

NORDISKE STUDIER I LEKSIKOGRAFI

Titel: Neologismer i norsk - Kartlegging av leksikalsk språkendring før og nå

Forfatter: Ruth Vatvedt Fjeld & Lars Nygaard

Kilde: Nordiska Studier i Lexikografi 10, 2010, s. 506-521
Rapport från Konferens om lexikografi i Norden, Tammerfors 3.-5. juni 2009

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.

RUTH VATVEDT FJELD & LARS NYGAARD

Neologismer i norsk Kartlegging av leksikalsk språkendring før og nå

This article presents a short history of Norwegian neography and some of the problems in the field of lexical development, as well as attempts of classification of neologisms. It also presents a new, semiautomatic method of mapping neologisms from newspaper corpora. Some of the findings show that the overwhelming amount of new words in Norwegian are inherited words, mostly composite, and that loan words all in all are very few.

Nøkkelord: neologismer, leksikalsk språkendring, neografi, Norsk Aviskorpus, automatisk nyordsregistrering

1. Innledning

Ordforrådet er det som endrer seg raskest i et språk. På 20 år kan omfattende leksikalske endringer skje. Ifølge Robert og Clarence Barnhart (Barnhart et al. 1990) er neologismer et språks åndedrett. At et språk får inn nye ord, er et tegn på at det lever og tilpasser seg nye kulturelle og sosiale forhold og bruksområder. Vil man vite noe om et språks utvikling, må man først og fremst kartlegge dets nye ord og uttrykk, neologismene. Slik kartlegging er en del av et språks neografi, som innebærer å registrere et språks nye lingvistiske enheter i det samfunnet der språket brukes. Neografi innebærer samtidig en kartlegging av generelle kulturelle endringer. Nye ord og uttrykk speiler endringer i den kulturen språket fungerer i og skal tjene som uttrykksmiddel innenfor, så studier av neologismer har to sider, én rent lingvistisk og én kulturell.

Neografien ble lenge forsømt i nordisk leksikografi, og det er ikke tilfeldig, da språkforskning og språkrokt i Norden mye har dreid seg om å etablere moderne nasjonalspråk og å verne språket mot påvirkning utenfra. I stedet for å kartlegge neologismene har puristiske holdninger ført til fornektning av nylånte ord eller nyordsdannelser som ikke følger tradisjonelle mønstre. Innsatsen har vært å fortie eller fortrenge dem ved å oppfordre til bruk av ”gode gamle ord”. Mest uttalt har denne fornektningen og nasjonale iveren vært på Island og Færøyene (jf. f.eks. Thomsen 2002: 107), og resultatet har blitt at det er stor forskjell mellom de ordene som faktisk er i bruk, og de ordene som står oppført i ordbøkene

og regnes som korrekte. I Norge har Språkrådet hatt som offisiell målsetting å finne såkalte avløsere for nyordene, det vil si man har konstruert nye leksemer etter innarbeidet mønster med tradisjonelt inventar. Motviljen mot fremmed språkpåvirkning og ønsket om å bevare det nasjonale har vært sterk i den offisielle norske språkrøkten, som også ellers i det nordiske språkområdet. Språklig purisme og nasjonalisme har vært oppfattet som to sider av samme sak.

Men slike språkholdninger har ikke vært framtreddende bare i Norden. Avhengig av nasjonal innsats og "språkrøkt" har tempoet for opptak av nye ord i et språk variert sterkt. Lenge var for eksempel tysk språkrøkt også svært puristisk, men der har det skjedd en dramatisk endring i det lingvistiske klima de tre senere tiårene, og i moderne tysk tar man uten motforestillinger inn nyord og fremmed ordmateriale, først og fremst fra engelsk. Det samme har skjedd med dansk, og dansk språk har nå flere lånord fra engelsk enn svensk og norsk. Åpenheten mot engelsk har ført til en merkbar endring av dansk ordforråd, og det har skjedd på noen få tiår.

Nasjonal ordbeskyttelse var ukjent i norsk fram til etableringen av den nye, norske nasjonalstaten i 1814, da nordmennene også ble opptatt av å ha et eget, norsk språk. Men før det fantes en nasjonal språkreise og en offisiell språkrøkt i Norge, flommet lån fra Mellom-Europa over de norske språkbrukerne og fornyet ordforrådet i stor grad. Det er ikke noen enestående hendelse. Mange språk har i perioder vært så åpne for lån at de etter en sterk kulturell og språklig påvirkning utenfra kan regnes som blandingsspråk, engelsk er ofte nevnt som eksempel på det. Norsk og andre skandinaviske språk kan vel også regnes med her, etter den sterke påvirkningen fra tysk i seinmiddelalderen (Simensen 1995).

Moderne norsk bokmål er også i vår tid under sterk påvirkning utenfra, kanskje sterkere enn moderne nynorsk, da nynorskbrukerne jamt over er mer språkpolitisk bevisste. Vårt inntrykk er likevel at det er stor forskjell på norsk og dansk språksituasjon, fortsatt er hovedmengden norske ord rotnorske eller lån fra mellomeuropeisk for lenge siden. Mange er likevel opptatt av faren for at norsk skal bli ødelagt av engelsk ved at vi nærmest drukner i en bølge av engelske lånord. Vi har derfor ønsket å undersøke nærmere mengden av nye ord som kommer inn i norsk, og hva slags nye ord vi tar inn, samt å utvikle en ny metode for generell kartlegging av nyord i norsk. Ved hjelp av store elektroniske aviskorpus og statistiske beregninger mener vi å ha kommet et godt skritt i retning av å klare det.

Vi antar at de fleste nyord tidlig blir brukt i avistekster, og slikt materiale har vi fått tilgang til gjennom prosjektet *Norsk Aviskorpus*, som er ledet av Knut Hofland ved Aksis, Universitetet i Bergen. En undersøkelse av neologismer i dette materialet er utført ved Institutt for lingvistiske og nordiske studier, Uni-

versitetet i Oslo ved Ruth Vatvedt Fjeld og Boye Wangensteen i samarbeid med Kaldera språkteknologi ved Lars Nygaard, der vi forsøker å beskrive hva som skjer med ordforrådet i det moderne bokmålet, hva slags nye ord som kommer inn i det moderne avisspråket, hvor de kommer fra og hvordan de tas opp i språket.

