



Eit lemmaregister for Ivar Aasens Norsk Ordbok med dansk forklaring i databaseformat

Oddrun Grønvik

Hvad Ordformerne især vedkommer,
da er de hertil hørende Enkeltheder saa mange,
at de ikke let kunne afhandles andensteds end i en Ordbog.
(A1864 § 384)

1 Innleiing

Ivar Aasens *Ordbog med dansk forklaring* (A1873) er ei grunnleggjande verk i studiet av norsk språk. Ordboka ligg føre som papiordbok¹ og dessutan som elektronisk tekstdokument i Språksamlingane ved Universitetet i Oslo².

Det har lenge vore eit mål å laga ein elektronisk versjon av A1873 som er (a) ein sjølvstendig database, der oppslagsord, artikkeltekst og geografiske kjelder er søkbare kvar for seg; (b) søkbar via moderne rettskriving og (c) lagd til rette for bruk i språksamlingane for nynorsk. Tilrettelegging for bruk i språk-

1) Med latinske typar frå 2003.

2) Ein versjon på cd som vart utarbeidd til Ivar Aasen-jubileet i 1996 er knapt brukande lenger.

samlingane inneber at alle ord i A1873 som kan reknast som oppslagsord med tilknytning til ein gjeven definisjon, skal lenkjast til ein artikkel i ordregisteret for språksamlingane for nynorsk – Metaordboka³.

Denne artikkelen skal dels handla om føresetnadene for at ein slik analyse- og oppkopplingsoperasjon blir gjennomført, dels om nokre innsikter i Aasens normeringsprinsipp som kjem for dagen når ein ser nærare på den delen av ordtilfanget som ikkje har fått plass som oppslagsord for ordartiklar.

2 Kva funksjon har A1873?

Ivar Aasens *Norsk Ordbog med Dansk Forklaring* (A1873) set fram eit utkast til eit standardisert skriftspråk bygt på norsk talemål frå samtida, med 1840 til 1870 som sentral dokumentasjonsperiode. Utkastet byggjer på omfattande innsamlingar av samtidig talemål, med påfølgjande analyse for å få gjeve skriftspråket eit einskapleg og vitskapleg rettvist uttrykk⁴. Ønsket om einskap, klårleik og systematisk samanheng er klarast sett fram i *Norsk Grammatik* (A1864), som, i tillegg til sitt eigenverd, må reknast som eit forarbeid til A1873. Dette ønsket går att under alle punkt i A1864: *Lydlære, Ordformer, Bøiningsformer, Orddannelse og Sætningslære*. Det same gjeld for dei to tillegga, *Om landskabsmaalene* og *Om landssproget*.

Det som skjer når eit talemål, uttrykt gjennom ulike dialektar, får eitt standardisert skriftleg uttrykk, er at ein gjev eit register av ordformer status som lemma, representert ved ei eller fleire grunnformer som er tilordna dei eigenskapane det aktuelle språk-systemet krev at eit lemma skal ha. Lemmaet er i utgangspunktet ein abstraksjon. Alle standardiserte skriftspråk er resultat av slike abstraksjonsprosessar, men etter måten få standardskriftspråk har fått form gjennom ein så komprimert og grundig dokumentert abstraksjonsprosess som Aasens utkast til norsk landsmål. For å forstå handteringa av ordtilfanget i A1873 må ein ta utgangspunkt i kriteria i A1864. Dette vil generelt gje betre forståing av korleis eitt standardskriftmål får form. Med ei kriterieliste for hand blir

3) <http://www.edd.uio.no/perl/search/search.cgi?appid=7&tabid=571>

4) Jf A1873 Fortale p. VII og A1864 § 382.

det også lettare å jamføra standardiseringsprosessen mellom ulike språk, og slik kan denne gjennomgangen også ha overføringsverd, endå om noka slik jamføring ikkje vil bli prøvt her.

3 Kvifor betra tilgjenget til A1873?

Det er eit utbreidt syn at A1873 fortener å bli brukt meir, både innanfor forskning og som referansepunkt for allmenn språkinteresse. A1873 har eit vell av opplysningar under ei litt tett overflate. Ideelt burde alle ordsøk for norsk språk ta turen innom A1873, som ein første sjekk.

Så kva er dei viktigaste grunnane til at folk i dag ikkje slår opp så ofte i A1873? Det er nok at ordboka ikkje finst på nett, og vidare at Aasens rettskriving skil seg ein del frå moderne (ny) norsk. For å bøta på dette blir det no laga ein versjon av A1873 som er (1) ein strukturert database, der oppslagsord, artikkeltekst og geografiske kjelder er søkbare kvar for seg; (2) søkbar via moderne rettskriving og (3) samordna med dei andre komponentane i språksamlingane for nynorsk.

Direkte tilgjenge til ordboksteksten som søkbar database vil også lyfta nytt stoff fram i dagen. A1873 er ekserpert, grundig og fleire gonger, som del av språksamlingane ved UiO⁵. Bruk av A1873 i redigeringa av Norsk Ordbok har likevel synt at tidlegare, manuelle ekserperingar ikkje er fullstendige. Redaktørane må derfor alltid bruka grunnteksten både til kjelde og kvalitetskontroll⁶ og det blir mykje (til dels forvirrande) blading i papirordboka, eller søking att og fram i e-teksten. Berre det å kunna sjå fleire artiklar samstundes og i parallell vil gjera A1873 mykje lettare å bruka.

