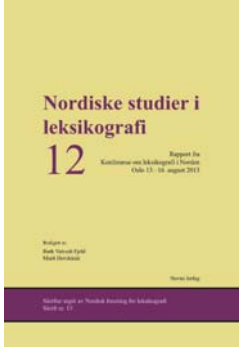


NORDISKE STUDIER I LEKSIKOGRAFI

Titel:	Samvirket mellom ordbank og ordbok	
Forfatter:	Oddrun Grønvik og Christian-Emil Ore	
Kilde:	Nordiske Studier i Leksikografi 12, 2013, s. 139-158 Rapport fra Konferanse om leksikografi i Norden, Oslo 13.-16. august 2013	
URL:	http://ojs.statsbiblioteket.dk/index.php/nsil/issue/archive	

© Nordisk forening for leksikografi 2014

Betingelser for bruk 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.

Søgbarhed

Artiklerne i de ældre Nordiske studier i leksikografi (1-5) er skannet og OCR-behandlet. OCR står for 'optical character recognition' og kan ved tegngenkendelse konvertere et billede til tekst. Dermed kan man søge i teksten. Imidlertid kan der opstå fejl i tegngenkendelsen, og når man søger på fx navne, skal man være forberedt på at søgningen ikke er 100 % pålidelig.

Samvirket mellom ordbank og ordbok

Oddrun Grønvik og Christian-Emil Ore

In August 2012 the web edition of *Nynorskordboka* was launched with a revised orthography for Nynorsk and many new functions, directed at improving user friendliness and information transmission about the Nynorsk standard language. Until then, part of speech information and (abbreviated or coded) inflection patterns were hard coded as part of the dictionary text. In the new web edition, the dictionary database is linked to the Word Bank for Nynorsk. The entry head is generated from the Word Bank, as are schema showing inflected forms written in full. The Word Bank for Nynorsk has become an efficient tool for maintaining and transmitting orthographic information on Nynorsk. Use of the revised web edition has increased considerably in the last year.

1. Oppgradering, oppkopling og kvifor

Handordbøkene Bokmålsordboka (BOB) og Nynorskordboka (NOB) har ei lang fortid som ordbøker på nett. BOB og NOB kom i andre utgåve i 1993, og vart publiserte som gratis-resursar¹ i 1994 frå INL.² Nettsida for handordbøkene fekk noverande utsjånad i 2009,³ med ei viss oppgradering for funk-

sjonar (Hoel 2010:10–11). Teksten frå 2005- og 2006- utgåvene stod uendra.

Frå 2009 vart det arbeid med å leggje inn teksten i handordbøkene i det databasesystemet som Norsk Ordbok nyttar. Det vil sin tur gje betre ordbøker både på papir og på nett. Arbeidet vart utført ved ILN, UiO,⁴ med medverknad frå Språkrådet.⁵ Oppgraderinga omfatta (1) korrekte fullformer i bøyingsskjema i artikkelhovudet, (2) fullskriving av forkortingar i definisjonar og bruksdøme og (3) endringar i databasestrukturen for BOB, NOB og Norsk ordbank. Endringane i databasestrukturen omfatta inndeling av ordartiklar etter artikkeltype, kopling av tilvisingar, og komplettering av informasjon i artikkelhovudet (oppslagsdelen). Alle ordartiklar skulle ha med dei same kategoriane for informasjon om ordklasse og bøying.⁶

Føremålet med oppgraderinga var (1) standardisering av normeringsformatet for einskildord (same opplysningar, betre systematikk) og (2) ei meir fullstendig og opplysende formidling til folk flest om rettskriving og bøying i norsk.

Språkrådet hadde i tillegg eit ønske om arbeidssparing. Informasjon om gjeldande rettskriving var i 2009 spreidd på Norsk ordbank, handordbøkene, og årsmeldingar og sakspapir frå Språkrådet. Det var stor overlapping, men også ein del divergens. Om all rettskrivingsinformasjon for norsk kunne

-
1. På www.edd.uio.no/ordboksoek/ordboksoek.html.
 2. Institutt for nordistikk og litteraturvitskap, UiO (1990–2002).
 3. <http://nob-ordbok.uio.no/perl/ordbok.cgi>.
 4. ILN, som INL vart del av frå 1.1.2003.
 5. Språkrådet og UiO er felles eigarar av BOB og NOB på nett.
 6. Jf. brev 22.12.2010 frå styringsgruppa for Norsk ordbank til Språkrådet.

samlast på ein stad, og hentast derifrå av t.d. utgjevarar av språkverktøy (ordlister, retteprogram osv.), ville det spare Språkrådet for tidkrevjande kontrollarbeid på uferdige produkt.

