med fleire feil enn det som faktisk førekjem i maskinomsetjing mellom nynorsk og bokmål. Dermed blir tala misvisande, utan at annotatørane kan lastast for det.

Lærdom

Begge problema, ulik evaluering og mild evaluering, må Språkrådet ta på si kappe. Instruksjonen var lite spesifikk og overlét det aller meste til annotatørane. Det vart ikkje gjennomført noka pilotundersøking med tanke på samkøying og rettleiing. Vi tek desse problema til vitende og vil prøva å løysa dei betre i framtida.

Konsekvensen av måleproblema

Dei omtalte problema påverkar ikkje konklusjonane og råda i kapittel 8 og 9. Råda er ikkje baserte på dei ovannemnde gjennomsnittsskårane, men på kor mange setningar som inneheld uakseptable feil eller lyte. Mengda av setningar av dette slaget i materialet gjev meir enn nok grunnlag for å gje råd om bruken av desse verktøya.

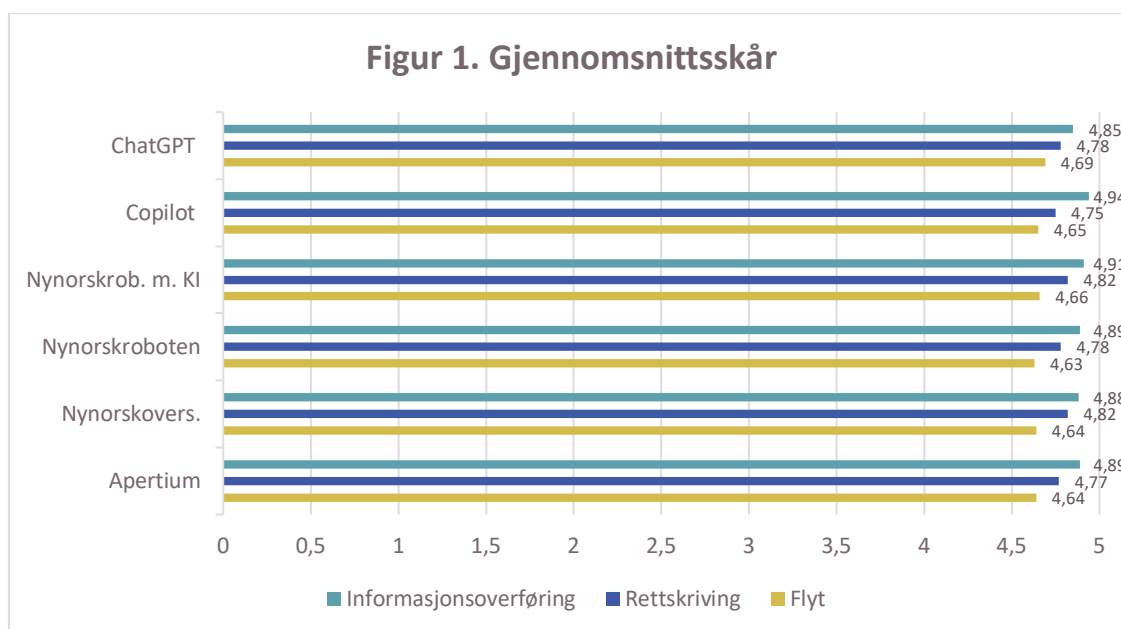
6 Resultat

Nedanfor presenterer vi først gjennomsnittstala (6.1), deretter talet på setningar med feil og lyte (6.2). Det siste målet er det viktigaste. Figur 3 og tabell 8 inneheld dei mest opplyssande tala.

Analysane er gjennomførte i Python ved hjelp av Jupyter Notebook. Dataa er behandla med *pandas*, og dei statistiske analysane er utførte med *SciPy*, *statsmodels* og *krippendorff*.

6.1 Gjennomsnittsskår

Figur 1 viser gjennomsnittsskåren for kvart verktøy i tre dimensjonar.



Skilnadene mellom verktøya innanfor kvar dimensjon er forsvinnande små. Skilnadene mellom dimensjonane er tydelegare og verkar systematisk: Alle verktøya skårar lågast på flyt og høgast på informasjonsoverføring. Men ein kan ikkje utleia så mykje av det. Det tyder fyrst og fremst at det er lenger mellom setningar med misvisande omsetjing enn det er mellom setningar med dårleg språk. Det er ikkje uventa når språka er såpass like.

Ein kunne tenkt seg at god flyt heng saman med tvilsam informasjonsoverføring, vel å merka hos maskinar (sjå særleg ChatGPT-tala). Men det vil vera ei overtolking av små skilnader. Dessutan er det ingen tvil om at ord-for-ord-omsetjing òg er semantisk risikabelt, som vi skal sjå fleire døme på i kapittel 7.

Nærare om «flyt»

Det er i dimensjonen *flyt* at skilnadene er aller minst, og ingen av dei er signifikante. Det vil seia at det ikkje er grunnlag for å sjå skilnadene som noko anna en utslag av slump.

Tabell 1. Flyt, gjennomsnitt

Verktøy	Skår	n
Apertium (AP)	4,64	732
Nynorskoversetteren (NNO)	4,64	729
Nynorskroboten (NR)	4,63	731
Nynorskroboten med KI (NRKI)	4,66	722
Copilot (CO)	4,65	726
ChatGPT (CH)	4,69	732

Kruskal–Wallis-test: H-statistikk 3,47; p-verdi: 0,63

Ingen signifikante skilnader

Nærare om rettskriving

Skilnadene i dimensjonen rettskriving er litt større, men heller ikkje dei er statistisk signifikante, med eitt unntak: Nynorskroboten med KI (høgast skår) mot Copilot (lågast).

Tabell 2. Rettskriving, gjennomsnitt

Verktøy	Skår	n
Apertium (AP)	4,77	733
Nynorskoversetteren (NNO)	4,82	730
Nynorskroboten (NR)	4,78	730
Nynorskroboten med KI (NRKI)	4,82*	722
Copilot (CO)	4,75*	727
ChatGPT (CH)	4,78	733

Kruskal–Wallis-test: H-statistikk 13,46; p-verdi: 0,02

*Dunn-test, p-verdi signifikant: NRKI mot CO: 0,025

Nærare om informasjonsoverføring

Skilnadene i dimensjonen informasjonsoverføring er dei største. Likevel er berre tre skilnader signifikante: 1) Copilot (best) mot ChatGPT, 2) Copilot (best) mot Nynorskoversetteren og 3) Nynorskroboten med KI (best) mot ChatGPT.

ChatGPT skil seg altså negativt ut i dimensjonen informasjonsoverføring. Nynorskoversetteren kjem litt dårlegare ut enn Copilot, men vi tvilar på at Nynorskoversetteren eigentleg skil seg frå dei andre regelbaserte verktøya her. Vi trur det er meir sannsynleg

at dei regelbaserte verktøya generelt skiplar meininga oftare enn Copilot, om det då er nokon reell skilnad utanfor dette materialet.

Tabell 3. Informasjonsoverføring, gjennomsnitt

Verktøy	Skår	n
Apertium (AP)	4,89	730
Nynorskoversetteren (NNO)	4,88*	730
Nynorskroboten (NR)	4,89	730
Nynorskroboten med KI (NRKI)	4,91*	722
Copilot (CO)	4,94*	725
ChatGPT (CH)	4,85*	733

Heile tabellen: Kruskal–Wallis-test: H-statistikk 24,19; p-verdi: 0,00002

*Dunn-test, p-verdi signifikant: CO mot CH 0,0002, CO mot NNO 0,01, NRKI mot CH 0,01

Nærare om dei to verktøytypane

Her er ein tabell der verktøya er slått saman i grupper etter kva teknologi dei er baserte på.

Tabell 4. Resultat etter verktøytype

Verktøytype	Rettskriving	Flyt	Informasjonsoverføring
KI-baserte utan NRKI	4,77*	4,67	4,90
KI-baserte med NRKI	4,79	4,66	4,90
Regelbaserte med NRKI	4,80*	4,64	4,89
Regelbaserte utan NRKI	4,79	4,63	4,89

*Mann–Whitney U-test: p-verdi 0,02

Flyt-verdiane er litt betre hos dei KI-baserte verktøya, men det er utelukkande på grunn av ChatGPT. Skilnadene er heilt marginale og ikkje signifikante.

Det same gjeld tala for dimensjonen rettskriving, som er litt betre hos dei regelbaserte verktøya. Den einaste signifikante skilnaden går mellom dei regelbaserte og dei KI-baserte verktøya når Nynorskroboten med KI blir gruppert saman med dei regelbaserte. Men skilnaden mellom desse konstruerte gruppene er òg liten.

Merk at dei regelbaserte verktøya truleg ville ha gjort det mykje betre i denne dimensjonen om konsekvent formval hadde vore teke med i rekninga.

Når det gjeld informasjonsoverføring, er det uråd å sjå nokon systematisk skilnad mellom regelbaserte og KI-baserte verktøy. Ein viktig grunn er sprik mellom dei KI-baserte verktøya. ChatGPT dreg ned snittet (sjå tabell 3).

Grunnlaget for å skilja mellom dei ulike verktøya og verktøytypene er altså anten tynt eller ikkje til stades. Små skilnader, målefeil, skort på signifikans og låg semje mellom annotatørane gjer at vi ikkje kan leggja vekt på tala under 6.1. Det viktigaste funnet er at skilnadene truleg er små.

6.2 Eit betre mål: talet på akseptable og uakseptable setningar

Vi skal no ta føre oss eit nytt mål: prosent setningar som krev retting eller anna forbetring. Strengt teke er grunnlaget for utrekninga (n totalt) setnings*observasjonar*, ikkje setningar, men når talet er stort, kjem det på eitt ut.

Her er skårfordelinga i kvar dimensjon:

Tabell 5. Flyt, fordeling

Verktøy	skår 1	skår 2	skår 3	skår 4	skår 5	n totalt
Apertium	1	3	39	176	513	732
Nynorskoversetteren	0	4	40	173	512	729
Nynorskroboten	1	6	35	178	511	731
Nynorskroboten med KI	0	3	36	167	516	722
Copilot	0	4	32	181	509	726
ChatGPT	0	2	28	166	536	732

Tabell 6. Rettskriving, fordeling

Verktøy	skår 1	skår 2	skår 3	skår 4	skår 5	n totalt
Apertium	0	4	24	109	596	733
Nynorskoversetteren	0	1	14	100	615	730
Nynorskroboten	0	3	19	116	592	730
Nynorskroboten med KI	1	1	15	91	614	722
Copilot	1	2	16	137	571	727
ChatGPT	1	1	20	114	597	733

