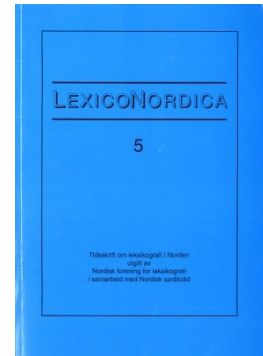


# LexicoNordica

Titel: Ordboksstruktur och etymologisk information  
Forfatter: Bo Svensén  
Kilde: LexicoNordica 5, 1998, s. 149-170  
URL: <http://ojs.statsbiblioteket.dk/index.php/lexn/issue/archive>



© LexicoNordica og forfatterne

## 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 LexicoNordica (1-16) 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.

*Bo Svensén*

## Ordboksstruktur och etymologisk information

Most monolingual general dictionaries can be described as basically synchronic, but some of them also include a diachronic component providing etymological information. This contribution gives a survey of different aspects of dictionary structure relevant for the presentation of etymological information in monolingual dictionaries, applying the conceptual and terminological framework used in *Nordisk leksikografisk ordbok* [Nordic Lexicographic Dictionary].

### 1. Inledning

En enspråkig allmänordbok är normalt en synkronisk ordbok, dvs. den beskriver språkets tillstånd vid en bestämd tid. I de flesta synkroniska enspråkiga allmänordböcker är denna tid identisk med ordbokens samtid, dvs. ordboken är vad man kallar en synkronisk samtidsordbok. Vissa sådana ordböcker för allmänt bruk avviker från den synkroniska principen genom att de innehåller en diakronisk komponent i form av en **etymologiangivelse**.

Det har på det hela taget inte publicerats särskilt många specialstudier rörande etymologisk information i enspråkiga ordböcker. Drysdale (1979) diskuterar några allmänna frågor, delvis i anslutning till Heller (1965), vars uppsats tar sikte på "five major flaws in the etymological practice of modern dictionaries". Drysdale (1989) ger en kortfattad översikt av handboks-karaktär, medan Seebold (1982) jämför ett antal tyska standardordböcker. Ingen av dessa, möjligen med undantag för Seebold, behandlar dock ämnet särskilt systematiskt.

Presentationen av etymologisk information är på många sätt en intressant fråga från strukturell synpunkt, och det gäller då såväl etymologiangivelsens plats och funktion i ordboksstrukturen som själva angivelsens interna struktur. Många av de begrepp som behandlas i *Nordisk leksikografisk ordbok* (NLO 1997) avser just ordbokens struktur, och denna uppsats är tänkt att samtidigt utgöra en demonstration av hur man med hjälp av NLO:s begreppsapparat kan beskriva strukturen hos en enskild informationstyp.

### 2. Adresseringsstruktur

Ett textelement i en ordboksartikel har i en given situation en av två olika funktioner: antingen utgör det en utsaga om ett annat textelement, eller också är det självt föremål för en sådan utsaga [1]. I exemplet utgörs själva utsagan av "fra. *sobre*, lat. *sobrius* nykter", och det som utsagan handlar om är lemmat "**sober**".

[1] **sober** ... [fra. *sobre*, lat. *sobrius* nykter]

Kombinationen av två textelement som står i en sådan relation till varandra kallas i NLO en **bearbetningsenhet**. Det är ofta fruktbart att betrakta förhållandet mellan bearbetningsenhetens två delar som relationen mellan ett **tema** och ett **rema**, där i detta fall "**sober**" är tema och "fra. *sobre*, lat. *sobrius* nykter" är rema. Det som i en sådan relation utgör rema kallas **angivelse** (i exemplet en etymologiangivelse), och det som utgör tema kallas angivelsens **adress** [2]. Den vanligaste adressen för en angivelse är lemmat (som i exemplet), och i det här fallet säger man att det föreligger en **adressering** från etymologiangivelsen till lemmat.

adress	angivelse
tema	rema

[2] **sober** ... [fra. *sobre*, lat. *sobrius* nykter]

Ordningföljden och relationerna mellan adresser och angivelser kallas **adresseringsstruktur**. Etymologiangivelsens adressering är intimt förknippad med dess plats och funktion i ordboksstrukturen.

### 3. Etymologi och makrostruktur

Den struktur som är av störst betydelse för etymologiangivelsens adressering är den som brukar kallas **makrostruktur** och som i NLO definieras som ordningföljden och relationerna mellan de lemmen som ingår i ordbokens lemmalista. Den enklaste formen av makrostruktur är **glattalfabetisk makrostruktur**, där varje textblock (**artikelblock**) bara innehåller ett lemma och där alla lemmen i hela ordboken är uppförda i strängt alfabetisk ordning. Där är det som regel inte några problem med adresseringen: det föreligger en **explicit adressering** från etymologiangivelsen till det enda lemmat i artikelblocket [3].

[3] **fideikommiss** *-et, -er* (histtenl.) eiendom som ... [lat. 'betrodd (gods)'; av *fides* troskap, og perf. pt. av *committere* overdra] (Aschehoug og Gyldendal 1992)

Men det är inte helt oproblematiskt med adresseringen i alla ordböcker som har glattalfabetisk makrostruktur. Det förekommer exempelvis att man anger etymologin vid bara ett av flera besläktade ord och överlämnar åt användaren att räkna ut vilka andra lemman som etymologi-angivelsen möjligen kan tänkas vara tillämplig på [4]. I sådana fall skulle man möjligen kunna tala om en **implicit adressering**, avsedd enbart för den som vet vilka ord som är besläktade.

- [4] <sup>1</sup>**perfek't** adj. ...  
 HIST.: sedan 1663; av lat. *perfectus* 'färdiggjord; fulländad', till **per** och *facere* 'göra'; jfr **parfait**

---

**perfektion** ...

HIST.: sedan 1651

**perfektionism** ...

HIST.: sedan 1947

**perfektionist** ...

HIST.: sedan 1950-talet

(NEO 1995–96)

Det är också vanligt med makrostrukturer där det kan förekomma flera lemman i ett och samma artikelblock. Det första lemmat i varje sådant artikelblock kallas i NLO **ingångslemma**, de övriga kallas **sublemman**. Om därvid samtliga lemman i ordboken är sorterade strängt alfabetiskt, sägs ordboken ha en **nischalfabetisk makrostruktur** [5].

