

Test av skriveprogrammer på norsk

mars 2026

Språkrådets test av skriveprogrammer på norsk

mars 2026

Sammendrag	2
Innledning.....	3
Kravet i opplæringsloven	4
Metode	6
Mer om de enkelte leverandørene	6
Automatisk og menneskelig evaluering.....	9
Sammenlikning med metoden i tilsvarende tester i 2025.....	10
Testtekstene	10
Evalueringsmål (metrikker).....	12
Resultater	15
Analyse av resultatene	18
Hovedfunn.....	18
Drøfting og implikasjoner.....	21
Referanser	23

Sammendrag

Skriveprogrammer er mye brukt i skolen og i arbeidslivet, og de spiller en viktig rolle i tekstarbeid på norsk. Opplæringsloven fra 2024 inneholder krav om at skriveprogrammer som brukes i skolen, skal fungere på bokmål og nynorsk. Språkrådet har testet de mest brukte programmene for å undersøke om de oppfyller dette kravet. Denne rapporten viser resultatene fra denne testen.

Språkrådets konklusjoner, kort oppsummert:

- Ingen av stavekontrollene i skriveprogrammene Språkrådet har testet, innfrir kravene som opplæringsloven stiller, heller ikke de som er mest brukt i norsk skole.
- Stavekontrollene fungerer noe dårligere på nynorsk enn på bokmål.
- Stavekontrollene overser mange feil og markerer i flere tilfeller korrekt skrevne ord som feil.
- Når stavekontrollene markerer noe som feil, foreslår de ofte alternativer som ikke er riktige.
- Endringene fra forrige test er marginale.
- De nye stavekontrollene Språkrådet har testet for første gang i denne runden, står ikke noe tilbake for stavekontrollene i de etablerte skriveprogrammene fra Microsoft, Apple og Google.

Innledning

I den nye opplæringsloven, som trådte i kraft 1. august 2024, er det tatt inn en egen bestemmelse om skriveprogrammer, § 15-4 andre ledd. Loven stiller likevel ingen krav om formell godkjenning av skriveprogrammer som skal brukes i skolen, slik tilfellet er for norskspråklige ordlister og ordbøker. Slike ordlister og ordbøker skal etter lovens § 15-4 første ledd godkjennes av Språkrådet.

Hverken opplæringsloven eller forarbeidene til loven gir noen presis beskrivelse av hva som menes med et skriveprogram. I stortingsproposisjonen står følgende:

Uttrykket «skriveprogram» omfatter ulike former for læringsverktøy og tekstbehandlingsprogram med tekststøtte, rettefunksjoner o.l. (Prop. 57 L (2022–2023), s. 631)

For både 2025 og 2026 har Språkrådet fått følgende oppdrag fra Kultur- og likestillingsdepartementet:

[...] bidra til at aktører som utvikler og tilbyr språkteknologiske verktøy og tjenester, er kjent med de kravene som stilles til slike verktøy og tjenester i offentlig bruk

[...] bidra til at de får innspill og dokumentasjon som bidrar til at produktene de tilbyr, har en tilfredsstillende språklig kvalitet

[...] Språkrådets rådgivings- og veiledningsvirksomhet [skal] omfatte hvilke språkpolitiske krav og anbefalinger og hvilke språklige aspekter det offentlige bør ta hensyn til i planlegging, utvikling, anskaffelse og bruk av språkteknologi og kunstig intelligens. (Tildelingsbrev 2026)

På bakgrunn av kravet i opplæringsloven § 15-4 og Språkrådets oppdrag fra Kultur- og likestillingsdepartementet har Språkrådet i mars 2026 gjennomført testing av stave- og grammatikkontrollen i sju ulike produkter. Vi har valgt å teste nettopp stave- og grammatikkontrollene fordi det er enkelt å påvise om disse støtter både bokmål og nynorsk, og om de følger rettskrivingen.

Dette er den tredje testen (se også Skriveprogramtest 2025) av de tre skriveprogrammene som til sammen har om lag 90 prosent av markedet i den norske skolen (LNK-undersøkelsen 2024):

- *Dokumenter (Docs)* fra Google
- *Word* fra Microsoft
- *Pages* fra Apple

I tillegg har vi testet fire andre produkter:

- *Writer* fra LibreOffice, en komplett, åpen kildekode-basert kontorpakke med samme bruksområder som Microsoft Office
- *Korrektoren* fra Lingtechapps, norsk korrekturverktøy, upublisert medio juni 2026
- *IntoWords* fra Vitec, program for lese- og skrivestøtte
- *Skoleskrift* fra Karin Ohlis AB, svensk skriveprogram kun for Apple-produkter

Hensikten med testen har vært å finne ut om stave- og grammatikkontrollen (heretter kalt stavekontrollen) i skriveprogrammene oppfyller opplæringslovens krav. Språkrådet ønsker å

- bidra til at stavekontrollene blir bedre, ved å opplyse leverandørene av programmene om lovkravene og hva de innebærer, samt om utfordringene de enkelte stavekontrollene har
- opplyse skoler og myndigheter (dem som lovkravene er rettet mot) om hvor godt eller dårlig stavekontrollene fungerer, slik at det kan tas opplyste valg som ikke resulterer i at opplæringsloven brytes, og slik at både bokmåls- og nynorskelever kan få best mulige skriveprogrammer
- vise de store aktørene at det er mulig å utvikle en pålitelig stavekontroll (med tanke på at en eller flere av de mindre aktørenes stavekontroller fungerer bedre)

Testene vi gjennomførte i 2025, viste at hverken Word, Dokumenter eller Pages da innfridde opplæringslovens krav til skriveprogram. Leverandørene fikk individuelle tilbakemeldinger. Vi gjorde noen erfaringer etter den runden med tester:

- Vi fikk lagt press på leverandørene og vist kvantitativt at skriveprogrammene ikke er gode nok.
- Vi opplevde at leverandørene ønsker å arbeide for at skriveprogrammene deres oppfyller lovkravet.

Kravet i opplæringsloven

Andre ledd i opplæringsloven § 15-4 lyder slik:

Skolen skal bruke skriveprogram som støtter både bokmål og nynorsk, og som følger offisiell rettskriving. (Opplæringsloven 2024)

Begrepet «offisiell rettskriving» er ikke nærmere forklart i lovteksten eller forarbeidene til loven, men i forarbeidene til språkloven er det understreket at offisiell rettskriving omfatter både regler for staving og bøyning av ord, og formelle skriveregler, for eksempel tegnsetting og bruk av stor og liten forbokstav (jf. Prop 108 L (2019–2020), s. 153).

Hverken lovteksten eller forarbeidene til loven presiserer nærmere *hvor godt* et skriveprogram må fungere for at opplæringslovens krav skal være oppfylt. Språkrådet oppfatter ordlyden «følger offisiell rettskriving» som et absolutt krav, det vil si at

stavekontrollen skal oppdage og markere alle feil, og at den ikke skal feilmarkere der det ikke er noen feil. Kravet kan synes strengt, men når Språkrådet vurderer ordbøker og ordlister til godkjenning for skolebruk, påtaler vi alle feil vi finner, og venter at de blir rettet opp. Det er en forutsetning for at disse ressursene skal få godkjenning. Det er ikke urimelig å tro at stavekontrollene vil bli en alvorlig konkurrent til ordbøkene som en veileder for rettskriving. Vi mener derfor at det ikke bør stilles lavere krav til kvaliteten på stavekontroller enn til kvaliteten på ordbøker, selv om loven ikke krever godkjenning av skriveprogrammer som skal brukes i skolen.