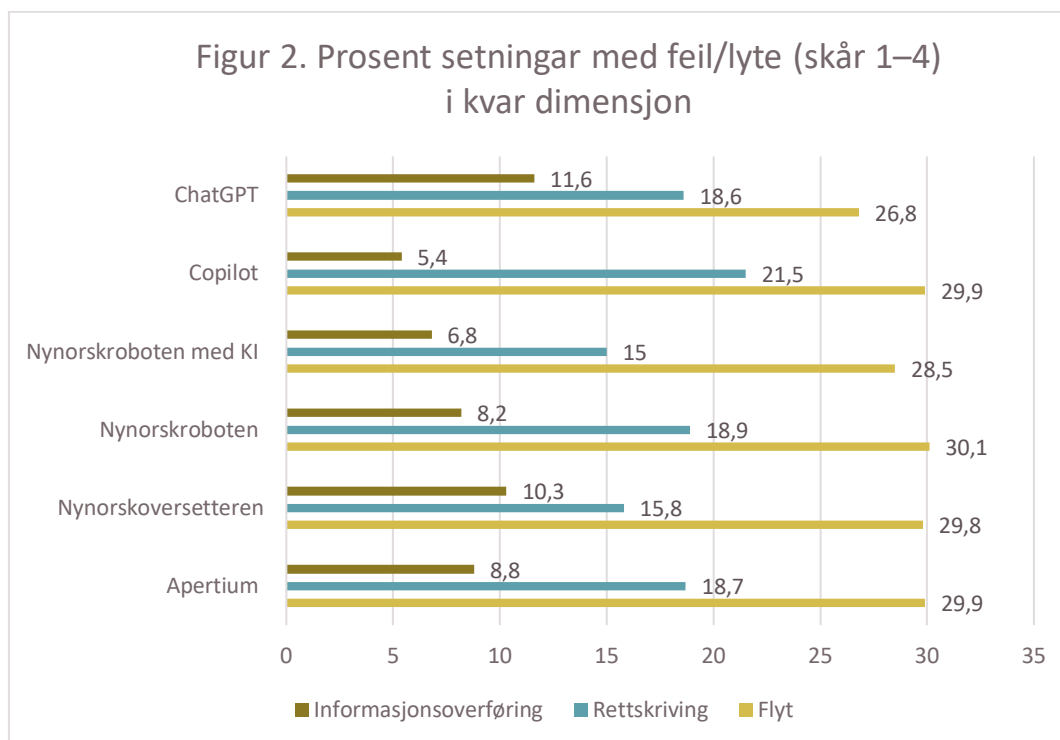
Tabell 7. Informasjonsoverføring, fordeling

Verktøy	skår 1	skår 2	skår 3	skår 4	skår 5	n totalt
Apertium	0	3	9	52	666	730
Nynorskoversetteren	0	1	10	64	655	730
Nynorskroboten	1	1	13	45	670	730
Nynorskroboten med KI	1	2	8	38	673	722
Copilot	0	0	2	37	686	725
ChatGPT	1	5	12	67	648	733

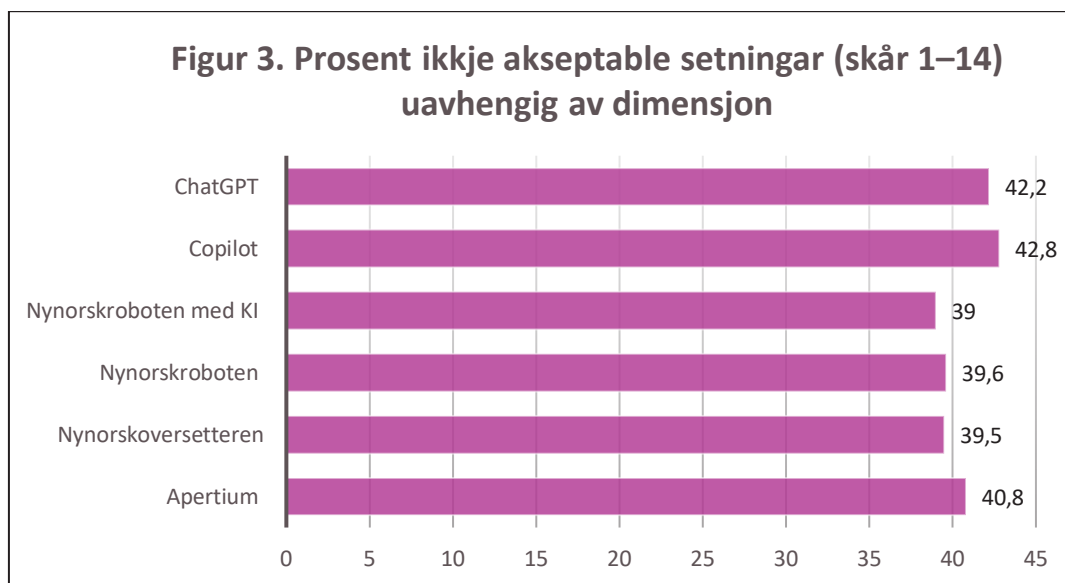
Resultatet er svært høgreskeivt eller topptungt. Knappt nokon setningar har fått dårlegare karakter enn middels. Ved fyrste augnekast ser det ut som eit godt resultat. Men denne synsmåten er ikkje rett. Dette vi ser ovanfor, er skeivleik (bias) knytt til sjølve metoden og gjennomføringa.

Stikkprøver viser at dei aller fleste setningane som har fått lågare skår enn 5 i ein eller fleire dimensjonar, inneheld klåre feil og lyte. (Sjå vedlegg 2, der eit utval av døma som har fått 4 av 5 poeng, er markerte med talet 4.) Rett nok er somme setningar straffa utan at vi kan sjå nokon opplagd grunn til det, men døma på ustraffa setningar med lyte er fleire.

Kvalitetsmålet som er brukt i figur 2 nedanfor, er altså mykje meir informativt enn gjennomsnittsskåren.



Det skjuler seg berre éin signifikant skilnad her, mellom Nynorskroboten med KI og Copilot i dimensjonen rettskriving (21,5 prosent mot 15 prosent, p-verdi 0,02). Men sidan grensene mellom dimensjonane er uklære, er det meir informativt å sjå på samle-tala i figur 3 nedanfor, som ikkje tek omsyn til kva dimensjon setningane er straffa i. (Merk at tala i figur 3 ikkje er summen av tala i figur 2, sidan nokre setningar er straffa i fleire dimensjonar.) Dei aller fleste setningar som har fått under 15 poeng, har feil eller lyte som bør rettast.



Tabell 8. Prosent ikkje akseptable setningar

Verktøy	prosent 1–14	tal 1–14	n total
Apertium	40,8	300	735
Nynorskoversetteren	39,5	290	735
Nynorskroboten	39,6	291	735
Nynorskroboten med KI	39,0	287	734
Copilot	42,8	314	734
ChatGPT	42,2	310	735

Ingen signifikante skilnader

Skilnadene er ikkje uventa etter det vi har sett til no: Dei to Apertium-baserte verktøya kjem likt ut og eit hakk over Apertium åleine. Nynorskroboten med KI-støtte har det lågaste feiltalet (39 prosent). Dei regelbaserte verktøya ser ut til å produsera fleire akseptable setningar enn ChatGPT og Copilot, men ikkje nok til at nokon av skilnadene er statistisk signifikante.

Skilnadene mellom verktøya er overskygde av hovudfunnet: Ein svært stor prosent av dei maskinomsette setningane krev manuell kontroll og redigering. Merk at tala i tabellane uttrykkjer haldninga til gjennomsnittsannotatøren i undersøkinga. Den minst

strengje annotatøren opererer i spennet 20,4–32,3, den strengaste i spennet 47,8–55,1. Det er rimeleg å trekkja den slutninga at minst ein tredel av setningane i ein tekst som er maskinomsett til nynorsk, krev etterarbeid. Sidan ein ikkje kan vita kva setningar det gjeld, må heile teksten kontrollerast. Kort sagt er maskinomsetjingar til nynorsk i beste fall utkast. Kvaliteten ligg langt under det ein ville eller burde ha godteke av ein menneskeleg omsetjar.

Vi går no over til å visa kva som skjuler seg bak skår 1–14.

7 Kva skjuler seg bak tala?

Vi har ekserpert ei rekkje feil og problem frå setningar som annotatørane har straffa, sjå vedlegg 2. I dette kapittelet tek vi føre oss eit utval.

Feiltypane er adhockategoriar som har utkrystallisert seg i gjennomgangen av materialet. Dei er ikkje knytte til evalueringsdimensjonane.

Etter pilteikna (>) står det maskinomsette formuleringar. Etter likskapsteikna (=) står det framlegg til korrekte omsetjingar.

7.1 Bokmålsrestar

Alle verktøya lèt det stå att restar av uomsett tekst: *forutseande* (AP), *opplevdes* (NNO), *De europeiske fellesskap* (NR, NRKI), *uforutsigbar* (CO), *uaktsom* (CH).

Slike feil er grove, men stort sett lette å få auga på og retta.

7.2 Ordval: bokmålsprega

Lista over slike feil er lang. Her fire døme:

- *noko ytterlegare å anføra* (AP, NNO, NR) / *oppføra* (NKRI)
= *noko å leggja til*
- *for så vidt gjeld* (AP, NR, NNO, NRKI)
= *når det gjeld*
- *uaktsomheit og taushetsplikt* (CO)
= *aktløyse og teieplikt*
- *domar, kjennelser og beslutningar* (CH)
= *domar, orskurdar og andre avgjerder*

Vi ser mellom anna at samtalerobotane ikkje meistrar grunnleggjande administrativt vokabular, endå fasiten ligg lett tilgjengeleg på norske offentlege nettsider.

7.3 Ordval: galen ekvivalent

Nokre slike feil er opplagde for dei fleste, til dømes denne:

- Bestemmelsene i første og andre ledd gjelder tilsvarende for tjenesteytere hvis tjeneste består i å gi tilgang til et kommunikasjonsnettverk
> ... tenesteytarar *viss teneste* består i å ... (AP, NR, NRKI, NNO)
> ... tenesteytarar *der tenesta* går ut på å ... (CH)
= (t.d.): ... tenesteytarar *som har som oppgåve* å ...

Legg merke til at verktøya feilar på ulike måtar. Feilen hjå dei regelbaserte er av det mekaniske slaget og lett å få auga på. Feilen hjå ChatGPT er liknar meir på dårleg tekst-binding i menneskeskapt tekst.

Nokre feil kan vera vanskelege å identifisera for andre enn dei mest kompetente språk-brukarane. Døme:

- *blir avskjediget* (om embetsmenn)
> *blir oppsagt* (AP, NR, NNO)
= får avskil
- en kraftavtale inneholder *en adgang*... til (i avtaletekst)
> *tilgang* ... til (AP, NR, NNO, NRKI)
= eit høve
- *berøre muligheten for*
> *røre høvet for* (CH) / *røre ved moglegheita for* (CO)
= påverkar høvet/moglegheita

Slike feil kan i verste fall få gjennomslag i offentleg prosa og føra til skeivutvikling av nynorsken.

7.4 Ordval: andre kortslutningar og oppfinningar

Andre kortslutningar med ekvivalentar er jamt fordelte mellom verktøytypene i materialet. Ei viss arbeidsdeling kan skimtast: Dei regelbaserte feiltolkar, medan ChatGPT diktar opp heilt nye ord:

- *betroelser de mottok* > *fortrulege ting dei fekk* (AP, NNO, NRKI)
- *praiemarkedet* > *praisemarknaden* (CH)

Skal ein oppdaga slike feil, er det ofte ikkje nok å lesa grundig korrektur; ein må òg jamføra med kjeldeteksten.

7.5 Ordklassefeil o.l.

Det er særleg dei regelbaserte verktøya som bommar på ordklasse:

- *veglovde* (for veglova) (AP, NR, NNO)
- *på eies grunn* (NNO)
- *som eigar eller driver* (NR)

Desse feila er alvorlege, men stort sett lette å få auga på.

Ein litt annan og meir overraskande feiltype er endring av forkortinga *PAH* til *Pah* (AP).

7.6 Genusfeil og galen substantivbøying

Ingen verktøypaar er frie for genusfeil:

- *ein* følgje (AP, NNO, NR)
- *ei* påstand (CH, CO)
- *ei* utnyttingsløyve og seks *leteløyver* (CO)
- *eitt* marknad (CH)

Det er mange feil av dette slaget i materialet.