- [5] **drag draga dragare**  
**dragdjur**  
**dragé**  
**dragen dragfri**  
**dragg dragga**  
**dragig dragkamp dragkedja dragkärra**  
**dragningskraft dragningslista**  
**dragon**

Om det i ordboken görs avsteg från den strängt alfabetiska ordningen på så sätt att det bara är ingångslemman som obligatoriskt är sorterade strängt alfabetiskt, sägs ordboken ha en **nästalfabetisk makrostruktur** [6].

- [6] **drag draga dragare dragdjur dragen**  
**dragfri dragig dragkamp dragkedja**  
**dragkärra dragningskraft dragningslista**  
**dragé**

**dragg dragga**  
**dragon**

Bakom båda strukturtyperna ligger åtminstone delvis en pedagogisk avsikt, nämligen att visa vilka ord som är etymologiskt besläktade. I princip bör det då inte bli någon konflikt mellan makrostrukturen och etymologistrukturen. När det gäller den etymologiska informationens utformning har lexikografen två vägar att gå, beroende på hur pass noggrant han vill redovisa etymologin för vart och ett av de lemman som ingår i artikelblocket. Vid noggrann redovisning måste som regel varje lemma i blocket förse med etymologiangivelse. I sådana fall finns det ingen principiell skillnad gentemot den glattalfabetiska principen när det gäller adresseringen: varje etymologiangivelse är explicit adresserad till ett enda lemma [7].

- [7] **ambiva'lens** ... [fra nylat. *ambivalentia* afl. af latin *ambivalens*, se *ambivalent*] – **ambiva'lent** ... [fra latin *ambivalens* (gen. *-valentis*) med doppelt værd ...]  
(Politikens Nudansk ordbog 1995)

Men i många fall vill man antingen spara utrymme eller anser att användarna inte är betjänta av detaljerad etymologisk information vid varje lemma. Då brukar artikelblocket få en enda etymologiangivelse, som då kan sägas vara adresserad till samtliga lemman i blocket. Den fungerar som ett slags minsta gemensam nämnare, och det är uppenbart att precisionen inte kan bli lika stor i sådana fall [8].

- [8] **mo5tor... motori5k ... motorisera ... moto5risk ... motori5m ... motori5st** ... [lat. *movēre* (perf. part. *mo5tus*) sätta i rörelse. Jfr *m o t i o n*, *m o t i v*]  
(Svenska och främmande ord 1974)

Som ett mellanting mellan [7] och [8] kan man betrakta en adresseringsstruktur där etymologiangivelsen visserligen är gemensam för hela textblocket men där olika delar av angivelsen är adresserade till olika lemman i blocket [9].

- [9] **stëreo-** ... **stereofonisk** ... **stereometri5** ... **stereoskop** ... [grek. *stereo5s* hård, fast + *fonë* ljud, *mëtron* mått, *skopëin* skåda ...]  
(Svenska och främmande ord 1974)

Man kanske väntar sig att etymologiangivelsen, om den skall vara adresserad till samtliga lemman i textblocket, skall ligga sist i detta.

Men så är det långt ifrån alltid, och i sådana fall måste man rimligtvis tänka sig etymologiangivelsen som adresserad även till sublemmana [10].

- [10] **CROIRE** [krwar] v. tr. (lat. *credere*; 980) ... → **croyable** adj.  
 (v. 1100) ... → **croyant, e** adj. et n. (1190) ... → **croyance** n.  
 f. (1361) ... → **incroyable** adj. (1490) ... → **incroyablement**  
 adv. (1495) ... → **incroyant, e** adj. et n. (1890) ...  
 → **incroyance** n. f. (1836) ...

(Lexis 1989)

Till den bristande precisionen i den citerade etymologiangivelsen bidrar också den i fransk lexikografi vanliga metoden med s.k. **regroupement**, som bl.a. innebär att det i ett artikelblock kan ingå etymologiskt besläktade avledningar och sammansättningar även från andra delar av alfabetet.

En variant på denna struktur föreligger i sådana artikelblock där ingångslemmat får en fullständig redovisning men där sublemmana enbart beskrivs med avseende på form och ordklassstillhörighet. Sublemmana är som regel avledningar av ingångslemmat, och deras underordnade ställning i förhållande till detta framgår ofta av att de är satta med mindre teckengrad. I ett sådant fall måste den etymologiangivelse som avslutar beskrivningen av ingångslemmat rimligen uppfattas som adresserad även till sublemmana [11].<sup>1</sup>

- [11] **flu+id** ('flu:ɪd) *n.* **1.** a substance ... – *adj.* **2.** capable of ...  
**5.** ... [C15: from Latin *fluidus*, from *fluere* to flow] – 'flu+id+al  
*adj.* – flu+'id+i-ty or 'flu+id+ness *n.* – 'flu+id+ly or 'flu+id+al+ly  
*adv.*

(Collins 1979)

#### 4. Etymologi och förhållandet makrostruktur–mikrostruktur

Liksom makrostruktur avser ordningsföljd och relationer mellan lemman, avser **mikrostruktur** ordningsföljd och relationer mellan de

---

<sup>1</sup> Strukturtypen är ganska vanlig i anglosaxiska ordböcker, och en subartikel av detta slag kallas där **run-on entry** (se t.ex. Collins 1979:xvii). Vi har ingen bra nordisk term för detta begrepp. Visserligen förekommer *run-on entry* i NLO, men termen används där som engelsk ekvivalent till *strekartikkel*, vilket är missvisande: en strekartikkel inleds av ett lemma som innehåller en representationssymbol (ett tilde, ett bindestreck e.d.), och som framgår av exemplet kan en *run-on entry* även inledas av ett fullständigt utskrivet lemma.

angivelser som är adresserade till ett lemma. Det innebär att homonymi manifesterar sig i makrostrukturen och polysemi i mikrostrukturen, och den princip man i en viss ordbok har valt att följa när det gäller förhållandet homonymi–polysemi har därför stor betydelse för etymologistrukturen.

Traditionellt har man definierat homonymi efter språkhistoriska kriterier. Det har inneburit att likalydande ordformer med olika betydelse och olika ursprung har betraktats som homonymer och i ordböcker presenterats i olika artiklar. Med detta historiska synsätt blir det normalt inte någon konflikt mellan etymologistrukturen och ordboksstrukturen i övrigt [12].