På bakgrunn av det ovenstående legger Språkrådet følgende forståelse av ordlyden i lovbestemmelsen til grunn:

- At skriveprogrammene støtter både bokmål og nynorsk, betyr at funksjonene i skriveprogrammene skal virke like godt på begge språkene. Både bokmål og nynorsk må være med i lista over korrekturspråk.
- At skriveprogrammene følger offisiell rettskriving, betyr at de
 - følger reglene for staving og bøyning av ord, slik de er gitt på ordbøkene.no (også kalt «standardordbøkene»)
 - følger formelle skriveregler om tegnsetting, bruk av store og små bokstaver m.m.
 - følger norske regler for skriving av sammensatte ord

For stavekontrollen i skriveprogrammene betyr dette nærmere bestemt

- at kontrollen må markere alle feil i staving og bøyning, alle brudd på formelle skriveregler om tegnsetting og bruk av store og små forbokstaver m.m. og alle brudd på norske regler for skriving av sammensatte ord
- at kontrollen *ikke* skal markere skrivemåter som er godtatt i normen, som feil, heller ikke skrivemåter som er uvanlige
- at nynorskkontrollen ikke skal akseptere bokmålsord i nynorsktekst, og vice versa
- at minst ett av de alternative forslagene som kommer opp når kontrollene markerer noe som feil, skal være korrekt

Det siste punktet må sees på som et liberalt krav. Les mer om dette i metodedelen.

Metode

Basert på kriteriene for en vellykket stavekontroll slik de er skissert ovenfor, ville vi måle

- hvor god stavekontrollen er til å identifisere faktiske feil
- hvor ofte forslagene fra stavekontrollen er riktige
- hvor god stavekontrollen er til å ikke feilmarkere korrekt språk

Vi laget korte testtekster som vi matet inn i skriveprogrammene. Noen av tekstene hadde med hensikt høy konsentrasjon av feil, slik at vi kunne se om stavekontrollene identifiserte feilene og kom med gode forslag til alternativer. Andre tekster inneholdt korrekte former som vi mistenkte kunne være utfordrende for stavekontrollen å identifisere fordi formene er mindre vanlige i bruk. Vi kaller disse for «feller».

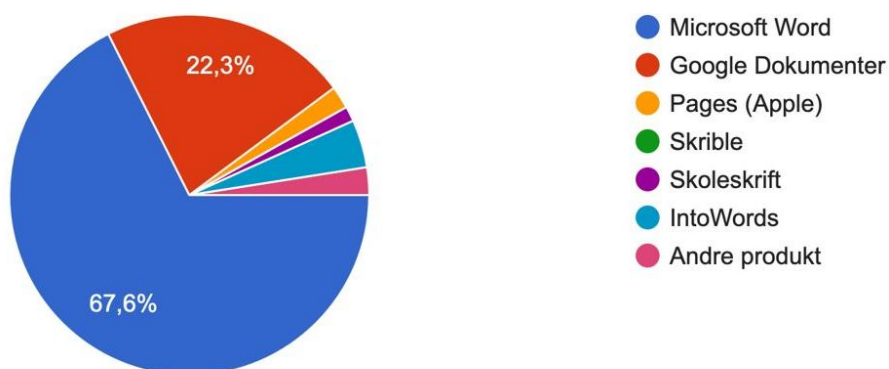
Resultatene i denne rapporten er sammenlikninger av markeringene fra stavekontrollene opp mot fasit.

Vi kunne testet flere aspekter ved stavekontrollene, for eksempel hastighet eller brukertilfredshet. Det har vi valgt å ikke gjøre i denne undersøkelsen, blant annet fordi det ikke er relevant for lovkravet.

Mer om de enkelte leverandørene

Kva skriveprogram bruker elevane dykkar mest til skulearbeid i dette skuleåret?

444 svar



Figur 1: Resultater fra undersøkelsen til Landssamanslutninga av nynorsk kommunar, LNK, 25. september 2024

Da vi skulle velge ut hvilke skriveprogrammer vi skulle teste, tok vi utgangspunkt i leverandørene som er nevnt i en undersøkelse gjort av Landssamanslutninga av

nynorsk kommunar (LNK) i 2024 (se figur 1). Word, Pages og Dokumenter har alene en markedsandel på ca. 90 prosent i norsk skole. I alle programmene limte eller skrev vi inn teksten. Vi ba alle leverandørene hjelpe oss med å få tilgang til det samme grensesnittet til skriveprogrammet/stavekontrollen som en norsk skoleelev ville fått.

Nedenfor lister vi opp overordnet informasjon om programmene vi har testet, i tillegg til enkelte opplysninger som kan være relevante for brukerne av de ulike programmene.

Microsoft Word

Testdato: 23. februar 2026

Versjon: nettversjon på office.com

Word er det absolutt mest brukte skriveprogrammet i norsk skole (LNK-undersøkelsen 2024). Programmet kommer ofte til sluttbrukeren som del av en større Office-pakke. Vi brukte nettversjonen av Word, siden leverandøren mente at oppdateringer i stavekontrollen blir lansert der først. Stavekontrollen i Word er ved siden av Dokumenter den eneste av dem vi har testet, som skiller mellom røde og blå feilmarkeringer. De blå skal i utgangspunktet markere ulike typer grammatiske feil, mens de røde skal vise stavefeil.

Apple Pages

Testdato: 27. februar 2026

Versjon: iPadOS 26.3, Pages 14.3

Pages er en skriveprogram-app fra Apple som gjerne er forhåndsinstallert på for eksempel nettbrett som brukes i skolen. Pages bruker en stavekontroll som er felles for hele operativsystemet, altså er den ikke appspesifikk. Vi erfarer at enkelte iPad-brukere bruker andre skriveprogrammer enn Pages, for eksempel Apple-tilpasset Microsoft Word.

IntoWords

Testdato: 25. mars 2026

Versjon: skrivebordsversjonen

IntoWords er et skrive- og lesestøtteprogram levert av selskapet Vitec. Det er ikke først og fremst et skriveprogram, og vi testet det gjennom funksjonaliteten «Grammatikk». Det finnes også en nettleserversjon av programmet, men vi lastet ned skrivebordsversjonen med språkpakken «Norsk», som inneholdt både bokmål og nynorsk.

Korrektoren (obs: ikke tilgjengelig medio juni 2026)

Testdato: 18. mars 2026

Versjon: beta, nettversjon (egenutviklet redigeringsprogram) i nettleseren Google

Korrektoren er et upublisert korrekturverktøy utviklet av Lingtechapps (Torbjørn Nordgård). Intensjonen er at produktet ikke skal være et eget skriveprogram, men en såkalt «plug-in» som kan kobles på andre skriveprogrammer. Utvikleren sier at produktet er myntet på profesjonelle brukere og vil kreve en avgift. Vi fikk tilgang til et midlertidig grensesnitt som vi kunne bruke til å gjennomføre testene våre.