7.7 Andre bøyingsfeil

Det er ikkje så mange samsvarsbøyingsfeil, men mange nok til at ein kan slå fast at ingen av verktøya er til å lita på:

- *påbod* [fleirtal] *som blir/vert gjeve* (AP, NNO, NR, NRKI)
- *stor* medvit (NRK)
- *uaktsam* lovbrod (CO)

7.8 Anaforisk pronomen o.l.

Eit trykklett *den* som viser tilbake i ein tekst på bokmål, skal oftast omsetjast med *han* og *ho/henne* på nynorsk. Verktøya kjenner til denne regelen. Men dei regelbaserte verktøya bommar fleire stader på bøyingsforma

- i kva grad *henne* når måla (AP, NNO, NR)
= ... *ho* ...

Både regelbaserte og KI-baserte verktøy kan gjera alvorlege feilkoplingar:

- Når den domfelte eller kriminalomsorgen begjærer løslatelse på prøve, fremmer påtalemyndigheten saken for tingretten, som avgjør den ved dom
> ... som avgjer *han* ved dom (AP)
= ... *henne* ... (dvs. saka)

I dømet ovanfor har Apertium truleg late pronomenet retta seg etter det siste substantivet (*retten*) i staden for det rette substantivet (*saka*). I dømet nedanfor har Copilot late seg villeia. Kjeldesetninga er tvitydig og må truleg tolkast av eit menneske:

- Selv om Norge har en befolkning med gjennomgående god livskvalitet, er den ujevnt fordelt ...
> ... er *ho* ujamnt fordelt ... (CO)
= ... *han* ...

Dette er alvorlege feil som kan vera vanskelege å oppdaga i ein vanleg korrektur. Stilbrot av følgjande slag finn vi i begge verktøygruppene:

- En lovbrøyer som er utilregnelig ... kan ved dom overføres til tvungent psykisk helsevern når denne har ...
> ... når *denne* har ... (AP, NNO, NRKI, CO)
= ... *han eller ho* ...

«Denne» er elles ikkje heilt rett brukt på bokmål her, og det ville ein profesjonell omsetjar ha sett.

7.9 Enkel bunden form

Denne grove stilfeilen førekjem i begge verktøygruppene:

- *Denne rett skal* ... (AP, NR)
= ... *retten* ...
- *dei forbod som* ... (CO, CH)
= ... *forboda* ...

7.10 Føresett egedomsord

På nynorsk skal eit egedomsord som står føre substantivet, ha trykk. Det er skilnad på **hunden** *min* og *min* **hund**. I det siste uttrykket er det alltid trykk på *min*. Der ordstillinga no ofte er snudd i forvaltningsnynorsk (som i *å vareta sine interesser*), peikar spora mot maskinomsetjing.

Det er nok uråd å laga fullstendige reglar for når eit egedomsord skal ha trykk og ikkje. Ein må forstå konteksten. Tidlegare har vi registrert denne feilen hos regelbaserte verktøy, men no ser vi at KI-baserte verktøy merkeleg nok gjer same feilen:

- barn har rett til respekt for *sitt privatliv* (CO)
= ... *privatlivet sitt*

Dessutan finn vi kortslutningar hos begge slags verktøy der eit føresett egedomsord står til fleire påfølgjande ledd i kjeldeteksten:

- Bedre *rettssikkerhet* for disse barna forutsetter også oppfyllelse av deres behov for, og rett til, å få informasjon, medvirkning og støtte
 - > ... av *deira behov for*, og rett til, å få ... (CO)
 - > ... av *behova deira for*, og rett til, å få ... (AP, NNO)
 - > ... at behova og *rettane deira for* ... (CH)
 - = ... at dei får oppfylt behovet for og retten til å få ...

7.11 Genitivs-s

Det er ikkje så mange tilfelle av uomsett genitivs-s i materialet. Her er eit døme frå kvar verktøygruppe:

- *menneskes helse* (AP, NNO, NR)
= *helsa til folk/menneske*
- *barnets relasjonelle vanskar* (CO)
= t.d. *dei relasjonelle vanskane barnet har*

7.12 Garpegenitiv og andre mislukka omskrivingar

Det kan sjå ut til at samtalerobotane er ansvarlege for dei fleste mest tvilsame garpegenitivane. Døme:

- *dei siste åra sin debatt* (CO)
= *debatten dei siste åra*
- *folketrygda sine ytingar* (CO)
= *ytingane frå folketrygda*

Copilot er overrepresentert her.

Bruk av preposisjon er ofte ein betre omskrivingsmetode enn garpegenitiv, men verk-tøya brukar metoden for mekanisk, med grove feil som følgje:

- *i bestemmelsens forstand*
> *i forstanden til føresegna* (AP)
= *etter (definisjonen i) føresegna*

7.13 Genitivskortslutningar: preposisjonssamanstøyt

Ein vanleg feiltype er omskriving med preposisjon i nærleiken av preposisjonsfrasar.

- *lovbryterens forhold til rusmidler*
> *forholdet til rusmiddel til lovbytaren* (NRKI)

= t.d. *lovbrytaren sitt forhold til rusmiddel*

7.14 Genitivskortslutningar: berre eitt ord omskrive

Eit viktig kjenneteikn på maskinomsetjing er at ikkje alle ord i ein konstruksjon i kjeldespråket har vorte med over i ein konstruksjon i målspråket:

- *arbeidstakernes [...] helse og velferd*
> *helsa til arbeidstakarane og velferd* (AP, NNO, CO)
= *helsa og velferda til arbeidstakarane*

Det er overraskande at samtalerobotane ikkje løyser dette problemet betre enn dei regelbaserte verktøya.

7.15 Passiv: ikkje omskriven

Presens passiv etter bokmålsmønster er ikkje uvanleg i omsetjingane, og det ser ut til at regelbaserte og KI-baserte gjer like mange feil. Dei regelbaserte greier altså ikkje å analysa setningane, og dei KI-baserte har etter alt å døma dårlege førebilete. Eit problem for begge verktøytypane er at det finst st-verb (refleksive og resiproke verb) som faktisk har st-former i presens. Verktøya ser ikkje skilnaden.

Døme:

- *myndighet [...] tillegges den som*
> ... *tilleggjast* den som (NNO)
= *myndighet (e.a.) blir lagt til den som / skal leggjast til den som ...*
- *[i så fall] spørres partene om de ønsker dette*
> ... *spørst* partane (AP, NNO, NR, NRKI)
= ... *blir partane spurde / skal partane spørjast ...*
- *den som overtrer ... straffes*
> ... *straffast* (CH)
... *blir straffa / skal straffast ...*

Dette er grove feil som i mange tilfelle berre har fått mild straff (4 poeng).

7.16 Passiv + partisipp

Kombinasjonen av passiv og partisipp er eit døme på stiv målføring som ofte bryt med stilidealet i nynorsk. Alle verktøya omset om lag slik:

- [utleieren kan] kreve dekket nødvendige utgifter
>... krevja dekte nødvendige utgifter

På naturleg nynorsk som i vanleg talemål brukar ein heller at-setning i slike tilfelle:
krevja at nødvendige utgifter blir dekte eller krevja å få dekt nødvendige utgifter.

7.17 Passivrelaterte kortslutningar

Ein vanleg feiltype hos dei regelbaserte (men ikkje dei KI-baserte) verktøya oppstår når det logiske objektet for eit verb i passiv inneheld fleire ord. Då blir ofte berre det eine med i den omsette konstruksjonen:

- Som avfall regnes ikke avløpsvann og avgasser
> Som avfall *blir* ikkje avlaupsvatn *rekna og* avgassar (AP)
- Når arbeidstakeren er alene om omsorgen, doubles antall stønadsdager
> ... blir talet på dobla stønadsdager (NNO)

Dette mønsteret kan ein venja seg til å vera på vakt mot, men slike kortslutningar som dette (flytting av parentes) er det vanskeleg å bu seg på:

- Den som overtrer § 22 første ledd, straffes som regel a) med bot ... b) med bot og ...
> ... *som regel a straffa*) med bot ... b) med bot og ... (NNO, NR)

Det same gjeld uutgrunnelege feil av dette slaget hos KI-verktøya:

- [det] kan oppstå fare eller voldes ulempe for andre ...
> *volde ulempe* for andre ... (CH)
- det han eller hun eier ... eller senere erverver ...
> ... eller seinare *blir erverva* (CO)

7.18 Ekvivalent i syntaktisk kortslutning

Denne typen feil kan sjå ut til å vera mest typisk for regelbaserte verktøy.

Døme: *å begå en forbrytelse* kan ofte omsetjast med *å gjera seg skuldig i eit lovbrøt*, men når vi set uttrykket inn i ein annan syntaktisk samanheng, kan det bli heilt misvisande:

- når [en lovbrøyer] har begått eller forsøkt å begå et lovbrudd som ...
> ... har gjort eller *prøvd å gjera seg skuldig* i eit lovbrøt som ... (AP, NRKI)

= .. *har gjort eller prøvd å gjera eit ...*

Det finst tallause variantar av slike feil. Skal ein oppdaga alle, må ein granska heile teksten nøye.

Eit døme til, henta frå ChatGPT:

- jordøkosystemene i unionen ... er i dårlig tilstand og under stadig forringelse
> ... er i dårleg tilstand og *stadig vert forverra* (CH)
= (til dømes:) ... og *vert stadig* meir forringa ...

7.19 Andre kortslutningar

Det fyrste dømet kan sjå ut som døme på samsvarsbøyingsfeil, men feilen er ikkje av det slaget menneske plar gjera:

- er et tiltak truffet av hensyn til noen bestemt persons interesse
> ... nokon *bestemde* ... (AP, NR, NRKI, NNO)
- ... slitte og preget av tiden
> ... slitte og *preg* av tida (CH)

7.20 Manglande retting

Retting og klårgjering høyrer med i omsetjing av statleg prosa. Det er noko dugande omsetjarar gjer spontant. Verktøya kan difor godt lastast for ikkje å vera i stand til å retta dei mest opplagde feila i kjeldene, som til dømes denne:

- betroelser de mottok mens de var en godkjent prest/forstander i et registrert trossamfunn
> ... *dei var ein godkjend* ... (AP, NNO, NRKI)
= (t.d.) ... *dei var godkjend* ...

7.21 Utelating og overforenkling

Det kan sjå ut til at dei KI-baserte verktøya går lengst i å retta feil i førelegget (sjå førre punkt), men samtidig går dei gjerne for langt i (umotivert) omskriving. Det gjeld i alle fall ChatGPT. Her er eit døme som kanskje ikkje ville ha vorte oppdaga i korrekturen:

- antallet vigslere bør stå i et rimelig forhold til tros- eller livssynssamfunnets behov for vigslere
> talet på vigslarar bør stå i eit rimeleg forhold til behovet i *samfunnet* (CH)
= ... i *trus- eller livssynssamfunnet*

7.22 Anna manglande omskriving

Ord-for-ord-omsetjing er kanskje den største systemfeilen ved maskinomsetjing frå bokmål til nynorsk. Ein kan ikkje venta at regelbaserte verktøy skal makta å snu om på heile setningar, men det er grunn til å venta at KI-baserte verktøy skal gjera nett det, sidan dei er trente på mønster i autentisk tekst. Innimellom glimtar samtalerobotane til, men i det store og heile er resultatet eit vonbrot. Det omsette materialet inneheld mange setningar som er venstretunge, passivt formulerte, substantivorienterte eller berre generelt krunglete. Alle verktøya har ytt tilskot til denne samlinga av stilbrot.

Nokre døme er sette opp nedanfor, og fleire står i vedlegget. Det finst endå fleire i sjøelve materialet, og mange har ikkje fått «nedsett karakter» i det heile. Ein av grunnane til det er at mildare variantar av problemet ikkje kan stemplast som direkte galne. Ein annan grunn kan vera at ein ikkje lenger ventar noko betre av nynorsk som er omsett frå bokmål, etter fleire tiår med maskinomsett nynorsk frå det offentlege.

