

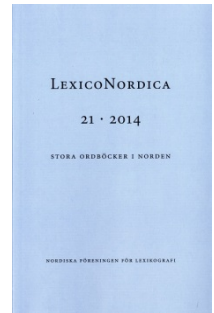
LexicoNordica

Titel: *Tanums store rettskrivningsordbok*
– det norske bokmålets fasit

Forfatter: Øystein Eek & Boye Wangenstein

Kilde: LexicoNordica 21, 2014, s. 19-38

URL: <http://ojs.statsbiblioteket.dk/index.php/lexn/issue/archive>



© LexicoNordica og forfatterne

Betingelser for brug af denne artikel

Denne artikel er omfattet af ophavsretsloven, og der må citeres fra den. Følgende betingelser skal dog være opfyldt:

- Citatet skal være i overensstemmelse med „god skik“
- Der må kun citeres „i det omfang, som betinges af formålet“
- Ophavsmanden til teksten skal krediteres, og kilden skal angives, jf. ovenstående bibliografiske oplysninger.

Tanums store rettskrivningsordbok – det norske bokmålets fasit

Øystein Eek & Boye Wangenstein

The authoritative spelling dictionary of Norwegian Bokmål has for 75 years been *Tanums store rettskrivningsordbok*. The manuscript has recently been converted into a simple, but well tagged database. Redesigned digital and printed versions are planned for 2014. The flexible spelling system of Norwegian Bokmål will be clearly visible in the new user interface, with full documentation of the variant form and inflection systems, and with numerous cross-references. The publication might inspire the Language Council of Norway to reconsider some peculiarities of the standardized Bokmål. The editors of the Tanum dictionary also invite interested parties to a discussion about the selection of lemmas in a modern and comprehensive spelling dictionary of a language where new compounds emerge every day.

1. Innledning

Den autoritative rettskrivningsordboken for norsk bokmål har gjennom 75 år vært *Tanums store rettskrivningsordbok* (heretter kalt Tanum). I denne artikkelen vil vi presentere konverteringen av ordbokbasen og tilretteleggingen for ny utgivelse. Konverteringen og datarettingen er det mest omfattende databaseprosjektet som har vært fullført i Kunnskapsforlaget i senere år, i et vellykket samarbeid mellom den danske dataleksikografen Jens Erlandsen, tittelansvarlig ordbokredaktør Boye Wangenstein, Kunnskapsforlagets porteføljeansvarlige redaktør Øystein Eek, samt forlagets IT-utviklere Per Halvor Tryggeseth og Jonas Knutsen.

Prosjektet har en eldre database som grunnmateriale, så det har

ikke vært nødvendig å starte med et bokmanus, men denne databasen har en komplisert elementstruktur og er ikke fullt oppdatert i samsvar med rettskrivningsendringene i 2005. Siste bokutgave har derfor vært en viktig referanse. Hensikten med prosjektet har vært å gi det meget omfattende innholdet i Tanum en så enkel struktur som mulig, likebehandle alle valgfrie former i bokmål, og knytte ordleddene i sammensatte ord til riktig grunnord gjennom lenking. Ved å publisere denne databasen i digital og trykt form vil vi både gi det norske språksamfunnet et forbedret leksikografisk hjelpemiddel og bidra til å rette søkelyset mot en del spesielle forhold ved den norske bokmålsrettskrivningen.

2. Historikk

Tanum utkom første gang i krigsåret 1940. Arbeidet hadde da pågått i ca. fire år, og med et imponerende resultat, når man tenker på at dataverktøy ikke fantes, og at det ikke var en redaksjonell stab til rådighet for redaktørene. Den opprinnelige redaktør var professor Jacob Sverdrup (1881–1938), men han døde før arbeidet var fullført og ble etterfulgt av lektor Marius Sandvei (1905–1993). Rettskrivningsordboken var dermed påbegynt før 1938-rettskrivningen var på plass, men boken ble utgitt som en «bibel» for det nye bokmålet kort tid etter reformen. Den imøtekom tydelig et stort behov, og fikk ikke noen virkelig konkurrent før *Bokmålsordboka* (1986).

Det var utkommet rettskrivningsordbøker i Norge også tidligere, men disse var av langt mindre omfang. Historien kan føres tilbake til 1885, da Aars' *Retskrivningsregler* (som i tillegg inneholdt en ordliste) ble autorisert av departementet. Senere kom flere nye utgaver, men dette var små trykksaker der korte ordlister tjente til eksemplifisering av rettskrivningsreglene, jf. Aars & Hofgaard (1907). Etter språkreformen i 1907 kom de første egentlige ordlister

til skolebruk med offentlig godkjenning. I 1910 utkom Ivar Alnæs' *Norsk uttale-ordbok*, som ikke inneholdt bøyningsformer, men hadde hele 14 000 oppslagsord. Den første litt større ordboken med bøyningsformer og korte forklaringer var S. Juell Tønnessens, som kom i to utgaver, 1910 og 1926, altså før og etter 1917-reformen. Den mest utbredte rettskrivningsordboken med offisielt stempel etter 1917 var trolig Krogsrud & Seip (1924).