4 Ordboksteksten

4.1 Papirutgåver

A1873 er ei hovudkjelde for dokumentasjonen av moderne norsk. Men papirordboka er i dag truleg lite brukt utanfor krinsen av

5) NLO nynorsksetlane

6) Grønvik et al (2014) s. 331 ff kap. 10.7.

fagspesialistar. Dette har fleire grunnar. Papirversjonane 1873–1985 er sette i fraktur, ei skriftform som dei færreste no les lett. Opplaga har aldri vore store etter dagens mål⁷. Ei nyare papirutgåve frå 2003 er sett i latinske typar, men er ikkje identisk med den opphavlege teksten, av di Aasens tilleggsartiklar i 1873-utgåva er innarbeidde i teksten utan å vera markerte⁸. Skal ein vita nøyaktig kva Aasen skreiv må ein derfor framleis gå attende til 1873-utgåva, eller til opptrykk av den.

4.2 Elektroniske tekst

A1873 finst også som elektronisk tekst. Ordboka vart innskriven og korrekturlesen ved Språksamlingane, Universitetet i Oslo, framfor hundreårsminnet om Aasens død i 1996. Ordboka vart publisert i ein cd med Aasen-tekstar. Desse tekstane er seinare lagde ut på nettet av Nynorsk Kultursentrum⁹. Sjølve teksten i elektronisk form finst også som del av Språksamlingane ved Universitetet i Oslo.

4.3 Status elektronisk tekst

Men endå om teksten ligg føre og er søkbar i elektronisk form, er han ikkje lettbrukt for dagens norskkunnige. Dette kjem seg av (1) rettskrivinga i oppslagsord og døme, som skil seg ein del frå moderne nynorsk. (2) Vidare er framstillingsforma kompakt, i tråd med papirordboka sitt mål om å få med mest mogleg informasjon på minst mogleg plass. (3) Aasens ordbokskonvensjonar er naturleg nok også delvis forelda. Foreldinga gjeld særleg namneformene i dei geografiske heimfestingane. Skrivemåten for stadnamn i Noreg er mykje endra sidan A1873 kom ut, mellom anna som følgje av Aasens arbeid med ordtilfanget¹⁰. Dei administrative inndelingane som Aasen refererer til, er også endra. (4) Organiseringa av kategoriar i ordartikkelen i A1873 er for si tid mønstergyldig systematisk, men likevel ikkje så systematisk gjennomført som i moderne databaseprodukt.

7) Totalt opplag frå og med 1873 er noko over 15 000. Jf. Grepstad (2014) nr. 16 (1000), 1006 (ikkje oppgjeve), 1018 (10 000), 1045 (2000), 1049 (2000), 1074 (ikkje oppgjeve).

8) Kruken og Aarseths utgåve 2003 s. XXXI

9) <http://www.aasentunet.no/?module=Articles&action=Article.publicOpen&id=4435>

10) Indrebø (1927: 57).

Den viktigaste skilnaden er at ei ordbok på Aasens tid vart skriven som ein einaste, samanhengande tekst, og føreset slik lesing. Dagens ordbøker blir i langt høgare grad redigerte med tanke på at kvar artikkel skal stå seg som eit sjølvstendig dokument, slik bruk på nettet krev. I dette tiltaket er det likevel snakk om å letta tilgjenget til Aasens tekst, utan å gjera endringar.

4.4 A1873 i databaseformat

Hausten 2014 vart A1873 lasta inn i eit førebels, enkelt databaseformat¹¹, som tillet søk og sortering på oppslagsord og på detaljar i den xml-koda artikkelteksten. A1873 er ikkje publisert som søkbar database enno.

Det at A1873 no er søkbar, gjer det mogleg å uttala seg med større visse om eigenskapar ved ordboka enn ein før kunne gjera. A1873 har om lag 38 000 ordartiklar. Alle tal lenger ut i denne artikkelen skriv seg frå søk i den førebelse databaseversjonen av A1873.

5 Kva bør vera søkbart som oppslagsord i A1873?

5.1 Kriterium og kriteriumsuttrykk

Ein viktig føresetnad for å få god utnytting av ordboksteksten er at alle ordformer i A1873 som bør vera søkbare som oppslagsord, blir sorterte ut så enkelt og effektivt som mogleg frå sjølv tekstmassen. Manuell fullekserpering er for arbeidskrevjande, vil ta for lang tid, og blir uansett mindre funksjonell ei xml-koda utgåve vil vera.

Ein kan nærma seg denne prosessen frå to kantar, og båe er naudsynte. For det første må ein isolera dei kriteriuma Aasen sjølv formulerer i A1864. Dette vil gje haldepunkt for å slå fast korleis dei norske ordformene i den danskspråklege artikkelteksten vart vurderte av Aasen sjølv.

For det andre må sjølv ordboksteksten analyserast, med tanke på å sortera ut formlar i redigeringspråket som skil ulike

11) Den noverande databaseversjonen av A1873 er skapt etter analyse ved Christian-Emil Smith Ore, Eining for digital dokumentasjon, Universitetet i Oslo, med noko bistand frå underskrivne.

ordformer i grupper, og vonleg svarar til dei kategoriane Aasen sjølv formulerer. Sidan A1873 er ein gjennomarbeidd og godt planlagd ordbokstekst, kan ein ha god von om å finna slike formar, og at dei er systematisk og konsekvent brukte. Men tekstlesingsprinsippet som ligg under, gjer at ein også må rekna med avvik.

5.2 Standarddefinisjon på oppslagsord – og brukt på A1873

Med *oppslagsord* meiner ein til vanleg artikkeloverskrifta for ein ordartikkel, som er tilordna funksjon som samlepunkt for kategoriane rettskriving, ordklasse, bøyingsparadigme, etymologi og tyding. Oppslagsordet får status som lemma, representert ved ei eller fleire grunnformer.