- [12] **sky I** subst. ...: *der er ikke en s. på himlen ...* [fællesnord. ...; besl. m. *skjul, skygge I–II ...*]  
**sky II** subst. ...: (stivnet) kødsaft ... [20. årh.; fra fransk *jus ...*]  
 (Politikens Nudansk ordbog 1995)

Annorlunda blir det när man lämnar det historiska betraktelsesättet och låter andra kriterier bestämma ordboksstrukturen. Den i fransk lexikografi vanliga metoden med s.k. **dégrouperment** bygger snarast på semantiska kriterier och innebär att likalydande ordformer med samma ursprung och i övrigt samma formella egenskaper men med olika (huvud)betydelse behandlas som homonymer, inte som delbetydelser hos samma ord. Mikrostrukturen avlastas, och makrostrukturen blir mer komplicerad [13]. I exemplet ser man på hänvisningarna från **2** till **1** och från **3** till **1** hur det har uppstått en diskrepans mellan det etymologiska perspektivet och ordboksstrukturen.<sup>2</sup>

- [13] **1. COLONNE** ... n. f. (lat. *columna*; v. 1100) **1.** Support vertical de forme cylindrique ...  
**2. COLONNE** ... n. f. (de *colonne* 1; 1660) Alignement vertical de chiffres ...  
**3. COLONNE** ... n. f. (de *colonne* 1; 1680) **1.** Alignement de personnes les unes derrière les autres ...  
 (Lexis 1989)

Det mest extrema fallet av diskrepans mellan etymologistruktur och ordboksstruktur är förmodligen det som representeras av *Nationalencyklopedins ordbok* (NEO 1995–96). I NEO har man tillämpat den

<sup>2</sup> Det är dock att märka, att hänvisningarna under **2** och **3** inte lyder "voir *colonne* 1" utan "de *colonne* 1". Det innebär att man här får veta något om ordets betydelseutveckling, nämligen att betydelsena **2** och **3** har utvecklats ur betydelse **1**.

s.k. lemma-lexemmodellen, som enkelt uttryckt kan sägas innebära att likalydande ordformer som tillhör olika ordklasser eller i övrigt har olika formella egenskaper betraktas som homonymer och presenteras i separata artiklar. I strukturellt avseende bortser man helt från den historiska aspekten [14].

- [14] **1byrå** [by´ - el. by` -] subst. ~n ~ar ...  
 • en (finare) förvaringsmöbel ...  
 HIST.: sedan 1788; av fra. *bureau* '(tygklädd) skrivpulpet; skrivrum; kontor', urspr. 'grovt tyg', till *bure* 'vadmal'
- 2byrå** [by´ - el. by` -] subst. ~n ~er [byrå´er]  
 • avdelning som sköter viss typ av tjänster inom offentligt organ ...  
 HIST.: sedan 1801; av samma urspr. som **1byrå**  
 (NEO 1995–96)

Det omvända fallet gäller synen på vad polysemi är. Enligt lemma-lexemmodellen innebär polysemi olikhet i semantiskt avseende mellan likalydande ordformer som tillhör samma ordklass och i övrigt har samma formella egenskaper. Även här bortses i strukturen helt från det historiska. Makrostrukturen avlastas, medan mikrostrukturen blir mer komplicerad. Resultatet blir i många fall ganska överraskande kombinationer av betydelseenheter [15].

- [15] **skall`a** verb ~de ~t ...  
**1** ljuda starkt och (ngt) ekande ...: *slagorden ~de* ...  
 HIST.: sedan 1682; till **skall**  
**2** hårt träffa ngn's huvud med den egna skallen ...: *han ~de sönder motståndarens näsa* ...  
 HIST.: sedan 1913; till **skalle**  
 (NEO 1995–96)

Detta innebär att etymologin genomgående måste anges på delbetydelsenivå, och när det gäller den stora mängd av fall där delbetydelseerna faktiskt *har* samma ursprung anger man etymologin under en av betydelseerna och hänvisar dit från de övriga [16].

- [16] **lätt** adj., neutr. ~  
**1** som har ringa vikt ...  
 HIST.: före 1520; fornsv. *lätter* 'lätt; lättsinnig; lättfärdig'; gemens. germ. ord, besl. med bl.a. grek. *elakhys* 'liten; obetydlig; låg'; jfr **lunga**  
 ---  
**2** som sker eller används med ringa kraft ...



HIST.: före 1520; se **lätt 1**

---

**3** sorglös och obekymrad ...

HIST.: före 1520; se **lätt 1**

---

**4** som inte orsakar några större problem ...

HIST.: före 1520; se **lätt 1**

---

**5** som är av ringa omfattning ...

HIST.: sedan 1751; se **lätt 1**

(NEO 1995–96)

## 5. Etymologi och mikrostruktur

Förhållandet mellan etymologistruktur och mikrostruktur är förhållandevis okomplicerat. Angivelsen kan placeras i stort sett på två ställen i artikeln, sist [17] eller först [18].

- [17] **fideikommi5ss** *-et, -er* (hist., utenl.) eiendom som ikke kan avhendes, men som må gå i arv etter visse fastsatte regler [lat. 'betrodd (gods)'; av *fides* troskap, og perf. pt. av *committere* overdra]

(Aschehoug og Gyldendal 1992)

- [18] **DISCRIMINATION** ... n. f. – 1870 ; lat. *discriminatio* «séparation» **1.** PSYCHOL. Action de distinguer ... **2.** COUR. Le fait de séparer un groupe social des autres ...

(Petit Robert 1993)

Det normala är ju att etymologiangivelsen är adresserad till lemmat. Till det normala hör också att ett polysemt lemma har en enda etymologiangivelse och att den är adresserad till lemmat i alla dess delbetydelser.<sup>3</sup> Men det finns också fall där en del av en för hela lemmat gemensam etymologiangivelse är specialadresserad till en delaspekt av lemmat, exempelvis en delbetydelse [19] eller en enstaka ordförbindelse [20].

- [19] **konso5ll** *-en, -er* **1** fremspring på vegg ... **2** vegghylle (under

---

<sup>3</sup> Den etymologistruktur som blir resultatet av en tillämpning av lemma-lexemmodellen utgör här alltså ett undantag, se [16] ovan.

speil) **3** (mindre) veggmøbel **4** enhet ... hvor instrumenter og brytere ... er samlet [fr., bet. 4 fra eng.]