Frå språkvitskapleg og leksikografisk side var det ønskeleg med ei grundig oppgradering både av Ordbanken for nynorsk og NOB. Gjennomarbeiding av kategorisystem, paradigme-register og innhald ville gje eit heilt anna oversyn over nynorsk-rettskrivinga enn det som låg føre i 2009, og dermed eit mykje betre grunnlag for å vurdere føresetnader for og verknader av den påtenkte rettskrivingsrevisjonen for nynorsk.

Parallelt med oppgraderingsarbeidet i ordboksdatasane førebudde Språkrådet ei ny rettskriving for nynorsk. Den vart sanksjonert av Kulturdepartementet 20.9.2011.⁷ Språkrådet fekk i oppgåve å få den nye rettskrivinga kunngjord via NOB før skulestart 2012. Oppgradering av NOB vart derfor prioritert framfor BOB.

2. Kva er Norsk ordbank

Norsk ordbank (bokmål og nynorsk) er ein database over grunnord, bøyingsmønster og kopling mellom dei. Koplingane kan innehalde informasjon om normering og tidsrommet normeringa gjeld for. I tillegg inneheld Norsk ordbank ein full-formsgenerator. Norsk ordbank har eksistert sidan 2001⁸ og vart

7. Brev frå Kulturdepartementet til Språkrådet 20.9.2011.

8. <http://www.edd.uio.no/prosjekt/ordbanken/>.

i si tid laga slik at den i tillegg til annan bruk (Hagen og Nøklestad 2010:55 f.) kunne eigne seg til å vere lager og vedlikehaldspunkt for rettskrivinga av bokmål og nynorsk. Mens BOB og NOB alltid har vore eit offisielt talerøyr for Språkrådet, har ikkje Norsk ordbank vore det før i 2012. Dersom Norsk ordbank også skulle spare Språkrådet for dobbeltarbeid, og alle andre for avskrift, måtte Ordbanken koplast til ordboksdatabasane slik at artikkelhovudet på nett vart riktig generert for kvar artikkel. I neste runde må dette formatet tilpassast ein ny prenta versjon av handordbøkene.

Samankopling av ordbank og ordboksdatabase var dermed sett som den mest rasjonelle måten å (1) kunngjere den nye nynorskrettskrivinga, (2) rasjonalisere vedlikehaldet av rettskrivinga, (3) leggje til rette for generering av oppslagsdelen i ordbøker og ordlister frå Ordbanken. I resten av denne artikkelen vil vi konsentrere oss om nynorskdelen av Norsk ordbank. Termen «Ordbanken» vil derfor bety nynorskdelen av Norsk ordbank dersom ikkje anna er sagt. Datateknisk er dei to delane like og har lik oppbygging, jf. avsnitt 2.1 nedanfor.

2.1. Ordbankens oppbygging

Ordbanken er basert på ei utviding av eit stavekontrollsystem som IBM laga på 1980-talet (Engh 1993). Hovudtanken er å lenke kvar grunnform til eitt eller fleire bøyingsmønster som i sin tur kan generere alle moglege bøyde former av grunnforma. Denne prosessen inneber at det blir generert teoretisk moglege, men udokumenterte bøyde former. Mellom dei udokumenterte formene vil det vere somme former som verkar

så avvikande at dei kan vekkje reaksjonar mellom språkbrukarane⁹. Desse formene treng ein i språkteknologiske analyseverktøy, medan dei ikkje trengst i t.d. system for skrivestøtte. Der vil ein helst sjå berre dei allment godtekne formene. For å unngå moglege, men uakseptable former, og sikre at alle sett med bøygde former svarer til offisiell rettskriving, er det lagt tilleggsinformasjon til lenka mellom grunnform og bøyingsparadigme.

Tilleggsinformasjonen gjeld kategoriane normeringsstatus og tidsperiode då normeringsstatusen er eller var gyldig. Det er dermed mogleg å nytte Ordbanken til å lage eit historisk register over norsk rettskriving. Ordbanken har no informasjon for tidsperioden 1996 til 2013. Ein kan, om ein vil, generere ei normert rettskriving for kvart av åra i perioden.