Substantivisk:

- Forsvaret må være bevisst sin avhengighet av sivile leverandører
> ... medvite *avhengnaden sin av* sivile leverandører (AP, NNO, NRKI)
> ... medviten om *si avhengigheit* av sivile leverandører (CO)

Berre ChatGPT har løysinga, nemleg leddsetninga «om kor avhengig det er av ...».

Passivt og bakvendt:

- Departementet kan megle i saken etter at partene er gitt adgang til å uttale seg
> ... etter at partane *er gjeven tilgang til* å uttala seg (AP, NR, NRKI, NNO)
= ... partane *har fått høve til* ...

Venstretungt:

- > *Til medlemmer i trygda mellom 18 og 67 år som har varig redusert inntekts-
evne med minst 50 pst. på grunn av sjukdom, skade eller lyte, vert uføretrygd
ytt* (NR)
= *Det vert ytt uføretrygd til medlemmer ...*

Det siste dømet kan òg tolkast som utslag av ein aversjon mot å innleia setningar med «Det er/blir/vert ...». Det er noko som er typisk for både maskinomsetjing og KI-generert tekst (jf. rapporten «[Språket i KI-genererte tekster](#)»).

Oppstylta:

- Finner kommunen ikke å kunne ta hensyn til innsigelsen, skal statsforvalteren ordinært mekle mellom partene
> *Finn kommunen ikkje å kunna ... (o.l.)*
= (t.d.) *Dersom kommunen vel å ikkje ta omsyn til motsegna ...*

Ingen av verktøya hadde andre løysingar her.

8 Konklusjonar

- 1) Alle verktøya produserer grunne omsetjingar (ord for ord).
- 2) Det er språklege feil eller lyte i over ein tredel av setningane. Feila er av alle slag.
- 3) Regelbaserte og KI-baserte verktøy kjem nokså likt ut i denne granskinga, men dei regelbaserte har kvalitetar som ikkje er nærare undersøkte.
- 4) Språklova slår fast at statsorgana skal bruka korrekt og klårt språk. Sidan omsetjingsverktøya gjer talrike og uføreseielege feil, kan ikkje dette kravet oppfyllest utan full manuell gjennomgang av dei maskinomsette tekstane. Ein maskinomsett tekst må reknast som eit utkast.
- 5) Sidan feila i tekstane er av alle slag, kan dei ikkje rettast ved hjelp av ei enkel sjekkliste. Statsorgana må difor ha tilgang til fullgod nynorskkompetanse.

9 Råd om forsvarleg bruk av omsetjingsverktøy i staten

9.1 Manuell kontroll

Det viktigaste denne granskinga fortel, er at dei to største problema med maskinomsetjing (grunn omsetjing og talrike feil og kortslutningar) er *felles for verktøytypene*. Til dømes greier ingen av verktøya å praktisera enkle tommelfingerreglar om vanlege grammatiske konstruksjonar som passiv og genitiv fullt ut.

Same kva verktøy ein brukar, er det difor heilt avgjerande med grundig etterarbeid utført av menneske. Dette etterarbeidet krev fullgod kompetanse i målspråket nynorsk. Kompetansekravet er òg eit lovkrav til alle statsorgan ([språkløva § 16](#)).

9.2 Manuell omskriving

Ei formelt feilfri ord-for-ord-omsetjing er ikkje godt nok. Omsetjing og etterarbeid må gjerast på målspråket sine premissar og vera basert på godt fagleg skjønn.

God omsetjing går ut på å endevenda og omkalfatra setningar for å få eit mest mogleg idiomatisk resultat. I maskinomsetjing dominerer ord-for-ord-omsetjing, med innslag av omsetjing konstruksjon for konstruksjon. Setningar blir sjeldan skrivne om til optimalt målspråk

Ingen av verktøya på marknaden kan måla seg med kompetente omsetjarar som arbeider på tradisjonelt vis. Difor burde alt omsetjingsarbeid ideelt sett overlatast til slike omsetjarar. Men menneskelege omsetjarar kan ikkje tevla med maskinar om volum og tempo. Difor brukar mange omsetjarar i dag omsetjingverktøy sjølve til å gjera grovarbeidet. Statsorgana og utviklarane av omsetjingsverktøy bør då leggja godt til rette for at omsetjarane får den tida og dei ressursane dei treng for å skriva om til god nynorsk.

9.3 Mogleg arbeidsdeling mellom regelbaserte og KI-baserte verktøy

Det er vel kjent at samtalerobotar hallusinerer når dei produserer tekst. Omsetjing frå bokmål til nynorsk er ikkje noko unnatak. Robotane kan i nokre tilfelle omsetja meir idiomatisk enn regelbaserte verktøy, men ein veit aldri når kreativiteten tek overhand, og det kan vera vanskeleg å oppdaga kvar meininga er endra. Det skjerpar kravet til etterarbeid.

Regelbaserte verktøy er meir føreseielege enn samtalerobotar, òg i måten dei gjer feil på. Vi meiner difor det er tryggare å bruka eit regelbasert verktøy enn ein samtalerobot til grunnarbeidet, altså til å laga utkast. Ein kan gjerne be ein samtalerobot om hjelp med omskriving, men det bør gjerast ved sida av, slik at ein har full kontroll heile tida.

Det finst òg andre grunnar til å bruka regelbaserte verktøy (baserte på Apertium):

- Leverandørane rettar seg etter norsk språkpolitikk.
- Verktøya har fungerande normvalinnstillingar. Resultatet er truleg ein nynorsk med færre konsekvensfeil enn det ein får med KI-verktøy.
- Leverandørane tilbyr oppfølging av kundane.
- Ein kan få tilpassa produktet fagleg, med termbasar o.a.
- Ein kan få kurs i bruk av verktøya i tillegg til nynorskkurs.

Føremonen med KI-produkta er sjølvsagt at dei er gratis. Men det aukar òg faren for ukontrollert og ukvalifisert bruk.

Språkrådet meiner det er mest tilrådeleg å bruka eit Apertium-basert verktøy som hovudverktøy så lenge ein kan rekna med at punkta ovanfor gjeld. Vi føreset då at utviklarane arbeider med å forbetra innstillingsmenyane (sjå bolk 3). Vi føreset òg at statsorgana nyttar tilboda om kurs og oppfølging dersom dei kjøper inn verktøya.

9.4 Tilrådde metodar

Mange har etterlyst lister over kva feil og problem ein må vera særskilt på vakt mot. Svaret er at ein må vera på vakt mot alle, og at det ikkje er råd å laga ei komplett liste over feller. Den lista med 22 feilkategoriar som vi har lagt ved her, har liten verdi som verktøy for dei som ikkje har den tekstkompetansen som trengst, og dei som har den kompetansen, treng ikkje lista. Kort sagt: Det å bruka sjekklister i staden for medarbeidarar eller konsulentar med brei kompetanse er ikkje ein tilrådd metode.

Språkrådet meiner dette er dei to beste omsetjingsmetodane i dag:

Den beste metoden

1) Ein person med solid kompetanse i nynorsk omset manuelt og skriv om heile setningar der det trengst.

2) Ein like kompetent person kontrollerer resultatet manuelt – og eventuelt automatisk i tillegg.

Med manuell omsetjing *meaning for meaning* i staden for *ord for ord* blir resultatet meir idiomatisk enn med maskinomsetjing. Ein annan føremon med manuell omsetjing er at gode omsetjarar av eige tiltak rettar feil, løyser flokar og gjer teksten klårare og meir eintydig.

Den beste maskinbaserte metoden

- 1) Originalteksten blir omsett med eit regelbasert verktøy av ein person med solid nynorskkompetanse som har fått opplæring i bruk av verktøyet. Den regel-omsette teksten blir brukt som utkast.
- 2) Omsetjaren kontrollerer resultatet nøye og skriv om til meir idiomatisk nynorsk der det trengst. Omsetjaren kan bruka KI-støtte i tillegg, men set KI-inspirerte løysingar inn manuelt.
- 3) Ein like kompetent person kontrollerer resultatet manuelt – og eventuelt med digitale verktøy i tillegg.

10 Råd om utvikling av omsetjingsverktøy i framtida: integrerte ressursar

Vi reknar med at utviklarane av verktøya arbeider vidare for å minimera problema med grunn omsetjing og kortslutningar. Det er diverre lite som tyder på at desse problema kan løysast fullt ut med heilautomatisk omsetjing, anten omsetjinga er regelbasert, KI-basert eller begge delar.

Språkrådet etterlyser difor arbeid med utvikling og integrering av det som kan kallast halvautomatisk omsetjing. Det vi saknar, heilt konkret, er verktøy i brukargrensesnittet som kombinerer desse funksjonane:

- 1) varslings om tvilstilfelle, det vil seia det når det trengst menneskeleg vurdering (til dømes av kva synonym eller syntaktisk konstruksjon som høver best i samanhengen)
- 2) kopling til andre språkressursar som gjer det mogleg for eit menneske å avgjera desse tvilstilfella effektivt (ordbokartiklar, termlister, grammatiske forklaringar)

I dag er det svært uvisst kor effektivt maskinomsetjing til nynorsk eigentleg er, om vi tek kvalitet, naudsynt etterarbeid og kompetanseutholding med i rekninga. Verkeleg menneske–maskin-samarbeid ved hjelp av dei nemnde funksjonane kan redusera problema med låg kvalitet og minkande kompetanse og dermed auka effektiviteten i vid tyding.

Vedlegg 1: Utdrag av instruksen til annotatørane

Vurderinga av omsetjinga

Ei god omsetjing tilfredsstiller desse kriteria: 1) Meiningsinnhaldet er intakt. 2) Omsetjinga følgjer rettskrivinga (skrivemåtar, bøyning, teiknsetjing osv.). 3) Språket er godt og idiomatisk òg på andre måtar. Du skal evaluere kvar setning i tre dimensjonar som svarer til kriteria ovanfor: 1) informasjonsoverføring, 2) rettskriving, 3) flyt.

Før du evaluerer, skal du vurdere om omsetjinga har lyte som skriv seg frå meir eller mindre uløyslege problem i bokmålsteksten, til dømes uklår bodskap. Om du meiner det er urimeleg å laste «omsetjaren» for lytet, skal du ikkje trekkje for det, men i staden hake av i boksen «Problem med bokmålssetninga». Ikkje bruk denne boksen til noko anna enn slike problem, og ikkje hak av for noko som du vel å trekkje for.

I alle dimensjonane går karakterskalaen frå 1 til 5. Karakteren 1 er reservert for dei verste tilfella, der lyte øydelegg inntrykket heilt. Det kan vere eitt stort lyte, fleire små eller noko anna. Karakteren 5 er reservert for setningane med dei beste løysingane.

NB: Det er viktig at du set karakter i dei tre dimensjonane *uavhengig av kvarandre*. Det vil til dømes seie at du set karakter for flyt uavhengig av informasjonsoverføring og formelle feil. Tenk deg at feila i dei andre dimensjonale er retta!

Nedanfor får du ei viss rettleiing om karaktersetjinga. Rettleiinga er vag med vilje, for inntrykksbasert evaluering kan ikkje gjerast til ein eksakt vitskap. Du avgjer sjølv kor alvorleg ein feil eller eit lyte er, og kor mykje det skal trekkje ned. Du avgjer òg korleis karakterane mellom 1 og 5 skal brukast.

Set karakter mest mogleg intuitivt utan å ta omsyn til normalfordeling eller korleis du trur dei andre annotatørane evaluerer. Det avgjerande er ikkje nett kva karakter du set på kvar setning, men at du prøver å vere om lag like streng heile tida. Finn eit system/mønster og hald fast ved det. Stol på røynsla og språkkjensla. Merk: Bruk kunnskapen (intuisjonen) din om kva som jamt over er rekna som god nynorsk.