Alle artikkeloverskrifter i A1873 har sjølv sagt status som oppslagsord, men spørsmålet er om det finst meir. Da må ein sjå forbi sjølv oppslagsordet for ordartikkelen, og vurderer om det finst ordformer inne i artikkelteksten som Aasen har tiltenkt status som lemma. Eit hovudspørsmål blir då kva Aasen tenkte om dei indre strukturane i ordtilfanget som heilskap. Korleis kan ein resonnera om samanhengane mellom ulike ord? Kva høyrer saman og bør presenterast samla? Ei drøfting av dette finn ein i A1864, i kapitla om ordlagning. Desse spørsmåla er også nemnde i til to tillegg til A1864¹².

I A1864-kapitlet om ordlagning vel Aasen å nærma seg ordtilfanget både frå form og frå tyding. Tilnærminga ut frå form (§§ 250–266) fører over i drøftingar av ulike slag avleiing og samansetning, med rotavleiing først, deretter suffiksavleiing og så samansetning¹³. Frå leksikografisk synsvinkel peikar denne tilnærminga mot alfabetisk artikkelordning. Ei alfabetisk ordning av ordtilfanget for eit språk som norsk syter for at dei korte grunnorda med eventuelle rotavleiingar kjem samla og først, deretter avleiingar og samansetningar. I ei ordbok kjem ein alfabetisk ordna ordbolk dermed som hovudregel med det viktigaste først, det meir perifere lenger ut.

Tilnærminga ut frå tyding (§§ 267–300) tek for seg sam-

12) *Om landskabsmaalene og Om landssproget.*

13) Aasens klassifisering av avleiingselement skil seg noko frå dagens. Såleis er ord som byrjar på *and-* eller sluttar med *-dom* rekna som samansetningar (§ 258–259), men Aasen peikar på at tyding er avbleikt og bruk ikkje lenger sjølvstendig.

bandet mellom avleiingsform og tyding, og fører over i drøfting av ulike tydingsrelasjonar og korleis slike relasjonar knyter seg til ordform¹⁴. For leksikografar peikar denne tenkjemåten over mot tydingsordninga inne i artikkelen; dersom ei avleiingsending alltid fører til at rota blir tillagd ei gjeven tyding, så bør den tydinga kanskje ha fast førsteplass. Dømet som Aasen nemner først er verbalsubstantiv (på *-ing*, *-nad* osv.) som inneber substantive-ring av sjølve verbhandlinga.

På makronivå vil ei semantisk tilnærming til ordtilfanget ha som sitt leksikografiske motstykke ordbøker skipa som ontologiar, anten det er i form av tesaurusordbøker, elektroniske ordnett eller andre former for omgrepstyrt klassifisering av ordtilfanget. Men også i den alfabetiske ordboka har tilnærming til ordtilfanget ut frå omgrepsordning ein plass, innanfor og i tilknytning til definisjonar. Overordna omgrep blir hyperonym i definisjonar, ord med same tyding og bruk kan bli førte som synonym, og avleiingar og samansetningar som ikkje blir handsama i eigne ordartiklar, kan førast som døme på ordlagingssevne og dermed utdjupa beskrivinga av grunnordet som er oppslagsord for artikkelen. Både tilnærmingar til ordlagingsord er relevante for leksikografien. Det at Aasen er sterkt oppteken av sambandet mellom form og tyding på eit overordna nivå, og av organisering av ordtilfanget ut frå tyding, gjer at ein bør sjå ikkje berre oppslagsorda, men også artikkeltekstane, som fora for presentasjon av ordtilfanget.

5.3 Prinsipp for lemmautval og lemmapresentasjon

Korleis valde Aasen ut dei orda som skulle få eigne ordartiklar, og korleis skilde han ut det ordtilfanget som fekk ei meir kurso-risk handsaming?

Alt Aasen tok med i A1873, var dokumentert i samlingane hans. Det kan vera tenleg å tenkja seg Aasens ordsamlingar som eit korpus, innsamla over dryge tretti år, som syner talemålet som ein mangearta masse av ordformer. Kvar ordform har i utgangspunktet preg av lokal fonologi og av kontekst i bruk. Fonologien kan knytast til stad (geografisk heimfesting), brukskonteksten fortel om bøyning og ordstilling, og dermed indirekte om ordklasse

14) Jf. § 267 Anm., der Aasen drøftar den systematiske tydingsskilnaden mellom kortavleiingar og *-ing*-avleiingar frå verbrøter.

og tyding. Somme ordformer er mykje brukte, andre sjeldne. Ord-
lengd kan i teorien variera frå eitt fonem, attgjeve med ein eller
to bokstavar (/i/, /aa/), til lange samansetningar¹⁵. Somme
bøyingsparadigme er fullstendig dokumenterte, andre ikkje.
Somme avleiingsgrupper er godt utbygde rundt ei grunnform,
andre ikkje. Somme grunnformer er produktive i samansetningar,
andre ikkje. Dette kan sjåast som ein serie graderte skalaer for
meir eller mindre systemrett uttale, meir eller mindre systemrett
samsvar med bøyingsformer eller andre nærskylde ordformer,
meir eller mindre identifiserbar tyding. Vilkåret for å kunna
gje eit ord status som lemma er styrt av kva informasjon som ligg
føre i høve til dei ulike skalaene. Såleis kan eit ord vera full-
dokumentert i form, men likevel ikkje mogleg å ta med i ei ord-
bok av di tydingsinformasjonen ikkje er god nok.