Figur 1 viser redigeringsgrensesnittet i Ordbanken. Toppinja inneheld identifiserande informasjon. Under til venstre er tabellar for normeringsstatus og paradigme med status. Fullformstabell er nedst til venstre. Til høgre frå toppen ser ein tabell med lenke til ordboksdatasane, under tabell for grammatisk merking i IBM-tilfanget og tidlegare utgåver av NOB.

Kvar artikkel i Ordbanken har ein unik kombinasjon av ei grunnform pluss eitt eller fleire bøyingsparadigme pluss statushistorie i høve til grunnform og per paradigme. Ordbankartiklane er ikkje skilde på homografar etter dette punktet, slik dei kan vere i ei ordbok (t.d. ut frå ordhistorie). Norsk er som kjent rikt på homografe grunnformer. Dermed er det mogleg at éin ordbankartikkel kan vere lenka til fleire ordartiklar i ei ordbok. (Grønvik og Ore 2013:250 f.).

9. Døme på slike former er gradbøygde former av somme adjektiv, som **interessantare* – **interessantast*.

The screenshot shows the Ordbanken interface for the lemma 'lyge'. The main window displays the following information:

- Lemma-id:** 56435
- Funksjon:** lemma
- Grunnform:** lyge
- Registrer:** Tekst | Kommentar/logg
- Normering:** A table showing the norming process from 1996 to 4000.
- Paradigmer:** A table showing inflectional paradigms for the verb 'lyge'.
- Bøyingsdata:** A table showing the full forms and tags for the verb 'lyge'.
- Kilder:** A list of sources including NN_ORDBOK, NOB_2013, NOB_2012, and NN_ORDBOK.
- Alternativ pa:** A list of alternative forms including IBM 199 and NN_O Vs.
- Ny ortoklump:** A section for new orthographic groups.

The 'Bøyingsdata' table is as follows:

P...	N...	Fullform	Tag	Fra dato	Til dato
198	1	lyge	verb inf <intrans1> <trans11/for> <trans: 1996	4000	
198	2	lyga	verb inf <intrans1> <trans11/for> <trans: 1996	4000	
198	3	lyg	verb pres <intrans1> <trans11/for> <tran 1996	4000	
198	4	lygast	verb inf pass <intrans1> <trans11/for> <t 1996	4000	
198	5	laug	verb pret <intrans1> <trans11/for> <tran 1996	4000	
198	6	loge	verb perf-part <intrans1> <trans11/for> · 1996	4000	
198	7	loge	adj <perf-part> nøyt ub eint <intrans1> < 1996	4000	
198	8	logen	adj <perf-part> m/f ub eint <intrans1> <t 1996	4000	
198	10	logne	adj <perf-part> bu eint <intrans1> <trans 1996	4000	
198	11	logne	adj <perf-part> fl <intrans1> <trans11/fo 1996	4000	

Figur 1: Ordbanken – redigeringsgrensesnittet

I både Ordbank og ordboksdatabasane er det lagt stor vekt på dataintegritet. Det er knytt identitetsnummer til artiklar, i NOB også til definisjonar og døme. Bøyingsparadigma i Ordbanken har egne kodar, tre tal eller bokstavar.

2.2. Ordbankens omfang og innhald

Når det gjeld realinnhald, har Ordbanken for nynorsk om lag 113 000 artiklar, noko meir enn NOB, som har om lag 100 500 oppslagsord fordelte på rundt 85 000 innhaldsartiklar (normalartiklar). Dette er den delmengda av nynorsk ordtilfang som er eksplisitt normert. Ordbanken for nynorsk har om lag 657 000 fullformer i alt. Av desse er 509 000 fullformer etter 2012-rettskrivinga. Fullformene er genererte på grunnlag av 1578 paradigme. Etter 2012-rettskrivinga er 518 paradigme med i rettskrivinga. Rettskrivingsrevisjonen som tok til å gjelde 1.8.2012, har dermed ført til at meir enn 1000 ulike bøyingsparadigme og 150 000 ordformer (som er variantar) er gått ut or nynorskrettskrivinga.

