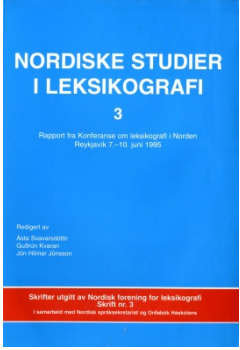


# NORDISKE STUDIER I LEKSIKOGRAFI

|            |   |   |
|------------|---|---|
| Titel:     | Nationalencyklopedins ordbok - några karakteristiska typer av information   |  |
| Forfatter: | Sven-Göran Malmgren   |   |
| Kilde:     | Nordiske Studier i Leksikografi 3, 1995, s. 275-282<br>Rapport fra Konferanse om leksikografi i Norden, Reykjavík 7.-10. juni 1995    |   |
| 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 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.

Sven-Göran Malmgren

## Nationalencyklopedins ordbok — några karakteristiska typer av information

In this paper, some characteristics of a new defining dictionary of Swedish, *Nationalencyklopedins ordbok* (NEO; The dictionary of the National Encyclopedia), are presented. The three-volume dictionary, comprising about 62 000 lemmas, will appear in 1995–96. Partly, it is based on another dictionary of Swedish, *Svensk ordbok* (SOB; A dictionary of Swedish). However, there are several new — or partly new — informational categories in the NEO, accounted for in the paper. These categories are (1) etymologies (for about half of the lemmas) and years of the first known occurrence (for every lemma and every numbered sense); (2) valency; (3) morphological analysis; (4) word formation; (5) musical accent, and (6) references to synonyms, antonyms, co-hyponyms, etc.

### 1 Inledning

Vid Institutionen för svenska språket vid Göteborgs universitet pågår sedan några år arbetet på en större svensk definitionsordbok, *Nationalencyklopedins ordbok* (NEO). Ordbokens tre band om sammanlagt ca 2000 sidor och med ca 62 000 lemman är planerade att utkomma hösten 1995, våren 1996 och hösten 1996. En CD-ROM-version av ordboken förbereds också. Ordbokens namn förklaras av att den ges ut i anslutning till den stora svenska *Nationalencyklopedin* (och av samma förlag, Bra Böcker); i bokhyllan kommer de tre ordboksbanden att se ut som lika många band av encyklopedin.

Arbetet på NEO utgår från den lexikala databas som utgjorde grunden för den 1986 utgivna *Svensk ordbok* (SOB). I synnerhet bygger betydelsebeskrivningarna på SOB:s betydelsebeskrivningar, om man bortser från några tusen nya ord i NEO (se vidare Allén & Malmgren 1993). NEO är emellertid inte en ny upplaga av SOB utan en ny ordbok, med bl.a. åtskilliga informationskategorier som inte fanns representerade i SOB och på något undantag när inte heller i någon annan allmän ordbok över nutidssvenskan. Nedan presenteras de viktigaste av dessa informationskategorier jämte några problem som vi konfronterats med vid utarbetandet av ordboken.<sup>1</sup>

---

<sup>1</sup>Arbetet på ordboken leds av professor Sture Allén med artikelförfattaren som biträdande projektledare. Övriga medarbetare är eller har varit Sture Berg, Bertil Engdahl, Birgitta Ernby, Carl von Garaguly, Sven Lövfors, Christian Sjögreen (systemering), Lennart Tegner och Maria Toporowska Gronostaj. Dessutom har ett stort antal experter engagerats för kontroll av olika typer av uppgifter i ordboken.