Når tilfanget så skal ordnast til ei lemmaliste som skal
leggjast fram i ei ordbok, seier det seg sjølv at korte ord med god
geografisk distribusjon og etter måten høg frekvens får prioritet.
Om dei i tillegg er produktive både i høve til avleiingar og saman-
setningar, får dei ein sentral plass både som kjerne i ordklynger
i ordboka og som representantar for ordlagingsystemet i det ak-
tuelle språket. Det er slike ord – dei ”store orda” – den vitskaplege
leksikografen kan seia mest om¹⁶.

I A1864 kjem det direkte og indirekte fram – gjennom
kommentarar og gjennom ordning av tilfanget – korleis Aasen ser
på generelle eigenskapar ved ordtilfanget. Han har ein klar
preferanse for korte ord med sterk forankring i språket både dia-
kront og synkront, og med sterkt potensial i avleiing og saman-
setning. Dette grunnjev han dels med systemomsyn; slike ord
eksemplifiserer systemet og dannar slik sett modellar for ordla-
ging som kan brukast om att i den vidare utviklinga av ordtil-
fanget (A1864: 346 f). Eit anna omsyn er behovet for transparens.
Aasen er svært oppteken av at ordtilfanget skal vera spontant
forståeleg og effektivt for meinigmann. Då må ordlagning for nye
og ukjende omgrep i den grad det er mogleg gå ut frå det kjende,

15) Om ein tek utgangspunkt i Nynorskorpuset, ser det ut til at det går ei grense for
ordlengd i norsk på omkring 35 bokstavar og 10–12 stavingar. Så lange samansetnin-
gar verkar utprega skriftspråklege og vil neppe dukka opp ofte i talemålsdokumenta-
sjon for norsk.

16) ”Det blev derfor nødvendigt at vrage alt det, som kunde vrages, og at affærdige
de ubetydeligste Ord saa kort som muligt, men derimod behandle de vigtigere Ord
med en større Flid og Udførlighed.” (A1873: XI)

og byggja på den omgrepsverda som er kjend, og den allmennmenneskelege førestellingsevna som set oss i stand til å tolka nye former og kombinasjonar utifrå assosiasjon med det vi alt veit (A1864: 348 og A1873 Fortale: XII).

Aasen filtrerte heile ordformmassen gjennom dei kriteria han hadde bestemt seg for og gjort greie for i A1864. Som resultat av denne sorterings- og filtreringsprosessen sat han att med eit stort tal grunnformer der han var sikker på kva rettskrivingsforma – mønsterforma – måtte bli i A1873, både i oppslagsform og bøyingsparadigme. Når dei gjeld oppslagsord for ordartiklar, har han med få unntak vore sikker på korleis normert form bør sjå ut, med ordklasse og eventuell bøyning. Dei få unntaka – markerte med ”(?)” etter oppslagsordet ¹⁷ – har for det meste blitt stadfest som rette i ettertid, ettersom meir tilfang er kome til.

5.4 Kva omfattar normering?

I moderne norsk blir nytt ordtilfang, som rutine i normeringsprosessen, tilordna normert form, somtid med alternativ, og likeins ordklasse og bøyingsmønster. Til vanleg er dette enkelt, for det ligg føre eit omfattande system av mønster og førebilete. Eventuelle tvilstilfelle får ei meir omfattande handsaming¹⁸, men tvilen blir ikkje synleggjort for språkbrukarane i normeringsvedtaket.

Som nemnt ovanfor ser det ut som om Aasen har vore sikker på analysen og dermed oppslagsforma for alt utan under ein promille av oppslagsorda for ordartiklane. Ordklasse er sløyfa berre unntaksvis. Berre ord med systemavvik, dvs. uregelrett bøyning i høve til dei mønsterforma som er sette opp i A1864, har eksplisitte bøyingsopplysningar – om lag 10 prosent. Det at normering av et lemma i norsk implisitt omfattar systemtilpassing ned til bøyingsskjema, er såleis noko som går attende til Aasens praksis i A1873.

Det er rimeleg å tenkja seg at dess mindre sikker Aasen var på kva mønsterforma burde vera, dess nærare ville han halda seg til den dokumenterte talemålsforma, dersom han i det heile fører opp ordforma i A1873.

17) Om lag 300, av om lag 38 000 artiklar.

18) Jf. Språkrådets ”Retningslinjer for normering” 01.09.2015 pkt 4.

5.5 Lemmautval ut frå føremål

Aasens føremål med A1873 var å prova identiteten til norsk språk ved å gje ein einskapleg presentasjon av det heimlege ordtilfanget. Dette skjer i ei samfunnsstode der det blir reist tvil om det i det heile finst noko sjølvstendig norsk språk, og der norsk talespråk står under press frå framandspråkleg skriftkultur, særleg dansk¹⁹.

Når det gjeld lemmautval, legg derfor Aasen – i denne rekkjefølgja – vekt på (1) dokumentasjon frå norske målføre (synkron talemålsdokumentasjon), (2) (synkront) slektskap med andre nordiske språk, (3) slektskap med andre germanske språk, (4) form i norrønt som svarar til grunnforma eller er nærskyld med henne (A1864: § 377–378). Om alle desse opplysningane ligg føre og heng saman, finst det grunnlag for å etablera ei leksikalsk mønsterform. Mønsterforma svarar til oppslagsforma i ordboka A1873, med ordklasse og eventuelt bøygde former²⁰.

