


# NORDISKE STUDIER I LEKSIKOGRAFI

Titel:	Atypiske lemmer i Det Norske Akademis Store Ordbok - Noen betraktninger om deres plass og funksjon i en moderne diakron og deskriptiv ordbok	
Forfatter:	Hanne Lauvstad	
Kilde:	Nordiska Studier i Lexikografi 11, 2012, s. 396-404 Rapport från Konferens om lexicografi i Norden, Lund 24.-27. maj 2011	
URL:	<a href="http://ojs.statsbiblioteket.dk/index.php/nsil/issue/archive">http://ojs.statsbiblioteket.dk/index.php/nsil/issue/archive</a>	

© Nordisk forening for lexicografi

## 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 lexicografi (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.

# Atypiske lemmaer i Det Norske Akademis Store Ordbok

## Noen betraktninger om deres plass og funksjon i en moderne diakron og deskriptiv ordbok

*Hanne Lauvstad*

*Det Norske Akademis Store Ordbok* ('The Norwegian Academy's Dictionary'), an extended and updated version of *Norsk Riksmålsordbok* ('The Norwegian Riksmål Dictionary', vols I–IV 1937–57, vols V–VI 1995) will be a descriptive, diachronic dictionary covering the Norwegian Bokmål and Riksmål, and, as such, one of two national dictionaries for the Norwegian language. As modern Bokmål has its origins in both Old Norse and Danish, the dictionary's vocabulary poses certain challenges to lemmatization. The dictionary contains several archaisms and dialect words, i.e. diachronic and diatopic variants, among them several atypical lemmas, which have to be adapted and standardized as part of the dictionary's macrostructure. Examples illustrate how this problem may be solved in the light of the basic criteria for the lemmatization: etymology, part of speech, gender and inflection, style nuance and usus.

*Nøkkelord:* lemmatisering, diatopisk og diakron varians, makrostruktur

### Bakgrunn

Det Norske Akademis Store Ordbok (NAOB) skal publiseres digitalt i 2014 og er en utvidelse og modernisering av Norsk riksmålsordbok (NRO). NRO utkom i fire bind i årene 1937–57, pluss to supplementsbind i 1995. NAOB skal være en deskriptiv, diakron ordbok for bokmål og riksmål. Redigeringspråket er moderat bokmål. NRO er hovedsakelig basert på ekserpter fra norsk skjønnlitteratur fra 1814 og fremover. Vi arbeider med å supplere basen med ord fra nyere skjønnlitteratur og sakprosa og dessuten med å øke tilfanget av nyord, bl.a. anglisismer, fagord som er gått inn i allmennspråket og slang. Dessuten skal det eksisterende materialet moderniseres språklig og innholdsmessig der det er behov for det. Endret syn på språket og strengere krav til leksikogra-

fisk metode er også grunner til å revidere det ordboksmaterialet som stammer fra NRO.

NRO har en lemmabestand som bl.a. består av en rekke av det jeg i tittelen har valgt å kalle *atypiske lemmaer*. Vår utfordring i NAOB er å definere lemmatiseringsprinsipper<sup>1</sup> som også passer til disse atypiske lemmaene, og å gjennomføre dem. Hovedprinsippet for lemmatisering i NAOB er etymologi.<sup>2</sup> Men også faktorer som gjelder form (ordklasse og genus) eller semantikk (f.eks. at betydningen er radikalt forskjellig fra den opprinnelige) vektlegges, jf.

**ask**<sup>3</sup> (subst., m.) (av norr. *askr*)

I 1. (bot.): 'løvtreart i oljetrefamilien, med motsatte, ulikefannede blader; lat. *Fraxinus excelsior*';

2. (koll.): 'virke av dette treet'

II 1. (dial.): 'eske'; 2. (om norrøne forh.) hulmål for væske = 4 boller | jf. *bolle* og *juste*

Merk at dette er satt som to forskjellige lemmaer i NAOB. Lemmatiseringssystemet er imidlertid ikke helt konsekvent i NRO. Til tross for ordklasseskifte har det f.eks. vært vanlig å anføre adjektiv- el. adverbavledninger under et tilhørende substantiv, sannsynligvis for å spare plass i papirordboken NRO. Dette er nå endret i NAOB, slik at ny ordklasse gir nytt lemma, jf.