I 2005 vart ordklassenemningane for norsk reviderte ¹⁰ og dei endra ordklassenemningane tilrådde til i bruk i skulen (Johannesen og Hagen 2007). Ein del ord vart omklassifiserte etter dei nye nemningane, som byggjer på ordklassesystemet i Norsk referansegrammatikk (NRG:20 f.). Etter oppgraderinga av NOB i 2012 kjem ordklassenemningane frå Ordbanken. I Ordbanken var systemet i NRG nytta fullt ut, og det måtte dermed justerast noko for å samsvare med 2005-vedtaket. I tillegg måtte det først inn nemningar på ordklasseplass for forkortingar, avleiingsprefiks, forledd i samansetningar og symbol. Dette er språklege einingar som har artiklar i alle allmennordbøker for norsk, men som fell utanom det vanlege ordklassesystemet, av di dei er ordlagingselement (forledd i samansetningar, avleiingsprefiks), eller av di dei representerer ein særskild, konvensjonalisert bruk av ord og ordlag (forkortingar,

10. Utdanningsdirektoratet og Språkrådet (2005).

symbol). Ordbøker tek med desse artikkeltypane av di det ofte er spørsmål etter korrekt skrivemåte.

2.3. Paradigme og vising av bøyingsmønster

Revisjonen av nynorskrettskrivinga 2011–2012 føresette ein full gjennomgang av bøyingsparadigma i Ordbanken før endringane vart innførte. Det vart avgjort at all variasjon måtte liggje på paradigmenivå eller over. Variantar i høve til bøyings-skjema innanfor eit og same paradigme vart fjerna.¹¹ Dersom eit verb t.d. har to fortidsformer, føreset dette to sjølvstendige paradigme. Går den eine paradigmet ut i samband med rettskrivingsendring, kan dette uttrykkjast ved sluttdato for status «normert», medan det andre står uendra.

Figur 2 syner eit døme på korleis Ordbankens paradigmeskjema – som vart utarbeidde som del av fullformsgeneringa omkring midten av 1990-talet – er nytta til å gje bøyingsinformasjon i NOB på nett. Skjemaet har både fullformer og informasjon om det grammatiske systemet. Har eit ord fleire bøyingsparadigme, blir desse viste på kvar si linje i skjemaet. Merk at standardskjemaet for nynorsk verb-bøyning har to infinitivsformer, både e-infinitiv og a-infinitiv. I staden for ein ekstra kolonne har vi valt å setje dei under kvarandre.

11. Skjemaet for verbparadigme hadde t.d. separat femininumform for verbaladjektiv og sams form for maskulinum og femininum.

Bøying i samsvar med 2012-rettskrivinga:

lyge	Infinitiv	Presens	Preteritum	Presens perfektum	Imperativ
v.	å lyga å lyge	lyg	laug	har loge	lyg

lyge	Perfektum partisipp				Presens partisipp
	Hankjønn/hokjønn	Innkjønn	Bunden form	Fleirtal	
v.	logen	loge	logne	logne	lygande

Figur 2: Bøyingsinformasjon frå Ordbanken om verbet *lyge*, inkludert skjema for samsvarbøying av verbaladjektivet.

Ei eittspråkleg papirordbok er berar av mykje implisitt informasjon. Redaksjonen kan spare plass ved å sløyfe det som synest opplagt for morsmålsbrukarar. Plasskrevjande informasjon som gjeld store kategoriar kan stuvast inn i forord og brukarrettleiingar. Men i ei elektronisk ordbok er dette korkje naudsynt eller ønskeleg. Det er best at informasjonen kan bli vist saman med ordartikkelen. Alle paradigme må vere eksplisitt registrerte på alle relevante grunnformer. Det vart naudsynt å lage nokre få nye paradigme for nynorsk. Kategorien «blanda bøyning» av svake verb har t.d. fått eigne paradigme knytte til dei rette grunnformene, med kode «v3». ¹² Paradigmesystemet i Ordbanken byggjer på rettskrivingssystemet til IBM. Føremålet med IBM-systemet var praktisk, det byggjer ikkje på nokon spesifikk morfologisk teori.

12. v1 = a-verb, t.d. *rope - ropar - ropa - ropa*. v2 = te-verb, t.d. *rope - roper - ropte - ropt*. Verb med bøyning etter båe desse mønstera kan også ha eit kombinert paradigme, *rope - ropar - ropte - ropt*. Dette paradigmet er gjort eksplisitt i 2012-utgåva av NOB på nett og har fått koden v3.