Karakterar for INFORMASJONSOVERFØRING (intakt bodskap på ord- og setningsplan)

- 1 for setning der viktige delar av bodskapen har gått tapt eller meininga er forvrengd på viktige punkt

- 2–4 etter ditt skjønn
- 5 for setning der budskapet i originalen etter alt å dømme er intakt; ingen naudsynt informasjon har gått tapt, og ikkje noko er lagt til

Ei uklår setning kan omsetjast med noko som er tilsvarande uklårt, eller klårgjerast med ei sannsynleg tolking. Begge delar kan vere fullgode løysingar. Du må vurdere det ut frå samanhengen. Du har òg naudutgangen «Problem i bokmålsteksten».

Karakterar for RETTSKRIVING (formelle feil)

- 1 for setning som er svært alvorleg skjemd av feil
- 2–4 etter ditt skjønn
- 5 for setning som er feilfri eller berre har ubetydelege feil

Du kan òg trekkje for inkonsekvensar innanfor setninga, men berre dei grovaste (same ord eller klårt same kategori). Anna er stilbrot som går under Flyt.

Karakterar for FLYT (idiomatisk nynorsk)

- 1 for setning som er svært prega av uidiomatisk ordlegging
- 2–4 etter ditt skjønn
- 5 for setning som er lytelaus (treng ikkje vere den best tenkjelege løysinga, men må vere fullt ut idiomatisk/akseptabel)

Hugs å ikkje trekkje for lyte som allereie er fanga opp i dei andre dimensjonane ovanfor.

I dimensjonen Flyt skal du fokusere på den språklege kvaliteten på den omsette setninga, bortsett frå det du alt har trekt for. Det kan gjelde alt frå skral setningsbygning via uidiomatisk ordval til grove stilbrot.

Klønede og unødige omstendelege formulerte setningar skal straffast for dårleg flyt sjølv om kvar formulering i og for seg er god nynorsk.

Nærare om plassering

Normbrot eller berre dårleg nynorsk?

Nokre grammatiske feilskjer (til dømes tvilsame genitivar eller genitivsomskrivingar) kan det vere vanskeleg å plassere. Vil du rekne dei som brot på rettskrivingsreglar eller grammatiske reglar, trekkjer du for dei (berre) under rettskriving. Er dei meir diffuse, trekkjer du (berre) i dimensjonen *flyt*. Vel *flyt* når du er i tvil.

Ordval

Det kan òg vere vanskeleg å plassere ordval i rett dimensjon. Uheldige val langs nynorsk–bokmåls-aksen fell i tre kategoriar:

1) Bokmålsord som ikkje er normerte i nynorsk (dvs. ikkje står i Nynorskordboka eller på lista over ord som kan godtakast i nynorske ordlister). Desse skal du (berre) rekne som formelle feil (rettskrivingsfeil). Det er noko tilfeldig kor grensa er sett for bokmålsord i nynorsk, så du må sjølv avgjere kor grov feilen er, og då bør du sjå på samanhengen ordet står i. Ord som er vanskelege å skifte ut med noko som fungerer godt, bør vurderast mildare enn heilt unødvendige bokmålsord.

2) Tillatne «bokmålsord». Desse vurderer du (berre) under flyt, som meir eller mindre idiomatisk nynorsk, ut frå di kjensle for kva som er typisk eller tradisjonsrett. Ikkje slå ned på «bokmålsord» som har vorte heilt vanlege i nynorsk. Du kan likevel trekkje noko for kombinasjonar som verkar heilt umotiverte, særleg denne: bokmålsnære ord brukte i staden for meir vanlege nynorskord i ei setning der ein òg finn markert tradisjonelle nynorskord (eller omvendt: uvanleg tradisjonelle nynorskord i ei setning som strittar imot slike ord).

3) «Nynorskord» som ikkje gjev heilt den same meininga som bokmålsordet i samanhengen. Slike ekvivalensproblem trekkjer du for berre under informasjonsoverføring. Er det andre problem med ordet, kan du trekkje fleire stader.

Om å ta omsyn til setningslengd

Lengda på setningane varierer kraftig. Du kjem til å lure på om eit visst lyte skal trekkje ei lang setning like mykje ned som ei kort setning. Svaret på det er nei. Du skal evaluere setninga som heilskap. Då vil eit lyte automatisk påverke inntrykket av dei stutte setningane mest. Det er greitt.

Det absolutte feiltalet er altså mindre viktig enn det relative feiltalet (feilprosenten). Men merk at du ikkje skal prøve å rekne ut nokon feilprosent (lyte dividert med ordmengd). Det er for tidkrevjande, og det vil dessutan føre til at lange setningar med stor avstand mellom lyta slepp litt for billig unna. Små, stygge tuver skal trass alt kunne velte lange lass.

Du kan ikkje løyse problemet med setningslengd matematisk. Du får berre leggje deg på ein mellomting så intuitivt du kan. Prøv å vere nokolunde like streng i vurderingsmåten frå økt til økt. Den einaste fasiten er kor negativt du som lesar faktisk ville ha reagert på omsetjinga.

Om originaltekstane og omsetjingane

Tekstane er av tre slag: utdrag frå lover, utdrag frå NOU-ar og utdrag frå planar og andre tekstar frå offentlege etatar.

Omsetjingsprogramma er instruerte til å omsetje etter ein nokså konservativ rettskrivingsmal (tradisjonelle former innanfor 2012-rettskrivinga og den tidlegare læreboknormalen). Men du skal ikkje vurdere om programma har følgd denne ortografiske målen. Vanleg nynorsk midt i leia er like godt. Vurder kvar setning for seg som god eller dårleg forvaltningsnynorsk, slik du kjenner denne nynorsken.

Når det gjeld ordlegging og syntaks, har programma fått instruks om å halde seg til traust statsnynorsk. Det kan du gjerne leggje til grunn.

Vedlegg 2: Lengre liste med døme på feil i omsetjingane

Nedanfor følgjer ei rekkje døme på omsetjingsfeil hjå dei seks maskinomsetjingsverktøya som vart evaluerte. Døma er ekserperte frå tre utval av setningar: dei 60 setningane som har fått lågast skår, 100 av dei setningane som har fått mindre eller lik 3 poeng i minst éin dimensjon av minst ein av annotørane, og 225 setningar som har fått til saman 14 av 15 poeng av minst ein av annotørane.

Merk at ikkje alle omsetjingar av kvar kjeldesetning er med i granskinga eller i utvalet nedanfor! Når omsetjinga til eit visst verktøy ikkje er med i dømesamlinga, kan fråværet koma av 1) at verktøyet har omsett setninga rett, 2) at verktøyet ikkje er prøvd på den aktuelle setninga.

Føremålet med denne dømesamlinga er berre å peika på kva slags og kor store feil verktøya (av alle slag) gjer. Det er ikkje teke med meir kontekst enn strengt naudsynt. Kategoriane er laga for høvet på pragmatisk grunnlag, og skilja mellom dei er ikkje skarpe.

Nokre omsetjingar er markerte med talet 4 etter forkortinga av verktøynamnet. Dette er eit utval (basert på stikkprøver) av omsetjingar som berre er trekte med eitt poeng i éin dimensjon av ein av annotørane. Poenget med denne markeringa er å visa at 4 poeng i dei fleste tilfelle i røynda er ein strykkarakter.

1 Bokmålsrestar

Apertium-baserte

- enhver skal ... > *enhver* skal ... AP
- foreligger > *foreligger* AP
- forholdstallene > *forholdstallene* AP
- forurensningsmyndigheten > *forurensningsmyndigheten* AP
- straffelovgivningen > *straffelovgivningen* AP
- forutseende > *forutseende* AP 4
- en selvhostet open-source grunnmodell > ... *sjølvhostet* ... AP

- for så vidt gjelder offentlig veg, dekkes utgifter ved tiltakene som vegutgifter ...
> ... *dekkes* ... AP, NNO, NR
- opplevdes > *opplevdes* NNO
- De europeiske fellesskap > *De* ... NR, NRKI
- høygradig psykisk utviklingshemming > *høygradig* ... NNO, NR, NRKI
- sammenholdt ... med > *samanhaldt* ... med (NRKI)

KI-baserte

- uforutsigbar > *uforutsigbar* CO 4
- Kravet må være framsatt innen rimelig tid > ... *framsatt* ... CO 4
- betroelser > *betroingar* CO
- Finner kommunen ikke å ... > *Finner* kommunen ikkje å ... CO
- innsigelsen > *innsigelsen* CO
- Den som har gjort vesentlige forberedelsestiltak med sikte på å utføre en handling som vil utgjøre en overtredelse kan ved dom forbys å ... > ... *forbys* å ... CH, CO
- uaktsom > *uaktsom* CH
- Celler må også «vite» hvor i en sammensatt (flercellet) organisme de befinner seg, slik at de kan skru av og på de riktige genene > ... (*flercella*) ... CH 4

2 Ordval: bokmålsprega

Døma er både innanfor og utanfor norma. Sjå òg *Anna manglande omskriving*

Apertium-baserte

- bestemmelsene > *bestemmelsene* AP
- å forholde seg passiv kan i noen tilfeller innebære å «støtte» ... > å *forhalda seg* passiv ... AP

- Embetsmenn står i embetet til de har nådd aldersgrensen, eller til de søker avskjed eller blir avskjediget > ... *avskjed* eller ... NNO

Begge verktøytyper

- forhold som antyder ulovlig virksomhet > ... *antydar* ... AP, NRKI
- enhver tjeneste som består i å ... > ... som *består* i å ... AP, NNO, NR, NRKI
- for så vidt gjelder offentlig veg, dekkes utgifter ved tiltakene som vegutgifter > *for så vidt gjeld* ... AP, NR, NNO, NRKI
- forventet samlet inntjening > *forventa* ... AP, NRKI, NNO
- antar forliksrådet at det kan være grunnlag for ... > *antek* ... AP, NNO 4, NR, NRKI
- har noe ytterligere å anføre
> ... *ytterligere å anføra* NRKI
> ... *ytterligere å oppføra* AP, NNO 4, NR
- skriftlig underretning > ... *underretning* AP, NR, NRKI
- de øvrige > *dei øvrige* NRKI
- de siste ti årene forut for fastsettingstidspunktet > ... *forut* for ... NR, NRKI, NNO
- for så vidt skaden ikke skyldes forhold som også eieren hefter for > *for så vidt skaden ikkje* ... NNO, NR, NRKI
- er et tiltak truffet > ... *treft* AP, NR, NRKI, NNO, CH
- det forventes begrenset permanent arealbruk ... > ... er *forventa* ... AP, NNO, NR, NRKI, CH
- har *opptrådt* uaktsomt > har *opptredd* ... AP, NNO, NR, NRKI; har *opptrådt* ... CH, CO
- ... at [følgen] kunne *inntre* > ... *tre inn* AP, NNO, NR, NRKI, CH; ... *inntreffe* CO
- ... slik at det gir mulighet for kontakt og kommunikasjon med andre arbeidstakere i virksomheten > ... *moglegheit* for NNO, NRKI, NR, CO

- Avtalepartene, som tar i betraktning at inngåelsen av denne avtale ikke på noen måte skal berøre muligheten for noen EFTA-stat til å tiltre De europeiske fellesskap ...> ... som *tek i betraktning* at ... *moglegheita for* nokon Efta-stat til å ... AP, NNO, NNR, CO
- døde dyr > døde dyr (ALLE)

