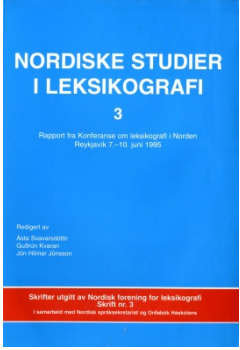


NORDISKE STUDIER I LEKSIKOGRAFI

| | | |
|------------|---|---|
| Titel: | Et arkiv over forandringer i moderne norsk ordforråd og dets betydning for norsk leksikografi |  |
| Forfatter: | Boye Wangenstein | |
| Kilde: | Nordiske Studier i Leksikografi 3, 1995, s. 413-420 Rapport fra Konferanse om leksikografi i Norden, Reykjavík 7.-10. juni 1995 | |
| URL: | http://ojs.statsbiblioteket.dk/index.php/nsil/issue/archive | |

© Nordisk forening for leksikografi

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.

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.

Et arkiv over forandringer i moderne norsk ordforråd og dets betydning for norsk leksikografi

Since 1968 a project for registrating lexical changes in Norwegian has been going on at the Section of Lexicography, University of Oslo. The number of excerpts is now about half a million, covering new words, new meanings of words, new phraseology and various characteristic lexical features in modern Norwegian. The greatest part of the material is computerized, and each excerpt is provided with a code which gives information about word class, gender and various lexical and syntactic features. The topic of this paper is furthermore what the material has been used for, particularly its importance for dictionary making in Norway the last 20 years, and finally how useful material of this kind is in dictionary making.

1 Innledning

Vi leksikografer i Norge blir litt misunnelige på kollegaene våre i andre nordiske land når vi hører om deres innholdsrike elektroniske tekstkorpus med millioner på millioner av ord som de kan boltre seg i når de lager sine ordbøker. Språkdata i Gøteborg har tekstkorpus med flere titalls millioner ord. Dansk Ordbog kan øse fra et på over 40 millioner. Et hovedmål for leksikografien i Norge må bli å komme etter på dette området, og vi kan nå konstatere at det er ting på gang også i vårt land. Men det er langt fram, og dette foredraget skal handle om noe av det vi har hatt å hjelpe oss med til nå.

Ved Avdeling for leksikografi ved Universitetet i Oslo (UiO) har en fra 1968 registrert forandringer i ordforrådet i moderne norsk allmennspråk med særlig vekt på nyord. Tallet på ekserpter er nå rundt en halv million. Jeg skal si litt om hvordan arkivet over det innsamlede materialet er bygd opp og hva det er blitt brukt til, særlig i hvilken grad det er blitt benyttet under utarbeidelsen av norske ordbøker. Til slutt vil jeg gi en vurdering av i hvilken grad et slikt ekserpert materiale egner seg som grunnlag for ordbøker.

Det var daværende ordboksredaktør Dag Gundersen ved Det norske litterære ordboksverk (som seinere ble en del av Norsk leksikografisk institutt (NLI) ved UiO, og nå er en del av Avdeling for leksikografi) som satte i gang prosjektet, som til daglig har gått under navnet Nyordsprosjektet. Prosjektet fikk støtte fra NAVF de første årene før NLI ble opprettet.

Innsamlingen av materiale er organisert på den måten at folk som er interessert, får et gratisabonnement på en avis eller et tidsskrift mot at de streker under nye ord, nye bruksmåter av ord, og ellers slikt de synes viker av fra vanlig språkbruk eller finner språklig interessant på annen måte. Tallet på frivillige ekserptorer har variert fra vel 30 til snaut 20. Opplegget ligner etter det jeg forstår, på det som Dansk Ordbog har med sine "sporhunde". I tillegg til disse frivillige hjelperne leser og streker de tilsatte ved avdelingen så mye avis- og tidsskriftstoff som de har tid til ved siden av andre gjøremål. I en kort veiledning for ekserptorene, står det at hensikten med innsamlingen er "å bidra til beskrivelsen av det norske

språk. Senere kan materialet brukes til praktiske formål som ordbøker eller til å fastslå hva som bør regnes for rett eller galt". Det er derfor slik at dette materialet inneholder mye interessant språkstoff utover akkurat nyord.