Utvikleren tok selv kontakt med Språkrådet i 2025 for å få svar på en del spørsmål han hadde i forbindelse med utviklingen. Grensesnittet vi fikk tilgang til, hadde enkelte åpenbare mangler. Blant annet ble innlimt tekst håndtert på en uhensiktsmessig måte (stavekontrollen krevde stor forbokstav etter hvert linjeskift). Dette påpekte vi for utvikleren allerede i 2025. En del av alternativforslagene ble presentert som én streng, uten mellomrom mellom forslagene.

Årsaken til at vi har valgt å publisere resultatene for Korrektoren, selv om den ikke er et tilgjengelig produkt, er at vi ønsker å vise de største aktørene at det ikke er en umulig oppgave å lage en stavekontroll som er i samsvar med norsk rettskriving. Resultatene i neste del viser at på grunn av dedikert arbeid fra utvikleren gir Korrektoren det klart beste resultatet på våre tester.

LibreOffice (Writer)

Testdato: 4. mars 2026

Versjon: 26.2

LibreOffice er i dag det mest aktuelle alternativet med åpen kildekode (open-source) til Microsoft Office som skrivebordspakke. Pakken inneholder blant annet et klassisk skriveprogram. LibreOffice kan lastes ned og brukes gratis. Dette prosjektet er en arvtager til OpenOffice, som tidligere dominerte denne nisjen. Så vidt Språkrådet erfarer, er ikke OpenOffice like aktuelt lenger, og vi klarte heller ikke å laste ned en norsk språkpakke for OpenOffice.

Et problem i stavekontrollen til LibreOffice er at det enkelte ganger er veldig mange alternativforslag, opptil 15 stykker. Selv om noen av dem er riktige, forsvinner noe av nytten ved alternativforslag når det er så mange.

[Alle kan bidra](#) i utviklingen av LibreOffice-produkter.

Skoleskrift

Testdato: 27. februar 2026

Versjon: iPadOS 26.3, Skoleskrift 2

Skoleskrift 2 er et enkelt skriveprogram med bokstavlyd, og vi lastet ned appen via Apples Appstore. Appen fungerer kun på Apple-produkter. Vi merker oss at stavekontrollen oppfører seg påfallende likt som stavekontrollen i Apple Pages.

Vi vil løfte fram to overordnede problemer/begrensninger vi støtte på under testingen. For det første skrudde ikke stavekontrollen seg på hvis vi limte inn tekst, så vi måtte skrive inn alt manuelt. For det andre fikk vi ikke koblet ut autokorrekturefunksjonen, så vi måtte manuelt følge med på om ord ble rettet eller ikke. Adferdsmønsteret for når autokorrekturen slo seg på, og når det bare var stavekontroll (rød strek), forble uklart for oss. Dette gjorde resultatene noe usikre.

Google Dokumenter

Testdato: 20. mars 2026

Versjon: docs.google.com

Vi erfarer at Googles skriveprogram Dokumenter i skolen benyttes gjennom plattformen Google for Education. Etter kontakt med leverandøren forstår vi det slik at det er samme stavekontroll som benyttes på tvers av plattformer, så vi testet Dokumenter i den versjonen som er tilgjengelig via docs.google.com.

Stavekontrollen i Dokumenter er ved siden av Word den eneste av dem vi har testet, som skiller mellom røde og blå feilmarkeringer. De blå skal i utgangspunktet markere ulike typer grammatiske feil, mens de røde skal vise stavfeil.

Automatisk og menneskelig evaluering

Språkteknologiske verktøy kan evalueres automatisk, men også manuelt av mennesker. I denne testen har vi valgt menneskelig testing, fordi det mangler en løsning for automatisk testing av stavekontroller på norsk, og fordi eventuelle mangler i en stavekontroll kan identifiseres allerede på små kunstige datasett. Det er dermed ikke spesielt ressurskrevende å evaluere verktøyene manuelt. Vi lener oss på Språkrådets egen fagkompetanse som fasit.

I språkteknologifaget som helhet er automatisk evaluering, såkalte *referansemålinger* (benchmarks), standarden. Fordelen med automatiske evalueringsmetoder er at de er etterprøvbare, og at det er mulig å teste nye modeller og programmer uten ekstra kostnad når det automatiske systemet først er laget. En referansemåling består av datasett som det testes på, og utvalgte evalueringsmål.

Nedenfor følger en kort beskrivelse av en tilnærming til (automatisk) referansemåling av stavekontroller på islandsk og et datasett som kanskje kunne vært brukt til å utvikle en referansemåling for norsk. Disse er tatt med som inspirasjon til mulig framtidig arbeid

med å utvikle referansemåling og som begrunnelse for hvorfor vi mente at å lage en manuell test var den beste løsningen.

Forskere ved Universitetet på Island (Háskóli Íslands) publiserte i 2025 et datasett ment for referansemåling av stavekontroller og grammatikkontroller for islandsk språk (Ármannsson et al. 2025). Testsettet består av parallellstilte setninger, der to og to setninger er identiske med det unntaket at den ene inneholder en feil, mens den andre ikke gjør det (fasit). En stavekontroll eller språkmodell testes ved at man mater inn setningen som inneholder feil, og sammenlikner det stavekontrollen eller språkmodellen gir tilbake, med fasitsetningen. I teorien kan da en ny stavekontroll eller språkmodell evalueres med kun et tastetrykk. Denne automatiske evalueringsmetoden har likevel et grunnleggende problem ved seg: Den gir bare rom for én fasit per setning, og all utdata som skiller seg fra fasiten, vil trekke ned skåren. Referansemålingen egner seg derfor dårlig for den norske språksituasjonen, der det finnes flere alternative og likestilte skrivemåter for mange ord.

For øyeblikket finnes det altså ingen formålstjenlig referansemåling for norsk stavekontroll, hverken for bokmål eller for nynorsk. Norsk andrespråskorpus (ASK) (en samling tekster på norsk skrevet av andrespråksbrukere, med merknader fra en lærer) har blitt konvertert til et parallelldatasett som likner på det i Ármannsson et al. (Jentoft 2023), men a) det er bare én fasit per setning, b) datasettet er ikke gjennomgått i detalj av et menneske, og c) feilene er ikke markert på ordnivå, men på setningsnivå. Derfor kan vi ikke bruke dette datasettet i denne undersøkelsen.

Sammenlikning med metoden i tilsvarende tester i 2025

Vi har i hovedsak benyttet samme metode og evalueringsmetrikker som i testene i 2025 (Skriveprogramtest 2025). Men vi har gjort mindre endringer i hvilke feiltyper vi måler, og selvsagt byttet ut testtekstene.

Siden testtekstene er byttet ut og «vanskelighetsgraden» kan være ulik, velger vi å ikke legge veldig stor vekt på å sammenlikne resultatene fra 2025 og 2026. Endringene er likevel tatt med i resultatdelen. Det mest sentrale vil uansett være å undersøke om programmene oppfyller lovkravet i dag.

Testtekstene

For å sammenlikne kvaliteten på skriveprogrammene med kravet i opplæringsloven har vi laget et sett med tekster som vi bruker til å se hvor godt stavekontrollene presterer i møte med feil og feller. I den ene teksten har vi lagt inn feil av ulike slag. Den andre teksten er feilfri, men inneholder en del ord som også har andre tillatte skrivemåter enn dem vi har brukt (se vedlegg).

