

MEIJERBERGS ARKIV

FÖR

SVENSK ORDFORSKNING

UTGIVET

AV

STYRELSEN FÖR MEIJERBERGS INSTITUT

VID GÖTEBORGS UNIVERSITET

GENOM

ANNA HELGA HANNESDÓTTIR

42

GÖTEBORG 2016

FRAMTIDENS LEXIKOGRAFI

Rapport från ett symposium i Göteborg 5 oktober 2012

Ingår som del 42 i Meijerbergs arkiv för svensk ordforskning

© Författarna och
Meijerbergs institut för svensk etymologisk forskning
Göteborgs universitet
Box 200
SE-405 30 Göteborgs universitet
Omslagsbild: Andrey VP/Shutterstock

ISSN 0348-7741
ISBN 978-91-87850-01-1

Responstryck AB, Borås 2016
Sättning: Sven Lindström

Innehåll

<i>Lena Rogström</i> Förord.....	1
<i>Lars Borin</i> Lexikografi för maskiner och lexikografi för människor.....	9
<i>Anna Helga Hannesdóttir</i> Tvåspråkig lexikografi – ny teknik, nya funktioner. Den digitala ord- boken som parallellkorpus.....	29
<i>Pär Nilsson</i> Teori i praktiken. Om prototyp teorin och SAOB.....	59
<i>Caroline Sandström</i> Perspektiv på svensk lexikografi i Finland med <i>Ordbok över Finlands svenska folkmål</i> och <i>Finlandssvensk ordbok</i> som exempel.....	75
<i>Emma Sköldberg & Christian Mattsson</i> Ful- och finlexikografi? Om ordboksverksamhet i Sverige i dag och i morgon.....	111

Förord

Alla har vi ett slags förhållande till ordböcker. De flesta av oss har använt skolordlistor i olika färger, och i en del hem finns de kanske kvar som ett minne från den tiden då man fick behålla skolböckerna. Ofta lever de en ganska anonym tillvaro i någon hylla men ibland tar man kanske fram dem för att slå upp något ord. Utan att ha bevis för påståendet vågar jag säga att ganska många ordlistor och ordböcker som värmer hyllor i skolor, hem och på arbetsplatser är ganska gamla. Det anses inte alltid så påkallat att skaffa den senaste upplagan. Kärnordförrådet är ju oftast detsamma, stavningen och böjningen likaså. Gamla ordböcker blir liksom aldrig helt inaktuella.

För den vanlige brukaren skiljer sig heller inte de olika upplagorna av t.ex. SAOL så mycket från varandra även om de senaste upplagorna faktiskt har förändrats en hel del gentemot sina föregångare. Ibland är kanske förändringarna av större vikt för språkvetaren än för allmänheten, särskilt om de handlar om morfologiska principer och inte om själva ordförrådet. För den vanlige användaren finns det troligen viktigare saker i världen att uppmärksamma än en ny avstavningsprincip. Därför går ofta ganska stora omarbetningar av ordböcker användaren förbi, som i stället förblir sin gamla stillsamma ordlista trogen, i stället för att köpa en ny upplaga. På så sätt invaggas man i en – felaktig – tro att lexikografi skulle vara en ganska händelselös verksamhet och förbehållen de verkliga entusiasterna.

Vid trettioandra sammankomsten för svenskans beskrivning i Karlstad 2011 avslutades en av dagarna med en paneldebatt om ordböckernas funktion som kulturbärare. Deltagare var lexikografer från olika ordböcker: Sven-Göran Malmgren (SAOL, Svensk ordbok), Anki Mattisson (SAOB), Birgitta Lindgren (Nyordboken), Caroline Sandström (Ordbok över Finlands svenska folkmål) och Eva Thelin (Svenskt dialektlexikon). Jag hade förmånen att få vara moderator. Vi diskuterade ordböckernas funktion som kulturbärare, svårigheten med att kombinera både läsekretsens krav på många ord och tydliga enkla förklaringar med vetenskaplig stringens, tiden det tar att göra en ordbok samt all information som man faktiskt kan få fram ur ordböcker förutsatt att man vet hur de är uppbyggda och att man rätt kan tolka alla punkter, streck, typsnitt och förkortningar (= läs förordet!). Publiken var ganska stor och flera kom efteråt och sade att det *faktiskt* hade varit mer intressant att lyssna på än vad de hade trott från början.