Mens materialet de første to-tre årene ble skrevet ut på sedler, begynte en i 1971 de første forsøkene på å skrive inn og lagre det elektronisk. Samtidig ble det utarbeidet et system av bokstavkoder, slik at det innsamlede materialet kan deles inn i emner, og en får muligheter til ikke-alfabetiske søkeinn ganger og kryss-søk i det.

Den første plassen i koden er forbeholdt markering av ekserptets ordklasse og når det gjelder substantiv, samtidig kjønn på ekserptet. Den andre viser om det ekserperte ordet er sammensatt eller ikke. Er det usammensatt, markeres det med en *x*. Hvis det er sammensatt, blir ordklasse og eventuelt kjønn på førsteleddet angitt. Videre kan en så føye til så mange tilleggskoder som en måtte ha bruk for. Hver av dem består av to bokstaver, og angir språklige særtrekk ved det aktuelle ekserptet:

asylant, *cxsi* = usammensatt felleskjønnsord dannet med suffiks
brevgiro, *cn* = sammensatt felleskjønnsord med et nøytrumsord som forledd
eldreomsorgstjeneste, *ccsfks* = sammensatt felleskjønnsord, sammensatt felleskjønnsord som forledd og komposisjonsfuge *s*.
hyttebygge, *vcrg* = retrograd verbdannning med et felleskjønnsord som forledd
ostehøvelprinsipp, *ncsfbb* = sammensatt nøytrumsord, sammensatt felleskjønnsord som forledd, brukt billedlig
versting, *cxsc* = usammensatt felleskjønnsord lånt fra svensk

De tilleggskodene som blir brukt, kan deles inn i 12 grupper:

1. Koder for sammensatte ord (f.eks. for bindestreksord, komposisjonsfuge, sammensatt forledd o.fl.)
2. Koder for affiks, avledninger og analogidannelser
3. Koder for lån fra andre språk
4. Koder for ordklasseovergang (også overgang fra affiks til selvstendig ord)
5. Koder for allusjoner, bildebruk, stilfigurer
6. Koder for lydlige trekk
7. Koder for påfallende bøyingsformer, bøyning, ortografi og ordbruk
8. Koder for forkortinger og forkortede uttrykksmåter
9. Koder for fraser og andre ordsammenstillinger
10. Koder for passiv og refleksiv form av verb
11. Koder for syntaktiske trekk
12. Andre koder (f.eks. koder som markerer overgang fra abstrakt til konkret betydning, ekserpt som står i anførselstegn, slang o.fl.)

Kodesystemet kan utvides etter hvert som behovet måtte melde seg, og det utgjør et ganske finmasket nett som vi anser som et godt hjelpemiddel til å kartlegge viktige utviklingstrekk i moderne norsk.

Med god hjelp av Dokumentasjonsprosjektet er nå også seddelmaterialet lagt inn og lagret elektronisk, og en er i ferd med å legge alt inn i en database med f.eks. muligheter for forskjellige kryss-søk. Foreløpig er dette bare en mindre prøvebase, men den skal utvikles etter hvert, og tanken er at den med tiden skal bli tilgjengelig på Internett.

2 Hva materialet er blitt brukt til

I veiledningen som ble skrevet for ekserptorene for 27 år siden, ble det sagt at materialet bl.a. skulle brukes til å fastslå hva som bør regnes for rett eller galt. Den funksjonen har det også hatt. Når Norsk språkråd skal normere et nyord, eller vurderer å forandre skrivemåten av et ord, kontakter de svært ofte vår avdeling for å høre hvilke skrivemåter vi har registrert i våre arkiver før de tar en avgjørelse. Nyordsmaterialet utgjør m.a.o. en viktig del av det empiriske grunnlaget for språknormeringsvedtak.

Mange av dem som tar vår emnekrets i leksikografi, henter naturlig nok stoffet til sin semesteroppgave fra vårt materiale. Som andre eksempler på hva materialet blir brukt til, kan jeg nevne en hovedfagsoppgave om komposisjonsfuge i sammensatte ord, og en omfangsrik avhandling om verb i passiv fulgt av perfektum partisipp, som kom i fjor. En svært stor del av materialet for moderne norsk i denne avhandlingen kunne hentes fra våre arkiv, ved å søke på koden *pp* (påfallende passiv).