Test A (full av feil)

Her har vi tekster som inneholder ulike typer feil. I dette eksempelet på nynorsk er feil markert i rødt:

- Kunstløp har aldri vore **noken** god **øvelse** for **Norge**.

Test/tekst A	bokmål	nynorsk
ca. 200 ord	36 innlagte feil	37 innlagte feil

Med test A tar vi sikte på å sjekke om stavekontrollen fanger opp det som ikke er i tråd med rettskrivingen. Her skal stavekontrollen markere alle innlagte feil i tekstene med farget strek. I tillegg sjekker vi at minst ett av de alternative forslagene som kommer opp ved feilmarkeringene, er korrekt.

De innlagte feilene i tekstene er av ulike slag: stavefeil, bøyingsfeil, samsvarsfeil, samskrivingsfeil, og/å-feil, feil i bruk av stor eller liten forbokstav, kommafeil, bruk av feil preposisjon, foranstilt eiendomspronomen (i nynorsktekst), bokmålsord (i nynorsktekst) og nynorskord (i bokmålstekst).

Noen av disse feilkategoriene faller strengt tatt ikke inn under rettskrivingen, slik vi har definert begrepet tidligere i rapporten. Vi har likevel tatt dem med fordi de representerer brudd på grammatiske normer som de fleste vil oppfatte som feil.

En feilkategori vi valgte å droppe i denne testrunden, er inkonsekvent bruk av *-a* eller *-e* som infinitivsending i nynorsktekst. Ingen av stavekontrollene har markert slik inkonsekvens som feil i tidligere tester, så vi så det som lite hensiktsmessig å ta det med nå. Men inkonsekvent bruk av infinitivsending faller inn under den strenge definisjonen av rettskriving, så vi vil ta feilkategorien inn igjen i framtidige tester.

Test B (full av feller)

Her har vi feilfrie tekster som inneholder en del former og skrivemåter som alle er tillatt innenfor rettskrivingen, men som er mindre brukt enn andre alternativer. Disse kaller vi *feller*. I dette eksempelet på bokmål er feller markert i grønt:

- Mange **blautdyr gjømmer** seg nede på **botnen**.

Test/tekst B	bokmål	nynorsk
ca. 140 ord	30 feller	30 feller

Med test B tar vi sikte på å sjekke at stavekontrollen ikke markerer feil der det ikke er gjort feil. Siden programmet ikke skal markere noe i det hele tatt, skal det heller ikke komme opp alternative forslag.

En av grunnene til å gjøre denne delen av testen er at vi mistenker at stavekontrollene kan favorisere hyppigere brukte former av ord, noe som vil kunne føre til en såkalt normativ ensretting av språket. Et eksempel kan være at *hoppa* rettes til *hoppet*. Enkelte brukere av skriveprogrammene mener å ha sett at programmene oftere markerer uvanlige, men riktige former de siste årene. Dette kan være et tegn på at KI og maskinlæring er involvert, med den konsekvensen at statistisk frekvens i datasett styrer programmenes atferd.

Evalueringsmål (metrikker)

Stavekontroll som oppgave kan sees på som en type informasjonsgjenfinning (information retrieval), en prosess der det gjelder å finne relevant informasjon i strukturerte eller ustrukturerte data. I vårt tilfelle er testtekstene ustrukturerte data, og informasjonsgjenfinningen går ut på å markere eller ignorere ord eller elementer. Alle evalueringsmålene vi bruker, er *feilbaserte*. Det vil si at vi bare ser på feilene eller fellene som er lagt inn i tekstene, når vi teller, og ignorerer ordene uten feil. Dette står i kontrast til *leksikalske* evalueringsmål, der alle ordene i teksten ville vært med i tellingen.

De vanligste evalueringsmålene i informasjonsgjenfinning er *gjenfinningsgrad* (recall), *presisjon* (precision) og ulike varianter av det sammensatte målet *f-skår* (Huyssteen et al. 2004). *F-skår* består av gjenfinningsgrad og presisjon. Alle evalueringsmålene vi bruker, har en skala fra 0 til 1, der 1 er best. I denne rapporten visualiserer vi ofte flere av disse skårene i samme graf, men det er viktig at leseren husker på at det er snakk om ulike mål på samme y-akse. Alle disse målene kan omgjøres til prosenter, der for eksempel 0,5 tilsvarer 50 prosent.

Test A (full av feil)

For å finne ut hvor mange av de faktiske feilene stavekontrollen klarer å markere, bruker vi evalueringsmålet *gjenfinningsgrad for feil* (error recall). Vi er ute etter å sammenlikne antallet feil som stavekontrollen har fått med seg, med det totale antallet feil i teksten. Vi deler antallet korrekte feilmarkeringer (true positive, TP) på det totale antallet feil i teksten (TP + false negative FN). Skalaen går fra 0 til 1, der 1 er best. Det er verdt å merke seg at feilmarkeringer på korrekt skrevne ord her blir ignorert.

- *Gjenfinningsgrad for feil*, eksempel: 10 feil i teksten, 8 av disse markert av stavekontrollen = $8 / (8 + 2) = 0,8$.

Test A (forslagene)

For å si noe om kvaliteten på forslagene fra stavekontrollen bruker vi evalueringsmålet *presisjon*, heretter kalt *forslagspresisjon*. Til dette formålet finnes det også mer finmaskede og strengere evalueringsmål (se Huyssteen et al. 2004), men i vår variant ser vi bare på om minst ett av forslagene fra stavekontrollen er korrekt. Vi deler antallet

markeringer som inneholder minst ett korrekt forslag (TP), på det totale antallet feilmarkeringer (TP + false positive FP). Skalaen går fra 0 til 1, der 1 er best.

- *Forslagspresisjon*, eksempel: 10 feilmarkeringer, 5 av disse har relevante og korrekte forslag til alternativ = $5 / (5 + 5) = 0,5$.

Fra skoleelevens perspektiv er det best om det ikke gis gale forslag i det hele tatt, men vi har ikke tatt hensyn til dette i vår evaluering. Dette er altså et ganske liberalt mål.

Test B (full av feller)

For B-testene bruker vi en tilpasset variant av evalueringsmålet presisjon kalt *feilpresisjon* (error precision) (Huyssteen et al. 2004). Vi er ute etter å sammenlikne antallet feilmarkeringer programmet har gjort, med det totale antallet feller i teksten. Vi deler antallet feller stavekontrollen korrekt har ignorert, altså antallet feller minus antallet markeringer (TP), på det totale antallet feller (TP + FP). Skalaen går fra 0 til 1, der 1 er best.

I evaluering av stavekontroller er det vanlig å bruke en annen type presisjon, der man ser på stavekontrollens feilmarkeringer og hvor mange av disse som er korrekte. Men dette fungerer ikke i vårt tilfelle, fordi tekstene våre ikke inneholder noen feil som skulle vært markert, og da måtte man hatt 0 i telleren.