2. Hva er et nytt ord?

Nyheter i et ordforråd kan være enten helt nye ord og uttrykk, eller ny betydning eller bruk av tidligere uttrykk. Man kan kanskje også regne eksisterende, men svært lite brukte ord som plutselig får sterkt økt bruk, som en type nyord. I hvert fall er en endring i frekvens en nyhet i et språks ordforråd. Et eksempel på en slik bruksendring er for tiden ordet *klima*, som har vært et kjent, men lite brukt fremmedord i norsk svært lenge. I dag er det imidlertid på alles lepper og på avisenes forsider nesten hver dag, ordet er blitt et allmennord. Fagtermer kan også plutselig få nytt liv, som *tsunami* gjorde det etter en stor naturkatastrofe i 2004. Ordet er lånt fra japansk og var kjent innen meteorologien også før det fikk den enormt økte bruken og ble et allmennord som dannet sammensetninger og avledinger. Det har nå fått en ganske generell og overført betydning 'plutselig og uventet katastrofe' i motsetning til den meteorologiske 'stor havbølge forårsaket av underjordisk jordskjelv'.

Nye ord og uttrykk regnes ikke lenger som neologismer når de er kommet i allmenn bruk og er blitt en del av allmennordforrådet. Da skal de ha sin plass i vanlige ordbøker. Det å være nyord er altså et overgangsfenomen. Alle nyord kan oppfattes som potensielle okkasjonalismer. Det er nye ord som oppstår i konkrete situasjoner og er tolkbare kun i sin spesielle kontekst, og som brukes én eller få ganger over kort tid og ikke går inn i språket som en varig del av ordforrådet. Bare om de blir brukt av mange og over tid, regnes de som nyord, som seinere igjen kan bli allmennord. Men svært mange nyord dør en ung død, de brukes en kort stund, så blir de overflødige eller unyttige. Det trenger ikke være noe feil med ordet som sådant, og heller ikke med referenten. Et typisk eksempel på et nyord med kort liv er ordet *tastafon*, som ble betegnelsen på telefonapparater uten ringeskive men med trykknapper for tallene i stedet. Et belegg fra *Sunnmørsposten* 1979 dokumenterer ordets bruk i det norske språket:

Slik ser det nye telefonapparatet vårt ut. Den offisielle betegnelsen er knappsats-apparat, men en publikumskonkurranse har allerede dømt vidunderet for "tastafon".

Nyordet forble et 1980-tallsord, og i nye utgaver av ordbøker bør dette ordet merkes som foreldet eller tidstypisk for den perioden. Tastafonene ble nemlig så populære at de raskt fortrenget de gamle dreieskiveapparatene, og det gamle ordet *telefon* ble brukt om *tastafonene* da alle telefonapparater ble tastafoner. Derimot fikk vi nyordet *fasttelefon* i tillegg til *mobiltelefon*, da den kom. Innen nyordsutviklingen ser vi altså at det ofte er negasjonen, det komplementære eller alternativet til et innarbeidet ord som blir nyord når et nytt fenomen skjer eller en ny innretning er oppfunnet. Det nye ordet tar ofte over hele betydningsfeltet som ble dekket av det innarbeidede, en tekstmelding via mobiltelefon kalles blant nordmenn flest i dag bare *melding*, og det blir automatisk forstått som *sms-melding*. Et nytt ord kan altså få to skjebner, enten blir det fort glemt og dør, eller det går inn i det allmenne ordforrådet. En ny betydning kan enten føre til polysemi eller den fortrenger den gamle betydningen helt.

Leksikalske nyheter er ofte lånt fra andre språk ved at de følger med oppfinnelser eller fenomener som kommer fra giverlandet, som *tsunami*. Ofte er det kulturelle fenomener som matrettbetegnelser vi låner, *pizza* er i dag en viktig del av norsk kosthold, men ordet er lånt fra italiensk. I norsk etterkrigstid har særlig USA hatt sterk kulturell påvirkning på landet vårt, og dermed er mange ord også blitt lånt inn fra amerikansk engelsk. Ordet eller uttrykket trenger ikke være rotekte engelsk, det kan stamme fra et annet sted eller språk, men har kommet til oss via engelsk og gjerne fått fonologiske og morfologiske spor etter det. Det gjelder særlig såkalte internasjonismer, som er dannet av røtter fra de klassiske språkene gresk og latin. Slike ord blir lånt og brukt i de fleste moderne språk. Et typisk eksempel er *televisjon*, som tidlig ble forkortet til *TV*. Mange slike lånord er som nevnt blitt forsøkt nasjonalisert eller erstattet av hjemlige ord, og disse oversettelseslånene er også nyord; ofte lever de parallelt med det opprinnelige lånet, slik også med *TV* og *fjernsyn*. Først etter at apparatene kom i vanlig bruk, kom språkrøkterne inn i bildet og laget den direkte oversettelsen *fjernsyn*, og dermed fikk vi to nyord i norsk for det samme begrepet. Det ser ut til at språkbrukerne oftest velger det enkleste, og den norske avløseren mangler en grei forkortelse, uttrykket *FS* har aldri vært foreslått eller forsøkt brukt som avløser for *TV*. I de rundt femti årene vi har hatt slike apparater, har betegnelserne skiftet noe, fra *televisjonsapparat* til *fjernsyn* til *TV* eller *tv*, også skrivemåten *teve* har vært foreslått. Det vi kan registrere i dagens språkbruk, er at kortformen *TV* er mest i bruk, og det er interessant, for svært mange nordmenn kan ikke forklare hva bokstavene står for, uten at de oppfatter det som noe problem. Betegnelsen *fjernsyn* er i prinsippet mer selvforklarende for nordmenn flest, men klarhet og gjennomsiktighet må ofte gi tapt for korthet i språkvalget til folk flest. Forkortelsen *TV* er leksikalisert i moderne norsk.