## 2 De nya informationskategorierna i NEO

De viktigaste nya informationskategorierna i NEO är:

- Etymologier; årtal för första kända belägg i svenskan
- Konstruktionsuppgifter vid alla verb och vid adjektiv och substantiv med bestämmingar av valenskaraktär
- Ordled ( $\approx$ morfologisk analys, av alla uppslagsord)
- Nomina actionis vid alla verb

Dessutom är informationen kraftigt utökad inom ytterligare två kategorier:

- Uttalet (bl.a. ges uppgifter om tonaccent)
- Hänvisningar mellan semantiskt besläktade ord, med åtskillnad mellan hyponymer, hyperonymer och kohyponymer (samt, givetvis, synonymer och antonymer)

Av följande exempel ur det nu tryckfärdiga första bandet framgår hur artiklarna i NEO är disponerade, bl.a. hur de nya informationskategorierna är inordnade i artiklarna. Observera "rubrikerna" i kapital som är avsedda att leda läsaren rätt om han bara vill läsa delar av artikeln. (Etymologin och årtalet för första belägg står under rubriken "HIST." sist i artiklarna.)

**abrupt** adj., neutr. ~

ORDLED: *ab-rupt*

• plötsligt inledd eller avslutad särsk. om viljestyrt förlopp {→**hastig**, **tvär 2**}: *säga ett ~ adjö; ett ~ svar*

HIST.: sedan 1843; av lat. *abruptus* 'tvärt avbruten'; jfr **ruptur**

**ackompanje**'ra verb ~de ~t

ORDLED: *ac-komp-anj-er-ar*

SUBST.: **ackompanjerande**, **ackompanjer- ing**; **ackompanjemang**

• spela ackompanjemang till viss huvudstämma etc.: *orkestern ~de kupletterna med säker stilkänsla*

BET.NYANS: överfört i fråga om tillkommande sinnesintryck e.d.: *vinerna som ~de var en torr sherry och ett halvtorr moselvin*

KONSTR.: ~ (*ngn* el. *ngt*) (*på ngt*)

HIST.: sedan 1779; av fra. *accompagner* 'ledsaga; ackompanjera'; jfr **kompani**

Fig. 1. Artiklarna abrupt och ackompanjera i NEO

Jag ska nu något kommentera de sex (helt eller delvis) nya informationskategorierna och samtidigt exemplifiera med hjälp av s.k. skärmskrifter från databasen. De sorteringar, i de flesta fall baklängessorteringar, som ligger till grund för exemplen kan när som helst göras av medarbetarna, och förhoppningen är att åtminstone vissa av dem blir möjliga att göra på CD-ROM-versionen av NEO.

## 2.1 Etymologier; årtal för första belägg

Etymologier ges i NEO för alla enkla ord och även för många avledda och sammansatta ord, då speciella skäl talar för detta. Oftast behöver endast (högst) en etymologi per lemma ges, men då lemmatiseringen i NEO görs på grundval av rent formella — icke-etymologiska — kriterier kan det förekomma att olika betydelsemoment inom ett och samma lemma har olika etymologier. Sammanlagt ges ca 27 000 etymologier.

Årtal för första kända belägg i svenskan ges för alla numrerade betydelsemoment, ca 68 000 stycken. Huvudkällan är *Svenska Akademiens ordbok* (SAOB) och, särskilt för den del av alfabetet som ännu inte täcks av SAOB (*t-ö* och en mindre del av *s*), SAOB:s samlingar (SAOS), som numera är mikrofilmade och direkt tillgängliga även i Göteborg. Andra viktiga källor är Söderwalls *Ordbok öfver svenska medeltidsspråket* och *Nyord i svenskan från 40-tal till 80-tal*. Dessa källor tillhandahåller tillsammans den stora majoriteten av uppgifterna om första belägg, men för flera tusen av uppslagsorden måste uppgifterna sökas på annat håll. Det gäller bl.a. många ord i början av alfabetet som är för nya för SAOB men för gamla för *Nyord i svenskan*, och det gäller ord som är för nya för *Nyord i svenskan*. I inte så fall har vi också funnit tidigare belägg än SAOB-beläggen med hjälp av institutionens (Språkbankens) egna korpusar, inte minst den stora Strindbergskorpusen. De ämnesexperter som engagerats för kontroll av artiklar om facktermer har också i många fall angett tidigare årtal för första belägg än någon av huvudkällorna.