Omfanget av Tanum, som opprinnelig het *Norsk rettskrivningsordbok* og ble utgitt av Johan Grundt Tanum, innebar et kvantesprang. Her skulle ikke bare selve rettskrivningen dokumenteres, men også mest mulig av det ordforrådet som kunne registreres gjennom bruk av mange ulike kilder, som redaktør Sandvei gjør rede for i forordet. En rekke ordbøker og andre oppslagsverk er benyttet,

[m]en det var ikke nok med ordbøker. Også alle slags spesialhandbøker er brukt [...]. En rekke store firmaer i ulike bransjer har velvilligst overlatt redaksjonen varekataloger, og et stormagasin som Steen & Strøm A/S har tilmed utarbeidd fortegnelse over alle sine varer for ordboka. Videre har foreninger og institusjoner hjulpet med å skaffe oppgaver over ord som brukes i arbeidslivet i Norge i dag. Dette store tilfang av moderne bruksord, som ellers er vanskelig tilgjengelig, har redaksjonen ekserpert, for at ordboka i særlig grad også skal bli høvelig for ervervslivet. (Sverdrup & Sandvei 1940:III–IV)

Tanum blir dermed også et interessant samfunnshistorisk dokument.

Norsk rettskrivningsordbok fikk påskriften «Godkjent av Kirke- og Undervisningsdepartementet til offentlig bruk og som hjelpebok i skolen». Språkrådets og dermed myndighetenes medvirkning har i senere år begrenset seg til å «se igjennom» reviderte

utgaver for å forvise seg om at normen er dekket på en fullgod måte.

Marius Sandvei samarbeidet fra 1974 med Bernt Fossetøl (1928–88) om redaktøransvaret. Fra 1983 overtok Kunnskapsforlaget ansvaret for utgivelsen, og i 1994 ble Boye Wangensteen (f. 1942) engasjert som redaktør.

Selve konseptet Tanum har bestått uendret i alle årene etter 1940, med bare mindre justeringer av artikkelstrukturen. Mange nye ord er kommet til, mens svært få ord er tatt ut, selv om de etter hvert er blitt lite aktuelle. Ordboken har dermed et noe mer diakront preg enn den opprinnelig hadde. Ord fra jord- og skogbruk, håndverk, industri, sjøfart og samferdsel slik disse næringene fremstod ved utbruddet av den annen verdenskrig, er fortsatt godt representert i 9. utgave fra 2005. I senere år har avis-korpuset ved Universitetet i Bergen (avis.uib.no) vært en hovedkilde for supplering av lemmautvalget. Nyordmaterialet som er utgitt i *Nyord i norsk* (Guttu & Wangensteen 2012), er blitt gjennomgått og brukt som ressurs. Allmennspråket er i prinsippet holdt oppdatert, og fagord som er brukt i allmennspråket likeså.

Den første digitale utgaven av Tanum kom på disketter i 1991, og en CD-ROM-utgave i 1997, begge på Kunnskapsforlaget. Disse digitale utgavene hadde komplettvisning av bøyningsformer, i samsvar med daværende rettskrivning. Da 9. bokutgave skulle utgis så raskt som mulig etter rettskrivningsendringene 2005, var det ikke ressurser til å oppdatere hele databasen med de nye bøyningsformene, så det ble valgt en enklere løsning der de formene som vises i bokutgaven, ble oppdatert, mens resten måtte ligge.

Tanum er omtalt i regjeringens Språkmelding fra 2008:

På bokmålssida representerer *Tanums store rettskrivningsordbok*, også utgjeven av Kunnskapsforlaget, eit godt supplement til Bokmålsordboka. [...] Med i alt rundt 300 000 ord er dette den største og mest omfattande

rettskrivingsordboka for bokmål. Ho er gjennomgått av Språkrådet i tråd med gjeldande rettskriving, har fullstendige opplysningar om alle tillatne former og varianter og er dermed ei sikker kjelde til den offisielle bokmålsrettskrivinga. (Stortingsmelding nr. 35 (2007–2008), kap. 8.4.4.7)

Ordboken har gjennom mange årtier etablert seg som standardreferansen for bokmål for offentlig sektor, for forfattere og journalister, for studenter etc. Den står på utallige kontorer landet over (om enn til dels i gamle utgaver), og boken har også av ordlisteforfattere vært brukt som kilde i tillegg til *Bokmålsordboka*.