Når ein set opp eit system av bøyingsmønster for norsk, står hovudvalet mellom å føre fleire alternativ på kvar paradigme-linje (komplekse mønster) eller ei einskild form for kvar linje (enkle mønster). I Ordbanken er alle paradigme i utgangspunktet enkle mønster. Men sidan ei grunnform kan vere tilordna fleire paradigme, kan ein sjå på kombinasjonar av paradigme knytte til ei grunnform som komplekse mønster.

Om ein definerer ord som paret <grunnord, bøyingsmønster>, vil talet på ord auke når talet paradigme aukar. I tabellen i figur 3 viser vi korleis det blir for nynorsk. Dei komplekse paradigma er laga ved å setje saman enkle paradigme. Til dømes kan «700» (m1)¹³ og «800» (n1)¹⁴ setjast saman til eit felles (m1, n1) med fleire alternativ på kvar linje.

Talet på paradigme stig når ein tel med komplekse paradigme (ein får både kvart einskilt og kombinasjonar av dei), men berre til ein viss grad. Talet på paradigme aukar kraftig når

	Alle		2012-rettskrivinga	
	tal ulike paradigme	tal ord	tal ulike paradigme	tal ord
Komplekse	755	111 545	580	103 650
Enkle	1 578	269 556	518	112 409

Figur 3: Paradigme og grunnformer i Ordbanken for nynorsk.

13. Paradigme nummer 700 = bøyingsmønsteret uttrykt i *hund* – *hunden* – *hundar* – *hundane* med NOB-kode m1.

14. Paradigme nummer 800 = bøyingsmønsteret uttrykt i *hus* – *huset* – *hus* – *husa* med NOB-kode n1.

alle sideformene skal gje opphav til egne paradigme. Men talet på grunnformer fell når ein tel med komplekse paradigme, for då blir oppslagsformer med meir enn eitt paradigme talde berre ein gong i staden for fleire.

Tabellen i figur 3 viser at 1 060 bøyingsparadigme og meir enn 150 000 rettskrivingsvariantar fekk sluttdato 2012 i samband med rettskrivingsrevisjonen for nynorsk. Han syner også at det store fleirtalet av paradigme som er unormerte no, var tillegg til hovudparadigma.

Endeleg må som nemnt alle grunnformer ha paradigme i Ordbanken, endå om grunnforma er ei einaste form. Føremålet er få vist rettskrivingsinformasjonen så ryddig som mogleg i nettutgåva av NOB, og bruke einsarta skjema så langt råd er.

2.3. Leddanalyse i Ordbanken og NOB

Tidlegare utgåver av NOB og BOB markerer deling av oppslagsord etter typografiske behov, i høve til avsnittsinndelinga i den prenta ordboka. Oppslagsord i same avsnitt er forkorta, og delestreken i første oppslagsord syner kva som er sløyfa i dei følgjande oppslagsorda.

Delemerke kan falle saman med komposisjonspunkt, men treng ikkje gjere det. Samansette ord i eine avsnitt har ikkje delemerke, medan avleiingar kan ha det. Deling av avleidd ord gjeld særleg importord. Døme på delingspraksis i NOB 2006 er vist i figur 4.

type	NOB 2006	NOB 2012
grunnord	rip el	ripel
avleiing	ralj ere	raljere
avleiing	ragg ete	raggete
	ransak e	ransake
samansetning	ramp estrek	rampe strek
samansetning	rakletre	rakle tre

Figur 4: Døme på delemerke i den trykte ordboka i 2006 og i den elektroniske ordboka i 2012.

I samband med oppgraderinga av NOB og Ordbanken var det ønske om å kunne vise leddeling etter språklege kriterium, det vil seie forledd og etterledd med ordklasse, og eventuell fuge. Det vart derfor laga ein leddelingstabell, som tillegg til Ordbanken for nynorsk. Leddelingstabellen byggjer på ei grov leddeling gjord med Paul Meurer sitt leddelingsprogram.

Fugeregisteret for norsk omfattar i tillegg til -e-, -s- osv. også nullfuge (figur 5, døme 1) og minussfuge (døme nr. 3). Ord som har leddeling med minussfuge, blir viste i ordartikkelen med sine to godkjende oppslagsformer, i dette tilfellet «trafikkaos el. trafikk-kaos».