Likväl har det hänt att vi i sökandet efter ett rimligt årtal för första belägg har sett alla tänkbara källor uttömda. Vi kanske har haft ett konkret belägg från 1980-talet, men vi har varit säkra på att ordet eller betydelsen funnits i svenskan sedan 1960-talet. I sådana (undantags)fall har vi, med stor tvekan, bestämt oss för att frångå kravet på ett konkret första belägg och tillåtit oss skrivningar som "trol(igen) sedan 1960-talet". I andra fall kan vi godta årtalet för det äldsta kända belägget men markera att ordet/betydelsen troligen är betydligt äldre med en skrivning av typen "åtm(instone) sedan 1971". Så gott som alla dessa problem uppstår vid ord eller betydelser som kommit in i svenskan relativt nyligen.

Fig. 2 visar årtal för första belägg och i ett par fall etymologier för ett antal ord med efterleden *-sång*.

|               |           |                          |
|---------------|-----------|--------------------------|
| minnesång 1/1 | 1848      | till ty. *Minne 'kärlek' |
| ottesång 1/1  | före 1520 | fornsv. *ottosanger      |
| vaggsång 1/1  | 1749      |                          |
| folksång 1/1  | 1810      |                          |
| kråksång 1/1  | 1843      |                          |

Fig. 2. Historiska uppgifter för ett urval ord med efterleden *-sång*

## 2.2 Konstruktionsuppgifter (valenser)

Alla verb i NEO — ca 9000 — har försetts med konstruktionsuppgifter av valenstyp. Det framgår alltså explicit om verben (eller de resp. betydelsemomenten av polysema verb) är intransitiva eller transitiva, och det ges även uppgift om ev. prepositionsobjekt, predikativ,

valensadverbial m.m. Om komplementen är fakultativa anges det med parentes. Det framgår också huruvida huvudorden i verbkomplementen är person- eller "sak"-betecknande ("ngn" resp. "ngt"), eller bådadera. Här nedan återges konstruktionsuppgifterna vid verben från *betunga* till *betyda*. (Frågan huruvida ett visst komplement hör till valensramen eller ej — med annan terminologi: om det är fyllnadsled eller tilläggsled — kan vara vanskelig att avgöra, men ett användbart testbatteri finns i Nikula 1986.)

|               |                     |
|---------------|---------------------|
| betunga 1/1   | & ngn/ngt (med ngt) |
| beltvinga 1/1 | & ngn/ngt (med ngt) |
| betvivla 1/1  | & ngt/att+SATS      |
| betyda 1/1    | & ngt/att+SATS      |
| betyda 1/2    | & ngt/att+SATS      |

Fig. 3. Konstruktionsuppgifter för verben *betunga*–*betyda*

Också adjektiven och substantiven har vid behov försetts med konstruktionsuppgifter. Här är problemet att dra gränsen mellan fyllnadsled och tilläggsled nog ännu svårare än vid verben. För adjektiven har dock en modell och ett par test utarbetats (se Malmgren 1987). Några av typerna kan exemplifieras med *artig* (*mot ngn*), *jämbördig* (*med ngn*) och *van* (*vid ngn/ngt*). En stor grupp substantiv som ofta får anses ta fyllnadsbestämningar är givetvis verbalsubstantiven, en annan grupp är "relationella" substantiv, t.ex. *bror* (*till ngn*). Av fig. 4 framgår konstruktionsuppgifterna vid ett antal adjektiv.

|               |                         |
|---------------|-------------------------|
| led 1/2       | & (mot ngn)             |
| ledlad 1/1    |                         |
| ledbruten 1/1 |                         |
| ledig 1/1     | & (för ngn) (att + INF) |
| ledig 1/2     | & (från ngt/ngn)        |
| ledig 1/3     |                         |
| ledsam 1/1    | & (för ngn) (att + INF) |