KI-baserte

- bevillingen > *bevillinga* CO
- barn og ungdom som er utrygge i relasjoner vil som en beskyttelse ofte reagere med å ta kontroll over situasjonen ... > ... vil som ein *beskyttelse* ... CO 4
- beslutningsmekanisme > *beslutningsmekanisme* CO
- uaktsomhet > *uaktsomheit* CO
- taushetsplikten > *taushetsplikta* CO
- ... dersom vedkommende var i en selvforskyldt utilregnelighetstilstand > ... ein sjølvforskyldt *utilreknelegheitstilstand* CO
- har ... unnlatt > ... *unnlate* CO
- for så vidt skaden ikke skyldes forhold som også eieren hefter for > ... *skuldast* ... CO
- samt > samt CO, CH (begge konsekvent)
- ... uforutsigbare ... konsekvenser > ... *uforutsigbare* CO, CH
- uvitenhet > *uvitenheit* CO, CH
- metallforbindelser og PAH-forbindelser > *metallforbindelsar* og PAH-forbindelsar CO, CH
- dommer, kjennelser og beslutninger > *domar, kjennelser og beslutningar* CH
- avtalepartene ... har besluttet ... > ... *beslutta* ... CH

- å videreutvikle Den høyere påtalemyndighets fagledelsesvirksomhet og få denne opp på et nivå som svarer til behovet > ... og *bringa* ho opp ... CH

3 Ordval: galen ekvivalent

Apertium-baserte

- de øvrige parter i konvensjonen > resten av partar i konvensjonen AP, NR
- ved vurderingen av gjentakelsesfaren skal det legges vekt på det begåtte lovbruddet ... > ... *det gjorde lovbrøtet* AP
- beredskapen skal stå i et rimelig forhold til sannsynligheten for akutt forurensning og omfanget av skadene og ulempene som kan inntreffe > ... *sjansane* for akutt ... skadane og ulempene som kan *henda* AP 4
- dersom en kraftavtale inneholder en adgang for kraftleverandøren til ensidig å gjøre endringer i avtalen ... > ... inneheld ein *tilgang* ... AP, NR, NNO
- embetsmenn står i embetet til de har nådd aldersgrensen, eller til de søker avskjed eller blir *avskjediget* > ... *oppsagt* AP, NR, NNO 4
- store mengder materialer og ressurser er bundet opp i det bygde miljøet > store mengder *materiale* ... NNO 4, NR 4
- EU har estimert at 60–70 pst. av jordøkosystemene i unionen, inkludert jordbruksjord, er i dårlig tilstand og under stadig forringelse > ... under stadig degradering NR
- ved beslutning treffes avgjørelser om a) saksbehandlingen som ikke etter loven krever kjennelse ... > ved *avgjerd* treffest avgjerder om a) saksbehandlinga som ikkje etter lova krev *rettsavgjerd* ... NR
- dommer, kjennelser og beslutninger > domar, rettsavgjerder og *avgjerder* NR, NNO

Begge verktøytypar

- forutgående medlemskap > *førre* medlemskap NRKI, NR, NNO
- den som driver virksomhet nevnt i første og annet punktum, skal forøvrig foreta nødvendig opprydding i området > ... skal *forresten* ... AP, NNO, NR, NRKI

- metallforbindelser og PAH-forbindelser > metallsamband og PAH-samband AP, NRKI
- det er ikke adgang til å inngå tidsbestemt leieavtale for bolig for kortere tid enn tre år > det er ikkje *tilgang* til å ... NRKI 4
- lovbrysterens øvrige adferd > den andre *åtferda til lovbyrtaren* AP; *resten av åtferda ... til lovbyrtaren* NRKI
- den som faktisk driver, bruker eller innehar eiendommen > ... *har eiegen*domen NNO, NR, NRKI, CO
- Bestemmelsene i første og andre ledd gjelder tilsvarende for tjenesteytere hvis tjeneste består i å gi tilgang til et kommunikasjonsnettverk > ... tenesteytarar *viss teneste* består i å ... AP, NR, NRKI, NNO > ... tenesteytarar *der tenesta* går ut på å ... CH

KI-baserte

- har besluttet å inngå denne avtale > har *avgjort* ... CO
- en informasjonssamfunnstjeneste er a) enhver tjeneste som ... > ... *kvar* teneste som ... CO
- for de aller fleste norske virksomheter gir en suveren modell minimal ekstra nytte i forhold til kostnad og risiko > ... *samanlikna med* kostnad og risiko CH
- berøre muligheten for > *røre ved* moglegheita for CO, *røre* høvet for CH
- (med logisk kortslutning:) medgåtte samlede utgifter > dei samla utgiftene som *er brukte* CH

4 Ordval: andre kortslutningar og oppfinningar

Apertium-baserte

- betroelser de mottok > fortrulege *ting dei fekk*... AP, NNO, NRKI
- pensjonsberettigede > *pensjonsrettmessige* NNO
- flere steder har fortettingen ... opplevdes som å forringe eksisterende by- eller bomiljø > ... som å *ringast* eksisterande by- eller bumiljø NNO
- den opprinnelige sendingens karakter må ikke forandres eller forringes > ... eller *ringast* NNO 4

- landskapstrekk > *landskapsstrekk* NR, NRKI

KI-baserte

- Embetsmenn står i embetet til de har nådd aldersgrensen, eller til de søker avskjed eller blir *avskjediget* > ... *avskilde* (CO, CH); ... *avskilte* (NRKI)
- ansvarsfrihet for visse overførings- og tilgangstjenester > ansvarsfrihøve ... CH
- Når arbeidstakeren er alene om omsorgen, doubles antall stønadsdager > ... *stønadsgåvedagar* ... CH
- uaktsomhet > *uaktsomskap* CH
- Reisen betales direkte til sjåføren, og særlig i praiemarkedet har kundene begrenset informasjon om prisen på turen før den er gjennomført > ... *praisemarknaden* ... CH 4

5 Ordklassefeil o.l.

Apertium-baserte

- PAH > *Pah* AP
- Loven skal fremme effektiv bruk av samfunnets ressurser > ... *framme* ... AP 4
- vegloven > *veglovde* AP, NR, NNO
- på eiers grunn > på *eiges* grunn NNO
- en ektefelles rett til å råde over det han eller hun eier > ... *eigar* NNO
- Enhver som helt eller delvis eier eller driver nett ... > ... *eigar* ... NR
- ulempe for eier eller bruker av grunnen > ... *eigar* eller *bruker* NNO, NRKI
- Departementet gir forskrifter om utforming og gjennomføring av ordningen, herunder om ... >... under *detta* om ... NRKI

6 Genusfeil og galen substantivbøying

Hankjønn/hokjønn

Apertium-baserte

- arbeidsavklaringspenger > *arbeidsavklaringspenger* Ap
- en enkeltvis og samlet vurdering > *ein ... vurdering* AP, NR, NNO
- en uforsettlig følge > *ein ... følge* AP, NNO, NR

Begge verktøytypar

- en riktig dimensjonert og profesjonell fagledelsesvirksomhet ... > *ein ... fagleiingsverksemd* AP, CO

KI-baserte

- ingen begrensning > *ingen* avgrensing CO
- noen bestemt persons interesse > ... *sin* interesse CO
- noen egen analyse > *ei eiga* analyse CH 4
- en påstand > *ei* påstand CH, CO

Inkjekjønn/anna

Apertium-baserte

- en selvhostet open-source grunnmodell ... > *et ...* Ap
- De europeiske fellesskap > *fellesskapane* NNO; ... *fellesskapar* NR, NRKI

KI-baserte

- Ved inngangen til 2024 var det tildelt en utnyttelsestillatelse ... og seks letetillatelser ...
... *ei* utnyttingsløyve ... og seks *leteløyver* CO
- medlemmet > *medlemmet* CO
- rettsgodene > *rettsgodene* CH, CO
- ett marked > *eitt* marknad CH 4

Alle kjønn

Begge verktøytypar

- Det omfatter flammehemmere, UV-stabilisatorer, PFAS-er, ftalater, bisfenoler, alkylfenoler og alkylfenoletoksilater, biocider ...
 - > *PFAS-er ... bisfenoler, alkylfenolar og alkylfenoletoksilater ...* AP
 - > *... PFAS-ar ... bisfenoler, alkylfenolar og alkylfenoletoksilatar ...* NRKI
 - > *... PFAS-ar, ftalatar, bisfenolar, alkylfenolar og alkylfenoletoksilar, biocidar ...* CO
 - > *... PFAS-stoff, ftalatar, bisfenolar, alkylfenolar og alkylfenoletoksilatar ...* CH

7 Andre bøyingsfeil

Samsvarsbøying av partisipp av sterke verb

Apertium-baserte

- de forbud og påbud som gis > ... blir/vert *gjeve* AP, NNO, NR, NRKI
- regler gitt av departementet > reglar *gjeve* ... AP, NNO, NR, NRKI
- Arbeidsavklaringspenger gis så lenge det er nødvendig ... > ... blir *gjeven* ... NNO, NR

Anna samsvarsbøying

Apertium-baserte

- § 20 annet til fjerde ledd > ... *anna* til fjerde ledd NNO 4
- virksomhet nevnt i første og annet punktum > ... *anna* punktum AP, NNO, NR, NRKI
- departementene ... bør ha stor bevissthet om ... > ... *stor* medvit ... NRKI 4

KI-baserte

- særskilt smittesanering > *særskilt* smittesanering CO 4
- uaktsomt lovbrudd > *uaktsam* lovbrøt CO

8 Enkel bunden form

Apertium-baserte

- Denne rett skal omfatte frihet til å ha meninger og til å motta og meddele opplysninger og ideer ... > Denne *rett* skal ... AP, NR

- Den høyere påtalemyndighet > Den høgare *påtalemakt* NRKI
- (med kortslutning; det eine ordet står att:) Enhver skal være oppmerksom på offentlig trafikkskilt, signal og oppmerking og skal rette seg etter de forbud og påbud som gis på denne måte > ... dei forboda og *påbod* ... AP, NNO, NR (men retta med NRKI)

KI-baserte

- De europeiske fellesskap > ... *fellesskapar* NR, NRKI; ... *fellesskap* CO, CH
- de forbud ... som > ... dei forbod ... som CO, CH
- Avtalepartene, som bekrefter den betydelige vekt de tillegger ... > ... den store vekt ... CH

9 Anaforiske pronomen o.l.

Den for han/ho

Apertium-baserte

- Forurensningsmyndigheten kan i forskrift eller i det enkelte tilfelle fastsette nærmere krav til avløpsledning, herunder om den skal være lukket og vanntett > ... om *den* ... AP, NNO, NR

KI-baserte

- ... å legge fram bevillingen eller dommen med attest for at den er endelig > ... *den* er endeleg CO

Han/ho for den, men med feil

Apertium-baserte

- Når den domfelte eller kriminalomsorgen begjærer løslatelse på prøve, fremmer påtalemyndigheten saken for tingretten, som avgjør den ved dom > ... som avgjer *han* ved dom AP 4
- Det er avgjørende for legitimiteten til den offentlige helsetjenesten at samfunnet har mulighet til å følge opp i hvilken grad den når sine mål > ... i kva grad *henne* når måla sine AP, NNO, NR
- Pass på regelmessige oppdateringer, spesielt ved vesentlige endringer i løsningen eller omgivelsene den brukes i > omgjevnadene *henne* vert brukt i NNO 4 (same feil hjå NNO andre stader)

KI-baserte

- Selv om Norge har en befolkning med gjennomgående god livskvalitet, er den ujevnt fordelt ... > ... er *ho* ujamnt fordelt ... CO 4

Denne for han/ho

Begge verktøytyper

- En lovbrøyer som er utilregnelig ... kan ved dom overføres til tvungent psykisk helsevern når denne har ... > ... når *denne* har ... AP, NRKI, NNO 4, CO
- Det kreves et tilstrekkelig antall personer for å videreutvikle Den høyere påtalemyndighets fagledelsesvirksomhet og få denne opp på et nivå som svarer til behovet > ... få *denne* opp på eit nivå ... NNO, NRKI, CO

10 Føresett eighedsord

Apertium-baserte

- det samme gjelder en utlending som etter sitt hjemlands lov ikke er myndig > ... etter *sitt heimlands* lov ... AP, NR, NRKI (CH: etter heimlandsretten)

KI-baserte

- barn har rett til respekt for sitt privatliv > ... *sitt privatliv* CO
- En tjenesteyter skal ved utøving av sin virksomhet alltid gi informasjon om ... > ... ved utøving av *si verksemd* ... CO

Begge verktøytyper

- Bedre rettssikkerhet for disse barna forutsetter også oppfyllelse av deres behov for, og rett til, å få informasjon, medvirkning og støtte >
> ... oppfyllelse av deira behov for, og rett til, å få informasjon, medverknad og støtte CO
> ... oppfyllelse av behova deira for, og rett til, å få informasjon, medverknad og støtte AP, NNO
> ... at behova og rettane deira for informasjon, medverknad og støtte blir oppfylte CH

I dei to siste døma er det ei syntaktisk kortslutning i tillegg: Verktøya ser ikkje kva ord *deres* står til.