**atlask** (subst., m.) (via mnty. *at(te)las*, *atlask*, fra arab. 'glatt'): 'glinsende stoff, særl. av silke, med en noe løs binding av trådene slik at renningen ses på et lengre stykke'

**atlaskes** (adj.) (avl. av *atlask*): 'som er laget av atlask'

EKSEMPEL: *en atlaskes kjole*

Noe fast system i NRO for når betydningen er så forskjellig at det er grunnlag for å opprette et nytt lemma med egen artikkel, er det heller ikke lett å få øye på. Det er f.eks. uklart hvorfor man har valgt å lemmatisere *bakke* (subst., m.) i betydningene 'lite serveringsbrett' og 'line [til å fiske med]' all den stund det siste skal ha en opprinnelig betydning 'brett som linen er viklet rundt' og derfor etymologisk tilhører samme lemma som det første. Begge ordene er varianter av *bakk* (subst., m.), som er lånt fra nty. *back* el. nederl. *bak* 'fat, trau'; oppr. fra romansk.

<sup>1</sup> Jf. Wolskis term *Lemmaverfabren* (Wolski 1989: 366).

<sup>2</sup> I pakt med Nordisk leksikografisk ordbok (NLO 173f.) bruker jeg her termen *lemmatisering* om det Wolski (med referansen til Wiegand 1983) kaller *Lemmaansetzen* (Wolski 1989: 365).

<sup>3</sup> Eksemplene fra NAOB er forkortet og redigert her.

## Problemer – og litt om årsakene til dem

NRO inneholder en rekke lemmaer som avviker fra standardordforrådet. For det første skyldes dette den diakront-deskriptive hensikten med ordboken. Det norske bokmålet har en sammensatt karakter med røtter i så vel norrønt som dansk skriftspråk. Når man vil dokumentere språket diakront, må man derfor ta hensyn til en rekke arkaismer, samtidig som det dukker opp mange dialektale ordformer som må innpasses i ordbokens makrostruktur. Årsaken til atypiskheten er i slike tilfeller altså enten diatopisk eller diakron. For det andre skyldes avvikene metoden. Lemmabestanden i NRO – og dermed også i NAOB – er blitt til ved ekserpering. Dette har medført at man har hatt en viss tendens til å feste seg ved rariteter, dvs. leksikalske enheter som i den litterære riksmåls-/bokmålstradisjonen har skilt seg ut fra det øvrige ordforrådet. Disse leksikalske enhetene er blitt opphøyet til atypiske lemmaer i NRO.

I NRO er mange leksikalske enheter som representerer slik ulik bruk – språkets *parole*, blitt behandlet atomistisk som egne hovedlemmaer, med til dels vilkårlig normering og med liten vekt på deres plass i språkssystemet som helhet, *langue*. Det forekommer ikke så helt sjelden at en leksikalsk enhet er definert med en lydlig variant av seg selv, f.eks. er den dialektale varianten *bløyg* blitt definert med den normerte formen *blei*, jf.

*blei* (subst., m.) (av norr. *bleðgi* 'flis'): 'kile; tynn, smal skive som man skjærer av noe'  
*bløyg* (subst., m.): se *blei*

Begge formene er norske og av samme etymologiske opprinnelse. Merk at lemmatiseringen er endret i NAOB i forhold til i NRO, slik at den første formen er representert med et hovedlemma i NAOB, mens den siste derimot er oppført som et henvisningslemma som er lenket til hovedlemmæet.