(Aschehoug og Gyldendal 1992)

[20] **rysk** ...

IDIOM: ...; *är du alldeles ~?* är du alldeles tokig?

HIST.: före 1520; fornsv. *ryzker*; till **ryss**; (*är du alldeles*) *rysk*

trol. urspr. till <sup>2</sup>**ruska**

(NEO 1995–96)

## 6. Etymologi och fördelningsstruktur

Ett annat strukturbegrepp som beskrivs i NLO är **fördelningsstruktur**. Med det menas strukturen hos den information som kan ges på olika ställen i ordboken: (a) i de enskilda artiklarna, (b) i översiktsartiklar och (c) i en självständig komponent i utanförtexten, dvs. före eller efter ordbokens huvudlemmalista. Exempler hittills har alla gällt enskilda artiklar, som är det vanligaste fallet.

Fallet med etymologisk information i översiktsartiklar kan sägas föreligga när exempelvis ett ofta förekommande ordbildningselement förklaras i en egen artikel med egen etymologiangivelse, till vilken hänvisas från etymologiangivelsen i ett stort antal enskilda artiklar [21].

[21] **-ist** brukt til å danne subst. for **1** person med en viss opptreden, levemåte ...: *antagonist*, ... *alkoholist* ... **2** dyrker av en viss vitenskap: *germanist*, *nordist* ... [fr. *-iste*, lat. *-ista*, gr. *-istes*]

**mona5rk** ... [gr.] **monarki5** ... [gr., av *monos* alene, og *arkhe* regjering] **mona5rkisk** ... [se *-isk*] **monarki5st** ... [se *-ist*] **monarki5stisk** ... [se *-isk*]

(Aschehoug og Gyldendal 1992)

Ibland kan det förekomma att ordbildningselementet visserligen förklaras i en egen artikel men att användaren förväntas klara sig utan hänvisning dit från de enskilda artiklarna [22].

[22] **pan-**, **pant-**, **panto-**, in composition, all [Gr. *pÇs*, *pÇsa*, *pÇn*, gen. *pantos*, *pÇsis*, *pantos*.]

**panorama** ... **panoramic** ... [Gr. *horÇma*, a view, from *horaein*, to see.]

(Chambers 1977)

Att det ges etymologisk information i ordbokens utanförtext är inte så vanligt, men bland de nordiska ordböckerna förekommer det exempelvis i *Politikens Nudansk ordbog* (15:e uppl. 1995), där det i vissa artiklar hänvisas till en språkhistorisk översikt. Men den verkliga klassikern på det här området är *American Heritage Dictionary* (1969). Ordboken är försedd med ett alfabetiskt appendix med cirka 1.500 uppslagsord bestående av indoeuropeiska rötter, med riklig exemplifiering av varje rots förekomst i olika indoeuropeiska språk. Till denna komponent i utanförtexten hänvisas från ett stort antal artiklar i ordboken [23].

[23] **con•ceal** ... [Middle English *concelen*, from Old French *conceler*, from Latin *concilĀre* : *com-* (intensive) + *cilĀre*, to hide (see **kel-4** in Appendix\*.)]

**kel-4** . To cover, conceal, save. **1.** Germanic *haljǫ*, the underworld (< “concealed place”), in: **a.** Old English *hell*, hell: HELL; **b.** Old Norse *Hel*, underworld, goddess of death: HEL. **2.** Germanic *\*halljǫ*, covered place, hall, in: **a.** Old English *heall*, hall: HALL; **b.** Old Norse *höll*, hall: VALHALLA; ... (American Heritage Dictionary 1969)

## 7. Etymologi och angivelsestruktur

Med **angivelsestruktur** menas här angivelsens interna struktur och de upplysningar som ges där. Även inom en angivelse kan man, så att säga på mikronivå, tala om tema och rema, adress och angivelse, och där är etymologiangivelsen något speciell, eftersom den kan bestå av hela kedjor av adresser och angivelser. När det gäller etymologisk information är det förmodligen en omöjlighet att upprätta en typologi över olika slag av angivelsestruktur, eftersom variationsmöjligheterna är i det närmaste oändliga. Däremot kan man göra några principiella iakttagelser och urskilja vissa slag av komponenter och beskriva relationerna mellan dem.

### 7.1. Utvecklingsledens struktur och relationer

I det enklaste fallet finner man som regel tre olika komponenter: uppgift om ursprungsordets språktillhörighet, uppgift om ursprungsordets form, uppgift om ursprungsordets betydelse [24].

[24] **plekter** ... [gr. *plektron* liten stav]

(Aschehoug og Gyldendal 1992)

Den här sekvensen kan vid behov upprepas flera gånger inom en och samma etymologiangivelse, och för varje gång förflyttar man sig bakåt i tiden mot ett äldre språkstadium, antingen inom samma språk eller i ett annat språk [25].

[25] **kosty'm** ...

**1** ...

HIST.: sedan 1873; av fra. *costume* med samma bet.; av ita. *costume* 'vana, bruk'; av lat. *consuetudo* 'vana, sed, bruk'; ...

(NEO 1995–96)

Hur många gånger sekvensen upprepas beror dels på hur många utvecklingsled som i verkligheten går att urskilja, dels på hur detaljerat lexikografen anser sig behöva redovisa ordets historia. Man kan hoppa över ett eller ett par utvecklingsled, som bedöms som mindre viktiga, för att snabbare nå fram till det äldsta kända ursprungsordet [26]; i exemplet har man valt att utelämna den form i vilken ordet närmast inlånades i svenskan, det franska *céramique*.

[26] **kerami'k** ...

HIST.: sedan 1884; bildn. till grek. *keramos* 'krukmakarlera; lervas; tegel'

(NEO 1995–96)

En annan möjlighet är att göra halt innan man har nått så långt tillbaka som till det äldsta kända ursprungsordet [27]; här har man inte gått så långt som till det latinska ordet *mola* 'møllestein'.

[27] **mola5r** I *-en, -er* kinntann ... [lat. (*dens*) *molaris* (tann) som minner om en møllestein]

(Aschehoug og Gyldendal 1992)

Inom varje utvecklingsled kan i sin tur en eller två av komponenterna utelämnas. Språkets namn utelämnas naturligtvis om utvecklingsledet i fråga inte innebär byte av språk [28], i övrigt är det oftast ursprungsordets betydelse som utelämnas [29].