3. Offisielle rettskrivningsordbøker i Sverige og Danmark

At Tanum er et forlagsprodukt, men samtidig har en offentlig anerkjent status som referanse for bokmålet, setter utgivelsen i en særstilling i Skandinavia. Dansk Sprognævn har sin egen *Retskrivningsordbog*, som utgis digitalt på deres eget nettsted og som applikasjon (app), som utlisensieres til andre publiseringsportaler etc. mot betaling, og som utgis i bokform av et dansk forlag (for tiden Alinea A/S) mot royalti til Sprognævnet etter anbud. Sprognævnet har en egen redaksjon for rettskrivningsordboken, med redaktører som bruker ca. 50 % av sin arbeidstid på dette prosjektet, med et redaksjonsråd og med kontakt med Sprognævnets representantskap for å få føringer i prinsipielle normerings- og redigerings spørsmål. Siste bokutgave kom i 2012. Knyttet til utgivelsen er en nyordtjeneste på Sprognævnets nettsted.

I Sverige eier Svenska Akademien rettighetene til *Svenska Akademiens ordlista* (SAOL), som er utkommet siden 1874 i 13 utgaver. Den har en redaksjon ved Göteborgs universitet med vide

fullmakter når det gjelder redigeringen. Bokutgaven publiseres av forlaget Norstedts, som har Sveriges eneste gjenværende ordbok-redaksjon i forlagsverdenen, i et langsiktig samarbeid. Det finnes både en enkel nettoutgave og en mer sofistikert nedlastingsutgave, samt en app, som utgis i regi av Svenska Akademien selv. Både i Sverige og Danmark er de digitale utgavene gratis for brukerne.

4. Den nye Tanum-basen

Kunnskapsforlaget gjorde i 2011 en avtale med Jens Erlandsens firma EMP om å konvertere den eksisterende basen i programmet GestorLex over til programmet iLex, med en betydelig endret elementstruktur. Planen var så å gjøre datarydding i iLex, og dernest konvertere den nye basen videre til Kunnskapsforlagets ordbok-program Dict, før endelig redigering og utgivelse. Boye Wangenstein og forlagets seniorredaktør Øystein Eek begynte sin del av dataryddingen høsten 2012, og fullførte denne vinteren 2014.

I utgangspunktet synes artiklene i Tanum å ha en enkel struktur, med oppslagsord, uttale ved avvik fra hovedreglene, ordklasse-opplysninger, korte bøyingsformer, stikkordsmessige ordforklaringer, samt utvalgte uttrykk. Sammensatte ord er samlet i såkalte reder under grunnordet, men et rede kan bestå av flere semantisk beslektede homografer, f.eks. substantivet og adjektivet *likvid*.

I den nye utgaven skal alle redene oppløses, slik at det blir enklere å finne oppslagsordene, og alle sammensatte ord skrives fullt ut. Bokutgaven får altså glattalfabetisk makrostruktur. I bokutgaven 2005 kan valgfriheten f.eks. være oppført slik:

laga|sild el. **lage-** el. **lågå-**, -a el. -en; - el. -er, -ene (eg. sild fra Lågen).

Her har redaktørene skilt de tre tillatte stammeformene fra hver-

andre. Ved å kombinere de valgfrie stamme- og bøyingsformene, får man følgende valgmuligheter:

lagasild, lagasilda/lagasilden; lagasild/lagasilder, lagasildene
 lagesild, lagesilda/lagesilden; lagesild/lagesilder, lagesildene
 lågåsild, lågåsilda/lågåsilden; lågåsild/lågåsilder, lågåsildene

Under **loft** finner man en rekke sammensetninger med *loft(s)*, bl.a. *glugg(e)*. Ordet kan altså hete *loftglugg*, *loftsglugg*, *loftglugge* og *loftsglugge*. Bøyingsformene finner man så under **glugg**, der *glugge* er oppført som variantform.