Ordet *faktiskt* har i *Svensk ordbok* (2009) definitionstillägget: *spec. om ngt som verkligen gäller trots yttre tecken på motsatsen*. Ordet sade en hel del om förutfattade meningar också hos etablerade språkvetare och deras inställning till ordböcker och lexikografi. Samtidigt väcktes också idén om en mindre konferens helt ägnad åt svensk lexikografi där man skulle kunna presentera en del ordboksprojekt och nya tankar men också lyfta en del mer problematiska frågor som berör de svåra förhållanden som ordboksproduktionen befinner sig i idag när lönsamheten är låg och bokförlagen inte vågar satsa på nya ordboksprojekt. I föreliggande volym finns de fem presentationer som gjordes under endagarskonferensen som hölls i Göteborg den 5 oktober 2012. Föredragen hölls av Lars Borin, Anna Helga Hannesdóttir, Pär Nilsson, Caroline Sandström samt Emma Sköldberg och Christian Mattsson.

Ordet *faktiskt* avslöjar som sagt en del om den generella inställningen till ordböcker och lexikografi. Ofta ger verksamheten ett lite tråkigt intryck och framkallar kanske föreställningar om personer med gott sittfläsk som tar sig an outsinliga räckor av ord som ska behandlas enligt viss systematik för att slutligen sättas samman i en bok i alfabetisk ordning där resultatet blir förvillande likt alla andra ordböcker av samma slag som givits ut under årens lopp. Skämten bland språkvetare är många om alla lexikografer som aldrig tog sig ur bokstaven S. Och varför ens bemöda sig om att ge ut ordböcker i pappersform nu när det finns det oändliga internet där allt kan laddas ner snabbt och gratis?

Ändå väcker ordböcker och ordlistor en viss uppmärksamhet när de ges ut i ny upplaga. Inte minst gäller detta SAOL där den förändrade ords-katten i varje ny upplaga genomgår grundliga inspektioner och genererar många insändare. I den 13 upplagan (2008) föranledde införandet av bl.a. orden *guss* och *keff* en diskussion mellan en av lexikograferna och en kritisk läsare i den statliga televisionens morgonsoffa. Så visst väcker ordlistor fortfarande uppmärksamhet, i alla fall när det gäller åsikter om vilket ordförråd den nya upplagan innehåller.

Just denna blandning av ordbokens grå vardaglighet och förgivettagna existens som emellanåt livas upp av en upprörd debatt om språklig utarmning när en ny upplaga dyker upp säger en del om hur viktiga ordböcker trots allt är i vår kultur och vilket värde vi tillmäter dem som språkliga normbevarare. Ordboken finns där som ett slags ordförrådets väktare som kan förvalta de ord, ordformer och betydelser som vi använder, och vi förväntar oss samtidigt att ordboken ska rida spärr mot nya främmande ord som i en strid ström försöker etablera sig i vårt språk. Fokus ligger oftast just på det ordförråd som ordboken innehåller, och många språkbrukare omfattar den föreställningen att de ord som inte finns i t.ex. SAOL heller inte finns i svenska språket.

Ordboken som normföreställning

Ordböcker har funnits i Sverige sedan 1400-talet och de har spelat en mycket stor roll i etableringen av en svensk språknorm. Genom ordböcker har vi befäst en svensk ordsfatt och skapat stadga i ett språkbruk som kunnat normeras inte minst när det gäller stavning och böjning. Med denna allt fastare språknorm har vi lagt grunden för de språkliga förutsättningar som är grundvalen för allas delaktighet i samhället. En gemensam norm gör det lättare att lära sig läsa och skriva, en gemensam norm gör det möjligt att lära sig att behärska det viktigaste verktyget för att göra sin röst hörd: det egna språket. Ordboken är därmed viktig för vår föreställning om vad språket är och för upprätthållandet av de regler vi mödosamt lärt oss genom åren. Inte konstigt då att en del blir upprörda när man tar bort ord ur ordböcker och ersätter dem med nymodigheter som kan kännas främmande.