Den leksikografiske utfordringen i arbeidet med den moderne og utvidede ordboken NAOB er å ta hensyn til diakron og dialektal variasjon i ordbokens lemmabestand, dvs. *parole*, samtidig som man velger ut hvilken form av en leksikalsk enhet (jf. Zgustas term *lexical unit*, Zgusta 1971: 249 ff.) som skal danne (hoved)lemma i ordboken, altså det som representerer *langue*.

## Diskusjon

### Varianter

Variantene i NAOB kan være a) ortografiske varianter av samme leksikalske enhet (f.eks. *frilans/freelance*, med norvagisert vs. opprinnelig engelsk ortografi), b) (ortografisk gjengivelse av) fonologiske varianter av samme leksikalske enhet

(f. eks. *trykk/trøkk*), der det siste er en dialektal form, som bl.a. hører hjemme i visse nokså faste kollokasjoner av muntlig karakter, som *få seg en trøkk* ('få hard medfart'). Man har også c) lemmaer uten tilsvarende lemmatiserte hovedformer i ordboken, f.eks. *alvskott* (subst., n.), '(i folketro) brudd el. brekk innvendig hos dyr', en leksikalisert sammensetning hvor sisteleddet ikke har noen parallellform 'skudd', som ellers ville være hovedformen i NAOB, men isteden den (sett fra moderat bokmålsstandpunkt) muntlige variantformen *skott*, som henger sammen med bruksområdet. Det finnes derfor ikke noe alternativ til hovedlemma.

### Kriterier for valg av hovedlemma

I arbeidet med NAOB må man ta stilling til når det er grunnlag for å slå sammen etymologisk beslektede varianter av en leksikalsk enhet til ett lemma og hvordan man skal behandle atypiske, enkeltstående varianter i ordboken. Kriteriene for å definere lemmaer som atypiske er enten *frekvens* eller *form*.

Dette er, med inspirasjon fra Bo Svenséns diskusjon av homonymi kontra polysemi (Svensén 2004: 116ff.), NAOBs kriterier for hva som bør være hhv. hovedlemmaer og henvisningslemmaer:

- etymologi
- ordklassetilhørighet
- genus og bøyning
- stilnivå og bruksområde

#### *Eks. 1: Normative og synkrone parallellformer i norsk bokmål med lignende stilvalør*

*ben* (subst., n.) fra da. *ben*; tilsv. norr. *bein* 'knokkel; legg (senere bet.)'

*bein* (subst., n.) av norr. *bein* 'knokkel; legg (senere bet.)'

Dette kan regnes som varianter av samme leksikalske enhet, men de er lemmatisert hver for seg i NRO med delvis overlappende betydninger og samlet under hovedlemmaet *ben*, med *bein* angitt som variant og som eget henvisningslemma i NAOB. Etymologien er parallell; den ene variantformen er nedarvet direkte fra norr. *bein*, den andre lånt inn fra dansk *ben*. I moderne bokmål blir disse variantformene imidlertid oftest oppfattet som likeverdige når det gjelder stilnivå og bruksområder. Den norske bøyningsformen i pl. på *-a* er nå blitt normen, tidligere ble den danske pl. formen *benene* brukt i norsk, men den er nå temmelig sjelden. La oss se nærmere på spørsmålet om stilvalør og normering.

*Eks. 2: Variantformer med ulik stilvalør og ulik vei inn i moderne bokmål*

**hu** (subst., m.) (fra da. *hu*; sa. ord som *hug*), (mest foreld. el. arkais. og i faste uttr.): 1. 'sinn; sinnelag; tanke(r)'

UTTRYKK:

*med berådd, velberådd hu* | se berå og velberådd

(*bli, være, gjøre*) gram i hu | se gram, adj.)