11 Genitivs-s

Apertium-baserte

- bortfall av arbeidsinntekt for yrkesaktive i forbindelse med barns eller barnepassers sykdom, barns funksjonshemming og pleie av nærstående i livets

sluttfase > ... *barns eller barnepassars sjukdom, barns funksjonshemming* ... AP, NNO, NR

- en risiko for menneskers helse og for miljøet > ... for *menneskes helse og for miljøet* AP, NNO, NR
- Den høyere påtalemyndighets fagledelsesvirksomhet > Den høgare påtalemaktas fagleiingsverksemd NNO
- lønnskostnadsandelen – dvs. lønnskostnadenes andel av faktorinntekten > lønnskostnadsdelen – dvs. *lønnskostnadenes del* av faktorinntekta NR 4
- Gravide og barn er spesielt sårbare for blyeksponering, og bly kan hemme barns intellektuelle utvikling > ... barns intellektuelle utvikling NR

Begge verktøytypar

- ... pleie av nærstående i livets slutfase > ... *livets slutfase* NRKI
- de siste årenes debatt ... har løftet fram noen svakheter ved siste tiårs fortetting i og rundt byer og tettsteder > ... ved *siste tiårs* fortetting ... NNO, NR, CO

KI-baserte

- reglene i lovens første, fjerde og femte del ... > reglane i *lovas* første, fjerde og femte del ... CO 4
- barnets relasjonelle vansker > *barnets* relasjonelle vanskar CO 4

12 Garpegenitiv og andre mislukka genitivsomskrivingar

Apertium-baserte

- forliksrådets virksomhet > forliksrådsverksemda AP
- arbeidets organisering, tilrettelegging og ledelse, arbeidstidsordninger, lønnssystemer, herunder bruk av prestasjonslønn, teknologi mv. skal ... > *organiseringa, tilrettelegginga og leiinga i arbeidet* ... AP, NR
- Å forholde seg passiv kan i noen tilfeller innebære å «støtte» i bestemmelsens forstand > ... *i forstanden til føresegna* AP
- en hindring utenfor utleierens kontroll > ... utanfor *kontrollen til* utleigaren AP, NRKI

KI-baserte

- De siste årenes debatt om arkitektur og kvalitet i nybygde nabolag ... > *Dei siste åra sin debatt ...* CO
- en hindring utenfor utleierens kontroll > ... utanfor utleigaren sin kontroll CH
- rett til folketrygdens ytelser > ... folketrygda *sine* ytingar CO
- Den høyere påtalemyndighets fagledelsesvirksomhet > Den høgare påtalemakta *si fagleiingsverksemd* CO
- Arbeidets organisering, tilrettelegging og ledelse, arbeidstidsordninger, lønssystemer, herunder bruk av prestasjonslønn, teknologi mv. skal ... > *Arbeidet si organisering, tilrettelegging og leiing ...* CO

13 Genitivskortslutninger: preposisjonssamanstøyt

Apertium-baserte

- Antallet vigslere bør stå i et rimelig forhold til tros- eller livssynssamfunnets behov for vigslere > ... forhold *til behovet til* trus- eller livssynssamfunnet for vigslarar AP

Begge verktøytypar

- Opphaverens rett til å få adgang til åndsverk i andres besittelse > Retten *til opphavaren til* å få tilgang til åndsverk i *innehaving til andre* NRKI
- Ved vurderingen av gjentakelsesfaren skal det legges vekt på det begåtte lovbruddet sammenholdt særlig med lovbrysterens øvrige adferd, sykdomshistorie, aktuelle psykiske tilstand og forhold til rusmidler > ... resten av *åtferda, sjukdomshistoria, aktuelle psykiske tilstanden og forholdet til rusmiddel til* lovbrytaren NRKI

14 Genitivskortslutninger: berre eitt ord omskrive

Apertium-baserte

- barnets eller barnepasserens sykdom > *barnets eller sjukdomen til barnepassaren* NNO, NR
- Ved vurderingen av gjentakelsesfaren skal det legges vekt på det begåtte lovbruddet sammenholdt særlig med lovbrysterens øvrige adferd, sykdomshistorie, aktuelle psykiske tilstand og forhold til rusmidler > ... den andre *åtferda til lovbrytaren, sjukdomshistorie, aktuelle psykiske tilstand og forhold til rusmiddel* AP

Begge verktøytypar

- ... avtaler som gir avgjørende innflytelse over et foretaks organer med hensyn til deres sammensetning, stemmegivning eller beslutninger > med omsyn *til samansetninga deira, røysting eller avgjerder* NRKI
- ... faktorer i arbeidsmiljøet som kan innvirke på arbeidstakernes fysiske og psykiske helse og velferd > ... den fysiske og psykiske *helsa til arbeidstakarane og velferd* AP, NNO, CO

KI-baserte

- Arbeidets organisering, tilrettelegging og ledelse, arbeidstidsordninger, lønssystemer, herunder bruk av prestasjonslønn, teknologi mv. skal ... > *Arbeidsorganisering, tilrettelegging og leiing*, arbeidstidsordninger, lønssystem, mellom anna bruk av ... CH

15 Passiv: ikkje omskriven

Sjå òg Bokmålsrestar, Passivrelaterte kortslutningar.

Apertium-baserte

- Dersom endringer i DNA-et kun er til stede i kroppens somatiske celler (alle andre celler bortsett fra kjønnscellene), vil ikke disse arves ved kjønnnet reproduksjon > ... vil ikkje desse *arvast* ... AP 4
- myndighet etter regelverket ... tillegges den som utpekes av reguleringsmyndigheten > ... *tilleggjast* den som blir ... NNO 4
- Bevis ... føres ved å legge fram bevillingen eller dommen ... > ... *førast* ... NR

Forveksling med st-verb:

- Ved vurderingen av om vilkåret er oppfylt, ses det bort fra perioder med tjeneste i internasjonale > ... *sest* det bort frå ... NNO 4
- Ved dom treffes avgjørelse av ... + ved beslutning treffes avgjørelser om ... > ... *treffest* ... NR (4)
- Sentralt for utvalget er at barn og unge møtes med respekt > ... *møtest* med respekt NR (4)

Begge verktøytypar

- avgjørelser som fattes av den avregningsansvarlige > avgjerder som *fattast* ... AP, NR, NRKI

Forveksling med st-verb:

- Antar forlikrådet at det kan være grunnlag for å avsi dom etter § 6-10, spørres partene om de ønsker dette ... *spørst* partane ... AP, NNO 4, NR, NRKI
- Det ses bort fra perioder med tjeneste i internasjonale organisasjoner ... > Det *sest* bort frå ... AP, NR, NRKI

KI-baserte

- Døde dyr ... pålegges nedgravd > ... påleggjast nedgravne NRKI; ... påleggjast nedgravd CO
- Den som overtrer § 22 første ledd, straffes ... > ... *straffast* ... CH

16 Passiv + partisipp

Apertium-baserte

- Saken søkes ferdigbehandlet i første rettsmøte > ... *vert søkt ferdigbehandla* ... NR

Begge verktøytyper

- Arbeidet skal søkes utformet slik at ... > ... skal *søkjast forma ut* slik ... NNO, NR; ... skal *søkjast utforma* slik ... CO, NRKI
- Døde dyr ... *pålegges nedgravd* > ... blir *pålagde nedgravne* NNO; ... *påleggjast nedgravne* NRKI; ... *påleggjast nedgravd* CO
- Er husrommet i dårligere stand ... , kan utleieren kreve dekket nødvendige utgifter til utbedring > ... krevja dekte (e.l.) ... AP, NNO, NR, NRKI, CO, CH

17 Passivrelaterte kortslutningar

Sjå òg forvekslingar med st-verb under «Passiv: ikkje omskriven».

Apertium-baserte

- Som avfall regnes ikke avløpsvann og avgasser > Som avfall *blir* ikkje avlaupsvatn *rekna og* avgassar AP
- Til barn under 20 år ytes det på visse vilkår barnepensjon dersom én eller begge foreldre er døde > ... blir det ytt på visse vilkår barnepensjon ... AP 4

- På samme måte straffes overtredelse av vilkår i enkeltvedtak i medhold av denne lov og brudd på forbud etter §§ 35 og 36 > På same måte *blir brot straffa på vilkår i einskildvedtak i medhald av denne lova og brot på forbod* etter §§ 35 og 36 AP
- Den som overtrer § 22 første ledd, straffes som regel a) med bot ved alkoholkonsentrasjon i blodet til og med 0,5 promille ... b) med bot og betinget eller ubetinget fengsel ... > ... *som regel a straffa*) med bot ... NNO, NR
- For at arbeidstakeren skal få rett til omsorgspenger fra arbeidsgiveren, må barnets eller barnepasserens sykdom dokumenteres med egenmelding eller legeerklæring > ... må ... *blir/vert* dokumentert med ... NNO, NR
- For hver post under immaterielle eiendeler skal økonomisk levetid og valg av avskrivningsplan opplyses > ... skal økonomisk levetid og val av avskrivningsplan *blir* opplyst AP, NNO, NR
- Når arbeidstakeren er alene om omsorgen, dobles antall stønadsdager > ... blir talet på dobla stønadsdagar NNO 4
- Levende dyr skal ikke omsettes, tas inn i dyrehold, flyttes eller settes ut når det er grunn til mistanke om smittsom dyresykdom ... > ... omsetjast, *vert teke* inn i dyrehald, *vert flytt* eller *sett ut* ... NR
- Gjennom inntektsrammereguleringen og KILE-ordningen ... gis nettselskapene økonomiske insentiver ... > ... *blir gjeven nettselskapa* ... AP

Begge verktøytypar

- For skade og ulempe voldt ved slike tiltak ytes erstatning fastsatt ved skjønn > ... *vert/blir* erstatning *ytt fastsett* ved skjøn NNO, NR, NRKI
- Ved kjennelse treffes avgjørelser a) som avviser en sak fordi vilkår for å ta den under behandling ikke er oppfylt, eller som fremmer saken etter særskilt behandling av en påstand om avvisning ...
> Ved *rettsavgjerd vert avgjort* a) saker som *vert avviste* fordi vilkår for å ta dei under behandling ikkje er oppfylt, *eller som vert fremja* etter særskild behandling av ein påstand om avvisning ... NRKI
- en hindring ... som utleieren ikke med rimelighet kunne ventes å ha tatt i betraktning

- > ... som utleigaren ikkje innanfor rimelege grenser kunne *ventast å ha teke i betraktning* på avtaletida eller å ha unngått eller overvunne følgjene av AP
- > ... som utleigaren ikkje *med rimelege grenser* kunne *ventast å ha teke i betraktning* på avtaletida eller å ha unngått eller overvunne følgjene av NRKI
- > (ekstra alvorleg:) ... som utleigaren ikkje med rimeleg grunn kunne *venta å ta i betraktning* på avtaletida eller unngå eller overvinna følgjene av CH

KI-baserte

- Fører av kjøretøy skal avpasse farten etter sted, føre-, sikt- og trafikkforholdene slik at det ikke kan oppstå fare eller voldes ulempe for andre ... > *volde ulempe* for andre ... CH
- Ekteskapet medfører ingen begrensning i en ektefelles rett til å råde over det han eller hun eier når ekteskapet blir inngått(,) eller senere erverver ... > ... eller seinare *blir erverva*
CO

18 Ekvivalent i syntaktisk kortslutning

Nokre av desse ekvivalentane kan vera korrekte isolert sett (semantisk), men alle er sette inn i galen syntaktisk samanheng.