Når ei samansett grunnform frå Ordbanken er registrert og tagga i leddelingstabellen, vil orddelinga følgje med til oppslagsordet for ordartikkelen i NOB. Førebels må kvar samansetning registrerast med leddeling, men eit framtdsmål er å gjere denne prosessen (halv)automatisk.

I ein leddelingsanalyse vil det alltid kome for dagen ein god del samansetningar der førsteleddet ikkje kan knytast sikkert til berre eitt anna oppslagsord, jf. tabellen i figur 5 siste linje, som

Grunn- form	For- ledd	For- ledd gram	Fuge	Etter- ledd	Etterledd gram.	Oppslags- form
raudvin	raud	Adj		vin	Noun+M+ Sg+Indef	raud*vin
raudvins- toddi	raudvin	Noun	s	toddi	Noun+M+ Sg+Indef	raudvins* toddi
trafikkaos	trafikk	Noun	- k	kaos	Noun+N+ Sg+Indef	trafikkaos
trillebag	trille-	Føre- ledd		bag	Noun+M+ Sg+Indef	trille*bag

Figur 5: Leddeling i nynorskordbanken

både kan knytast til verb og substantiv.¹⁵ I praksis spelar ordklasse på forledd lita rolle, for funksjonen til forleddet i ei samansetning er å modifisere eller spesifisere etterleddet, same kva ordklasse forleddet har opphav i.¹⁶

15. Verb < ‘bag som kan trille’, substantiv < ‘bag som er (som) ei trille’.

16. I denne typen samansetningar er forleddet før ofte rekna som verb, jf. Beito (1986) s. 141 og NRG s. 75. I leddelegssystemet for handordbøkene står spørsmålet ope, i staden for å vere bunde til ein usikker hypotese.

3. Oppslagsdelen i Nynorskordboka (NOB)

Oppslagsdelen i NOB 2012 omfattar minst eitt sett av kategoriane (1) oppslagsord, (2) ordklasse, (3) bøyning. Oppslagsdelen i ein ordartikkel kan ha meir enn eitt oppslagsord med formopplysningar. Det inneber at to eller fleire ordbankartiklar kan vere kopla til same artikkel i ordboksdatabasen.

I NOB 2006 mangla ordklasse og bøyning for alle samansette oppslagsord. Det mangla også merking tilsvarende ordklasse for kategoriane forkorting, avleiingsprefiks, forledd i samansetning og symbol. I Ordbanken er desse kategoriane utfylte for kvar grunnform, men det mangla eit system for å føre denne informasjonen over i NOB. Det systemet er no på plass i og med at artikkelhovudet blir generert frå Ordbanken.

3.1. Frå eitt artikkelformat til fleire artikkeltypar

NOB 2006 har eitt artikkelformat. Det er ikkje skilt formelt mellom tilvisingsartikkel på den eine sida og normalartikkel med definisjon på den andre. Ved førstegongsredigeringa på 1970–80-talet var det eit overordna mål å gje lik status til alle former av oppslagsord, ettersom ein den gongen ikkje visste sikkert kva former som var mest i bruk. Eit oppslag kunne omfatte meir enn fire oppslagsformer.

I Ordbanken førte formatsidestillinga mellom hovudartikkel og tilvisingsartikkel til eit høgt tal mange-til-mange-koplingar, med fire eller fleire ordbankartiklar kopla til like mange ord-

artiklar. Det kunne vere opp til 25 koplingar utan sikker peikar til den eine ordartikkelen som inneheldt definisjon og bruksdøme. Det var ein etter måten liten del av ordartiklane i NOB som hadde parallellkoplingar, men det var mange nok til å gjere sambandet mellom ordboksdatabase og ordbank uoversiktleg.

I NOB 2012 er det skilt mellom (1) normalartikkel, (2) tilvisingsartikkel, (3) forleddsartikkel (oppslagsform endar med bindestrek) og (4) etterleddsartikkel (oppslagsform tek til med bindestrek). Normalartikkelen har alle rettskrivingsvariantar kopla til seg som oppslagsformer. Tilvisingsartiklar skal no ha ei oppslagsform, og er knytte til rett normalartikkel med «hyperlenking» frå definisjonsfeltet. Dette gjev ryddigare tilhøve mellom ordbank og ordboksdatabase.