Det einaste ufråvikelege kriteriet er at det må liggja føre synkron talemålsdokumentasjon. Aasen tok med berre ordformer som var dokumenterte i tilfanget hans som talemålsord i samtida²¹, frå ei kjelde han leit på. Det kan vera verdt å ha i mente kor lite dette tilfanget må ha vore i laupande ord, jamført med dagens talemålsinnsamlingar. Aasen og hans samtidige måtte kompensera dette med dess grundigare systemarbeid, slik at dei kunne resonnera om former, samanhengar, opphav og tyding.

Aasen har også eit negativt utvalskriterium, som er knytt til føremålet med ordboka. Føremålet er (1) å dokumentera det norske kjerneordtilfanget, (2) å styrkja ordlaging på heimleg grunn (A1864: § 389). Han reknar derfor ikkje nyare importord, same kva opphavsspråk det er tale om, som norsk ordtilfang (A1864: 346). Dette grunngjev han med at dei nyare importorda er lite gjennomsynlege, lagar ugreie i ordlagingsystemet, og fortrengrer heimlege ord. Importord frå mellomalderen reknar han derimot som tilpassa norsk og ”fuldkommen indlemmede i Sproget.” (A1864: § 389).

19) ”Hovedsagen er, at man her har for sig et levende Sprog, som tilhører Landets Folk og staar paa sin egen Grund, saa at det ikke lader sig forklare som en Underafdeling af noget andet Lands Sprog.” A1864 § 377 Anm.

20) ”den fuldkomneste av dem optages som almeengyldig til skriftig Brug”. § 385.

21) A1873 Fortale p. X. Eit utval ord frå eldre målføreordsamlingar som er rekna som verdfulle, men synkront udokumenterte, er oppførte med kjelde i eige tillegg s. 973–974.

For kjerneordtilfanget, som er godt dokumentert, er etter måten eins i forma frå landsende til landsende, og har kognat i nærskylde språk, er det enkelt å fastsetja mønsterform²². Dess meir av idealdokumentasjonen som manglar, dess vanskelegare blir det å identifisera ei ordform, dvs. stipulera ei systemrett mønsterform (med bøyning). Eit anna og viktig poeng er at dess større variasjonen i talemålsformene er, dess meir etymologiserande vil mønsterforma bli, under elles like tilfelle. Døme som Aasen sjølv dreg fram, er: *fiwrelde* n ("Fyrveld, Fryvil"); *ljøs* n ("Ljøs, Løs, Jos, Jøs"); *ver* n ("Veder, Veer, Vær, Veir"); *nesle* ("Netla, Nesle, Nosle, Nasle") (A1864 § 385).

5.6 Lemma inne i ordartikkelen

Kvar ordartikkel har sitt oppslagsord, og slik sett kan ein setja lågaste moglege tal oppslagsord i ordboka lik artikkeltalet: Det er 38180 ordartiklar i A1873.

Men Aasens ordartiklar inneheld ofte fleire ordformer som fortener lemmastatus. Slike ordformer er (a) samansetningar med oppslagsordet for artikkelen som førsteledd, (b) samansetningar med oppslagsordet for artikkelen som etterledd, (c) avleiingar, (d) fullsynonym, (e) nærsynonym, (f) antonym, (g) ord som høyrer til same semantiske sett.

Dersom denne ordmengda har lemmastatus, og finst som oppslagsord for andre ordartiklar, er det ikkje så viktig å finna orda inne i artikkelteksten. Men om orda er nemnde i artikkelteksten og berre der, er det viktig å kunna få tilslag ved søk også i desse orda.

6 Tekstmarkørar

Ønsket er å få søkt opp og sortert ut alle ordformer som er potensielle oppslagsord i A1873 ved hjelp av ein programmert tekstanalyse. Når dette arbeidet først blir gjort, er det betre å ta til vurdering litt for mykje enn litt for lite.

Føresetnaden for vellykka programmert tekstanalyse er at det finst markørar i teksten som skil oppslagsordkandidatar frå annan tekst. Slike tekstmarkørar kan vera typografiske eigen-

22) Døme s. 343 Anm. 1. avsnitt

skapar, eller teikn eller tekst i omgjevnadene til oppslagsordkandidaten.

Optimal innhausting av oppslagsordkandidatar føreset konsekvent og eksklusiv bruk av tekstmarkørar. Eit døme kan vera at teiknet kolon (:) **alltid** blir følgt av eit bruksdøme, eller at ordet ”hertil” **alltid** blir følgt av døme på samansetning med artikkelens oppslagsord som førsteledd.

Slik ufråvikeleg konsekvens får ein aldri i eldre ordbøker. Dei har vedteken artikkelstruktur og faste konvensjonar, og dei kan vera svært systematisk tenkte. Men produktet er ein laupande tekst som må lesast som laupande tekst, om ein skal få med all informasjon. Ein kan derfor ikkje rekna med å få tak i alle oppslagsordkandidatar som sikre treff ved ein programmert analyse av teksten.

Det ein kan gjera, er å finna dei sikraste markørane, ta ut desse frå teksten i konkordansar, og så haka av for ordformene nær den aktuelle tekstmarkøren som skal reknast med som oppslagsordkandidatar. Dess meir heterogen markeringa av moglege oppslagsord er, dess meir krevjande blir prosessen.

6.1 Tekstmarkørar i A1873

A1873 har sine konvensjonar i prent. Desse omfattar eigenskapar ved satsen, og tekstmarkørar. Satseigenskapar omfattar satstype (fraktur og latinske typar) satsstorleik og uthevingsgrad (halvfeit og rettskrift). Som tekstmarkørar kan ein rekna spesielle teikn og ordlag som til vanleg har oppslagsordkandidatar i ein gjeven, fast posisjon i nærleiken. I A1873 som word-dokument er dei fleste av denne konvensjonane tekne vare på, men ikkje alle.