[28] **tinktur** ...

HIST.: sedan 1675; av lat. *tinctora* 'färgning', till *tingere* 'färga'

...

(NEO 1995–96)

- [29] **nukle'ar** ... [genom eng. *nuclear*, fransk *nucléaire* fra nylat. *nuclearius* afl. af latin *nucleus* kerne ...]  
(Politikens Nudansk ordbog 1995)

Ibland utelämnas i stället ursprungsordets form, och användaren får då inte veta om denna form sammanfaller med lemmaformen (som i [30]) eller om den inte gör det (som i [31]).<sup>4</sup>

- [30] **modus** ... [lat. 'mål, måte']  
(Aschehoug og Gyldendal 1992)

- [31] **dialekt** ... [gr. 'samtale, språk']  
(Aschehoug og Gyldendal 1992)

Ibland nöjer man sig med att bara redovisa den väg som ordet har tagit och utelämnar såväl ursprungsordets form som dess betydelse [32].

- [32] **hybrid** ... [fra., lat., grek.]  
(Svenska och främmande ord 1974)

De redovisade utvecklingsleden står i ett tema-remaförhållande till varandra. Led nummer 1 utgör alltså adress för led nummer 2, led nummer 2 är adress för led nummer 3 osv. [33].

- [33] **direktör** ... [fra fransk *directeur* af middellat. *director* afl. af latin *dirigere* ...]  
(Politikens Nudansk ordbog 1995)

Inom varje led föreligger också ett tema-remaförhållande mellan komponenterna. Det är förmodligen rimligt att betrakta ursprungsordets form som adress och de andra två komponenterna som angivelser adresserade till denna [34]. Exemplet bör därför utläsas: "Det norska ordet *plekter* kommer av ett ord med formen *plektron* som hör hemma i grekiskan och har betydelsen 'liten stav'."

- [34] **plekter** ... [gr. *plektron* liten stav]  
(Aschehoug og Gyldendal 1992)

På liknande sätt bör man också uppfatta ett utvecklingsled där någon av komponenterna har utelämnats [35]. Exemplet skall utläsas: "Det

---

<sup>4</sup> Ett sätt att med enkla medel eliminera den svårigheten beskrivs hos Svensén (1987:185).

norska ordet *dialekt* kommer av ett ord vars form ej anges men som hör hemma i grekiskan och har betydelsen 'samtale, språk'."

[35] **dialekt** ... [gr. 'samtale, språk']

(Aschehoug og Gyldendal 1992)

## 7.2. Sambandet mellan grundordets och lemmats betydelse

När slutligen det äldsta utvecklingsledet presenteras, har man i många fall kommit fram till ett ord (vi kan kalla det "grundord") som skiljer sig ganska mycket från lemmat till såväl form som betydelse. Efter det äldsta utvecklingsledet kan det därför förekomma ytterligare en typ av angivelse, som oftast är av encyklopedisk karaktär och som har till uppgift att visa sambandet mellan grundordets betydelse och lemmats betydelse [36]. För att klargöra detta samband kan man använda flera olika tekniker. I exemplet sker det genom att man återknyter till ett begrepp som redan är känt från grundordets betydelseangivelse ("sadel") och sätter det i relation till ett nyintroducerat begrepp ("tillfälligt nattläger"), som har till uppgift att få läsaren att associera tillbaka till lemmats betydelseangivelse ('barn fött utom äktenskapet'), varigenom det fungerar som förmedlande länk mellan grundordets betydelse och lemmats betydelse.

[36] **bastard** ...

---

2 barn fött utom äktenskapet ...

HIST.: sedan 1636; ur fornfra. *bastard*, eg. 'sadelbarn', till medeltidslat. *bastum* 'sadel' (sadeln användes som tillfälligt nattläger)

(NEO 1995–96)<sup>5</sup>

I mikrostrukturer där etymologiangivelsen är placerad i artikelns början förklaras alltså detta samband medan lemmats aktuella betydelse ännu är okänd för läsaren, vilket riskerar att åtminstone temporärt verka förvirrande och göra förklaringen något poänglös [37].

[37] **Karakt**, das ... [frz. carat ... griech. kerátion = Hörnchen, Same des Johannisbrotbaumes; diese Samenkörner wurden früher zum Wiegen von Gold u. Edelsteinen benutzt]: **1. Einheit für die**

<sup>5</sup> Den för hela artikeln **bastard** gemensamma etymologiangivelsen står i NEO under **1** men visas här för överskådlighetens skull under **2**.

*Gewichtsbestimmung von Edelsteinen ... 2. Einheit der ...  
Skala für den Goldgehalt einer Legierung ...*  
(Duden 1983)

### 7.3. Hänvisningarna och deras funktion

Etymologiangivelsen innehåller ofta också hänvisningar till besläktade ord, dels i samma språk som ordbokens objektspråk, dels i andra språk. Sådana uppgifter verkar kunna ha två olika syften.

Det ena syftet är rent pedagogiskt: att främja ordförståelsen och i bästa fall också ge läsaren en angenäm aha-upplevelse. Då fungerar hänvisningen normalt som en mer eller mindre löst påhängd (eller parentetiskt inskjuten) upplysning, som inte utgör ett nödvändigt led i själva etymologiangivelsen. Det är framför allt här som det brukar anges släktingar även i andra språk än ordbokens objektspråk [38]. (De ordboksexterna hänvisningarna står i exemplet med kursivstil, de ordboksinterna med spärrad stil.)

[38] **merkanti5l** ... [fra., lat. *merca5ri* driva handel. Jfr *Mercurius*, handelns gud; fra. *mercredi* onsdag (marknadsdag); sv. *marknad*. Jfr *k o m m e r s*, *m a r k e t e n t e r i*]  
(Svenska och främmande ord 1974)

Hit hör också det specialfall då man med utgångspunkt från ett gemensamt fornspråk anger motsvarigheter i sidoordnade språk [39].

[39] **hed** ... [fællesgerm.: oldnord. *heitr*, eng. *hot*, tysk *heiss*; ...]  
(Politikens Nudansk ordbog 1995)

Det andra syftet med uppgifter om besläktade ord är att de skall fungera som ersättare för utvecklingsled. En sådan hänvisning är inte löst påhängd eller parentetiskt inskjuten utan är mer integrerad i själva etymologiangivelsen. Ofta är den snarare av morfologisk än etymologisk karaktär och går från en avledning till det ord som denna är avledd av eller från en sammansättning till dennas olika led [40].