3.2. Samsvar mellom oppslagsdel og tydingsdel i NOB

I NOB 2012 må alle ordartiklar ha ein definisjon. Dette kan settast opp som ei likning:

Oppslagsdel = Tydingsdel [tyding 1, tyding 2 .. Tyding n]

NOB har frå før mange skjelettartiklar for samansetningar, som består av oppslagsordet og ingen definisjon. Dei står som type-døme. Dei blir ståande i NOB på nett med ordklasse og bøyingsparadigme, men dei blir også lenka til ei samansetningsrekke, anten under ordartikkelen for forleddet eller i eigen ordartikkel for samansetningar (forleddsartikkelen).

Alle oppslagsformer for ein artikkel skal kunne brukast for alle tydingane som er definerte i artikkelen. Kan dei ikkje det,

er dei ikkje rettskrivingsvariantar, men nærsynonym. Eventuelle restriksjonar i form – t.d. i bruk av fleirtalsformer for konkret tyding (t.d. kjærleik [på pinne]) og eintal (kjensla kjærleik) som indikasjon på abstrakt tyding – er knytte til definisjonen, ikkje til oppslagsforma. I lys av prinsippa bak ordnett vil det kanskje vere best å knyte bøyning til dei ulike definisjonane. Men det vil vere eit særst omfattande arbeid som ein ikkje har resursar til no.

3.3. Splitting av artiklar i NOB

I NOB 2006 har ein del ordartiklar jamstilte oppslagsformer som ikkje er ortografiske variantar, men fullsynonym eller delsynonym. Døme er samansetningar med *lege* m1 el. *lækjar* m1 (ulik avleiing og etymologi), *sjuke* f2 el. m1 el. *sykje* f2 (ulike genus), eller *verd* n og *verdi* m (ulike genus, ulik etymologi).

I NOB 2012 er det eit krav at ordformer som er alternative oppslagsformer til same ordartikkel, må vere rettskrivingsvariantar. Dei må ha same ordklasse(r), same etymologi og ei forlikeleg normeringssoge. Samansette oppslagsord til same artikkel må ha same leddeling. Elles blir generering av oppslagsdel frå Ordbanken altfor innfløkt. Tanken om generering av oppslagsdelen frå Ordbanken til ordboksdatatabasen har såleis også ført til ein langt strengare lingvistisk stringens i hand-saminga av ordtilfanget for nynorsk.

4. Ordbanken, ordboka og språket

Rettskrivingsordlister føreset ein omfattande implisitt kunnskap om tyding. Mellom morsmålsbrukarar og innanfor ramma av ein stabil skriftkultur har det vore rekna som sjølvstøtt at alle språkbrukarar er morsmålsbrukarar, og at alle derfor kjenner tydingane på ord i ordlista, eller har folk rundt seg som veit kva orda tyder. Men slik er det ikkje lenger. Det svært store talet med andrespråksbrukarar i skole og arbeidsliv gjer at sambandet mellom ord(form) og tyding må gjerast eksplisitt for det normerte ordtilfanget. Ein fullformgenerator som manglar lenke på artikkelnivå til ein ordboksdatabase med definisjonar og bruksdøme, fortel ikkje meir enn ei gloseliste gjer. Dermed blir fullformsgeneratoren utilfredsstillande som språkverktøy. Ordbanken kan brukast til å lage tydingsberande språkverktøy av di han er knytt til ordartiklane i ei gjennomarbeidd ordbok.

Det er koplinga til NOB som ligg til grunn for at ein kan skilje mellom I *glis* m ‘person’, II *glis* m1, n1 ‘hånflir’ og III *glis* n1 ‘stor, fin bil’. Sambandet til tydingsdelen gjer det klart at Ordbanken her må ha tre «glis»-artiklar, og ikkje to eller ein. Det same gjeld skiljet mellom I *rett* m ‘(tillaga) mat’ og II *rett* m1 ‘privilegium, fridom’. Utan tyding, bruk og ordhistorie kunne ein ikkje sagt sikkert at dette skulle vere to artiklar i Ordbanken, og ikkje ein artikkel med to paradigme.

5. Resultat eitt år etter

Handordbøkene på nett er svært mykje brukte i Noreg, men bruken er førebels lite kartlagd. Trass i at første nettversjon var lite for auga, kunne Språkrådet i årsmeldinga si i 2008 skrive at «I 2008 vart det på denne adressa gjort 14,8 millionar oppslag i Bokmålsordboka og 10,7 millionar oppslag i Nynorskordboka. Bruken er tredobla på fire år og aukar stadig.» (s. 17).