Aasen bruker ofte, men ikkje alltid, halvfeit sats på ordformer som skal framhevast i teksten. I den prenta teksten er det også brukt satsstorleik og sperring til å skilja typografisk mellom opplysningskategoriar. I dei elektroniske filene er variasjon i satsstorleik og sperring sløyfa. Bruksdøme er skrivne inn i kursiv (i staden for liten rettskrift), all annan tekst er vanleg eller halvfeit rettskrift. Den elektroniske teksten av A1873 har såleis halvfeit sats til utheving av oppslagsord, bøyingsformer, målføreformar, avleiingar, samansetningar (som ofte manglar det leddet som fell saman med oppslagsordet) og ord som er mål for tilvisingar. Ein kan ikkje nøya seg med å ta med alle ordformer som

finst i halvfeit sats. Då ville ein få med både for mykje og for lite.

Dermed må ein stø seg på tekstmarkørane, og fina ut kva dei skal tyda i kontekst av å sjå på korleis dei er brukte. Alle døme som er henta fram i denne artikkelen er tekne frå ei testutgåve av A1873 i databaseformat, som førebels ligg i utviklingsdatabasen til Eining for digital dokumentasjon ved UiO²³.

Døme på tekstmarkørar inne i ordartiklane er t.d. ”s.”, som tyder ‘se’. Ord i halvfeit som følgjer, er oftast mål for ei tilvising. Den enklaste typen er denne:

blenda, s. blanda.

Ein annan tekstmarkør, sekvensen ”hedder ogsaa” fortel at det neste ordet (ofte i halvfeit) kan vera eit fullsynonym.

breiddfull, adj. fuld til Bredden. Hedder ogsaa: **breidmett** (Tel.) og **breiddmælt**.

Ein annan metode for tilvising til fullsynonym er parentes som tek til med likskapsteikn.

Bleng, m. 1) Valle ... (= Raamysa)

Ordet ”hertil” står framfor døme på samansetningar og avleiingar, og rekkjer av slike ord blir gjerne avslutta med ”og fl.”. I slike rekkjer må somme element ekspanderast med føreledd tilsvarande oppslagsordet for ordartikkelen, som nedanfor:

Blekk n 2. ... Hertil: **Blekksmid**, **Blekklykt**, **-plata**, **-røyr**, **-spann**, **-trekt**, og fl.

Eit føremål med den vidare tekstanalysen er å sjå om ordformer som opptrer som mål for tilvisingar, som synonym eller avleiingar og samansetningar til oppslagsordet bør registrerast som potensielle lemma i A1873. Dersom ordformene finst i oppslagsordregisteret, er det dobbelregistrering og ikkje naudsynt å ta det med når det står framheva inne i ein annan artikkeltekst. Men om føringa inne i ein ordartikkel med eit anna oppslagsord, er ein måte å presentera ordtilfang på som (av ein eller annan grunn) ikkje har fått eigen ordartikkel, så er det viktig å få dei med i lemmalista.

23) Satsbiletet i dømet er kopiert frå databaseutgåva, og er som nemnt ei forenkling av satsbiletet i den prenta utgåva.

A1873 gjev rik dokumentasjon av formvariasjon attom oppslagsorda. Oppslagsordregisteret for den elektroniske A1873 skal (i denne omgangen) ikkje ha med målføreformer – det vil gje for mange dublettar. Talemålsformene kan for det meste skiljast frå potensielle oppslagsformer av di dei står rett før eller etter ein geografisk markør²⁴, oftast eit stadnamn, som regel i parentes. For å få skilt målføreformene frå ordformer som skal ha ein plass i lemmaregisteret, er det etablert eit register over (forkorta) stadnamn i teksten. Dette registeret omfattar mange fleire enn dei forkorta stadnamna i forkortingslista i den prenta boka²⁵.

7 Organisering av ordtilfanget i A1873

7.1 Lemma med rolle som fullsynonym

Ein artikkel kan romma fleire ord som fortener oppslagsstatus. Hovudsystemet i A1873 er at oppslagsforma representerer einaste lemma i ordartikkelen. Men det hender at ordartikkelen endrar fokus ved definisjonen, frå å handla om oppslagsordet, til å handla om omgrepet. Det er som om Aasen har prøvt å kombinera den alfabetiske formmodellen for ordbøker med den ontologiske omgrepsmodellen. Då kan ordartikkelen innehalda eitt eller fleire fullsynonym til oppslagsordet, t.d. med ei anna geografisk heimfesting. Somme fullsynonym har eigen artikkel på alfabetisk plass, men ikkje alle. Dermed må ein, om alle definerte ord i A1873 skal vera søkbare, også få med desse orda.

Vanlege redaksjonelle formlar framfor synonym er ”Hedder ogsaa”, ”Andre St.”, og ”(=”. Nedanfor følgjer døme på bruken og kommentar til status for dei potensielle synonyma.

Formelen ”Hedder også” har 1 028 tilslag i ordartiklar.

vimutt, adj. fuld af Nykker eller besynderlige Indfald (Vim).

Hedder ogsaa **vimen**, og tildeels **vimren**.

A1873 har ikkje eigne artiklar for *vimen* og *vimren*.

gløymsam, adj. glemsom. Hedder ogsaa gløymen, **gløymisk** (Østl.), **gløymsken**, el. gløymskjen (B. Stift).

24) Termen ”hemfesting” er nytta i nynorsk leksikografi = ‘Namn på stad der ordet høyrer heime’.