Selv om vi her konsentrerer oss mest om de fellene vi *bevisst* har lagt inn, vil vi påpeke at flere av stavekontrollene i test B også feilmarkerer ord som ikke var ment å være feller. Sistnevnte feilmarkeringer er også regnet med i TP-utregningen. Om vi hadde ignorert dem, kunne det ha ført til en alvorlig skjevhet i resultatene. En stavekontroll som gjør én feilmarkering (på en av fellene), vil komme dårligere ut enn en som gjør ti markeringer, dersom ingen av de ti blir gjort på noen av fellene. Denne usikkerheten er en av grunnene til at vi har vektet denne skåren lavere enn gjenfinningsgrad for feil (test A).

- *Feilpresisjon* (error precision), eksempel: 10 feller i teksten, (minimum) 7 av disse er unngått av stavekontrollen = $7 / (7 + 3) = 0,7$.

Språkrådet-skår

For å få et helhetsbilde og en totalskår for hver enkelt stavekontroll tar vi et vektet gjennomsnitt av gjenfinningsgrad for feil (test A, 60 prosent), forslagspresisjon (test A, 25 prosent) og feilpresisjon (test B, 15 prosent). Den viktigste oppgaven for en stavekontroll er å finne og markere det som er feil, og derfor er gjenfinningsgrad vektlagt mest. Forslagspresisjon er vektlagt noe mindre, både fordi vårt evalueringsmål er ganske liberalt, og fordi det er stor variasjon i hvor langt unna det korrekte svaret de ulike «gale» svaralternativene er. Vi vekter feilpresisjonen fra del B minst, siden stavekontrollene i den testen også har feilmarkert ord som ikke var ment å være feller,

noe som skaper noe usikkerhet. Vi tror denne vektingen samsvarer med den allmenne oppfatningen av hva de viktigste formålene med en stavekontroll er.

F-skår egner seg i vårt tilfelle ikke til å kombinere presisjon og gjenfinningsgrad, fordi målingene ikke er gjort på de samme dataene (test A og test B bruker ulike tekster).

Skåren går fra 0 til 1, der 1 er best.

- *Språkrådet-skår*, eksempel:
gjenfinningsgrad for feil (test A) = 0,8
forslagspresisjon (test A) = 0,3
feilpresisjon (test B) = 0,4
Språkrådet-skår = $(0,6 \times 0,8) + (0,25 \times 0,3) + (0,15 \times 0,4) = 0,62$.

Resultater

Testrunde 3 - bokmål (endring fra forrige testrunde i parentes)

		Test A: full av feil		Test B: full av feller
	Språkrådet-skår	Gjenfinningsgrad for feil	Forslagspresisjon	Feilpresisjon
M-Word	0,63 (+0,00)	0,53	0,68	0,97
A-Pages	0,56 (+0,14)	0,42	0,69	0,90
G-Dokumenter	0,70 (+0,04)	0,64	0,86	0,67
IntoWords	0,59	0,36	0,90	0,97
Skoleskrift	0,55	0,39	0,73	0,87
LibreOffice	0,61	0,44	0,81	0,93
Korrektoren	0,74	0,64	0,83	1,00

Tabell: Matias Jentoft, Språkrådet • Laget med Datawrapper

Tabell 1: resultater for bokmål, tabellformat

Testrunde 3 - nynorsk (endring fra forrige testrunde i parentes)

		Test A: full av feil		Test B: full av feller
	Språkrådet-skår	Gjenfinningsgrad for feil	Forslagspresisjon	Feilpresisjon
M-Word	0,54 (+0,06)	0,46	0,47	1,00
A-Pages	0,56 (+0,06)	0,41	0,67	0,97
G-Dokumenter	0,57 (+0,02)	0,54	0,58	0,70
IntoWords	0,60	0,46	0,71	1,00
Skoleskrift	0,56	0,41	0,67	0,97
LibreOffice	0,61	0,46	0,72	1,00
Korrektoren	0,75	0,68	0,80	0,97

Tabell: Matias Jentoft, Språkrådet • Laget med Datawrapper

Tabell 2: Resultater for nynorsk, tabellformat

Resultater BOKMÅL

■ Språkrådet-skår ■ Gjenfinningsgrad for feil ■ Forslagspresisjon ■ Feilpresisjon

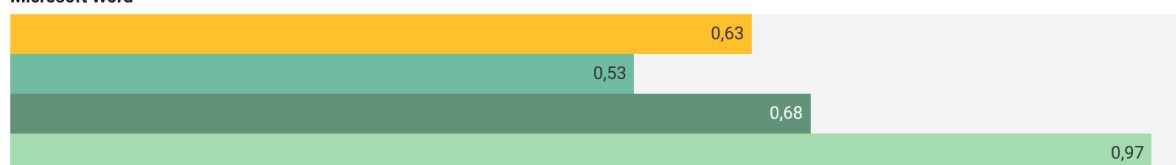
Korrektoren



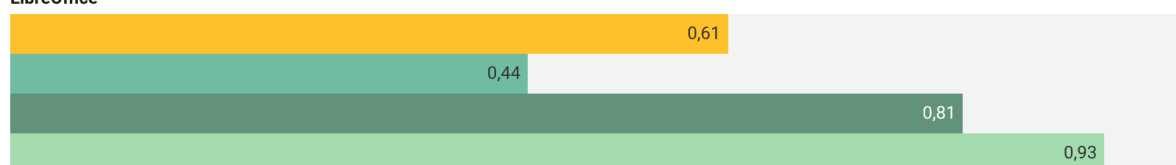
Google Dokumenter



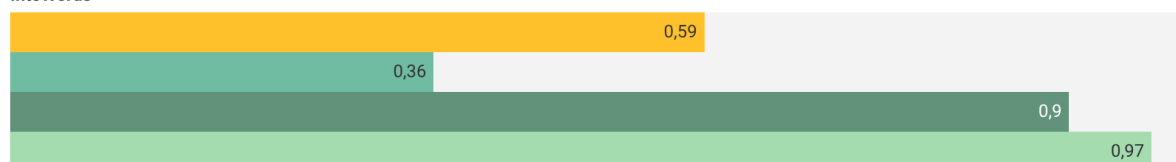
Microsoft Word



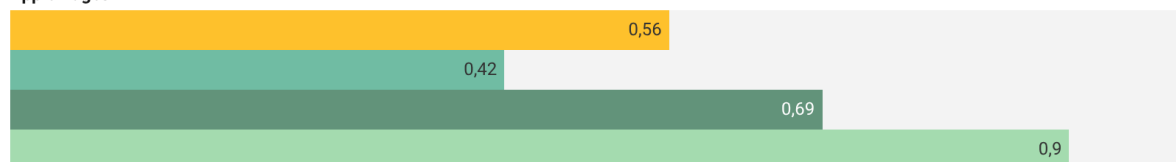
LibreOffice



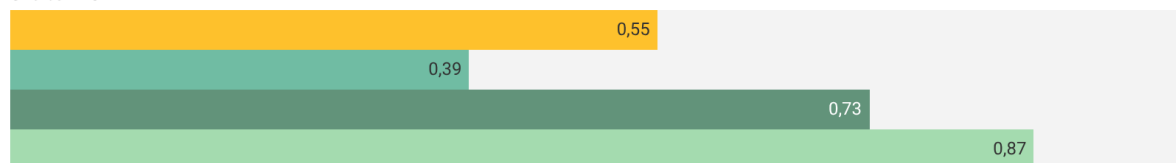
IntoWords



Apple Pages



Skoleskrift



Grafikk: Matias Jentoft, Språkrådet • Laget med Datawrapper

Figur 2: Resultater for bokmål, søylediagram

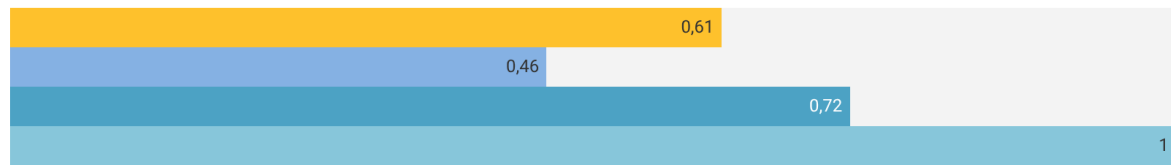
Resultater NYNORSK

■ Språkrådet-skår ■ Gjenfinningsgrad for feil ■ Forslagspresisjon ■ Feilpresisjon