*komme i hu* 'komme i tanker om; komme på; huske' | jf. ihukomme

*gå el. rinne noen av hu* (foreld. el. arkais.) 'gå noen av minne, glemmes av noen' | jf.

omhu

2. (særl. om elskovsfølelse) lyst; attrå

**hug** (subst., m.) (av norr. *hugr*; sa. ord som hu):

1. (dial. el. arkais.) 'sinn; sinnelag; tanke(r)' | jf. *hu*;

2. 'lyst; attrå' | jf. *hu*

Formene tilsvarer hverandre etymologisk, begge har betydningene 1. 'sinn; sinnelag; tanke(r)'; 2. 'lyst; attrå', men de har ulik brukshistorie i norsk (som hhv. arveord og lånord) og (noe) ulik stilistisk valør, dessuten ser det ut til at *hu* sjelden bøyes. Ellers virker det ut fra ordbokens litterære belegg som om begge variantformer kan brukes i fast uttrykk som *komme i hu*, mens idiomatiske uttrykk som *med velberådd hu* 'etter moden overveielse', *gram i hu* 'forbitret' og *rinne noen i hu* 'dukke opp i minnet' utelukkende knyttes til den danske formen. Også Bokmålsordboka (BMO), som ellers har valgt *hug* som hovedlemma, oppgir disse uttrykkene utelukkende med formen *hu*. Begge variantformer vil bli beholdt i NAOB som to lemmaer med etymologiopplysninger. Her er også eksemplene blitt stående igjen under hhv. *hug* og *hu* for å skille dem tydelig ut. Praksis for plassering av autentiske eksempler varierer for øvrig fra tilfelle til tilfelle, jf.

**brudd/brott** (subst., n.): Den første formen er lånt fra da. *brud* og dominerer i (moderat) bokmål, den andre er arveord fra norr. *brot*; begge formene er dannet til verbet *bryte*. Etymologi, genus og bøyning stemmer over ens, mens bruksområdene bare delvis overlapper; *brott* har flere dialektale betydninger som mangler belegg i formen *brudd*, bl.a. 'bratt skråning el. bakke', 'brenning (bølge(r) som bryter mot kyst el. skjær)', 'elvestryk', 'brus; knak, brak'; 'avbruket stykke, skår, særl. av glass el. steintøy; bruddstykke', (på klær) 'oppslag, brett el. kant, især som står ut'. Men de fleste betydninger er felles, f.eks. 'det å brytes el. være brutt', 'sted hvor man bryter stein e.l.', 'det å få (et bestående forhold) til å opphøre', 'anstøt (mot), krenkelse (av et løfte, en lov, en sedvane e.l.)', 'forbrytelse' (den siste betydningen er imidlertid sjelden når det gjelder formen 'brudd').

Stilvaløren må også sies å være forskjellig, i moderat bokmål vil *brott* oppfattes som en radikal eller dialektalt preget form. Her er det altså stilistiske og semantiske egenskaper som gjør at man anser *brudd* og *brott* som to forskjellige leksikalske enheter og oppretter dem som to hovedlemmaer til tross for felles etymologi.

*Eks. 3: Former uten stilistisk motstykke i bokmål men med beslektet lemma*

Visse arkaistiske former mangler motstykke i moderne bokmål. Et eksempel er **vovet**: (adjektivisk perfektum partisipp av *vove*, dansk form av *våge*), en variantform som har andre konnotasjoner i moderne norsk enn den mer umarkerte formen *våget*. Her kunne man eventuelt ha opprettet et eget lemma for adjektivet, siden de øvrige av verbets bøyingsformer er nærmest fraværende i moderne norsk. Men som diakron ordbok må NAOB også gi plass til og dokumentere en rekke tilfeller av eldre, litterær bruk av variantformen, som ville bli mindre synlig dersom alle sitater ble flyttet til artikkelen *våge*. Igjen er det stilpreg og brukshistorie som gjør at man, på tross av etymologi, opprettholder et eget lemma for *vove*. Imidlertid vil vi i NAOB, av hensyn til den marginale rollen verbet i denne danske formen spiller i moderne norsk språkbruk, forenkle NROs omfattende artikkel for *vove*, og den adjektiviske perfektum-partisippformen vil bli fremhevet med eget betydningsnummer og behørig belagt med autentiske eksempler, også fra den senere tid.