Dette oppsettet er lite egnet for en digital utgave, der både alternative stammeformer og bøyingsformer bør skrives fullt ut. I den gamle databasen finnes fulle bøyingsformer av også sammensatte ord, men de har ikke vært oppdatert. Vi har derfor måttet generere nye bøyingsparadigmer for usammensatte ord ut fra oppdaterte endelser, og koble ordleddene i sammensatte ord til riktig grunnord. Redene i Tanums bokutgave inneholder bare unntaksvis opplysninger om ordklasse og homograftilknytning for sisteledd av sammensatte ord. Det har vært et stort arbeid å foreta denne koblingen. I de tilfellene hvor det var bare ett oppslagsord sisteleddet kunne kobles til (altså ingen homografer), som ved *telegraf* (*bildetelegraf*, *billedtelegraf*, *branntelegraf* osv.), kunne koblingen gjøres maskinelt, men ved homografer måtte de behandles manuelt, som ved *råd* (*ambassaderåd*, *ansettelsesråd* osv.). Dette har vært tilfellet med mer enn 46 000 ord. Også mange førsteleddskoblinger er blitt feilaktige ved konverteringen. Redene i Tanum innledes ofte med grunnord av flere ordklasser med ensartet eller nesten ensartet skrivemåte (f.eks. substantivet *stemme* og verbet *stemme*), og analyseprogrammet har naturlig nok ikke kunnet koble sisteleddene til riktig førsteledd her. Disse redene har derfor måttet gjennomgås manuelt.

At også sisteleddene i sammensatte ord kobles til riktig

homograf, har ikke vært gjort tidligere i noen ordbok over norsk språk, så langt vi kjenner til. Vi mener derfor at arbeidet som nå har vært gjort med Tanum, representerer en leksikografisk nyvinning i Norge. I noen tilfeller har vi vært sterkt i tvil om hva som er mest plausible tolkning, og det er ikke alltid mulig å gi en entydig løsning. Et *profilbilde* kan være både et sidebilde (til *profil m*) og et tverrsnittbilde (til *profil n*).

Sammenlignet med *Bokmålsordboka*, der man får korrekt bøyning av sammensatte ord, får man i nye Tanum også sikker (så langt det er mulig) opplysning om hvilke ordledd ordet er sammensatt av. I tillegg er som nevnt ordmassen langt større.

Også redaksjonen for *Svenska Akademiens ordlista* har valgt en løsning for neste utgave der koblingen til riktig homograf vises, se Malmgren (2014).

I forbindelse med at systemet med hovedformer og sideformer i bokmål ble avskaffet i 2005, ble enkelte mindre brukte ordformer tatt helt ut av rettskrivningen, mens noen sideformer ble opphøyet til hovedformer. Av bøyingsformer fikk mange tidligere sideformer status som hovedformer.

De offisielle ordbøkene skal i utgangspunktet ikke gi noen føringer for hvilke former som er mest utbredt og dermed mest brukbare i språksamfunnet. Fra 1959 har myndighetene lagt til grunn for redigering av godkjenningspliktige ordlister at valgfrie former skal oppføres i alfabetisk rekkefølge. Tanums spesielle status har ført til at dette prinsippet har vært fulgt også her.

En naturlig konsekvens av dette har da vært at vi i datarydningen har etablert krysshenvisninger mellom alle valgfrie stammeformer, og at de samme utfyllende opplysninger har måttet tas med under alle variantene, for ikke å gi føringer for valg av former. Dette prinsippet utløser svært mange henvisninger i basen, pr. mars 2014 er det i alt ca. 340 000. Av disse gjelder ca. 53 000 kryssreferanser mellom ulike variantformer. Fra *lagasild* vil det være henvist til *lagesild* og *lågåsild*, fra *lagesild* til *lagasild* og *lågåsild*, fra

lågåsild til *lagasild* og *lagesild*, altså ved n variantformer $n \cdot (n-1)$ henvisninger. For sammensatte ord med valgfrihet i flere ledd blir antallet varianter og dermed henvisninger ganske høyt. Ved de aller mest kompliserte tilfellene har redaksjonen tatt sjansen på å foreta visse forenklinger, og gjøre rede for disse i en brukerveiledning.

Norsk ordbank ved Universitetet i Oslo – en database med fullformer og bøyingsinformasjon for leksemer i bokmål og nynorsk – er knyttet opp mot *Bokmålsordboka* og *Nynorskordboka* og er av Språkrådet anbefalt som referanse for korrekt rettskrivning. Basen er imidlertid ikke fullt oppdatert, og er omtalt slik i Språkrådets notat til Kulturdepartementet om Norsk ordbokpolitikk (03.03.2014):

Språkrådet har lenge hatt et ønske om å forenkle kontrollarbeidet med skoleordlister ved å bruke rettskrivningsinformasjonen i Norsk ordbank som hjelpemiddel i godkjenninga. Norsk ordbank, nynorsk og bokmål, er nå på vei til å bli godt oppdatert, og artikkelhodene, dvs. oppslagsform, ordklasse og bøyning, kan lastes ned gratis fra Språkbanken. Vi tenker oss at ordlisteprodusenter laster ned Ordbanken og bruker materialet (utvalgte deler av det) som basis for egne skoleordlister. Dermed vil Språkrådet bare trenge å vurdere opplegget for ordlista, og ellers ta stikkprøver. (Språkrådet 2014:23)