Usammensatte nyord kan lages av hjemlig ordmateriale, men det skjer forholdsvis sjelden, et eksempel er slanguttrykket *lættis*, som har rot i ordet *latter* eller *latterlig* tillagt avledningen *-is*, som er brukt i flere slanguttrykk. Det fins også eksempler på at de er laget av tilfeldig tilgjengelig materiale, som oppfinnelsen *nal*, som først var et konstruert varemerke (laget som omvendt stavemåte av personnavnet *Lahn*) og gikk over i allmennspråket som betegnelsen på en gummilist med håndtak til bruk ved rengjøring. Men den mest vanlige måten å lage nyord på, er komposita av innarbeidede enkle ordformer. Med ordene *TV* og *skjerm*, er det uproblematisk å lage nydanningen *TV-skjerm*. Den første gangen det ordet dukket opp i aviser eller i salgsreklame for fjernsynsapparater, var det ikke vanskelig å forstå selv om man aldri hadde hørt det før. Våre komposita kan i prinsippet bli så lange som helst: *TV-skjermssparer*, *TV-skjermsspareprogram* og så videre og gjør at vi i prinsippet lett kunne dekke alle behov for nye ord av eget ordmateriale. Likevel låner vi altså også en god del ord og uttrykk fra andre språk.

3. Hvorfor får vi nye ord?

Bo Svensén (2004) sier at et nytt ord som regel oppstår ved at et nytt begrep trenger en benevnning. At en ny term lages for et eksisterende begrep, er sjeldnere, om det ikke er i eufemiserende hensikt. Men det kan også være andre språklige behov, for eksempel lages ofte nye ord for eksisterende begrep i slang, som at *moro* ble kalt *gøy* blant byungdom en gang på 1900-tallet, og seinere gikk inn i det vanlige ordforrådet. Andre grunner til at et eksisterende ord blir byttet ut med et annet, kan være resultat av språkkrøkt, slik vi har ved såkalte avløserord, der et lånord blir forsøkt erstattet med et nyord laget av hjemlige rotord, for eksempel at *snowboard* blir (forsøkt) erstattet med *snøbrett*, eller *televisjon* med *fjernsyn*. Det samme kan skje i standardiseringsøyemed, der vi har fått flere uttrykk for samme begrep, og bare det ene blir anbefalt brukt. Hilke Elsen (2004) peker på at nye ord ofte lages for å oppnå spesielle virkninger i en tekst. Det er da gjerne slik at nyord lages etter ”nye” orddanningsregler, for eksempel at det er nye sammensetninger av lyder eller av morfemer. Det gir klangvirkninger eller bildemessige assosiasjoner, eller andre effekter som ikke har vært i språket tidligere. Ord lånt fra fremmede språk, som *tsunami* og *paparazzi* er gode eksempler på det. Et nyord kan også være lydlig sett lettere å uttale enn et innarbeidet uttrykk, eller det kan være lettere å forstå, og dermed fortrenge et tidligere innarbeidet ord. Årsakene til at nye ord og uttrykksmåter oppstår, kan altså være ganske mange, og ganske forskjellige. Nyord reflekterer orddanningspreferanser relatert til brukernes behov.

4. Klassifisering av neologismer

For at et ord eller uttrykk skal klassifiseres som neologisme, må det skille seg fra det eksisterende ordforrådet i et språk enten i form eller innhold da det ble tatt i bruk. Og det er et definisjonsspørsmål hvor lenge det skal regnes som nytt. I det enkelte nyordsprosjekt settes det derfor tidsavgrensninger, slik som for eksempel *Nyord i norsk 1945–1975* og *Nye ord i dansk 1955–1998*.

Neologismer kan sorteres i to typer etter form og innhold:

- A. neoforativ – et ord som ikke tidligere har eksistert i det aktuelle språket
- B. neosemantisme – et eksisterende ord som har fått endret, ofte spesialisert betydning

Vi kan også regne med en tredje type, som består av flere ord, der både form og innhold kan være nytt:

- C. neofrasem – et nytt uttrykk som består av flere ord, og enten inneholder neoforativ, eller en eksisterende frase med ny betydning.

En neoforativ kan være et innlånt ord eller et nykonstruert ord, mens et gammelt ord med ny betydning kalles neosemantisme. En neoforativ kan være uttrykk for et nytt fenomen, som *tv*, *pc*, men kan også være et nytt uttrykk for et innarbeidet begrep, som *renholder* for *vaskekone*, innført i eufemiserende hensikt, eller *snøbrett* for *snowboard*, innført i nasjonaliserende hensikt. Neoforativen har andre konnotasjoner enn det innarbeidede ordet, men ikke i grunnbetydning, og danner således et delsynonym til det.

Neofrasemer kan være helt nye flerordsuttrykk med neoforativ, som *sudden death* og *hard core*, eller bestå av eksisterende ord med ny betydning, som hilseformelen *ha det* for avskjed, eller *være på nett* i betydningen 'kan kontakt, følge med'. Slike neofrasemer blir imidlertid sjelden dokumentert i ord-bøker eller i grammatikker, da de ikke er interessante som orddanningsmønstre og lenge regnes som feil. Eksempel på neofrasem i norsk er *rundt forbi* i stedet for *rundt omkring* eller *ta en utskrift* med betydningen 'å skrive eller printe ut en tekst fra en datamaskin'. Neofrasemer av eksisterende ord betraktes av mange som slang eller feil, men når de gir mange tilslag i norske avistekster, bør også de før eller seinere vurderes som en del av det allmenne ordforrådet. Slik leksikalsk språkendring medfører ofte også syntaktiske endringer, for eksempel at argumentstruktur blir endret. Neofrasemet *tenke til å gjøre noe* irriterer mange og blir nok fortsatt rettet av mange norsklærere, dels fordi de ikke er dokumentert i faglitteraturen. Derimot blir nye såkalte ordgruppeleksemer, fraser med idio-

matisk innhold, som *høy sigarføring*, med betydningen 'være svært selvtilfreds og leve over evne' eller *på trynet* i betydningen 'idiotisk, meningsløst' gjerne registrert og dokumentert både i spesialordbøker og i allmennordbøker.