Fig. 4. Konstruktionsuppgifter för adjektiven *led*–*ledsam*

### 2.3 Ordled (morfologisk analys)

Uppslagsord med en flerstavig grundform eller böjningsform har försetts med en morfologisk analys, så när som på att morfem som inte innehåller en vokal inte beaktas. Denna analys innebär samtidigt en rekommendation om avstavning på konsekvent morfologisk grund i motsats till den blandning av mekanisk och morfologisk avstavning som hittills varit vanlig i

svenskan. (Ordboken är själv avstavad enligt denna morfologiska princip.) För att principen ska framgå så tydligt som möjligt analyseras vid substantiven och verben vanligen någon böjningsform som ofta har en stavelse mer än grundformen, i allmänhet bestämd form singular vid substantiven och presens vid verben. Sålunda t.ex. (i artiklarna **dag** resp. **jaga**): *dag-en* och *jag-ar* (eftersom *dag* och *jaga* inte skulle få några avstavningsmarkeringar).

En extra finess är att en stor morfemgräns anges vid sammansatta ord med hjälp av ett längre streck än divisen (t.ex. *buk-tal-ar-en*). Detta ger en extra upplysning om ordets struktur och samtidigt markeras därmed det bästa avstavningsstället. Att avgöra om sammansättning eller avledning föreligger, dvs. om långt och kort avstavningsstreck ska anges, är självfallet inte alltid lätt. I tveksamma fall har bl.a. ett betoningsskriterium använts. Sålunda har *geovetenskap* [jeˈ-] analyserats *geo-vet-en-skap-en* men *geografi* [-fiː] *geo-graf-in*.

Fig. 5 visar en skärmutskrift med ordledsanalyser av ett antal baklängessorterade ord (tecknet = står här för stor morfemgräns):

|                  |        |                               |                            |
|------------------|--------|-------------------------------|----------------------------|
| management 1     | subst. | Ingen: böjning; n-genus:      | manage · ment              |
| logement 1       | subst. | &et, plur. =, best. plur. &en | loge · ment · et           |
| detachment 1     | subst. | &et, plur. =, best. plur. &en | de · tache · ment · et     |
| signalement 1    | subst. | &et, plur. =, best. plur. &en | sign · ale · ment · et     |
| element 1        | subst. | &et, plur. =, best. plur. &en | ele · ment · et            |
| bränsleelement 1 | subst. | &et, plur. =, best. plur. &en | brän · sle=ele · ment · et |
| värmeelement 1   | subst. | &et, plur. =, best. plur. &en | värme=ele · ment · et      |

Fig. 5. Ordledsanalyser av ett urval ord med slutleden -ment

## 2.4 Nomina actionis

Vid alla verb ges uppgift om möjliga nomina actionis, i första hand verbalsubstantiv på *-(n)ing* och *-ande/-ende*, men också på *-else*, *-tion* m.m. om dessa avledningar är vanliga. Ett viktigt krav är att avledningen fortfarande ska beteckna verbhandlingen. Har avledningen t.ex. fått en konkret betydelse tas den inte upp. Fig. 6 visar uppgifterna om verbalsubstantiv vid ett antal baklängessorterade verb.

|                    |          |                              |
|--------------------|----------|------------------------------|
| mumifler\ a 1/1    | -ade -at | -ande, -ing; mumifikation    |
| teknifler\ a 1/1   | -ade -at | -ande, -ing                  |
| saponifler\ a 1/1  | -ade -at | -ande, -ing; saponifikation  |
| personifler\ a 1/1 | -ade -at | -ande, -ing; personifikation |
| unifler\ a 1/1     | -ade -at | -ande, -ing                  |
| typifler\ a 1/1    | -ade -at | -ande, -ing                  |
| verifler\ a 1/1    | -ade -at | -ande, -ing; verifikation    |