Det kunne dermed tenkes at Tanum ikke trengte sitt eget bøyingsverk, men kunne hente dette fra Norsk ordbank. Men ettersom Tanum-basen er mer omfattende enn Norsk ordbank (ca. 209 000 oppslagsord mot ca. 146 000 pr. april 2014, og inneholder mye av den samme informasjonen om oppslagsordene, er denne løsningen ikke aktuell:

	Norsk ordbank	Tanum-basen
Antall oppslagsord	146 342	209 125
Antall oppslagsord med liten forbokstav	139 650	202 950
Antall unike oppslagsord med liten forbokstav	49 707	111 258
Bøyningsformer av usammensatte ord	Ja	Ja
Bøyningsformer av sammensatte ord	Ja	Ja, via lenking til riktig homograf
Ordklasseopplysninger	Ja	Ja
Opplysning om normering	Ja	Ja
Krysshenvisninger til alternativformer	Nei	Ja
Uttale av irregulære ord	Nei	Ja
Referanse til bøyningskategori	Ja	Nei
Opplysning om appellativ/proprium	Ja	Delvis, ved at propriar ikke har ordklasse

Tabell 1: Fakta om Norsk ordbank og Tanum-basen.

5. Publisering av den nye basen

Gjennom mange år har det vært spådd at de trykte ordbøkernes tid er forbi, og at digitale hjelpemidler overtar. Nå er det fortsatt slik at det selges godt over 100 000 trykte ordbøker i Norge hvert år, men markedsandelen går stadig ned. Arbeidet med Tanum-prosjektet har vært basert på ønsket om å gjøre basen enklest mulig å redigere i fremtiden, og om å gjøre en ny digital utgave mest mulig brukervennlig. I en digital utgave på Kunnskapsforlagets digitale språktjeneste Ordnett trenger man ikke å lage et plassøkonomisk brukergrensesnitt slik som i en trykt bok. Forkortelser skal unngås eller gjøres klikkbare, bøyningsformer skal skrives fullt ut, med nødvendige ledetekster, og bøyningsopplysningene kan være dy-

namiske, dvs. at bøyningene av sammensatte ord kan hentes fra det usammensatte sisteleddet. For oppslagsordet *lågåsild* henter programmet bøyningen fra *sild*. Dermed blir det langt mindre sjanse for feil etter hvert som rettskrivningen blir endret.

I den planlagte bokutgaven vil det bli en stor utfordring å presentere alt innholdet mellom to permer. Basens nye innretning, der alle sammensatte ord skrives fullt ut, er langt mer plasskrevende enn rede-strukturen. Samtidig innebærer oppløsningen av redene at leserne mister opplysningen om hvilken av flere homografer førsteleddet i et sammensatt ord er dannet til. Vi må derfor sette inn et tall som viser dette, f.eks. **original²språk**, som viser at førsteleddet er adjektivet *original*, ikke substantivet. Så vil det i bokutgaven heller ikke være plass til å vise bøyningsformene av sammensatte ord fullt ut. Vi må nøye oss med å vise bøyningen av sisteleddet. Det vil altså stå **arboret** [arbor'et] S m/f (*arboretet; arboret/arboreter, arboreta/arboretene*), men **original²språk** S n [-språket; -språk, -språka/-språkene]. En ny utgave av Tanum i bokform vil bli langt mer omfangsrik enn de tidligere på grunn av de nye løsningene.

6. Databasen som ressurs for Kunnskapsforlaget

Den ferdige Tanum-basen gir et inventar av norske ordformer med grammatiske og andre opplysninger som blir stort og pålitelig. Den vil dermed være en viktig ressurs for Kunnskapsforlaget, både som fullformsbase som utnyttes i søk på Ordnett og som ressurs for andre ordbøker. En del av våre kunder etterspør grammatiske opplysninger til norske ekvivalenter i de tospråklige ordbøkene; disse vil nå kunne hentes dynamisk fra Tanum-basen. Ettersom ordmassen i Tanum er så stor, vil den utgjøre et arsenal for supplering av ordmassen i andre ordbøker. Samtidig finnes det ord i andre av våre norske ordbaser som ikke er med i Tanum, så

berikelsen kan også gå den andre veien. Tanum-basen kan også være interessant for utlisensiering til språkteknologiske prosjekter og andre formål. Den har allerede vært stilt til rådighet som ressurs for Iness, et stort språkteknologisk prosjekt ved Universitetet i Bergen.