5. Hvordan dannes neologismer?

Neoformativer er enten lån eller nyskapinger. De fleste neologismene i moderne norsk er ord eller fraser som er lånt fra andre språk, først og fremst fra engelsk. Nyskapingner dannes ved nye orddanningselementer, som ordene *rayon*, *nylon* etter mønster av *cotton* i engelsk. Ofte er de som her først varenavn som senere blir deproprialisert og går inn i allmennspråket. Eller det er vanlige ord som blir produktive ledd og settes sammen med andre orddeler med en spesiell betydning. Slik har forleddet *skap-* blitt produktivt i norsk (og dansk) i betydningen 'skjult, hemmelig' (Jarvad 1995: 226). Vi kan regne med tre hovedtyper etter orddanningsmåte:

- A. Skapte ord (kunstord): *dacron*, *nal*
- B. Lånte ord: *tape*, *tsunami*, *jogurt*
- C. Komposita laget av innarbeidede ord: *gassferje*

Type A er sjelden i norsk, som regel er det fordi et varemerke er blitt så vanlig og markedsledende at det får generisk bruk og er blitt deproprialisert. Type B kan komme inn i norsk språk enten akkurat som det var i giverspråket, eller det kan nasjonaliseres eller naturaliseres, i vårt tilfelle fornorskes, enten i skrivemåte eller uttale. Lån fra nærliggende språk blir ofte tatt direkte inn, som *tape*, som også har beholdt den engelske uttalen, eller det endres så det passer til vårt skriftsystem, som *tsunami* fra japansk og *jogurt* fra tyrkisk.

De vanligste neoformativer i norsk er av type C, og de aller fleste av dem er substantiv som er sammensetninger av gamle ord. Disse kan klassifiseres etter relasjonen mellom leddene (jf. Elsen 2004):

- A. determinativkomposita – med førsteledd som bestemmer etterleddet nærmere: *fredagsbønn*
- B. possessivkomposita: *rødtopp* 'person som har rød topp'
- C. kopulakomposita (dvandva), to likeverdige elementer: *våtkald*
- D. reduplikasjonskomposita, helt: *neinei*, eller delvis: *tripptrapp*
- E. tydeliggjørende komposita, særlig når en fremmed rot får hjemlig forledd: *enkeltindivid*, *geriljakrig*

6. Redigering av nyordbøker

Om man skal lage en nyordordbok, er det rimelig å velge kilder eller ordboksbase der man antar det er mange nyord. Det er først og fremst i faglitteratur i nye fagfelt, oversatte tekster, tekster som omhandler motefenomener og så videre. De fleste nyordbøker har hentet sitt materiale fra avistekster, noe som virker fornuftig, da det som oftest er gjennom meldinger om nyheter i et samfunn og om hva som skjer ellers i verden, rettet mot den store allmennheten at nyord introduseres i et språk. I tillegg kan man finne nye ord i reklametekster, varekataloger, film- og TV-tekster og i frie tekster publisert på Internett. Mange ord lever sitt eget liv lenge i muntlig språk og blir først fanget opp når de kommer inn i skriftspråket. Store talespråkskorpus vil derfor gi oss muligheter til å finne nye ord raskere enn før. Politikere og andre kjente personer mynter gjerne nyord, typiske eksempler i norsk samtid er at den norske kronprinsen brukte sammensetningen *adrenalinsint* i talen til sin brud i 2001, eller da lederen i det høyre-radikale Fremskrittspartiet i februar 2009 advarte det norske folk mot *snikislamisering*, som seinere også av noen blir kalt *snikislamifisering* – begge formene har mange tilslag i norske aviser. I løpet av året 2009 har *snik*- blitt et mye brukt forledd med den negative betydningen 'i skjul', delsynonymt med forleddet *skap*- (jf. Jarvad 1995), men med enda mer negativ konnotasjon. Det er imidlertid vanskelig å avgjøre om et nytt ord er representativt eller ikke, da det ofte bare er en viss tid at slike nye ord er frekvente. Ord som plutselig blir svært frekvente og så forsvinner ut igjen, kalles gjerne *moteord*. Om moteordene samtidig er nyord, er det vanskelig å vite om de skal tas med i ordbøkene eller ikke. Det er også vanskelig å vite om sammensetninger skal regnes som okkasjonalismer eller ekte neologismer. Eksempel på det er nettopp av typen *adrenalinsint*, som er brukt av en del med klar allusjon til kronprinsens tale. I dag finner vi cirka 60 belegg i *Norsk aviskorpus*, og det er ikke mye over et knapt tiår. Forleddet er ellers mye brukt i sammensetningen *adrenalinkick*, som har både lengre historie og større utbredelse. Begge hører hjemme i nyordbøker, men bare det siste i allmennordbøker. Lemmaseleksjonen er annerledes ved redigering av nyordbøker enn ved andre ordbøker.

Nyordbøker krever også spesielle informasjonstyper. Hvis man vil dokumentere leksikalsk språkendring, er det viktig å markere så nøyaktig som mulig når de nye ordene kom i bruk, men å vite hva som er førstebelegg, er ikke alltid lett. Det er greit når ordene myntes av språkbevisste kjendiser i situasjoner som blir dokumentert av andre grunner, som i eksemplene over, men som regel er ord og uttrykk brukt en god stund før de blir registrert, og dermed er det i høyden tidligbelegg man får tak i. Ofte kommer nyord ikke med i allmennordbøker for lenge etter at det har gjort sitt inntog i språket, og som regel kommer de

først inn i spesialordbøker, som slangordbøker, ordbøker over spesielle varieteter, lånordsbøker, fremmedordbøker eller andre spesialordbøker. For eksempel nevner Elsen (2004) at det ble utviklet et spesielt politisk ordforråd i det tidligere Øst-Tyskland. Dette ordforrådet ble først dokumentert mange år etter at staten ble oppløst, og da i spesialordbøker i stedet for i nyordsbøker mens de var i levende bruk.