Når nå nyordsmaterialet blir gjort lettere tilgjengelig for søking og annen utnyttelse elektronisk, håper vi at det skal bli bedre utnyttet til forskningsformål, både av hovedfagsstudenter og andre.

Når det gjelder ordbøker, har det knapt kommet en seriøs enspråklig ordbok i Norge de siste tjue årene som ikke i større eller mindre grad har gjort seg nytte av nyordsmaterialet, om vi ser bort fra reine fagordbøker. Jeg skal nevne de viktigste.

Nyord i norsk 1945–1975 (1982) var det norske bidraget i det nordiske samarbeidet med å registrere og vurdere tilveksten av nye ord i de nordiske språkene. Også når det gjaldt arbeidet med den danske og svenske nyordboka var nyordsmaterialet til hjelp når en skulle finne norske ekvivalenter til de danske og svenske nyordene.

Videre kan vi nevne den allmennspråklige, ettbinds *Riksmålsordboken* (1977) utgitt av Det Norske Akademi for Sprog og Litteratur. I forordet takker utgiverne NLI for adgang til instituttets arkiver. Denne ordboka kom i ny og atskillig større utgave i 1993 med tittelen *Norsk Illustrert Ordbok*.

Bokmålsordboka og *Nynorskordboka* var samarbeidsprosjekter mellom Norsk språkråd og Avdeling for leksikografi. Bøkene ble utarbeidet ved Avdeling for leksikografi, og en hentet selvfølgelig oppslagsordene bl.a. fra avdelingens store nyordssamlinger. Førsteutgavene kom i 1986 og de reviderte utgavene i 1993.

I mellomtida, i 1992, hadde også *Bokmålsordlista* kommet, en rettskrivningsordliste som bygde på *Bokmålsordboka*, men som det også var naturlig å supplere med nyord som hadde kommet til etter 1986, fra nyordssamlingene. I disse dager er den tilsvarende *Nynorskordlista* på det nærmeste ferdig.

I mange år har det vært arbeidet med supplementsbind til den fire bind store *Norsk Riksmålsordbok* (ferdig i 1957). Resultatet er blitt to supplementsbind som skal lanseres høsten 1995. Redaktøren har hatt sin arbeidsplass ved avdelingen, og vi blir igjen takket i forordet for den store hjelp og nytte han har hatt av avdelingens samlinger.

Et annet ordboksprosjekt, *Norsk anglisismeordbok*, har fått utskrifter av våre lånord fra engelsk og norske avløserord (kodet henholdsvis *ag* og *na*), som redaktørene har brukt særlig som utfyllende eksempelstoff.

Til spesialordbøker som *Bevingede ord* og Kunnskapsforlagets *Fremmedordbok* har også redaktøren hentet stoff fra nyordsmaterialet til nye utgaver.

3 Representativitet

En har forsøkt å gjøre utvalget av aviser og tidsskrifter representativt ved å spre det på både riksdekkende aviser og på lokale aviser. Alle landsdelene, Østlandet, Sørlandet, Vestlandet, Trøndelag og Nord-Norge er representert. På samme måte har en prøvd å få til et variert utvalg av allmenne tidsskrifter: *Naturen*, *Tidsskrift for samfunnsforskning*, *Filosofisk Tidsskrift*, *Tidsskrift for teologi og kirke*, *Geografisk tidsskrift* o.fl. Men når en er avhengig av frivillige og ulønnede medarbeidere, kan en ikke stå for hardt på kravene om den ideelle representativitet.

I veiledningen til ekserptorene er det angitt noen retningslinjer for innsamlingen. Som nye ord regner en ord som ikke står i ordbøker som *Tanums store rettskrivningsordbok* eller andre nyere norske ordbøker. Nå kan en vanskelig forlange at ubetalte medarbeidere skal bruke mye tid på å slå opp i ordbøker for å sjekke om et ord er ordboksført; det er ikke sikkert de har disse ordbøkene en gang. Da er det naturlig at de streker under ord som de oppfatter som nye, og at de tilsatte ved Avdeling for leksikografi siler det materialet som kommer inn.