7. Tanum-basen og bokmålsnormen

Bokmålets doble opprinnelse i det dansk-norske riksmål og i folkemålet og samnorskprosjektet fra 1930-årene ligger til grunn for den brede valgfriheten og det ofte dobbelte eller tredobbelte tilfanget av semantisk overlappende ordformer i bokmål. Riksmålets *lagesild* eksisterer side om side med dialektformene *lagasild* og *lågåsild*. I andre tilfeller er det vanskelig å se språkpolitiske motivasjoner for valgfriheten. Dette gjelder bl.a. bruken av valgfri fuge-s, som vi ser i *loftglugg/loftsglugg/loftglugge/loftsglugge*. Det er lett å se at språkbrukerne over det ganske land, som vil skrive et bokmål de kan personlig identifisere seg med, kan ha glede av valgfriheten i det første tilfellet, men mindre innlysende at dette er relevant i det andre eksempelet.

I Språkmeldingen fra 2008 gir Kulturdepartementet følgende føringer for fremtidig normering:

For bokmålet sitt vedkomande reknar departementet med – trass i at det framleis er etter måten stor valfridom på mange punkt i rettskrivinga – at den oppryddinga og forenklinga som vart godkjend i 2005 har vore tilstrekkeleg til at det så langt fram vi kan sjå, ikkje bør vera trong for større endringar. Dermed skulle det no vera lagt grunnlag for ein relativt stabil rettskrivingssituasjon i bokmål.

Truleg vil det også vera med på å sikra slik stabilitet at variasjonen i faktisk språkbruk i bokmål synest å vera klart

mindre enn den valfridommen som rettskrivinga opnar for.

Utover desse overordna prinsippa ønskjer ikkje departementet å leggja meir detaljerte og spesifikke føringar på ein framtidig normeringspolitikk. Det er viktig at ein slik politikk framleis har ei breiast mogleg språkdemokratisk forankring, og mange av dei normeringsprinsippa som vart utvikla av Norsk språkråd, vil truleg vera eit godt grunnlag å byggja vidare på. (Stortingsmelding nr. 35 (2007–2008), kap. 8.4.2.5)

I Språkrådets vedtekter, § 3, står bl.a. følgende: «Dersom faktisk språkbruk eller sterke systemomsyn tilseier det, kan Språkrådet vedta ny stavemåte og bøyning av tidlegare normerte ord og justeringar i skrivereglane.»

Det er altså god hjemmel for å foreta visse forenklinger, så lenge disse ikke er «større endringer», og det er god hjemmel for å undersøke den faktiske språkbruk som begrunnelse for justeringer. Vi er overbevist om at publiseringen av den nye Tanum-basen kan bidra til en fornyet diskusjon om forenkling av bokmålsnormen, basert på empiriske undersøkelser av de valg språksamfunnet gjør. I 9. utgave presenteres sammensetningen *høytlønnen* slik under oppslagsordet **høyt** el. **høgt** adv.:

høy(t)- el. **høg(t)** lønna el. løn(ne)t;

Under

høg adj., n. -t; komp. **høgere**, sup. **høgest** el.

høy adj., n. -t; komp. **høyere**, sup. **høyest**;

står det slik:

lønna el. løn(ne)t el. høgtlønna el. høgtløn(ne)t el. høyt-
lønna el. høytløn(ne)t

Begge disse oppføringene kommer midt inne i svært omfattende reder.

Når man legger alt dette sammen, blir det klart for bokmålsbrukeren at dette ordet kan skrives på 12 forskjellige måter, noe som gir $12 \cdot (12-1) = 132$ mulige kryssreferanser.

Hvis vi holder formene med *høg/høgt* utenfor, siden de brukes også på nynorsk, gir et enkelt Google-søk 26.03.2014 på disse formene i anførselstegn og på sider på norsk følgende resultat (inkl. tillatte bøyingsformer):

høytlønnet/-ede/-ete	17 029
høytlønt/-e	7 700
høytlønna	2 600
høylønnet/-ede/-ete	640
høylønt/-e	544
høylønna	58

Tabell 2: Google-treff på varianter av *høytlønnet*.

En rimelig forenkling ville da være å beholde de tre øverste, samt den av formene med *høg* som systematisk passer best med disse, nemlig *høgtlønna*. Alle rimelige språkpolitiske hensyn ville da ha vært ivaretatt, de 12 formene redusert til fire, og antallet krysshensvisninger til 12.

I andre språksamfunn vil selv fire valgfrie skrivemåter av samme ord oppleves som komplisert og kanskje også unødvendig, og det er grunn til å spørre seg om vi gjør språket mer eller mindre robust i det multinasjonale Norge i en situasjon der språket er under sterkt press fra engelsk, ved å opprettholde det hevdvunne systemet av valgfrihet.