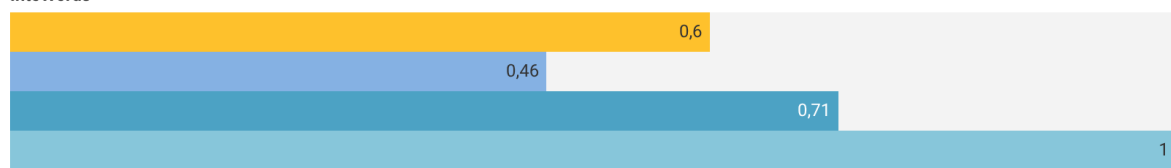
Korrektoren



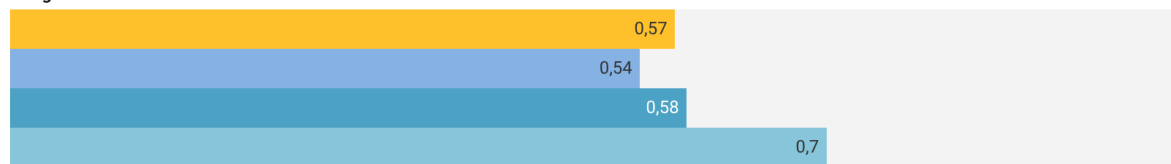
LibreOffice



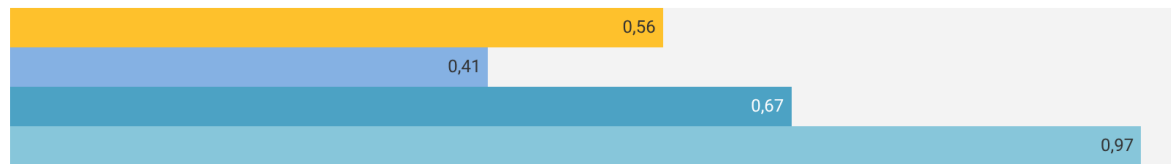
IntoWords



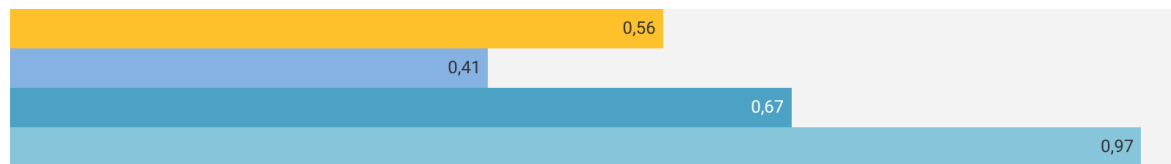
Google Dokumenter



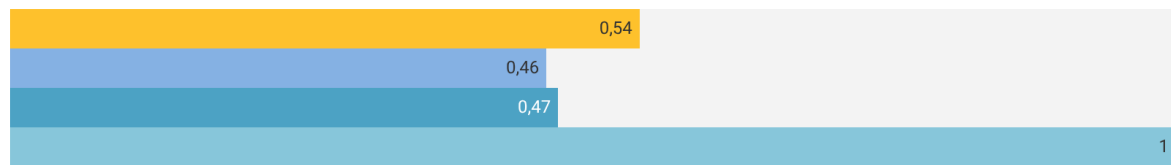
Apple Pages



Skoleskrift



Microsoft Word



Grafikk: Matias Jentoft, Språkrådet • Laget med Datawrapper

Figur 3: Resultater for nynorsk, søylediagram

Analyse av resultatene

Hovedfunn

Som nevnt i metodedelen må vi være litt forsiktige med å si så mye om utviklingen fra forrige testrunde, siden testdataene var ulike. Likevel har vi valgt å sammenlikne resultatene for Word, Pages og Dokumenter med tidligere testresultater, for vi vil demonstrere hvor lite som egentlig har skjedd med disse programmene, til tross for informasjonen og rådene de har fått fra Språkrådet.

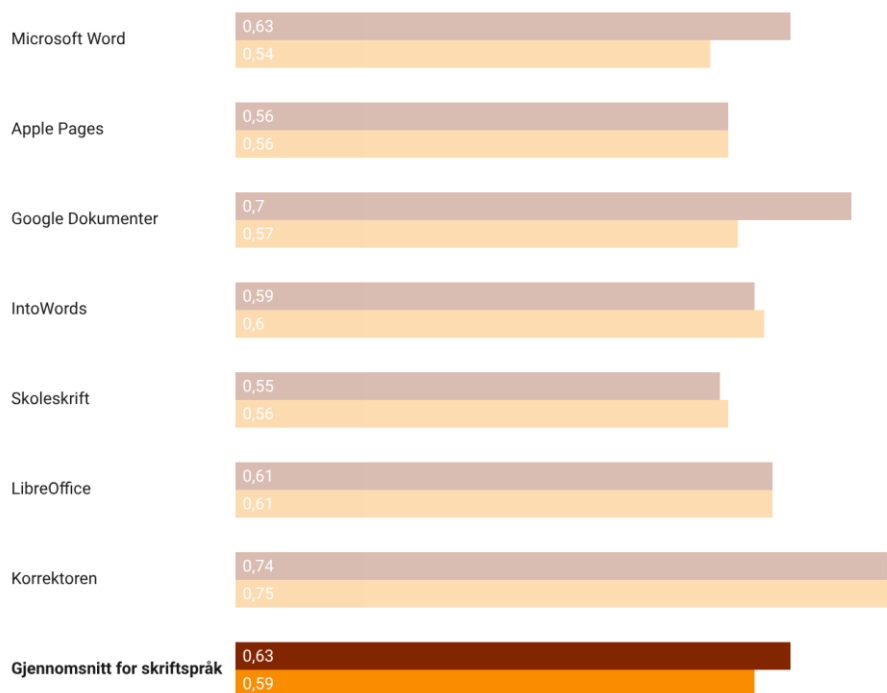
For Word, Pages og Dokumenter bekrefter resultatene i mars 2026 langt på vei resultatene vi fikk i testene vi gjorde i 2025. Dokumenter kommer nå totalt sett best ut av de tre for både bokmål og nynorsk, men når det gjelder feilmarkering av ord i feilfri tekst, kommer programmet desidert dårligst ut. Siden denne delen av testen er vektet relativt lavt, gir den ikke så store utslag på Språkrådet-skåren. Det er likevel verdt å merke seg at i disse feilmarkeringene foreslår Dokumenter i stor grad mer tradisjonelle skrivemåter enn dem som er brukt i ordene som blir markert. I nynorsktekst foreslås det ofte bokmålsalternativer. Dokumenter skiller seg dermed ut som det programmet som virker mest normativt styrende. For **nynorsk** har alle tre leverandørene fått en litt bedre Språkrådet-skår enn i forrige testrunde, men det er vanskelig å si om dette kan tilskrives faktisk forbedring i ytelse, eller om oppgaven kan ha vært «lettere». Språkrådet-skåren for de tre «store» er ellers bare marginalt endret fra tidligere tester.