Apertium-baserte

- Overføring til tvungent psykisk helsevern kan også idømmes en lovbrøyer som er utilregnelig etter § 20 annet til fjerde ledd når denne har begått gjentatte lovbrudd ... > ... kan òg *dømast til* ein lovbrøytar ... NNO 4
- En lovbrøyer ... kan ved dom overføres til tvungent psykisk helsevern når denne har begått eller forsøkt å begå et lovbrudd som ... > ... har gjort eller *prøvd å gjera seg skuldig* i eit lovbrøt som ... AP, NRKI
- Opphaverens rett til å få adgang til åndsverk i andres besittelse > ... *i innehaving til* andre NRKI
- Foreligger det grunn til mistanke om ... > *Ligg føre det grunn* ... > NNO, NR (men retta med NRKI)
- de øvrige parter i konvensjonen > resten av *partar* i konvensjonen AP, NR
- Utvalget har på denne bakgrunnen valgt å redegjøre kort for spørsmålet om ... > ... *å greia kort ut for* spørsmålet om NR

Begge verktøytypar

- straffes ... med bot og betinget eller ubetinget fengsel
> ... med bot og *vilkårsbunde eller fengsel utan vilkår* NNO, NR, NRKI
> ... med bot og *vilkårsbunden eller ubunden fengsel* CO

KI-baserte

- EU har estimert at 60–70 pst. av jordøkosystemene i unionen ... er i dårlig tilstand og under stadig forringelse
> ... er i dårlig tilstand og *vert ringare stadig* NRKI
> ... er i dårlig tilstand og *stadig vert forverra* CH

19 Andre kortslutningar

Her er nokre kortslutningar som ikkje er nemnde i andre bolkar. Dei tre fyrste kan minna om samsvarbøyingseil, men stikk truleg djupare.

Apertium-baserte

- Ved vurderingen av gjentakelsesfaren skal det legges vekt på det begåtte lovbruddet sammenholdt særlig med lovbrysterens øvrige adferd ... > ... *jamførte* særleg med ... AP
- Straffelovgivningen rammer bare forsettlige lovbrudd med mindre annet er bestemt > ... anna er *bestemde* AP
- er et tiltak truffet av hensyn til noen bestemt persons interesse > ... nokon *bestemde* ... AP, NR, NRKI, NNO
- (bortfall:) å vurdere den regnskapspliktiges framtidige utvikling > ... å vurderer *framtidige utviklinga* til ... AP, NNO

KI-baserte

- Andre kulturmiljø kan oppfattes mer slitte og preget av tiden > ... slitte og *preg* av tida CH 4

20 Manglande retting

Apertium-baserte

- En eventuell *beslutningsmekanismene* ... må også være annerledes enn for åpne data > Ein eventuell *avgjerdsmekanisme* ... må òg vera *annleise* ... (AP, NNO)

(Merknad: KI-verktøya rettar til eintal i substantivet, men CO og CH innfører genusfeil (*ei*), og NRKI har *annleise*.)

- Nivået på lønnskostnadene er generelt mer stabile over tid enn driftsresultatet > ... *stabile* ... NNO 4
- For Riksadvokatembetets del er det særlig behov for å ha tilstrekkelige kapasitet ... > ... *tilstrekkelege* NNO, NR
- (komma:) Enhver som helt eller delvis eier eller driver nett, produksjon eller organisert markeds plass, samt omsettere og sluttbrukere plikter å ... > ... *sluttbrukarar pliktar* å ... NNO, NR (men CO og CH rettar)
- (ulogisk omskriving av *i forhold til*:) Blant annet viser en samlestudie fra 2017 at det var 32–84 pst. høyere mikrobiell biomasse og biologisk aktivitet i økologisk drift sammenlignet med ikke-økologisk drift > ... *samanlikna med* ikkje-økologisk drift NRKI 4

Begge verktøytypar

- betroelser de mottok mens de var en godkjent prest/forstander i et registrert trossamfunn > ... *dei var ein godkjend* ... AP, NNO, NRKI
- (ulogisk/elliptisk:) Forurensningsmyndigheten kan i forskrift eller i det enkelte tilfelle fastsette nærmere krav til avløpsledning, herunder om den skal være lukket og vanntett > ... under dette om ... AP, NNO, NR / mellom anna om ... CO, CH

21 Utelating og overforenkling

KI-baserte

- EØS-loven: fortale til avtale om Det europeiske økonomiske samarbeidsområde (UD) > *EØS-lova – Fortale* CH
- Mange oppga også lite kunnskap om islam i befolkningen og en ensidig fremstilling av muslimer og islam i mediene som viktige årsaker > *Mange peikte òg på lite kunnskap om islam og ei einsidig framstilling av muslimar i media* CH
- Videre kan det å være innvandrere, kvinne og å bruke hijab [...] øke sannsynligheten for å bli utsatt for sammensatt diskriminering > *Samansette grunnlag – som det å vera innvandrar, kvinne og hijabbrukar – kan forsterka risikoen* CH

- 43 prosent av muslimene i utvalget oppga at de «ofte» eller «noen ganger» har opplevd at personer har gitt dem en følelse av at de ikke hører til i det norske samfunnet > ... har fått *kjensla av ikkje å høyra til i det norske samfunnet* CH
- Antallet vigslere bør stå i et rimelig forhold til tros- eller livssynssamfunnets behov for vigslere > Talet på vigslarar bør stå i eit rimeleg forhold til behovet i *samfunnet* CH
- De siste årenes debatt om arkitektur og kvalitet i nybygde nabolag har løftet fram noen svakheter ved siste tiårs foretting i og rundt byer og tettsteder > *Debatten om arkitektur og ...* CH

22 Anna manglende omskriving

Apertium-baserte

- (substantivisk, med kortslutning:) opphaverens rett til å få adgang til åndsverk i andres besittelse > ... *i andres innehaving* AP
- (substantivisk:) Ved vurderingen av om vilkåret er oppfylt, ses det bort fra ... > *Ved vurderinga av om ...* NNO 4
- (venstretungst:) Til medlemmer i trygden mellom 18 og 67 år som har varig redusert inntektsevne med minst 50 pst. på grunn av sykdom, skade eller lyte, ytes uføretrygd > *Til medlemer i trygda mellom 18 og 67 år som har varig redusert inntektsevne med minst 50 pst. på grunn av sjukdom, skade eller lyte, vert uføretrygd ytt* NR
- (lite aktivt:) Utmelding skal alltid kunne skje skriftlig > *Utmelding* skal alltid kunna skje skriftleg NR 4

Begge verktøytypar

- (passivt, med fleire slags feil:) Departementet kan megle i saken etter at partene er gitt adgang til å uttale seg > etter at *partane er gjeven tilgang* til å uttala seg AP, NR, NRKI, NNO
- (substantivisk med kortslutning:) Ved vurderingen av gjentakelsesfaren skal det legges vekt på det begåtte lovbruddet ... > *Ved vurderinga av gjentakingsfaren* skal det leggjast vekt på *det gjorde lovbrøtet* ... NRKI

- (substantivisk:) Forsvaret må være bevisst sin avhengighet av sivile leverandører > ... medvite *avhengnaden sin av* sivile leverandører AP, NNO, NRKI; ... medviten om *si avhengighet* av sivile leverandører CO
- (venstretungst:) Endringer som øker den maksimale hastighet eller ytelse på motorsykel (herunder moped) utover grenser fastsatt for vedkommende gruppe av motorsykler i medhold av annet ledd, må ikke foretas
> *Endringer som aukar [ord-for-ord-omsetjing] må ikkje gjennomførast* AP, NNO, NRKI
> *Endringer som aukar [ord-for-ord-omsetjing] må ikkje gjerast* CH, CO
- (substantivisk:) utvise nødvendig aktsomhet > utvisa naudsynt aktsemd AP, NNO; syta for naudsynt aktsemd CH
- (venstretungst med kommafeil:) Arten av forskning og utvikling som drives skal angis > *Arta av forskning og utvikling som blir driven skal angjevast* [e.l.] AP, NNO, NRKI (*oppgevast*), CH (*oppgivast*; rettar komma)
- (venstretungst og passivt:) Eiendom, bygning eller løsøre hvor det er mistanke om smittestoffer eller planteskadegjørere, kan pålegges rengjort, desinfisert eller destruert mv.
> [Ord-for-ord-omsett] *kan påleggjast reingjort, desinfisert eller destruert osb.* AP, NNO, NR, NRKI
> [Ord-for-ord-omsett] *kan påleggjast reinvasking, desinfisering eller destruksjon m.m.* CH
- (uklårt:) Blir godkjenning nektet uten at framleierens forhold gir saklig grunn til det ... > ... *forholdet til framleigaren* ... AP, NR, NNO; ... *forholda til framleigaren* ... NRKI; ... *framleigaren sitt forhold* ... CO; ... *framleigaren sine forhold* ... CH 4
- (substantivisk og uidiomatisk:) Det må fremmes bevissthet og forsvarsvilje for å styrke motstandsdyktigheten i befolkningen, redusere uakseptable sårbarheter og på sikt styrke robusthet og utholdenhet > *redusere uakseptable sårbarheter og på sikt styrkje robustheit og uthald* ALLE
- (oppstylda:) Foreligger det grunn til mistanke ... > *Føreligg / Ligg det føre* ... ALLE
- (oppstylda:) Finner kommunen ikke å kunne ta hensyn til innsigelsen, skal statsforvalteren ordinært mekle mellom partene > *Finn kommunen ikkje å kunna* ... (o.l.) ALLE

- (med passivproblem og alvorlege kortslutningar:) Dette gjelder likevel ikke så langt utleieren godtgjør at forsinkelsen eller mangelen skyldes en hindring utenfor utleierens kontroll som utleieren ikke med rimelighet kunne ventes å ha tatt i betraktning på avtaletiden eller å ha unngått eller overvunnet følgene av
 - > [Ord for ord-omsett] som utleigaren ikkje innanfor rimelege grenser kunne *ventast å ha teke i betraktning* på avtaletida eller å ha unngått eller overvunne følgjene av AP
 - > [Ord for ord-omsett] som utleigaren ikkje *med rimelege grenser* kunne *ventast å ha teke i betraktning* på avtaletida eller å ha unngått eller overvunne følgjene av NRKI
 - > [Ord for ord-omsett] som utleigaren ikkje med *rimeleg grunn* kunne *venta å ta i betraktning* på avtaletida eller unngå eller overvinna følgjene av CH