På samme måte vil nye betydninger eller nye bruksmåter av ord si mye for ekserptor. Og det vil til enhver tid være det som ekserptoren oppfatter som avvikende syntaktiske konstruksjoner, feil eller vaklende preposisjonsbruk han/hun vil streke under.

Jeg får legge til at det overfor medarbeiderne blir presisert at de ikke bare skal feste seg ved det som er "rart". De blir bedt om å ta med også f.eks. vanlige preposisjonsuttrykk og vanlige ord, særlig når de er brukt på en særlig opplysende og helst kortfattet måte, slik at de egner seg som ordbokssitat.

Men utvalget er basert på skjønn. Det som er avvikende for én, trenger ikke være det for en annen. Evner og kunnskap og hvor mye arbeid den enkelte er villig til å legge ned i materialinnsamlingen, varierer selvsagt mye. Det er derfor ikke vanskelig å finne svakheter metodisk sett ved en slik måte å drive materialinnsamling på.

Våren 1995 var det oppslag i media i Norge om en kvinneordbok som er under arbeid, og i den forbindelse ble det fokusert på at norske ordbøker mangler mange vanlige ord for mer typisk kvinnelige aktiviteter som håndarbeid, matlaging, kvinnelige klesmoter, sminke o.l. Det er pekt på at det ofte er menn som lager ordbøker, og det er menn som står for materialinnsamlingen.

Det første er ikke helt riktig, f.eks. har både *Bokmålsordboka* og *Nynorskordboka* også hatt kvinnelige redaktører. Men jeg har ikke problemer med å innrømme at det kan være skjevheter ved materialinnsamlingen. De fleste ekserptorene er menn, og det er rimelig å

tenke seg at mange av dem heller leser f.eks. teknisk stoff eller sportssidene i avisene og streker ord derifra enn fra mote- og matspalter. Dette er kritikk som vi må ta alvorlig og prøve å rette på i framtiden.

Men en slik materialsamling er ikke verdiløs fordi det kan reises metodiske innvendinger mot innsamlingen. Alt i alt tror jeg vi har greidd ganske godt å fange opp nye ord, betydninger, uttrykk og andre endringer som har gjort seg gjeldende i norsk allmennspråk de siste 25 åra. Men vi vil ikke få svar på hvor fullstendig eller eventuelt hvor mangelfullt materialet er uten å lage ett så fullstendig korpus som mulig for den samme perioden og sammenligne. Det blir det neppe aktuelt med det første. Problemet med representativitet har en forresten når en skal bygge opp store, elektroniske tekstkorpus også, men jeg innbiller meg at det tross alt er lettere å løse da enn når en er avhengig av frivillige ekserptorer.

4 Hvordan egner et slikt ekserpert materiale seg som grunnlag for ordbøker?

Jeg har sett på noen filer av det materialet som ble skrevet inn i 1994, til sammen nærmere 1000 ekserpter. Av disse valgte jeg vilkårlig ut 300 ved å ta de 20 øverste ordene på hver side fra en 15-siders utskrift av en liste over de ekserperte ordene.

Av de 300 ekserptene er det etter min vurdering ca. 85 (knappt 28%) som er interessante i ordbokssammenheng. Da tenker jeg på en vanlig ettbinds allmennspråklig ordbok. 63 (vel 21%) av dem vil bli vurdert som eventuelle nye oppslagsord, 19 (vel 6%) som nye betydninger av eksisterende oppslagsord, eventuelt som brukseksempel som viser betydningsnyanser, og 4 (vel 1%) har interesse, eller kan få interesse for normering. (Utvalget er for lite til at en kan legge stor vekt på disse prosentene.) Ikke alle ordene, betydningene eller betydningsnyansene er nye, men de er ikke med i 1993-utgaven av *Bokmålsordboka*. Heller ikke alle vil uten videre gå inn i ordboka, men de kan settes på venteliste mens en ser om de blir noe mer enn døgnfluer.