Fig. 6. Uppgifter om nomina actionis av ett urval verb på -fiera

## 2.5 Uttal

Jämfört med SOB är uttalsangivelserna i NEO mycket explicita. Strävan har varit att förut-sätta enbart de allmännaste uttalsreglerna som bekanta för användaren (jfr dock strax nedan om bitryck). Den viktigaste egentligt nya uttalsinformationen i NEO gäller tonaccenten, om vilken uppgifter av någon anledning inte brukar ges i allmänna svenska ordböcker (undantag: inlärningsordboken *Svenska ord*). Här som i andra sammanhang gäller naturligtvis att det som ges är ett riksspråksuttal. Sålunda ges accent 2 som enda (riksspråks)uttal vid de flesta sammansättningar, trots att många dialekter kan ha accent 1-uttal. Däremot försöker vi täcka variation inom riksspråksuttalet, vilket bl.a. resulterat i att nära tusen ord fått både accent 1- och accent 2-markering.

Fig. 7 visar uttalsmarkeringarna i databasen för en grupp ord som ofta visar uttalsvariation (här är > markering för accent 1 och < för accent 2).

|               |   |         |        |     |        |
|---------------|---|---------|--------|-----|--------|
| audlto<rlum   | 1 | subst.  |        |     |        |
| dormlto<rlum  | 1 | subst.  |        |     |        |
| kleresto<rlum | 1 | +subst. |        |     |        |
| konserverlum  | 1 | subst.  | -to<r- | el. | -to>r- |
| konslstorlum  | 1 | subst.  | -to<r- | el. | -to>r- |
| konvltorlum   | 1 | subst.  | -to<r- | el. | -to>r- |

Fig. 7. Uttalsuppgifter för ett antal ord på -orium

Viktigt är att ordboken redogör för oförutsägbara förändringar i uttalet från grundformen till böjningsformerna. Många ord har t.ex. accent 1 i grundformen och accent 2 i en eller flera av böjningsformerna. I de flesta fall går detta att beskriva med allmänna regler, men det finns en större undantagsgrupp: substantiv av 3:e deklinationen. Det vanligaste är här accent 2 i plural (*fil'mer*), men åtminstone i något hundratal fall förekommer accent 1 (*gre'ker*, *grän'der*). De senare fallen anges i NEO, och de allmänna reglerna ges i inledningen.

I ett avseende är uttalsangivelserna i NEO inte explicita: endast markering för huvudtryck ges, inte för bitryck. Inte heller ges alltid uttalsanvisningar för de stavelser som inte har huvudtryck och ev. är behäftade med uttalsvårigheter. Det är huvudsakligen sammansatta ord som på detta sätt inte får fullständiga uttalsmarkeringar. En anledning till detta var att det saknas en allmänt känd beteckning för bitryck i svenskan, en annan att fullständiga uttalsangivelser också skulle ha medfört krav på mycket vidlyftiga uppgifter om variantuttal (särskilt beträffande uttal av *o* i obetonad stavelse).

Med hjälp av ordbokens inledning och den morfologiska analysen bör dock även en icke-svensk användare av ordboken kunna komma fram till rätt uttal också av sammansatta ord. Antag att han slår upp artikeln **husbehov**, som har uttalsangivelsen [hu's-] och den morfologiska analysen hus-be-hov-et. Det långa strecket kan här fungera som en hänvisning till artikeln **behov**, och genom att summera uttalsuppgifterna på de båda artiklarna erhåller man — efter att ha förvandlat huvudtrycket på slutstavelsen i det självständiga ordet *behov* till bitryck på samma stavelse i efterleden *-behov* — det fullständiga uttalet av *husbehov*. Man måste dock också känna till regeln om accentflyttning i senare ledet av sammansättningar i

svenskan, men också denna regel är så gott som heltäckande (ord som *studiebidrag* uttalas t.ex. oftast med bitryck på sista stavelsen, trots att *bidrag* har huvudtryck på första stavelsen).