Den nye formatet frå 2009 var utvikla med skolen for auga, og resulterte i auka bruk. Forbetra funksjonalitet kom i 2012 for NOB, og har blitt godt motteke. Det var 28 millionar oppslag i NOB i perioden november 2012 til november 2013. Det er såleis ikkje tvil om at innsatsen har svara seg når det gjeld å gjere formidlinga av norma for nynorsk standardspråk meir tilgjengeleg og effektiv.

Litteratur

- Beito, Olav (1986): *Nynorsk grammatikk*. Oslo: Det Norske Samlaget.
- BOB= Wangensteen, Boye (2005): *Bokmålsordboka. Definisjons- og rettskrivingsordbok*. (1. utg. 1986, 2. utg. 1993, 3. utg. 2006, ny nettutgåve 2013). Oslo: Kunnskapsforlaget.
- Engh, Jan (1993): *Linguistic Normalisation in Language Industry. Some Normative and Descriptive Aspects of Dic-*

- tionary Development. I: *Hermes, Journal of Linguistics* 10, 53–64. <<http://download2.hermes.asb.dk/>>
- Grammatiske termer til bruk i skoleverket* (mai 2005) (oppdatert 17.03.2006). Tilråding fra Norsk språkråd og Utdanningsdirektoratet.
<<http://www.sprakradet.no/upload/152/Gramterm.pdf>>
- Grønvik, Oddrun & Christian-Emil Smith Ore (2013): What should the electronic dictionary do for you – and how?. I: I. Kosem, J. Kallas, P. Gantar, S. Krek, M. Langemets & M. Tuulik (red.): *Electronic lexicography in the 21st century: thinking outside the paper. Proceedings of the eLex 2013 conference, 17–19 October 2013, Tallinn, Estonia*. Ljubljana/Tallinn: Trojina, Institute for Applied Slovene Studies/Eesti Keele Instituut, 243–260.
- Hagen, Kristin & Anders Nøklestad (2010): Bruk av et norsk leksikon til tagging og andre språkteknologiske formål. I: *LexicoNordica* 17, 55–72.
- Hoel, Jan (2010): Edle årganger i ny tapning. I: *Språknytt* 2010 nr. 3, 10–11.
- Johannessen, Janne Bondi & Kristin Hagen (2007): Nye grammatiske termer i skoleverket. I: *Språknytt* 2007 nr. 2, 24–30.
- NOB = Marit Hovdenak, Laurits Killingbergtrø, Arne Lauvhjell, Sigurd Nordlie, Magne Rommetveit & Dagfinn Worren (2006): *Nynorskordboka* (1. utg. 1986, 2. utg. 1993, 3. utg. 2001, 4. utg. 2006, ny nettutgåve 2012). Oslo: Samlaget.
- NRG = Jan Terje Faarlund, Kjell Ivar Vannebo & Svein Lie (1997): *Norsk referansegrammatikk*. Oslo: Universitetsforlaget.
- Språkrådet (2008): *Språkrådet 2008*. [Årsmelding]. Oslo: Språkrådet.

GRØNVIK OG ORE

Lenker

Norsk ordbank: <http://www.edd.uio.no/prosjekt/ordbanken/>
Bokmålsordboka og Nynorskordboka på nett:
<http://nob-ordbok.uio.no/perl/ordbok.cgi>

Arkivmateriale

Brev 22.12.2010 frå styringsgruppa for Norsk ordbank til
Språkrådet.

Brev 20.9.2011 frå Kulturdepartementet til Språkrådet.

Christian-Emil Smith Ore
førsteamanuensis, dr.scient.
faglig leder Eining for digital dokumentasjon
Institutt for lingvistiske og nordiske studium
Universitetet i Oslo
Postboks 1123 Blindern
NO-0317 Oslo
c.e.s.ore@iln.uio.no

Oddrun Grønvik
hovudredaktør, Dr. Litt. H. C. University of Zimbabwe
Norsk Ordbok 2014
Institutt for lingvistiske og nordiske studium
Universitetet i Oslo
Postboks 1021 Blindern
NO-0315 Oslo
oddrun.gronvik@iin.uio.no