Det er videre verdt å merke seg at de fire stavekontrollene vi nå har testet for første gang, står seg godt sammenliknet med de tre «store». Korrektoren kommer til og med best ut av alle, og med ganske klar margin i nynorsktekst. Det viser at det er fullt mulig å lage stavekontroller som er bedre enn dem Microsoft, Apple og Google har kunnet tilby til nå.

Bokmål og nynorsk

Språkrådet-skår

■ bokmål
■ nynorsk



Laget med Datawrapper

Figur 4: Sammenlikning mellom bokmål og nynorsk per leverandør, søylediagram

Oppsummert:

- Resultatene for stavekontrollene på nynorsk er fortsatt dårligere enn for dem på bokmål. Det gjelder alle de tre «store», og forskjellen er særlig stor for Dokumenter og Word. For de fire nye er bokmåls- og nynorskresultatene langt jevnere. To av dem skårer best på nynorsk, og for én er resultatene helt like.
- **Gjenfinningsgrad for feil:** I bokmålstekst varierer andelen av innlagte feil som blir funnet, mellom 35 prosent og 65 prosent (cirkatall). Selv i det beste av programmene vil altså nesten fire av ti feil forbli uoppdaget. Det dårligste av programmene overser mer enn seks av ti feil. I nynorsktekst varierer andelen mellom 40 prosent og 70 prosent (cirkatall), men her er tallene for forslagspresisjon langt lavere, se kulepunktet om forslagspresisjon.
 - Det er vanskelig å se noe mønster i hvilke feiltyper som blir oppdaget og ikke, selv innenfor ett og samme program. Men ingen av programmene fanger opp kommafeil og bruk av feil preposisjon. Og/å-feil fanges bare opp av Korrektoren samt av bokmålskontrollene til Word og Dokumenter. Særskrevne sammensetninger fanges med få unntak i liten grad opp. Det

programmene er best til, er å avsløre ord som er stavet eller bøyd feil. Men også for disse feiltypene er programmene langt unna å innfri lovkravet.

- **Forslagspresisjon:** I gjennomsnitt blir det foreslått minst ett korrekt alternativ ved ca. 80 prosent av feilmarkeringene i bokmålstekst. I nynorsktekst er det tilsvarende prosenttallet 65, så her er forskjellen mellom bokmåls- og nynorskkontrollene markant. Word skiller seg negativt ut ved at det foreslås minst ett korrekt alternativ ved mindre enn halvparten av feilmarkeringene i nynorsktekst.
- **Feilpresisjon:** Bare én av stavekontrollene (Korrektoren) unngår helt å markere for feil i den feilfrie bokmålsteksten i del B. I gjennomsnitt markerer stavekontrollene 10 prosent av fellene som feil i bokmålstekst. For nynorsk er tallene litt bedre. Tre av stavekontrollene (Word, IntoWords og LibreOffice) unngår helt å markere for feil. I gjennomsnitt markerer stavekontrollene ved 6 prosent av fellene i nynorsk. Dokumenter skiller seg negativt ut ved å markere for ca. 30 prosent av fellene i både bokmål og nynorsk. Dokumenters foreslåtte alternativer ved disse markeringene går dessuten tydelig i en mer tradisjonell retning i bokmålstekst, og i nynorsktekst er mange av de alternative forslagene på bokmål.
 - o Bokmålskontrollene har gjort til sammen 21 feilmarkeringer i feilfri tekst. Av disse 21 foreslår stavekontrollene en mer konvensjonell skrivemåte for 14 av dem. Dokumenter står alene for 7 av disse.
 - o Nynorskkontrollene har gjort til sammen 12 feilmarkeringer i feilfri tekst. Av disse 12 foreslår stavekontrollene et bokmålsalternativ for 3 av dem. Dokumenter står for 9 av de 12 feilmarkeringene og for alle de 3

Drøfting og implikasjoner

Resultatene av testingen Språkrådet har gjort, viser at Microsofts, Apples og Googles skriveprogrammer, som i 2024 ble brukt av ni av ti elever i den norske skolen, per mars 2026 fortsatt ikke oppfyller kravene i opplæringsloven § 15-4 andre ledd, hverken på bokmål eller på nynorsk. De tre aktørene tilbyr nynorske stavekontroller i skriveprogrammene sine, men når disse fungerer dårligere enn kontrollene for bokmål, innebærer det at nynorskelevene har dårligere språklige verktøy enn bokmålselevene. Dette er i strid med intensjonen bak formuleringen i § 15-4 andre ledd om at skriveprogrammene skal støtte både bokmål og nynorsk. Loven krever at skriveprogrammene skal følge rettskrivingen, men testresultatene viser klart at ingen av programmene oppfyller dette kravet. De er fremdeles langt unna. Så lenge det ikke finnes gode og konkurransedyktige alternativer, vil norske skoleelever få mangelfull hjelp til å lære seg rettskriving gjennom de skriveprogrammene de har tilgang til på skolen. I mange tilfeller vil de få feil informasjon om hva som er rett og galt.

Språkrådet ser at de tre skriveprogrammene ikke har blitt nevneverdig bedre siden Språkrådet gjorde sine første tester i begynnelsen av 2025. Etter disse testene hadde Språkrådet separate møter med representanter fra Microsoft, Apple og Google. I møtene gav Språkrådet tydelig informasjon om hva kravet i opplæringsloven § 15-4 andre ledd innebærer. Aktørene ble også orientert om hvordan skriveprogrammene deres gjorde det på testene. Dette ble fulgt opp med mer detaljert skriftlig dokumentasjon i etterkant av møtene. Den samme typen skriftlig dokumentasjon ble også sendt til aktørene etter den andre testrunden i juni 2025.

Resultatene for de fire nye stavekontrollene Språkrådet har innlemmet, er interessante på flere måter. De ligger jevnt over på nivå med de tre «store» i kvalitet, og resultatene er gjennomgående ganske like for bokmål og nynorsk. Korrektoren skiller seg ut som klart best av alle stavekontrollene som vi har testet, i både bokmåls- og nynorsktekst. Også Korrektoren er et stykke unna å oppfylle kravene i opplæringsloven, men den er et godt eksempel på at det er mulig å lage stavekontroller med langt bedre kvalitet enn dem som har vært tilgjengelige på markedet til nå. Det kan være verdt å merke seg for produsenter som (fortsatt) ønsker å være konkurransedyktige på skolemarkedet. Ikke minst kan det være verdt å merke seg for skoleeiere og andre som burde vært opptatt av å tilby norske skoleelever skriveverktøy som innfrir lovkravene.