25) Per dato er 185 ulike stadnemningar isolerte, irekna forkorta ordlag av typen ”Ng, St.” ‘Nogle Steder’)

Adjektivet *gløymen* har eigen artikkel i A1873. Dei andre adjektiva har det ikkje.

Formelen ”Andre st.” har 899 tilslag i ordartiklar. Føremålet må vera å opplysa om ei anna form av oppslagsordet som er utbreidd, kanskje så utbreidd at ho kunne ha vore brukt som oppslagsform.

Vundel, m. 1) en liden Visk el. Tot. Hall. Andre St. **Vindel**.
(Jf. Vondul).

flattelgja, v.a. (er, de) , tilhugge Tømmerstokke saaledes at Siderne blive flade. ”Flattægje”, Tel. Andre St. flathogga, flatrydja (flatry’, Nordl.).

Oppslagsforma *Vindel* har ordartikkel, men med ein annan, urelatert definisjon. Oppslagsforma *flatrydja* har artikkel og er synonymdefinert til *flattelgja*, medan *flathogga* manglar artikkel.

Formelen ”kaldes” og ”kaldes ogsaa” fungerer som varsel om nærskylde ord som fortener å nemnast, ofte samansetningar. Den har tilslag i 150 ordartiklar. Eit døme er

Bjorhus, n. Bæverhule, Hytte som Bæveren har bygget. ”Bjørhus”, Østerd. Ogsaa kaldet **Bjorgard** (Tel.), og **Bøverbrote** (Gbr.). Levninger deraf kaldes **Bjortuft**, Bjørtomt. Østerd.

I denne artikkelen kjem det som truleg er den vanlegaste nemninga for omgrepet først, for oppslagsordet er ikkje heimfest. Så kjem dei mindre utbreidde nemningane, og til sist ei anna, relatert samansetning. Ordet *bøverbrote*, her ført som synonym til *bjorhus*, har artikkel, men med ei anna tyding. Orda *bjorgard* og *bjortuft* har ikkje eigen artikkel.

Formelen med likskapsteikn i parentes, følgt av ei ordform – (= ...) – er oftast nytta som innleiingsmarkør for fullsynonym til oppslagsordet. Likskapsteikn-formelen har 2 994 tilslag i ordartiklar. Døme er

Vikemot (i’), n. Tiden da en Uge er fuldt udløben (= Aattedagsmot). Tel. og fl.

yta, v.a. (er, te) ... 2) støde, give et Stød eller Tryk (= skuva). Gbr. Hall.

Aattedagsmot har eigen ordartikkel. Det same har *skuva*. Ingen av desse artiklane viser attende til den relevante artikkelen ovanfor; synonymien er uttrykt berre her.

7.2 Mål for tilvisingsartiklar

I moderne ordbøker er ein tilvisingsartikkel eit likskapsteikn som peikar frå ei ordform til ei anna, og om ein slår opp på den andre, finn ein ein innhaldsartikkel om det tilviste ordet. Tilvisingsartikkelen blir brukt til å visa frå ei form av oppslagsordet til den forma som innleier innhaldsartikkelen; oppslagsordet finst i fleire former, eller det er naudsynt å visa frå ei bøygd form av eit oppslagsord til grunnforma. Slik blir tilvisingsartikkelen brukt også i A1873 høgre kolonne (kommentar i sjå tabell 1):

Tabell 1 Tilvisning frå talemålsform til innhaldsartikkel.

Tilvisingsartikkel	Kommentar
Greisla , s. Greidsla	Oppslagsordet er ei målføreform av ordet det er vist til, og føremålet er å minna om dei ortografiske prinsippa i A1873 (stomnutlyd på <i>-d</i> blir skriven)
flaug , fløi; s. fljuga	Oppslagsordet er ei bøyingsform av ordet det er vist til (preteritum av verb med oppslagsform i infinitiv)

Tabell 2 Tilvisning med implisitt føremål.

Tilvisingsartikkel	Kommentar
flatrydja v. a. s. flattelgja	Ordet det er vist til, er fullsynonym med oppslagsordet, og implikasjonen er at definisjonen av ordet det er vist til, også gjeld for oppslagsordet.
Baksterfløyg , m. s. Fløyg.	Ordet det er vist til, er etterledd i det samansette oppslagsordet. Definisjonen under <i>Fløyg</i> høver også med samansetninga <i>Baksterfløyg</i> . Implikasjonen er at oppslagsordet er eit meir eksplisitt fullsynonym til ordet det er vist til.
flakebringa , s. flaka, v. n.	Oppslagsordet inngår i eit fast uttrykk under <i>flaka</i> . Uttrykket er definert, ikkje adjektivet <i>flakebringa</i> . Implikasjonen er at førsteleddet i oppslagsordet har samanheng med ordet det er vist til.
seten , part. s. vanseten.	Oppslagsordet er partisippform av eit verb, nytta som etterledd i avleide adjektiv. Ordet som det er vist til, er definert, og definisjonen omfattar forklaring av partisippforma. Implikasjonen er at oppslagsordet svarer til etterleddet i ordet det er vist til.
Skjelm , m. Frøbælg, Ærtebælg. Gbr. Andre St. Skjelmer, s. Skolm.	Oppslagsordet har same konsonantisme som ordet det er vist til. Implikasjonen er at det er ein etymologisk samanheng, og i artikkelen for <i>skolm</i> gjer Aasen greie for uregelrett fleirtal og for formvariasjon. Men definisjonen til oppslagsordet <i>skjelm</i> tyder på at Aasen ikkje er sikker på om <i>skjelm</i> synkront skal reknast som formvariant av <i>skolm</i> eller fullsynonym til <i>skolm</i> . Han let derfor både artiklar stå med definisjon.