Å skrive definisjoner av nye ord i et språk er ofte vanskelig, særlig for betegnelser for abstrakter eller ukjente fenomener. Nyord krever andre definisjonsmetoder enn tradisjonelle ord fordi det ikke er så lett å finne mange belegg. Dessuten spriker bruken gjerne mellom forskjellige belegg før ordet er blitt godt innarbeidet. Som regel blir definisjonene mer ensyklopediske enn for innarbeidede ord. Noen ganger er det umulig å fastslå opprinnelsen, og da er det særlig vanskelig. To nye ord som er mye brukt blant norsk ungdom i dag, er adjektivet *loka* og verbet *å loke*, som sannsynligvis har samme rot. De brukes begge i mange forskjellige betydninger, men beleggene samler seg om 'gal, utafør' for adjektivet og 'å slenge rundt omkring uten mål' for verbet. Etymologien er uklar, den kan kanskje komme fra gammelnorsk eller fra italiensk, men de nye ordene er så forvansket at det ikke er mulig å fastslå med sikkerhet.

7. Tidligere nyordsregistrering i norsk

Det har lenge vært gjort registrering og utforskning av tyske lån i norsk språk innen norsk filologi (jf. Simensen 1995). I tiden før og under andre verdenskrig registrerte og dokumenterte Aasta Stene (1949) engelske lån i norsk, men undersøkelser av neologismer som sådanne har oss bekjent ikke vært bearbeidet som språkvitenskapelig tema i norsk før på midten av 1950-tallet. Det første som ble gjort for å kartlegge neologismer i moderne norsk generelt, var en del av et fellesnordisk samarbeidsprosjekt som var satt i gang av de nordiske språknemndene. Språknemndenes årbøker *Nordiske språkspørsmål* kom som en serie i perioden 1955–1967, og hver årgang hadde lister over registrerte nyord i dansk, norsk og svensk. Målet med prosjektet var å ”studera tendenserna i fråga om nytillskotten i de nordiska ordskattna i allmänhet och att pröva vad det kan finnas av möjligheter för en samordning på de punkter, där en sådan är önskvärd.” (K. G. Ljunggren, sitert i Leira 1982: 9). Målet var altså ikke å registrere nyordene for å finne avløserer eller fjerne dem, men å finne former som var mest mulig like i de nordiske språkene. Dette målet har dessverre i liten grad blitt nådd, for eksempel fikk vi på norsk *datamaskin*, på svensk *dator* og på dansk *datamat*, til tross for iherdig innsats fra de nordiske språknemndene for å finne en felles betegnelse på den teknologiske oppfinnelsen. Men registrering

av nyord ble et viktig og nyttig biprodukt. I det første heftet angis det at de norske nyordene er ekserpert i perioden mai 1955–september 1956. De er ført opp systematisk i fire grupper, her presentert med eksempler:

1. Nye enkeltord: *friskus* (Dagbl. 1955), *krashe* (Dagbl. 1955)
2. Nye avledninger av tidligere kjente ord: *gisper* (VG 1955), *spreking* (Dagbl. 1955)
3. Nye verb dannet til substantiv: *kneskure* (Arbbl. 1955), *køe* (Aftenp. 1955)
4. Sammensetninger – nye eller tidligere lite brukte: *hjemmehjelp* (Arbbl. 1955), *innsmett* (VG 1955), *samrør* (VG 1955), *skumplast* (Dagbl. 1955)

Dette er ord som tilsatte i Norsk språkråd samlet inn fra avistekster. Det har ikke vært mulig å finne noen dokumentasjon på metoder eller mål for denne innsamlingen, ordene ble nok for det meste registrert av dem som følte interesse for temaet, når de selv leste de avisene de pleide å lese. Det kan blant annet bety at dekningsgraden ikke er tilfredsstillende. Som et ledd i det nordiske samarbeidet om dokumentasjon av nye ord ble den første norske nyordsordboka gitt ut av Norsk språkråd, *Nyord i norsk 1945–1975*. Den er i sin helhet redigert av Vigleik Leira (1982). De nordiske nyordbøkene skulle være registrerende og deskriptive, men samtidig skulle de være normerende og rådgivende og gå inn i et felles prosjekt om nordisk språkrøkt. Fra 1975 fram til 2001 har Leira ut fra systematisk ekserpering av et representativt utvalg aviser bygd opp et eget arkiv over nye ord i norsk. Arkivet inneholder mer enn 30 000 ord (en ukjent del av disse er kopier av sedler i Nyordsmaterialet ved Universitetet i Oslo). Noen av de innsamlede nyordene har blitt publisert i bladet *Språknytt*, som ble utgitt av Norsk språkråd. Materialet har vært med i ordboksbasisen for oppfølgeren til nyordboka fra 1982, *Nye ord i norsk 1975–2005* (red. Guttu & Wangensteen u.u.), som ventelig kommer ut i 2010 og er et samarbeidsprosjekt mellom Universitetet i Oslo og Språkrådet. Ellers har selvsagt alle som har arbeidet med ordbøker over norsk, samlet inn nye ord og uttrykk mer eller mindre systematisk, slik at ordbøkene de laget best mulig dekket samtidens ordforråd. Det er imidlertid vanskelig å trekke ut nettopp disse ordene i en egen oversikt for dokumentasjon av forskjellige tiders nyordsmateriale, da måtte man sammenlikne ordbøker som var gitt ut i flere utgaver, noe som ville vært svært tidkrevende og sannsynligvis gi få resultater.