## 2.6 Hänvisningar mellan semantiskt besläktade ord

De flesta semantiska hänvisningarna är hänvisningar mellan kohyponymer (sammanlagt ca 30 000). Dessa hänvisningar markeras med en pil, riktad åt höger (se artikeln **abrupt** i fig. 1). Gränsen till vad som bedömts som synonymer är självfallet inte skarp, men allmänt kan sägas att vi valt att tolka synonymbegreppet ganska strängt; vi har alltså krävt mycket stor betydelselikhet för att ange synonymirelation. Motsvarande gäller för antonymirelationen.

Utom dessa relationer anges i många fall hyponymi- och hyperonymirelationer (med nedåtriktad resp. uppåtriktad pil). I artikeln **nederbörd** ges exempelvis hyponymerna *regn*, *snö* och *hagel*. I artiklarna om dessa ord ges visserligen inte motsvarande hänvisning till hyperonymen *nederbörd*, men detta framgår ändå indirekt genom att ordet *nederbörd* är huvudord i definitionerna. De totalt ca 35 000 hänvisningarna mellan semantiskt besläktade ord — i viss mån sålunda tillsammans med betydelsebeskrivningarna — torde sammantaget resultera i ett rätt finmaskigt semantiskt nätverk.

## 3 Avslutning: något om NEO:s målgrupp

De flesta av användarna av NEO kommer självfallet att ha svenska som modersmål. Men det är enligt min mening viktigt, inte minst med tanke på att Sverige numera är ett mångkulturellt land, att också se en stor enspråkig svensk ordbok som ett hjälpmedel vid avancerad svenskinläring. Det låter sig sägas, och sägs också ibland, att mycket av informationen i enspråkiga definitionsordböcker redan är känd för många av användarna, något som kan vara frustrerande för ordboksförfattarna (jfr Jonsson 1994). Om man, med en försiktig uppskattning, tänker sig åtminstone några tusen NEO-användare med annat modersmål än svenska, behöver man emellertid inte bekymra sig så mycket över eventuell redundans: då blir inte minst flera av de nya informationskategorierna (tonaccent, nomina actionis, ordledsanalyser, konstruktionsuppgifter) av särskilt värde. Då blir det också ytterst motiverat att skriva allmänna lexikonregler i inledningen — när man säger att ingen läser ordboksintroduktioner förbiser man rimligen denna grupp av användare. Och då kan man inte minst se etymologiuppgiften, i och för sig den uppgift som i många fall säkert har det största läsvärdet för svenska användare, som en ytterligare hjälp för icke-svenska användare att förstå ordets betydelse.

## Litteratur

- Allén, Sture/Sven-Göran Malmgren 1993: Om Nationalencyklopedins ordbok. I: *Mänsklig mångfald*. Göteborg: Humanistiska utbildnings- och forskningsnämnden, Göteborgs universitet, 103–111.
- Jonsson, Hans 1994: Betydelsebeskrivningens förutsättningar i en enspråkig ordbok. I: *Lexico-Nordica* 1, 89–105.
- Malmgren, Sven-Göran 1987: *Om adjektiven i LPS-projektets lexikon*. Rapporter från Språkdata. 23. Göteborgs universitet: Institutionen för språkvetenskaplig databehandling.



NEO = *Nationalencyklopedins ordbok* 1–3. (utkommer 1995–96.) Höganäs: Bra Böcker.

Nikula, Henrik 1986: *Dependensgrammatik*. Malmö.

Nyord i svenskan = *Nyord i svenskan från 40-tal till 80-tal*. 1986. Utg. av Svenska språknämnden. Sthlm: Esselte Studium.

SOB = *Svensk ordbok*. 1986. Utarbetad vid Språkdata Göteborgs universitet. Sthlm: Esselte Studium.

Svenska ord = *Svenska ord — med uttal och förklaringar*. 1992. 2. uppl. Sthlm: Norstedts.

Söderwall, Knut Fredrik 1884–1918: *Ordbok öfver svenska medeltidsspråket* 1–3. Lund: Berlingska.