Som tabell 2 viser, er det i predigitale ordbøker ofte ikkje skilt kvast mellom normalartikkel og tilvisingsartikkel. Tilvisingsartiklane kan vera av andre slag enn dei enkle formekvivalentane i tabell 1. I tabell 2 er de vist døme på tilvisingsartiklar i A1873, der oppslagsordet ikkje er ei form av ordforma det er vist til. I slike tilvisingar må ordboksbrukaren resonnera seg fram til føremålet med tilvisinga, og båe ordformer må vera søkbare.

7.3 Antonym

A1873 Formelen ”Modsat” har tilslag i 383 ordartiklar.

flattelgja, v.a. (**er, de**) , tilhugge Tømmerstokke saaledes at Siderne blive flade. .. Modsat: rundtelgja, magatelgja.

Lemmaet *rundtelgja* finst berre som antonym i denne artikkelen, medan *magatelgja* har eigen ordartikkel.

versna, v.n. (**ar**) , foverres, blive værre. Tildeels meget brugl., men lyder oftest som vessna; nogle St. væsna. (Modsat besna, batna).

Båe desse antonyma har egne artiklar.

7.4 Rekkjer med samansetningar og avleiingar

Formelen ”Hertil”, brukt ved avleiingar og samansetningar til oppslagsordet, har 2 488 tilslag i ordartiklar. Eit typisk døme er dette:

Vise, m. ... **3**) Blomst, Begyndelse til Frugt. Sfj. Yttre Sogn. Hertil Heggjevise, Raunevise, Jordbærvise, Molteviser og fl. Paa Sdm. kun: Natevise, dvs. Nøddeknop, Hunblomsten paa Hasselen.

Av samansetningane på *-vise* har to eigen artikkel, dei andre ikkje. Formelen ”Hertil” blir også brukt til å innleia det ein no ville rekna som ei rekkje med artiklar for samansetningar:

Visa, f. en Vise, Folkesang. (G.N.vísa: Vers, Strophe). Hertil **Visebok** (oo), f. Visebog. **Visefugge**, m. Bundt el. Pakke af Viser. (Tel.). **Viselund**, m. Melodie til en Vise. (Tel.) **Visenote**, m. og **Visestubbe**, m. Stump af en Vise. **Visevedstev**, n. enkelte Vers eller Stropher. (Hall.). Jf. Stev.

”Hertil” fører oftast, men ikkje alltid, fram mot rekkjer med samansetningar og avleiingar. Andre ordtypar som kan vera følgja eit ”Hertil”, er stadnamn²⁶ eller personnamn. ”Hertil” kan også peika mot faste uttrykk.²⁷

8 Lemma med uføreseieleg status i A1873

Gjennomgangen ovanfor syner at ein ikkje kan sjå av oppsettet i ein artikkeltekst om lemma som er tekne med i artikkeltekst som synonym, antonym, mål for tilvising, samansetningar eller avleiingar, har eigne artiklar på alfabetisk plass. Omtalen av desse kategoriane kan også innehalda implikasjonar som må tolkast ut av teksten. Om ein skal få utnytta informasjonen om ordtilfanget i 1873 må ein derfor kunna slå opp i opplysningskategoriane ovanfor som del av det elektroniske lemmaregisteret for A1873 i databaseformat, og heller kosta på seg nokre dobbeloppslag.

Denne gjennomarbeidinga av A1873 er per dato i prosess. Den førebelse innlastinga som gjev mykje av grunnlaget for denne artikkelen, gjev opp mot 60 000 tilslag i lemmaregisteret, dvs. ein og ein halv gong artikkeltalet. Det verkar derfor fornuftig å halda fram i den leia som det er arbeidd i til no.

LITTERATUR

Forkortingar

A1864: Aasen, Ivar (1965)

A1873: Aasen, Ivar (1918)

NLO: Norsk litterært ordboksverk

Publikasjonar

Aasen, Ivar (1965): *Norsk Grammatik*. Omarb. Udg. af ”Det Norske folkesprogs Grammatik”, 3die uforandr. Udg. (Opph. utg. Christiania 1864.) Oslo: Universitetsforlaget.

Aasen, Ivar (1918): *Norsk Ordbog med dansk Forklaring*. Fjerde uforandrede udgave. Kristiania: Vestmannalaget og Cammermeyers forlag.

26) Sjå under ”**Vik**(ii), f.”

27) Sjå under ”**Von** f.”

- Aasen, Ivar (2003): *Norsk Ordbog med dansk Forklaring*. Utg. ved Kruken, Kristoffer og Aarset, Terje. Oslo: Samlaget.
- Grepstad, Jon (2014): *Bibliografi over Ivar Aasens publiserte skrifter*. 2. utg. Nynorsk kultursentrum. www.aasentunet.no.
- Grønvik, Oddrun, Gundersen, Helge et al (2014): *Redigeringshandbok for Norsk Ordbok 2014*. <http://no2014.uio.no/eNo/tekst/redigeringshandboka/redigeringshandboka.pdf>
- Indrebø, Gustav (1927): *Norsk Namneverk*. Oslo. Norli.
- Språkrådet (2015): *Retningslinjer for normering av bokmål og nynorsk*. Nynorskversjon. Pdf. Gjeldande frå 1.3.2015. <http://www.sprakradet.no/upload/Retningslinjer%20for%20normering%20NYNORSKVERSJON.pdf>.

Arkivtilfang

Metaordboka: <http://www.edd.uio.no/perl/search/search.cgi?appid=7&tabid=571>

Norsk litterært ordboksverk: *Setelsamling* (1921 – ca 2000). Papir.