[40] **konkurrence** ... [fra fransk *concurrance* afl. af latin *concurrere*, se *konkurrere*]  
(Politikens Nudansk ordbog 1995)

I många fall utgör hänvisningen hela etymologiangivelsen [41], och det förekommer också mellanformer [42].

[41] **sto`rslagen ...**

HIST.: sedan 1882; till <sup>1</sup>stor och <sup>2</sup>slå

(NEO 1995–96)

[42] **hypotek ...** [av *hypo-* og gr. *theke* det lagte; ...]

(Aschehoug og Gyldendal 1992)

## 8. Etymologi och åtkomststruktur

Med **åtkomststruktur** menas strukturen hos de indikatorer som leder ordboksanvändaren fram till önskad information. Man skiljer mellan **yttre åtkomststruktur** och **inre åtkomststruktur**: den yttre leder fram till det sökta lemmat och den inre till sökt information inom artikeln. Den yttre åtkomststrukturen blir knappast aktuell när det gäller etymologisk information, medan däremot den inre ger sig till känna på flera sätt.

En indikator som leder användaren fram till önskad information kallas i NLO **strukturindikator**. Man skiljer mellan **typografiska strukturindikatorer** och **icke-typografiska strukturindikatorer**. De förra består i att vissa teckensnitt (t.ex. Times, Helvetica), snittvarianter (t.ex. kursivstil, fetstil), teckengrader etc. används för att ange vissa funktioner hos textelement, medan de icke-typografiska består av tecken, i regel skiljetecken och symboler av olika slag.

Strukturindikatorerna kan när det gäller etymologisk information vara verksamma på olika nivåer. På den högsta nivån – i mikrostrukturen – brukar angivelsens början och slut normalt anges med hjälp av något slags parenteser: hakparenteser, runda parenteser eller spetsparenteser. Alternativt markeras angivelsens början mer explicit genom något slags symbol eller rubrik, ibland kombinerad med radskifte [43].

[43] **lam`pa ...**

HIST.: före 1520; fornsv. *lampa*; av lågty. *lampe* med samma bet.; av grek. *lampas* 'fackla', till *lampein* 'lysa'; jfr **lanterna**

(NEO 1995–96)

På närmast lägre nivå – i angivelsestrukturen – förekommer i exemplet två icke-typografiska strukturindikatorer: semikolon för att markera gräns mellan utvecklingsled som tillhör olika språk och komma för att markera gräns mellan utvecklingsled inom samma språk.

På den lägsta nivån – inom det enskilda utvecklingsledet – används i exemplet både icke-typografiska och typografiska strukturindikatorer. Enkla citationstecken anger att de textelement som de inramar utgör en betydelseangivelse adresserad till ursprungsordets form inom samma



utvecklingsled. Ursprungsordets form kännetecknas i sin tur av snittvarianten kursivstil, medan rak stil utan citationstecken anger metaspråk och fet stil med mindre teckengrad betecknar adressen för en ordboksintern hänvisning.

## 9. Etymologi och textförtätning

**Textförtätning** (eller **kondensering**) innebär att ett textelement förkortas med hjälp av ett visst förfarande. Det ursprungliga textelementet kallas i NLO **fulltext**, och resultatet av textförtätningen benämns **kondensat**. Fulltexten kan vara antingen faktiskt existerande eller endast potentiell (Wolski 1989:959 f.). I det förra fallet sker textförtätningen genom bearbetning av en föreliggande textversion till ett mer kortfattat framställningssätt; i det senare fallet utarbetas den kortfattade versionen mer eller mindre direkt, i regel genom ett standardiserat förfarande enligt riktlinjerna i en redaktionshandbok. Det är det sistnämnda fallet som skall diskuteras här.

Det är viktigt att påpeka, att textförtätning i princip inte får innebära informationsförlust: kondensatet skall ge samma information som fulltexten fast mer kortfattat. Användaren skall utifrån kondensatet kunna åstadkomma en användbar **dekondenserad text** – dvs. rekonstruera en tänkbar fulltext – enbart med hjälp av sin kunskap om de konventioner som gäller för ordbokens presentationssätt.

Det finns olika slag av textförtätning förfaranden (jfr Wolski 1989:961 ff.), av vilka dock inte alla blir aktuella när det gäller etymologisk information. Liksom strukturindikatorerna kan de beskrivas som verksamma på olika nivåer.

Om en angivelse ingår i flera bearbetningsenheter, dvs. har flera adresser, måste detta betraktas som resultatet av ett slags textförtätning; i den tänkta fulltexten måste ju varje adress antas ha sin egen angivelse. På makrostrukturnivå är detta möjligt i nisch- och nästalfabetiska makrostrukturer (jfr ovan avsnitt 3), där det ju kan förekomma att en etymologiangivelse är adresserad till flera lemman i ett och samma artikelblock (jfr [8] ovan). Den s.k. implicita adresseringen (jfr [4]) kan däremot inte betraktas som ett slags textförtätning, eftersom det i detta fall är omöjligt att åstadkomma en dekondenserad text där varje lemma har en egen etymologiangivelse.

I princip samma typ av textförtätning förekommer även på mikrostrukturnivå, nämligen då en och samma angivelse är adresserad till samtliga delbetydelser hos ett polysemt lemma. För att detta skall bli möjligt måste angivelsen placeras så att säga "på neutral mark", utanför

delbetydelsena. Här kommer man kanske i första hand att tänka på sådana formella och grammatiska angivelser som vanligen förekommer i artikelns början (uttal, böjning, ordklassstillhörighet), men det är självklart att även etymologiangivelsen måste räknas hit, antingen den är placerad före delbetydelsena (jfr [18]) eller efter dessa (jfr [19]). Denna typ av textförtätning är dock en omöjlighet i ordböcker som tillämpar lemma-lexemmodellen (jfr [15]).

Övriga typer av textförtätning är verksamma på angivelsenivå eller inom enskilda utvecklingsled. Två av dem blir sällan aktuella när det gäller etymologisk information:

- (a) beläggförkortning (redaktionell förkortning av ett autentiskt exempel),
- (b) ersättande av (en del av) lemmatecknet med en representations-symbol (tilde, bindestreck e.d.) i exempel, böjningsuppgifter o.d.