Det største prosjektet for innsamling av nye ord i norsk (bokmål), *Nyordsprosjektet* ved Universitetet i Oslo, ble ledet av Dag Gundersen. Innsamlingen har resultert i en database med over 300 000 belegg registrert av frivillige ekserptører som representerer forskjellige deler av landet. Ekserptørene fikk gratis avisabonnement mot å registrere og sende inn ord og uttrykk de reagerte

på i avisene, ut fra en kort veiledning. Et av rådene var å legge spesielt merke til ord og uttrykk som sto i anføringstegn, eller som hadde malfemismeadvarselen ”såkalt” foran seg, for å vurdere dem som nyordskandidater. Ekserptene ble arkivert i et seddelarkiv der det registrerte nyordet ble ordklassekodet og gitt en normert form. Dessuten ble det lagt til en frivillig kode for orddanningstype og semantiske nyutviklinger. Fra tidlig på 1980-tallet ble ekserptene skrevet inn direkte i en elektronisk database, og den delen av materialet ble dermed søkbart. Etter som prosjektet skred fram, ble dessuten kodesystemet ganske utbrodret og komplisert. Da det ble avsluttet i 2000, hadde de som behandlet ekserptene over hundre forskjellige valgfrie koder å forholde seg til, i tillegg til de obligatoriske for ordklassemarkering. Det var flere leksikografer som arbeidet med dette materialet, og det har vist seg at forskjellige personer har brukt kodesystemet noe ulikt. Materialet er derfor ikke helt konsistent kodet, og det må tas i betraktning ved bruken. Men det er en viktig dokumentasjon over tilveksten og endringene i ordtilfanget i norsk bokmål for siste del av 1900-tallet med svært gode observasjoner, særlig over neosemantiseringer og neofrasemer, som ellers er dårlig belagt. Nyordsdatabasen består av 266 314 sitater, de fleste fra 1968 til 1972, de eldste er fra 1920, de yngste fra 1999. Den eldste delen av kartoteket ble seinere også digitalisert og hele materialet er nå fritt tilgjengelig for interesserte språkforskere.¹

Boye Wangensteen (1995: 417) presenterer noen resultater fra *Nyordsprosjektet*. Etter en punktanalyse av 1 000 utvalgte ekserpter konkluderer han med at 28 prosent av det undersøkte materialet var utnyttbart i leksikografisk dokumentasjon av allmennspråket; cirka 21 prosent var mulige nye oppslagsord, 6 prosent var gamle ord med ny bruk eller betydning, 4 prosent var eller kunne bli interessante for normering fordi de representerte brudd med gjeldende rettskriving. For dokumentasjon av neologismer er rimeligvis nytteprosenten mye høyere. Prosjektet ble avsluttet av flere grunner. Registrering ved enkeltpersoners lesing og ekserpering var kostbart og møysommelig som metode. Resultatene var også usikre, blant annet kunne man ikke vite hva som ikke ble registrert av ordforrådsendringer. Det som ble ekserpert, gjenspeilte ekserptørens varierende kunnskaper om og interesse for språkendringer, i tillegg til skjevhet i utvalg av ekserperte tekster. Materialet har selvsagt vært viktig grunnlag i redigeringen av de to norske nyordbøkene, og ved utarbeidelsen av *Bokmålsordboka* (1986).

Neologismeprosjektet under *Norsk Aviskorpus* er på mange måter en videreføring og modernisering av *Nyordsprosjektet*, der materialet er store aviskorpus og registreringen foregår automatisk ut fra statistiske vurderinger.

1 http://www.dokpro.uio.no/bokmaal/nyord/nyord_ramme.html/

8. Manuell vurdering av automatisk genererte nyordskandidater fra norske nettaviser

Ved Universitetet i Bergen har man siden 1998 og fram til i dag bygd opp et stort tekstkorpus ved automatisk innsamling av norske avistekster som er publisert på Internett, kalt *Norsk Aviskorpus* (<http://avis.uib.no/>). Det totale korpuset er i 2009 på over 700 millioner ord. Ut fra dette materialet er det laget en nyordsbase ved at det hver dag kjøres ut lister over potensielt nye ord. De nye tekstene i *Norsk Aviskorpus* blir for hvert døgn sjekket mot allerede registrerte ordformer, og det produseres en liste over ord som ikke fantes i korpuset fra før. Søkene gir rundt 200 nyordskandidater hver dag, og disse kandidatene ble i en periode gått manuelt gjennom og vurdert som aktuelle lemmaer i *Bokmålsordboka*, eller som interessante nyord generelt (Wangensteen 2002). Det viste seg at arbeidet var relativt tidkrevende i forhold til fangsten. De fleste kandidatene var stavefeil, grafiske avvik og annen støy. Ellers var det mange navn, både på personer, steder, firmaer og varemerker, som ikke er interessante for allmennspråklige ordbøker. Dessuten var det mange nyord som ble vurdert som typiske okkasjonalismer med liten interesse for leksikografisk arbeid. Typiske eksempler er *muskelbil*, *notevenderske*, *morderhund* og *låveleder*. I skrivende stund er *pepperkakevandal* et aktuelt ord i norske aviser, siden en ung mann har gått amok i en pepperkaeutstilling i Bergen, og gir nå seks tilslag i *Norsk Aviskorpus*, men vil neppe overleve julen 2009. Den reelle fangsten nyord ved manuell gjennomgang av de automatisk genererte listene av nyordskandidater var bare cirka 2 prosent. Med en slik metode blir dessuten verken neofrasemer eller neosemantiseringer registrert. Vi har derfor vært på utkikk etter andre innsamlingsmetoder.