Her får vi nøye oss med noen eksempler fra de forskjellige gruppene.

Nye oppslagsord:

asylant (ordet er blitt svært frekvent i norsk etter at flyktninger, særlig fra det tidligere Jugoslavia, begynte å søke tilflukt i norske kirker for å unngå å bli sendt tilbake av norske myndigheter)

brevgiro (alle nordmenn med bankkonto har fått tilbud om brevgiro)

børstraktor (spøkefull betegnelse på de store privatbilene med firehjulstrekk som er blitt så vanlige)

frislipp (økonomisk frislipp: f.eks. om norske bankers hemningsløse utlånsvirksomhet i 80-årene)

kinaisering (= flytting av arbeidsplasser til lavkostland)

krematør (norske ordbøker har ikke ført opp noe nomen agentis til *kremere* og *kremasjon*)

lobbying (*lobbisme*, *lobbyvirksomhet* er ordboksført før)

pubertil (norske ordbøker har ikke noe adj. til pubertet, så det fyller et tomrom i leksikon. *Svenska Akademiens Ordlista* har *pubertal*, og vi har også belegg

på *pubertistisk* i nyordsmaterialet, så her trenger redaktøren å undersøke litt nærmere hvilken form ordet bør få i norsk)
tagge, tagger (spraye merker på fasader, trikker o.l.; person som tagger)

Nye betydninger og betydningsnyanser:

A4 (= ordinær, vanlig; i sms. *A4-barnehage*)
flyktning (i sms. som *avgiftsflyktning, skatteflyktning* = person som flytter til et annet land for å unngå avgifter og skatter)
på golvet (*de på salgsgolvet* = de vanlige selgerne; tidlige bare om folk som arbeider på fabrikkgolvet, til forskjell fra ledelse og funksjonærer)
harehjerte (= hjerteflimmer)
house party (= ungdomsfest i nedlagt fabrikklokale)
nullstille (overført, om person: *være nullstilt*)
spikre (= fastsette: *programmet er spikret*)
økse (= redusere, skjære ned: *økse budsjettene*)

Former som har interesse for normering:

føyk (pret. av *fyke*; i offisiell rettskrivning bare *fauk* el. *føk*)
spott (eng. = kort innslag eller opptreden; i offisiell rettskrivning *spot*)
et spise (= mat, kost; i offisiell rettskrivning *en spise*)
et eksistensial (bare ordboksført i fl. *eksistensialer*, og bare i *Norsk fremmedordbok*)

Kaster vi et blikk også på resten av de ca. 1000 ekserptene fra 1994, finner vi bl.a. noen forledd og etterledd som er særlig produktive.

I dette lille materialet fins det f.eks. 23 sammensetninger med *miljø-* som forledd: *-ansvar, -bølge, -charter, -evaluering, -forbedring, -hazop, -hensyn, -innpakning, -konsekvens, -krav, -ledelse, -merke, -merking, -parameter, -profil, -proteksjonisme, -reklame, -revisjonskompetanse, -riktig, -standard, -strategi, -ulempe, -vennlig*.

Nå mener jeg ikke at alle disse sammensetningene skal inn i ordboka, men ord som *miljøbølge, -hensyn, -innpakning, -krav, -merking, -riktig, -ulempe* vil nok bli vurdert i forbindelse med neste utgave. Av disse 23 sammensetningene er det bare *miljøvennlig* som står i *Bokmålsordboka* 1993.

Etterleddet *-kåt* (= sugen på, som har lyst til) er også produktivt: *bade-, demokrati-, innsparings-, markeds-, påske- (påskeskåt aprilsol), sol- (solkåte nordmenn), utlåns-*.

I *Bokmålsordboka* 1993 er det bare *PR-kåt* som er med, og det er neppe noen av disse sammensetningene her som vil få plass i ordboka. Men vi bør få med denne utvidede betydningen under oppslagsordet *kåt*.