Opplæringsloven stiller ikke krav om formell godkjenning av skriveprogrammer som skal brukes i skolen. Det at kvaliteten på skriveprogrammene fortsatt er dårlig snart to år etter at loven trådte i kraft, reiser et spørsmål om det burde vært et krav om at også skriveprogrammer skal være godkjent for å bli brukt i skolen, slik tilfellet er for ordbøker. Det ville trolig lagt et langt mer effektivt press på aktørene som utvikler

skriveprogrammer, om det måtte en formell godkjenning til for at produktene deres kan brukes i norsk skole.

Det er likevel viktig at skolemyndigheter, skoleeiere og lærere holder seg orientert om kvaliteten på skriveprogrammene som blir brukt, slik at de kan unngå å bryte loven ved å ta i bruk programmer som det er dokumentert at ikke holder mål.

Referanser

Ármansson et al. 2025

- *Playing by the Rules: A Benchmark Set for Standardized Icelandic Orthography*, Ármansson et al., NoDaLiDa 2025 ([lenke](#)).

Huyssteen et al. 2004

- *Evaluating Evaluation Metrics for Spelling Checker Evaluations*, Huyssteen et al. 2004 ([lenke](#)).

Jentoft 2023

- *Grammatical Error Correction with byte-level language models*, masteroppgave Matias Jentoft. Universitetet i Oslo, 2023 ([lenke](#)).

LNK-undersøkelsen 2024

- Rapport: *Opplæringslova og skriveprogram, ei undersøking blant rektorane i grunnskulen*, LNK (Landssamanslutninga av nynorskkommunar), 25. september 2024.

Opplæringsloven 2024

- Opplæringsloven § 15-4 *Ordlister, ordbøker og skriveprogram* ([lenke](#)).

Prop. 57 L (2022–2023)

- Lov om grunnskoleopplæringa og den vidaregåande opplæringa (opplæringslova)

Prop. 108 L (2019–2020)

- Lov om språk (språklova)

Skriveprogramtest 2025

- Språkrådets tidligere test av skriveprogrammer på norsk ([lenke](#))

Tildelingsbrev 2026

- Statsbudsjettet 2026 – Språkrådet – tildelingsbrev ([lenke](#)).

Vedlegg: Tekster brukt til test av stavekontroller i mars 2026

Test A – bokmål

I **desse** dager går de **Olympiske** vinterlekene av stabelen i **italia**. Det har blitt mange **gull medaljer** på de **Norske** utøverne. Men ikke alle har **lykkets** like godt. En og annen alpinist har fått seg en ufrivillig skogstur. Slikt gjør ikke **Danskene**, som heller griller seg en **flæskesteik** i målområdet.

Skihopperene våre har imponert, men det **gjeld** mest kvinnene. **Hopp skiene** deres har nok hatt **betre** fart enn **mennes sine**. **Dei** fleste utlendingene er fortsatt sure på grunn av **juksen** med skidressene på **vm ifjor**.

En underholdende **øvelsen** er **skyskiting**. Der skal det skytes stående, liggende, **treff**, og bom. Kommer man i strafferunden er det kroken på døra. Dessuten gjelder det **og** ikke treffe tilskuerne med skuddene. **Sjåvvinistiske** tilskuere kan rett nok **vere** irriterende, men de fortjener kanskje ikke å **døy** av den grunn.

Kunstløp har aldri vært **noken** god øvelse for Norge. I alle fall ikke siden **Sonia** Henie herjet i isen for snart hundre år siden. **Ho** tjente seg rik på **kromspringene** sine **å** fikk **heider** og ære på **kjøpen**. **Rikdomen** er utbredt i Norge **idag**, men den kan nok ikke **sporest** tilbake til kunstløp, er jeg redd.

Test B – bokmål

Alkohol smakte godt, men et glass **mjølk** kunne være et godt alternativ. Det ødela ikke **hjernesellene**. Etter noen **dryge** enheter med **champagne** var **erfaringen** at **trettheta** fort **seig** innpå. **Fotballspillere** **skøyt** ofte **høgt** over mål etter en runde på byen.

Forsking viser at en **skei** tran hver kveld gjør deg mindre **sjuk**. Du blir dessuten ikke så **kvitblek** i **huda**. Om du i tillegg **bur** ute i **skauen**, drikker rent **vatn** fra bekken og spiser **graut** så det tyter ut av ørene på deg, kan du ha godt håp om et lykkelig liv.

I naturen finner vi mye rart. I havet lever både **kval** og **ulk**. Mange **blautdyr** **gjømmer** seg nede på **botnen**, og en og annen **makk** kan ha **forvilla** seg ned i vannet. I **grisgrendte** strøk møter vi både **gjæser** og **uroskråker**. De sistnevnte slåss ofte så **fjørene** fyker.

Test A – nynorsk

I disse dagar går dei Olympiske vinterleikane av stabelen i Italia. Det har blitt mange gull medaljar på dei Norske utøvarane. Men ikkje alle har lyktast like godt. Ein og annan alpinist har fått seg en ufrivillig skogstur. Slikt gjer ikkje Danskane, som heller griller seg ei flæskesteik i målområdet.

Skihopperane våre har imponert, men det gjeld mest kvinnene. Hopp skiene deira har nok hatt betre fart enn mennenes sine. Dei fleste utlendingane er fortsatt sure på grunn av juksen med skidressane på VM ifjor.

Ein underhaldende øvelsen er skyskiting. Der skal det skytest ståande, liggande, treff, og bom. Kommer man i strafferunden er det kroken på døra. Dessuten gjeld det og ikkje treffe tilskuerane med skota. Sjøvvinistiske tilskodarar kan rett nok være irriterande, men de fortener kanskje ikkje å dø av den grunn.

Kunstløp har aldri vore noken god øvelse for Norge. I alle fall ikkje sidan Sonia Henie herja i isen for snart hundre år sidan. Ho tente seg rik på krumspringa sine å fekk heider og ære på kjøpen. Rikdomen er utbreidd i Noreg idag, men han kan nok ikkje sporast tilbake til kunstløp, er eg redd.

Test B – nynorsk

At alle har eit søskenbarn på Gjøvik, er vel kjent. Men korleis ser denne skapningen ut? Og korfor ser han slik ut? Han har ei sykje som gjer at han har dobbelsyn på det eine griseauga sitt. Skjegget ligg flekkvis på haka hans, og det er makk i naseborene. Vettet er det så som så med.

Slektingen bur på eit kjøkken utan utlaup for vatnet. Ein einsleg kaffikjele står på benken. I bestestua er det stor nød. Fleire modne blyantsteikningar heng på veggen over husalteret. Ein ringeljod skjer i øyret. Ein betre venn kunne ein nok ha tenkt seg.

Kva gjer smaksdommaren dersom løken er for sterk? Eller viss ertra er for rund? Om det bløder i tannkjøttet fordi svorden har vore for sprø? Me trur dommaren er viljug til å legge lokk på heile saka, ein vakker eftan i mai.

(Innlagte feil er markert med rødt, innlagte feller er markert med grønt.)