Däremot är det i etymologiangivelser vanligt med följande typer:

- (c) bruk av (standardiserade) förkortningar,
- (d) utbyte av textelement mot angivelsesymboler (dvs. symboler som själva utgör angivelser),
- (e) förkortad presentation av formuleringsalternativ med hjälp av parentestecken m.m., främst i språkprov och betydelseangivelser,
- (f) utelämnning av textelement.

Ofta förekommer alla fyra slagen av textförtätning i en och samma etymologiangivelse [44].

[44] **Ho|ri|zo5nt**, der; -[e]s, -e [lat. horizon (Gen.: horizon-tis) < griech. horízōn (křklos) = begrenzend(er Kreis, Gesichtskreis), zu: horízein = begrenzen, zu: hóros = Grenze, Grenzstein, Ziel] ... (Duden 1983)

En tänkbar dekonserverad text (i svensk översättning) kunde i detta fall vara: "Uppslagsordet härstammar närmast från ett latinskt ord som har formen *horizon* (i genitiv *horizontis*), som i sin tur härstammar från ett grekiskt ord som har formen *horizon* och som uttalas med trycket på andra stavelsen och med lång vokal i sista stavelsen och har betydelsen 'begränsande'; *horizon* kan uppträda tillsammans med ett ord som har formen *kyklos* och som uttalas med trycket på första stavelsen, och hela uttrycket har då betydelsen 'begränsande krets, begränsande synkrets'; *horizon* är avlett av ett ord som har formen *horizein* och som uttalas med trycket på andra stavelsen och har betydelsen 'begränsa'; *horizein*

är avlett av ett ord som har formen *horos* och som uttalas med trycket på första stavelsen och har betydelsen 'gräns, gränssten, mål'."

Typiska standardiserade förkortningar (typ c) i etymologiangivelser är givetvis språktillhörighetsangivelserna, i exemplet representerade av "lat." och "griech."

Exempel på bruket av angivelsesymboler (typ d) är att vissa ursprungsord är försedda med prosodiska angivelser i form av dels accenttecken, som ersätter textelement med innebörden 'uttalas med trycket på denna stavelse', dels längdstreck över vokaler, som ersätter textelement med innebörden 'uttalas lång'.

I det först angivna grekiska utvecklingsledet förekommer förkortad presentation av formuleringsalternativ med hjälp av parentestecken (typ e) såväl i ursprungsordet som i dess betydelseangivelse.

Utelämning av textelement (typ f) har i exemplet skett dels i början ('Uppslagsordet härstammar närmast från'), dels vid de förkortade språktillhörighetsangivelserna ("lat." betyder inte enbart 'latin' utan snarare 'ett latinskt ord' e.d.). På liknande sätt har utelämning skett vid prepositionen "zu", där fulltexten måste antas vara 'som är avlett av' e.d.

Bruket av prepositioner i etymologiangivelser är för övrigt en intressant fråga, och vissa ordböcker har här utvecklat ett mer eller mindre komplicerat kodsysteem. I NEO har man gått ganska långt på den vägen: "av" före språktillhörighetsangivelsen anger att ordet har inlånats direkt från språket i fråga (jfr ovan [4]), "ur" att ordet ytterst härstammar från detta språk men att ett eller flera mellanliggande språk har utelämnats (jfr [36]), "efter" att ordet är ett översättningslån från språket i fråga, medan "till" anger en mer allmän relation (i regel avledning eller sammansättning; jfr [15]). Vid arvord används ingen preposition alls (jfr [16]).

Till de ovan beskrivna sex typerna av textförtätning på angivelsenivå borde egentligen läggas en sjunde, som inte nämns av Wolski (1989), nämligen

(g) bruk av strukturindikatorer.

Det nyss citerade exempel [44] är ganska ovanligt genom att det helt saknar typografiska strukturindikatorer (snittvarianten rak mager stil kan i det här fallet inte betraktas som en strukturindikator, eftersom den används till alla slags textelement oavsett deras funktion). Mest påfallande är kanske att man har avstått från det vanliga förfarandet att markera ursprungsordens form med kursivstil (motsvarande fulltextens 'som har formen'); i stället använder man den icke-typografiska strukturindikatorn kolon för att markera att ett språkprov följer ("Gen.:"

resp. "zu:").<sup>6</sup> Andra icke-typografiska strukturindikatorer som bidrar till textförtätningen är likhetstecknet (motsvarande fulltextens 'som har betydelsen') och tecknet < (motsvarande fulltextens 'som härstammar från').

Det finns ytterligare en företeelse som rätteligen också borde räknas som ett slags textförtätning men som inte heller den nämns av Wolski (1989), nämligen

(h) bruk av hänvisningar.

Att hänvisningar i många fall används för att spara utrymme är ett välkänt faktum (jfr Svensén 1987:188–192). När det gäller textförtätning intar de dessutom något av en särställning eftersom det då oftast blir fråga om att rekonstruera en sökt fulltext i tre steg. Man kan exempelvis tänka sig ett utgångsläge där en tysk ordbok i etymologiangivelsen till lemmat *Geschwister* hänvisar till *Schwester* [45].

[45] **Ge'schwi-ster** ... [< ahd. *giswestar*, Kollektivbildung zu ahd. *swester*, Ø *Schwester*; ...]

(Deutsches Wörterbuch 1980)

Första steget blir nu att dekondensera hänvisningen "Ø *Schwester*" till 'se vidare under uppslagsordet *Schwester*'. En sådan hänvisning skapar förväntningar om att det i etymologiangivelsen till *Schwester* skall ges information om ursprunget till det senast anförda ursprungsordet, det fornhögtyska *swester*, vilket också visar sig vara fallet [46].

[46] '**Schwe-ster** ... [< mhd., ahd. *swester* < got. *swistar* < idg. *\*suesor* „Schwester”]

(Deutsches Wörterbuch 1980)

Som ett andra steg i dekondenseringen får man nu i artikeln *Geschwister* i stället för hänvisningen "se vidare under uppslagsordet *Schwester*" insätta den utförligare information som ges under *Schwester*, nämligen "< got. *swistar* < idg. *\*suesor* „Schwester”". Det tredje steget blir att dekondensera den insatta texten, och resultatet blir "som härstammar från ett gotiskt ord med formen *swistar*, som i sin tur härstammar från en rekonstruerad indoeuropeisk form *suesor*, som betyder 'syster'".