9. Semiautomatisk innhenting av nyord fra norske nettaviser

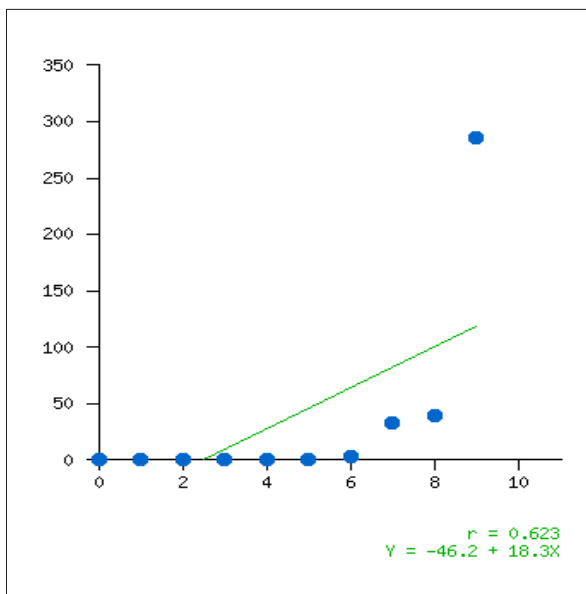
Norsk Aviskorpus er fortsatt vurdert som det beste utgangspunkt for registrering av neologismer i norsk. Hver dag blir de nye tekstene i aviskorpuset sjekket mot allerede registrerte ordformer, og resultatet blir publisert som en liste over nye ord i korpuset siste døgn. Ifølge prosjektets hjemmeside er sammenligningsgrunnlaget omfattende og utgjør mengden av ord som finnes i all elektronisk tekst som er tilgjengelig ved Aksis (bl.a. inkludert alle fullformene i *Bokmålsordboka*). Av de 200 000–250 000 løpende ord som daglig legges til i tekstkorpuset, er 1 000–1 500 nye ordformer.

Ønsket var å finne en mindre arbeidskrevende metode for å sortere ut ekte neologismer fra støy og navn i moderne norsk avisspråk, samt en måte å finne

neofrasemer og neosemantiseringer på. Vi mener å ha utviklet en semiautomatisk metode for dette, som ble utprøvd sommeren 2009.

Norsk Aviskorpus hadde i mai 2009 registrert cirka 2 millioner forskjellige ordformer, 1,3 millioner av disse var ikke registrert tidligere og ble betraktet som nyordskandidater. Ved hjelp av tre ulike filtre som sorterer ordformene i nyordsdatabasen, ble antallet nyordskandidater redusert fra 1,3 millioner til 3 465. Det første filteret silte automatisk vekk proprier, det vil si ord med stor forbokstav som ikke står etter punktum eller først i en setning. Det andre filteret fjernet ordformer med frekvens under en (arbitrær) terskelverdi. Slik sortering er det lett å gjøre, og har også vært utprøvd i liknende prosjekter. Etter disse to filtreringene hadde vi en liste på 17 277 ordformer, som fortsatt er en anseelig mengde å vurdere manuelt. Det ble derfor utviklet enda et tredje filter, som innebærer en metode som oss bekjent er ny.

Norsk Aviskorpus inneholder data for hvert år tilbake til 1998. Data innsamlet gjennom ti år gir oss muligheten til å vurdere enkeltordenes liv over tid. Filter 3 filtrerte vekk frekvente ord med konstant frekvens, mens ord med endring i frekvens ble ansett som interessante kandidater. Stigende eller synkende frekvens ble bestemt ved et statistisk standardmål, lineær regresjon med minste kvadraters metode. Ord med stigende frekvens over tiårsperioden ble ansett som potensielle neologismer. Den statistiske undersøkelsen kan framstilles grafisk, som viser skjæringspunkt og stigningskoeffisienten for neoforformativen *bloggen* (jf. figur 1).



Figur 1. Lineær interpolasjon for frekvensen til ordformen *bloggen*. Stigningskoeffisienten er 0,63.

10. Resultater og vurdering av resultatene

Den manuelle gjennomgangen besto i å gå gjennom 1 000 tilfeldig valgte ord av det som filterne tok bort. Denne gjennomgangen viste at det ved det første filteret var svært lite interessant for nyordsregistrering som ble borte. Bare rundt 1,6 prosent av ordene kunne betraktes som reelle nyordskandidater, og det er så lite at det ikke ville være arbeidet verdt å ta med. Noen av disse 16 ordene var nye forledd i sammensetninger, og nyordene er rimeligvis kommet gjennom som usammensatte. Vi konkluderer derfor at filter 1 fungerer i samsvar med forventningene.

I resultatene etter filter 2 fant vi cirka 6 prosent nyordskandidater, men også her var en stor del av disse sammensetninger der de usammensatte var gått gjennom filteret. Vi konkluderte derfor med at også filter 2, som fjernet frekvente ord som ikke hadde vist tilstrekkelig stigning gjennom tiårsperioden, fungerte godt nok.

Det som var igjen etter at filter 3 hadde virket og hadde fått sortert ut både frekvente ord og ord med stigende frekvens i tiårsperioden, ble gått gjennom ekstra nøyaktig. Vi fant at 300 av de 1 000 tilfeldig utvalgte ordene på den lista var reelle nyord, altså 33,8 prosent. Mange av disse ordene hadde nyproduktive for- og etterledd, og er derfor særlig interessante i en generell kartlegging av nyutvikling i ordforrådet. Vår konklusjon etter den manuelle gjennomgangen av det totale filtreringssystemet var at det fungerte bedre enn antatt. Totalt sett har metoden gitt gode resultater og store tidsbesparinger for kartlegging av leksikalsk språkendring.

Gjennomgangen av listene ga i tillegg til tradisjonell nyordsekserpering dessuten interessante funn og observasjoner om ordutviklingen i moderne norsk. Som et eksempel på det kan nevnes det engelske lånordet *krash*. Allerede i nyordslisten fra 1955 var verbet *å krashe* registrert. Fortsatt er stavemåten opprinnelig, ikke norvagisert, men allerede i 1956 har ordboken *Nyord i norsk* belegg på substantivet *krasj*, og A-tekst har belegg på verbet fra 1976. I tekstene fra 2009 er det mange belegg på både *krasj* og *krasje*, men også *kræsje* og *kræsje* dukker opp. Første belegg i A-tekst er fra 1972. Det vi kan se av nyordsmaterialet er at skrivemåten *krasj* brukes om sammenstøt i konkret forstand (*busskrasj*, *båtkrasj*), mens den mer fonetiske skrivemåten *kræsje* brukes i overført betydning (*børskræsje*).