*Eks. 4: Hapax legomena, unormerte dialektord og andre sjeldne leksikalske enheter*  
La oss se på noen eksempler på *hapax legomena*, unormerte dialektord o.a. sjeldne leksikalske enheter:

**abengslig** (adj.) (avl. av norsk dialektalt *avbendes*, *avbendeleg(e)* 'vanskelig tilgjengelig, bakvendt, avbaklig, uhåndterlig, overvettes'; se *-lig*), brukt med bet. 'urimelig; forkjært' og belagt med et autentisk eksempel fra P.Chr. Asbjørnsens *Huldreeventyr*. Lemmaet blir i NAOB sammenlignet (til slutt i artikkelkroppen) når det gjelder bruk og semantikk med de fonetisk lignende lemmaene *avbragdslig* og *abrenslig*.

**blodge** (vb.) (av norr. *blóðga*; avl. av *blóðugr*; se *blodig*; jf. *blogge*, *blygge* og *bløgge*) brukt dialektalt, med bet. 1. 'gjøre blodig; såre så blodet flyter', 2. 'slakte, stikke (fisk) slik at blodet renner av'. Lemmaet *blodge* forekommer alt i NRO og fremstår som et konstruktlemma, en restituert form i pakt med den etymologiske opprinnelsen, men uten autentiske belegg. I NAOB er dette valgt som hovedlemma. De lydlige, dialektale variantene *blogge*, *blygge* og *bløgge* er bevart som henvisningslemmaer av hensyn til *brukshistorien* (de autentiske eksemplene er bevart i artiklene som hører til henvisningslemmaene for å synliggjøre *usus*) og *brukerrelevansen* (en ordboksbruker kan ikke forventes koble

*blygge* til *blodge*). I NRO var alle disse variantene egne, isolerte (hoved)lemmaer, uten sammenkobling og hensyn til *langue*.

## Sammenfatning

I NAOB har vi følgende grunnlag for våre lemmatiseringsprinsipper:

- *representativitet og normativitet*: Vi velger som hovedlemma den variant som best samsvarer med den ortografiske formen for ordbokens språklige norm, i vårt tilfelle moderat bokmål. Om nødvendig oppretter vi et konstruktlemma. Normen kan imidlertid være vanskelig å avgjøre, jf. eksemplene *ben/bein* og *hul/hug*. I en deskriptiv ordbok som NAOB vil brukshistorie, frekvens og stilnivå tillegges vekt når man lemmatiserer. Etymologisk beslektede former med lignende bruk og stilpreg blir i den grad det er mulig samlet under et hovedlemma, eventuelt med henvisningslemmaer, som i en digital ordbok som NAOB lenkes til hovedlemmaene.
- *autentisitet*: Vi ønsker ikke å overnormere sjeldne varianter; jf. eksempelet *abengslig*. Denne fremgangsmåten bryter naturligvis med foregående prinsipper, men den gjelder vel å merke unntakstilfeller.
- *brukerrelevans*: Vi er opptatt av at brukeren er i stand til å finne den informasjon han eller hun leter etter. I denne hensikt bruker vi bl.a. digital lenking. Dessuten legger vi vekt på at lemmaet skal ha form som er slik at brukeren finner det igjen (jf. *blodge*-eksempelet).

Et problem vi ofte støter på i NAOB, er å avgjøre når vi skal samle litterære belegg (sitater) under hovedlemmaet og når vi bør la dem stå under henvisningslemmaet. Dette avgjøres av hvor og i hvilken grad vi ønsker å dokumentere og synliggjøre variantene. Hvis hovedlemmaet har få eller ingen belegg, blir sitateksempler flyttet dit, mens de gjerne får stå under henvisningslemmaet dersom det blir ansett som viktig å skille ut egne belegg på variantformen. Altså er det skjønn som avgjør hvor interessant det er å dokumentere bruken av en (underordnet) variantform.