8. Planer for videre redigering

I forordet til 1. utgave av Tanum (Sverdrup & Sandvei 1940:IV) står det bl.a. at

Ordboka vil gi opplysning om skrivemåte, kjønn og bøyning av flest mulig grunnord. Særlig er da mange særnorske ord, fagord og fremmedord kommet med. [...] så langt det har vært råd å få det til, [er] alle vanlige sammensetninger og avledninger tatt med.

En kan vel legge til at det er en god del *uvanlige* sammensetninger også, dvs. sjeldne. Videre står det i forordet at

Norsk *rettskrivningsordbok* legger fram et rikholdig ordtilfang uten å ville øve press i noen retning. Ordene står der enten de er nye eller gamle, særnorske eller halvdanske, heimlige eller fremmede. Så får den som bruker boka, selv velge.

Derfor finner vi i Tanum fortsatt mange ord fra f.eks. gammelt arbeidsliv (*brøte* 'velte tømmer i elva og fløte', *bakole* 'rem på seletøy'), fra bergverksdrift (*blotning* 'synlig del av formasjon', *jøpel* 'transportmaskin, hestevandring'), arkaiske ord (*kyrr* 'stille, rolig', *liktrå* 'spedalskhet'), en mengde fremmedord og fagord innenfor arkitektur (*peristyl* 'søylegang'), anatomi (*diafyse* 'mellomstykke på lange knokler'), medisin (*effluvium* 'utflod', *flatus* 'avgang av tarmgass') og fra en rekke andre fagområder.

Vi må stille spørsmålet om Tanum fortsatt skal være en ordsamling der en kan finne likt og ulikt (til glede for f.eks. kryssordløserne og scrabblespillere) eller om den skal strammes inn til en rettskrivningsordbok over dagens aktuelle ordforråd. Videre om

vi skal skille innholdsmessig mellom den elektroniske utgaven og bokutgaven.

En mulig vei å gå er å legge inn i ordmassen i Tanum et stort antall aktuelle ord fra avisarkivet i Bergen, men ikke ta med alle i bokutgaven, og ta ut av bokutgaven et stort antall uaktuelle ord.

Ovenfor er det opplyst litt om hvordan ordforrådet i Tanum ble samlet i slutten av 1930-årene. Dette var lenge før korpusleksikografiens tid. Et korpusbasert Tanum omkring 1940 ville ha gitt et langt magrere resultat. Uansett hvor stort tekstkorpus man hadde hatt tilgang til, ville mye av det som står i boken, ikke ha kommet med. Samlingen er derfor svært verdifull som et ordarsenal. Ordmassen i Tanum overgår langt det man finner i SAOL (125 000 i 13. utgave fra 2006) og *Retskrivningsordboken* (65 000 i 4. utgave fra 2012; dette tilsvarer antallet i *Bokmålsordboka*). Det er neppe noen som vil argumentere for en sanering som vil fjerne alt historisk og kuriøst ordstoff fra basen, men det er neppe heller noen som vil argumentere for å beholde alt i en fremtidig utgivelse. Her finnes for eksempel oppslagsord som vi som arbeider med redigeringen, ikke forstår betydningen av, og som vi heller ikke kan finne i noen tilgjengelige kilder. Det kan godt hende at slike ord skal beholdes i basen, men ikke publiseres. Språkrådet legger ingen føringer for den fremtidige redigeringen av ordforrådet i Tanum, men nøyer seg med å kreve at skoleordlister må dekke det sentrale ordforrådet i språket for å kunne bli godkjent (Stortingsmelding nr. 35 (2007–2008), kap. 8.4.4.5). Ordbokens redaksjon oppfatter at den har stor frihet til å ta ut dødt språkstoff. Redaksjonene for SAOL og *Retskrivningsordboken* har i senere utgaver gjennomført en betydelig sanering, uten protester. Ved å foreta en skjønnsom sanering og samtidig ta inn et større nyordmateriale, slik de andre skandinaviske redaksjonene også har gjort (for eksempel kom 4500 nye oppslagsord inn i 4. utgave av *Retskrivningsordboken*), vil Tanum få et mer moderne preg.

Vi har gjort oss tanker om Språkrådets rådgivende medvirkning i det fortsatte arbeidet med Tanum. Hovedsporet har vært at ordboken fortsatt skal være en forlagsdrevet autoritativ bokmålsordbok med offentlig anbefaling, i pakt med en langvarig tradisjon (jf. Språkmeldingens kap. 8.4.4.7). Vi har invitert Språkrådets fagråd 3, som har språknormering og språkobservasjon som sitt arbeidsfelt, til å bli med i en prosess der vi prøver å bli enige om føringer for utvikling av denne språkressursen på kort og lengre sikt. Vi vil spille inn forslag om visse forenklinger av bokmålsrettskrivningen, som fagrådet så kan ta opp til behandling i henhold til sine vedtekter. Ut fra Språkrådets notat til Kulturdepartementet mars 2014 (Språkrådet 2014) ser det ut til at man i Rådet for tiden vil prioritere utviklingen av Norsk ordbank. Dette trenger ikke å være til hinder for et fortsatt samarbeid om Tanum, men det gjør det mer usikkert om Rådet ønsker å engasjere seg mer aktivt i et prosjekt som det ikke selv styrer.