Ordet *park* har også utvidet sitt betydningsområde de siste årene (= område for en spesiell aktivitet), og vi finner den i sammensetninger som: *kunnskaps-, miljø-, nærings- og travypark*. Dette må nedfelles i ordboka både ved at en tar med noen sammensetninger av denne typen, og at en fører opp denne utvidede betydningen under *park*.

Når en arbeider med dette materialet, ser det både rikt og mangfoldig ut. Men problemet er at du ser hva du har, men ikke det du eventuelt mangler. Det er viktig å være klar over

de begrensningene et slikt ekserpert materiale har, og ikke gå ut fra at det forteller hele sannheten om utviklingen i ordforrådet.

Et eksempel: I dette materialet jeg har omtalt her, finner vi tre ord som betegner musikkformer innenfor vår tids ungdomskultur: *ambient*, *etno* og *hardcore*. Det fins mange flere slike musikkbetegnelser, noen har vi nok i nyordsmaterialet, andre ikke. For å oppfylle kravet om konsistens i ordboka, må en i tillegg finne fram til faglitteratur eller kontakte eksperter, f.eks. folk som dekker det som skjer på denne fronten, i avisene, slik at en får med i alle fall de viktigste betegnelse innenfor området.

5 Hva slags materiale er det ordboksredaktøren legger til side?

En god del av ekserptene er tatt med på grunn av syntaktiske særtrekk, og de har i mindre grad direkte interesse for ordboksredaktøren (jf. kodegruppe 11).

Svært mange av sammensetningene er tilfeldige sammensetninger laget for behovet i øyeblikket, og vil ikke bli tatt med i vår ordbok, f.eks. *bilvrakdyng*, *bondesøppel*, *brukt-boksalg*, *bunadsaktivist*, *budsjettkaos*, *forsikringsløs* (*forsikringsløse biler*), *en frida* (etter en populær norsk film om ei jente som heter Frida), *neonhumør*, *oberstbesøk*, *OL-skapt* (*OL-skapte arbeidsplasser*), *sannhetskommisjon*, *valgkampopplagt*.

Noen av ekserptene er fagord som blir for spesielle. Eks.: *brannflekksyke*, *porselenssyke* (betegnelser for krepsesykdommer).

Noe stoff fra lokalavisene er såpass lokalpreget (dialektord o.l.) at det ikke har interesse for en allmennspråklig standardordbok, f.eks. *bøvelsbygder* (nedsettende betegnelse på folk fra en spesiell bygd på Vestlandet), *noppemjøl* (lokal betegnelse på byggmjøl).

Noen av ekserptene er allerede ordboksført, f.eks. *etterbruk*, *blind alarm*, *bohemer* (i utvidet bruk: *bohemer i politikken*), *brikkene faller på plass*. Men de kan jo tjene som en bekreftelse på at redaktøren gjorde rett da han/hun tok inn ordet, betydningen eller brukseksemplet i ordboka.

Men dette materialet som ikke umiddelbart er så relevant i ordbokssammenheng, har selvfølgelig interesse for språkforskere i andre sammenhenger.

6 Konklusjon

Nyordsmaterialets bidrag til ordbøkene består i en komplettering når det gjelder nye ord, nye betydninger og betydningsnyanser, ny fraseologi og gode eksempler på hvordan ord brukes. Når en arbeider med et slikt ekserpert materiale, må en være klar over de begrensningene det har. Selv om det inneholder mye informasjon om utviklingen av ordforrådet i norsk, er det ikke fullstendig. Og jeg stiller meg stadig spørsmålet: Hva har vi gått glipp av sammenlignet med våre nordiske korpuskolleger?

Litteratur

Berulfsen, Bjarne/Dag Gundersen (red.) 1986: *Fremmedordbok*. Oslo: Kunnskapsforlaget.

Guttu, Tor 1993: *Norsk illustrert ordbok*. Oslo: Kunnskapsforlaget.

Kristensen, Kjeld 1994: Den Danske Ordbogs tekstkorpus og spORDhunde. I: Anna Garde/Pia Jarvad (red.): *Nordiske studier i leksikografi* II. København: Nordisk forening for leksikografi, 138–142.

Landrø, M. I./B. Wangenstein 1993: *Bokmålsordboka*. Oslo: Universitetsforlaget.