## Konklusjon

En diakron, deskriptiv ordbok som NAOB stiller leksikografen overfor visse utfordringer når det gjelder lemmatiseringen, utfordringer som er knyttet til diakron og diatopisk varians, altså ved det jeg omtaler som atypiske lemnaer. Vi har sett at etymologien er det viktigste kriteriet for hvilke av de atypiske lemnaene som blir hovedlemmaer og hvilke som blir henvisningslemmaer. Også ord-



klasse og genus er sentrale lemmatiseringskriterier. Imidlertid finnes det tilfeller hvor vi oppretter egne lemmaer for varianter med samme etymologiske opprinnelse. Da er det bruksmåter og stilnivå som avgjør. Spørsmålet blir hvor store forskjeller man kan godta når det gjelder bruk og stilvalør for å lemmatisere varianter samlet – eller velge å opprette forskjellige lemmaer.

Også språkpolitisk er dette en vanskelig balansegang. Det hersker ofte uenighet om de forskjellige variantenes stilvalør og hvilke som skal regnes som ortografisk hovedform. Eksempelvis vil mer konservative krefter mene at variantformer som *ben* og *bein* ikke er stilistisk likeverdige innenfor alle bruksområder. Også hensikten med ordboken virker inn på lemmatiseringen – i en deskriptiv ordbok som NAOB kan vi tillate oss å ta mer hensyn til diakron brukshistorie og stilistiske særegenheter enn det som er vanlig i en normativ ordbok. Som nevnt er også brukerrelevans viktig for oss. Samtidig er det vårt mål at lemmatiseringen er i pakt med språkssystemet som helhet, slik at etymologisk beslektede ordformer med felles bruksmåter og stilpreg blir samlet, mens lydlige varianter og andre variantformer henviser til en hovedform. Det digitale lenkesystemet blir et hjelpemiddel for å gjøre ordbokens makrostruktur oversiktlig og brukervennlig, siden man ved et tastetrykk kan gå fra henvisningslemmaer til hovedlemmaet, hvor all informasjon er samlet.

Hovedlemmaene representerer på denne måten *langue*, mens *parole* dels blir ivaretatt av henvisningslemmaene, dels av autentiske eksempler i artikkelkroppene. Inkonsekvensene når det gjelder lemmatisering i NRO skyldes antagelig at man der ikke alltid har skjelnet tydelig mellom *langue* og *parole*, noe som ikke bare viser seg i lemmatiseringen, men også på andre nivåer i ordboken, f.eks. når det gjelder utformingen av definisjonene. Men det er en annen historie.

## LITTERATUR

- BMO = Wangensteen, Boye (red.): Bokmålsordboka. Definisjons- og rettskrivningsordbok. 3. utg. Oslo 2005.
- NLO = Bergenholtz, Henning m. fl. (red.): Nordisk leksikografisk ordbok. 1997. Oslo.
- NRO = Norsk riksmålsordbok. B. I–IV 1937–57, B. V–VI 1995. Oslo.
- Svensén, Bo, 2004: Handbok i lexikografi. Ordböcker och ordboksarbete i teori och praktik. Andra, omarbetade och utökade upplagan. Stockholm.
- Wiegand, Herbert Ernst, 1983: Was ist eigentlich ein Lemma? Ein Beitrag zur Theorie der lexikographischen Sprachbeschreibung. I: Studien zur neuhochdeutschen Lexikographie. B. III. Utg. av Herbert Ernst Wiegand. Hildesheim, Zürich, New York (Germanistische Linguistik 1–4/82. S. 401–474).
- Wolski, Werner, 1989: Das Lemma und die verschiedenen Lemmatypen. Berlin, New York. I: Handbücher zur Sprach- und Kommunikationswissenschaft. B. 5.1. S. 360–371.

Zgusta, Ladislav, 1971: *Manual of Lexicography*. Haag, Paris. I: *Janua linguarum. Series maior* 29.

*Hanne Lauvstad*

Det Norske Akademis Store Ordbok, Kunnskapsforlaget

hanne.lauvstad@kunnskapsforlaget.no