11. Konklusjon og videre arbeid

Den manuelle analysen av de automatisk genererte listene over nyordskandidater viser at filtrene fungerte godt. Resultatet etter tre filtreringer viste at vel en tredel av ordene som har sluppet gjennom, er interessante å vurdere som neologismer (*antioksidant, birdie, ballbinge, blogg*). Dessuten ser det ut til at ordformene som har tilstrekkelig frekvens til å slippe igjennom filter 2, men ikke filter 3, ikke inkluderer mange nyord, og det tyder på at hypotesen om stigende frekvens bekreftes av studien. Kvalitetssikringen av filterresultatene er gjort ved manuell gjennomgang og vurderingskonstans av to erfarne leksikografer. Mer enn 30 prosent av de ordene som slapp gjennom alle tre filtrene, er på en eller annen måte neologismer. Til sammen viser filtrene at de gir gode resultater og store tidsbesparelser for ekserpering av neologismer. Det som må gjøres videre, er en næranalyse av resultatene av ordene som gikk gjennom alle tre filtrene ut fra avgrensning av nyordsbegrepet og bestemmelse av hvilken type ord som skal med i nyordsbasen.

De aller fleste nyordene i norsk avisspråk fra perioden 1998–2009 er komposita av vanlig type med innarbeidede sammensetningsledd, som *helseforetak, fredagsbønn, bredbånd/bredband* og *børsfest*. Nyord med lånt inventar tilpasses ofte raskt det norske ordforrådet, slik vi kan illustrere med lånordet *blog*, som har *blogge*, nominaliseringene *blogging, blogger* og *bloggings* ('blogginnlegg'), og en lang rekke komposita, som *bloggemne, blogginnlegg, blogginteresse* og *bloggprinsesse*.

Resultatene fra en statistisk nyordsregistrering kan også gi oss ny kunnskap om hvor neologismene i norsk kommer fra, da det ikke lenger er tilfeldig registrering av iøynefallende ordformer, men systematisk og heldekkende registrering. Den alminnelige mening og til dels offisielle påstand er at det norske språket oversvømmes av lånord fra engelsk. En foreløpig sortering av nyordsresultatet antyder imidlertid at bare 10 prosent er nyordene fra den siste tiårsperioden er lån fra engelsk, 88 prosent er vanlige norske sammensetninger, 2 prosent er lån fra andre språk, særlig svensk (*hets*), men også noen få fra italiensk (*scampi*), fransk (*bagett*) og latin (*missil*). Noen av nyordene er også neoforatismer som har kommet inn via slang, som *bananas, lættis* og *loka*.

Totalt sett har metoden gitt oppløftende resultater som gir inspirasjon til videre arbeid. For det første trenger vi å finne eldre data enn 1998. Det ser ut til at vi kunne oppnådd enda mer pålitelige resultater ved å studere en lengre periode enn ti år, siden mange neologismer har vært i bruk en god stund uten å ha blitt registrert. Her kan vi få hjelp fra A-tekst (<http://www.retriever.no/>), som har elektroniske tekster tilbake fra tidlig 1970-tall. Problemet er at det er

forholdsvis lite tekst. Men etter hvert som det utvikles bedre metoder for digitalisering av eldre trykte tekster, kan man tenke seg at grunnlagsmaterialet kan bli bedre. Dessuten har vi behov for å utvikle nye statistiske beregninger for å kartlegge neosemantismer ut fra kontekstendringer. Der kan muligens kollokasjonsanalyse ved hjelp av program som DeepDict (jf. Fjeld & Nygaard & Bick 2009) være til god hjelp.

Det er viktig å kartlegge utvidelsen av ordtilfanget i et språk. Det gir ikke bare oversikt over det rene ordforrådet, det gir også oversikt over hvordan hele ordinventarer oppfører seg og utvikler seg over tid. Det gir innsikt i orddanningsprosessen og endringer i dem, samtidig som det gir innsikt i kulturelle endringer i språksamfunnet.

LITTERATUR

- Barnhart, Robert K. & Steinmetz, Sol & Barnhart, Clarence L. (red.), 1990: Third Barnhart dictionary of new English. New York.
- Elsen, Hilke, 2004: Neologismen. Formen und Funktionen neuer Wörter in verschiedenen Varietäten des Deutschen. Tübingen.
- Fjeld, Ruth Vatvedt & Nygaard, Lars & Bick, Eckhardt, 2009: Semi-automatic retrieval of multiword phrases in a corpus of modern Norwegian. I: Computergestützte und korpusbasierte Methoden in der Phraseologie, Phraseografie und der Lexikografie, red. av Stefaniya Ptashnyk, Erla Hallsteinsdóttir & Noah Bubenhofner. Baltmannsweiler. S. 109–115.
- Jarvad, Pia, 1995: Kryptosammensætninger og skabsafledninger i virkeligheden og i ordbøgerne. I: Nordiske studier i leksikografi 3. Rapport fra Konferanse om leksikografi i Norden. Reykjavík 7.–10. juni 1995, red. av Ásta Svavarsdóttir, Guðrún Kvaran & Jón Hilmar Jónsson. Reykjavík. S. 221–231.
- Lera, Vigleik (red.), 1982: Nyord i norsk 1945–1975. Oslo.
- Simensen, Erik, 1995: Påverknaden på mellomnedertysk på norske diplom frå det 14. Hundreåret. I: Nordisk og nedertysk. Språkkontakt og språkutvikling i seinmellomalderen, red. av Ernst Håkon Jahr. Oslo. S. 55–80.
- Stene, Aasta, 1949: English Loan-Words in Modern Norwegian. Oslo.
- Svensén, Bo, 2004: Handbok i leksikografi. Ordböcker och ordboksarbete i teori och praktik. 2 uppl. Stockholm.
- Thomsen, Johnny, 2002: Præskription og deskription i nyere færøsk leksikografi. I: LexicoNordica 9. S. 107–118.
- Wangsten, Boye, 1995: Et arkiv over forandringer i moderne norsk ordforråd og dets betydning for norsk leksikografi. I: Nordiske studier i leksikografi 3. Rapport fra Konferanse om leksikografi i Norden. Reykjavík 7.–10. juni 1995, red. av Ásta Svavarsdóttir, Guðrún Kvaran & Jón Hilmar Jónsson. Reykjavík. S. 411–420.
- Wangsten, Boye, 2002: Nettbasert nyordsinnsamling. I: Språknytt 2/2002. S. 17–19.