Graden av textförtätning växlar självfallet från ordbok till ordbok, och olika typer av etymologiangivelser lämpar sig i olika hög grad för

---

<sup>6</sup> Dessutom måste användaren lära sig att ett ord som följer efter en standardiserad språktillhörighetsangivelse är ett språkprov ("lat. horizon", "griech. horízyn" etc.).

textförtätning. Avslutningsvis skall nu göras en jämförelse mellan två etymologiangivelser som ger ungefär samma information men som skiljer sig åt ganska mycket när det gäller graden av textförtätning [47, 48].

[47] e%*xàc't*<sup>2</sup> ... [f. L EX<sup>1</sup>(*igere act-* = *agere drive*)]  
(COD 1982)

[48] *exak't* ...  
HIST.: ...; av lat. *exactus* 'utförd; noggrann', till *exigere*  
'fullborda', till *ex-* och *agere* 'driva; handla'; ...  
(NEO 1995–96)

Den ordbok ur vilken det första exemplet är hämtat kan sägas göra skäl för sitt namn. En tänkbar dekonserverad text skulle i svensk översättning kunna lyda: "Uppslagsordet härstammar från ett latinskt verb vars infinitiv har formen *exigere* (se under uppslagsordet *ex<sup>1</sup>*) och vars supinstam har formen *exact-*, vilken är avledd av ett verb vars infinitiv har formen *agere* och betydelsen 'driva'." Kondensatet omfattar 35 tecken, vilket utgör 15 procent av antalet i den föreslagna fulltexten, som är 232.

NEO utmärks däremot av en förhållandevis låg grad av textförtätning, och en dekonserverad version av det senare exemplet skulle kunna lyda: "Uppslagsordet härstammar från ett latinskt adjektiv som har formen *exactus* och betydelsen 'utförd, noggrann' och som är avlett av ett verb med formen *exigere* och betydelsen 'fullborda', vilket är avlett av prefixet *ex-* (se under detta uppslagsord) och ett verb med formen *agere* och betydelsen 'driva, handla'." Kondensatet omfattar här 103 tecken, vilket utgör 34 procent av den tänkta fulltextens 307.

Den "metod" som här har använts kan självfallet inte göra anspråk på någon som helst exakthet, framför allt därför att man kan tänka sig ett nästan oändligt antal alternativa fulltexter. Men om två kondensat som ger ungefär samma information dekonserveras efter samma principer, borde resultatet åtminstone kunna bilda underlag för en grov jämförelse.

## Litteratur

### *Ordböcker*

American Heritage Dictionary 1969 = *The American Heritage Dictionary of the English Language*. Ed. W. Morris. New York: American Heritage Publishing Co. & Houghton Mifflin Co.

- Aschehoug og Gyldendal 1992 = *Aschehoug og Gyldendals Store norske ordbok*. Red. T. Guttu. 3. opplag. Oslo: Kunnskapsforlaget.
- Chambers 1977 = *Chambers Twentieth Century Dictionary*. Ed. A. M. Macdonald. Edinburgh: W. & R. Chambers.
- COD 1982 = *The Concise Oxford Dictionary of Current English*. 7th ed. Ed. J. B. Sykes. Oxford: Clarendon Press.
- Collins 1979 = *Collins Dictionary of the English Language*. Ed. P. Hanks. London & Glasgow: Collins.
- Deutsches Wörterbuch 1980 = G. Wahrig: *Deutsches Wörterbuch*. Mosaik Verlag.
- Duden 1983 = *DUDEN Deutsches Universalwörterbuch*. Hrsg. v. G. Drosdowski. Mannheim: Bibliographisches Institut.
- Lexis 1989 = *Dictionnaire de la langue française (Lexis)*. Paris: Larousse.
- NEO 1995–96 = *Nationalencyklopedins ordbok 1–3*. Höganäs: Bokför-  
laget Bra Böcker.
- NLO 1997 = H. Bergenholtz/I. Cantell/R.V. Fjeld/D. Gundersen/J.H. Jónsson/B. Svensén: *Nordisk leksikografisk ordbok*. Skrifter utgitt av Nordisk forening for leksikografi. Skrift nr. 4. Oslo: Universitetsforlaget.
- Petit Robert 1993 = *Le nouveau Petit Robert*. Ed. J. Rey-Debove/A. Rey. Paris: Dictionnaires Le Robert.
- Politikens Nudansk ordbog 1995 = *Politikens Nudansk ordbog & Sprogbrugsleksikon*. 15. udgave. Red. Ch. Becker-Christensen/P. Widell. København: Politikens Forlag.
- Svenska och främmande ord 1974 = E. Lindbäck/A. Milits/B. Svensén: *Svenska och främmande ord. Form, betydelse, ursprung*. Stockholm: Esselte Studium.

### Övrig litteratur

- Drysdale, P.D. 1979: Dictionary Etymologies: What? Why? And for Whom? I: *Papers of the Dictionary Society of North America 1979*. London, Ontario: Univ. of Western Ontario, 39–50.
- Drysdale, P.D. 1989: Etymological Information in the General Monolingual Dictionary. I: F.J. Hausmann/O. Reichmann/H.E. Wiegand/L. Zgusta (red.): *Wörterbücher, Dictionaries, Dictionnaires*. Berlin/New York: Walter de Gruyter, 525–530.
- Heller, L.G. 1965: Dictionary Etymology: Practice versus Theory. I: *American Speech* 40:113–119.

- Seebold, E. 1982: Die Erläuterung der Etymologie in den Wörterbüchern der deutschen Gegenwartssprache. I: H.E. Wiegand (red.): *Studien zur neuhochdeutschen Lexikographie II*. Hildesheim/New York: Olms, 189–223.
- Svensén, B. 1987: *Handbok i lexikografi. Principer och metoder i ordboksarbetet*. Stockholm: Esselte Studium och Tekniska nomenklaturcentralen.
- Wolski, W. 1989: Formen der Textverdichtung im allgemeinen einsprachigen Wörterbuch. I: F.J. Hausmann/O. Reichmann/H.E. Wiegand/L. Zgusta (red.): *Wörterbücher, Dictionaries, Dictionnaires*. Berlin/New York: Walter de Gruyter, 956–967.