Hvis Språkrådet mot formodning verken skulle ønske å engasjere seg aktivt for eller sette sitt stempel på Tanum i fremtiden, vil vi selvsagt stå friere til å redigere en stor rettskrivningsordbok for bokmål ut fra egendefinerte prinsipper. Alt tyder på at denne ressursen i overskuelig fremtid vil være mer omfattende enn databasen Språkrådet rår over, og dermed vil Tanum ha forutsetninger for å bli en nødvendig referanse for språksamfunnet også i fremtiden.

Litteratur

Ordbøker

Alnæs, Ivar (1910): *Norsk uttale-ordbok*. Kristiania: H. Aschehoug & Co. (W. Nygaard).

- Bokmålsordboka* (1986). Marit Ingebjørg Landrø & Boye Wangenstein (red.). Oslo: Universitetsforlaget. 3. utgave (2005) Boye Wangenstein (red.) Oslo: Kunnskapsforlaget. <<http://bit.ly/1muS6dp>> (mai 2014).
- Guttu, Tor & Boye Wangenstein (red.) (2012): *Nyord i norsk*. Oslo: Kunnskapsforlaget.
- Juell Tønnessen, S. (1910): *Retskrivningsordbok over det norske riksmåal*. Kristiania og Kjøbenhavn: Gyldendalske Boghandel Nordisk Forlag.
- Juell Tønnessen, S. (1926): *Norsk og fremmed-ordbok*. Oslo: Gyldendal Norsk Forlag.
- Krogsrud, Torgeir & Didrik Arup Seip (1924): *Norsk riksmålsordbok*. Kristiania: Steenske Forlag.
- Nynorskordboka* (2006). Marit Hovdenak et al. (red.). 4. utgåva. Oslo: Det Norske Samlaget. <<http://bit.ly/1nihodk>> (mai 2014).
- Ordnnett <<http://www.ordnett.no>> (april 2014).
- Retskrivningsordbogen* (2012). Anita Ågerup Jervelund, Jørgen Schack, Jørgen Nørby Jensen & Margrethe Heidemann Andersen (red.). København: Alinea A/S. <<http://www.dsn.dk/ro/ro.htm>> (mai 2014).
- Svenska Akademiens ordlista* (2006). Trettonde upplagan. Stockholm: Norstedts Akademiska Förlag. <<http://bit.ly/SPoES7>> (mai 2014).
- Sverdrup, Jakob & Marius Sandvei (1940): *Norsk rettskrivningsordbok*. Oslo: Johan Grundt Tanum.
- Tanums store rettskrivningsordbok*. 9. utgave. Boye Wangenstein (red.) (2005): Oslo: Kunnskapsforlaget.

Annen litteratur

- avis.uib.no = Norsk aviskorpus <<http://avis.uib.no>> (april 2014).
 Malmgren, Sven-Göran (2014): *Svenska Akademiens ordlista* gjennom 140 år: mot fjortonde opplagan. I: *LexicoNordica* 21 (dette bind).
- Språkrådet 2014 = *Norsk ordbokpolitikk*. (Notat fra Språkrådet til Kulturdepartementet 3. mars 2014) <<http://bit.ly/1nNiK3B>> (mai 2014).
- Språkrådets vedtekter <<http://www.sprakradet.no/Topppmeny/Om-oss/Vedtekter-for-Sprakradet/>> (mars 2014).
- Stortingsmelding nr. 35 (2007–2008): *Mål og mening. Ein heilskapleg norsk språkpolitikk*. Oslo: Det kongelige kultur- og kyrkjedepartement.
- Aars, J. & S.W. Hofgaard (1907): *Norske retskrivnings-regler med alfabetiske ordlister*. Kristiania: W. C. Fabritius & Sønner A/S.

Øystein Eek
 seniorredaktør, cand.philol.
 Kunnskapsforlaget
 Postboks 4432 Nydalen
 NO-0403 Oslo
 oystein.eek@kunnskapsforlaget.no

Boye Wangenstein
 ordbokredaktør, cand.philol.
 Kunnskapsforlaget
 Postboks 4432 Nydalen
 NO-0403 Oslo
 boye.wangenstein@
 kunnskapsforlaget.no