Men ordböcker, liksom alla andra kulturinstanser, utvecklas och förändras i enlighet med det samhälle de ska fungera i, vilket blir mer tydligt under vissa perioder. Just nu är vi inne i en sådan period. Från att ha varit en tryckt bok där alla ord vanligen står i alfabetisk ordning har den digitala eran nu öppnat nya möjligheter för ordboken, både vad gäller omfång, uppdateringar och sökmöjligheter. Förändringen har inte kommit plötsligt, men den är trots detta lite svår att hantera för de etablerade ordboksutgivarna och lexikograferna. Alla kan skriva en ordbok, alla har möjlighet att publicera den på nätet, principerna för hur orden ska presenteras behöver inte anpassas till att pappersåtgången ska bli så liten som möjligt. Vad innebär det för den etablerade lexikografiska verksamheten? Detta diskuteras i Emma Sköldbergs och Christian Mattsons artikel Ful- och fin-lexikografi där de lyfter frågor om vilka som producerar ordböcker, hur ordböcker kan finansieras, och hur man ska se på det faktum att man idag skapar ordböcker ihop på nätet. Benämningarna fin- och ful-lexikografi ansluter till uppfattningar om att vissa ordböcker skapade av de gängse lexikografiska aktörerna anses vara bättre och finare än den mer folkliga och fritt tillgängliga lexikografiska floran på nätet. Men det finns också en uppfattning om att ordböcker baserade på en stringent vetenskaplig metod i t.ex. betydelsebeskrivningarna är finare än ordböcker där definitionerna är enklare och mer pedagogiskt förklarade, och därmed mer begripliga för den vanlige ordboksanvändaren. Fint = svårt och akademiskt, fult = enkelt och folkligt. Också synen på en- och tvåspråkiga ordböcker tas upp, där man idag anser att enspråkiga ordböcker generellt sett har högre status än tvåspråkiga ordböcker som oftare – inte minst av ekonomiska skäl – är översättningar mellan olika språk där en stor del av den lexikografiska analysen inte är relevant för de ingående språken och där böckerna ofta är föråldrade eller bara omtryckta upplagor med nytt tryckår för samma

innehåll. Med utgångspunkt i dessa faktorer diskuterar författarna sedan en tänkbar utveckling för svensk lexikografisk verksamhet.

Den tvåspråkiga ordbokens status i framtidens lexikografi

Att den tvåspråkiga lexikografin bildat utgångspunkt för den enspråkiga är inget man tänker på när man idag diskuterar lexikografi. Men faktum är att den tidiga svenska lexikografin i huvudsak var två- eller flerspråkig och att de första systematiska beskrivningarna av svenskans språkssystem är de beskrivningar som finns av svenskan i de två- eller flerspråkiga ordböcker som gavs ut i Sverige från 1500-talet fram till 1850-talet. Då trycktes A.F. Dalins enspråkiga definitionsordbok *Ordbok öfver svenska språket*, som kan sägas bilda utgångspunkt för modern enspråkig lexikografi på svenskt språkområde.

Lexikografiteori för tvåspråkiga ordböcker grundar sig oftast på de ingående språkens funktioner som målspråk respektive källspråk i ordboken, plus att man också bör ta hänsyn till om ordboken är avsedd för förståelse eller produktion av det främmande språket. Det optimala utgångsläget för ett givet språkpar är egentligen att man konstruerar fyra olika ordböcker där varje språk fungerar som källspråk och målspråk för både L1 och L2-talare som vill använda ordböckerna för produktion eller förståelse. Alltså: en svensk som vill läsa tyska behöver en tysk-svensk ordbok med ett stort ordförråd som täcker in så mycket det går, även föråldrade ord eftersom svensken kanske vill läsa gamla, tyska böcker. Om svensken däremot vill skriva eller tala tyska själv, behöver svensken en aktuell svensk-tysk ordbok som kan erbjuda tyska ord och konstruktioner för produktion. För en tysk som vill läsa eller tala svenska behövs samma typ av ordbokspär, men då avpassade för en som har tyska som förstaspråk. Krångligt? Dyrt? Ja. Det är också därför som optimalt anpassade tvåspråkiga ordböcker saknas för de flesta av världens språkpar, undantaget de största. Dit hör inte svenskan.

Vanligen begränsar man därför ordboksparen till att omfatta två böcker: en där vardera språket är källspråk och en där samma språk är målspråk. Sedan får tysken använda den tysk-svenska ordboken som produktionsordbok när tysken vill skriva på svenska samtidigt som svensken använder samma ordbok för att förstå Goethe.

I dag, när man i större utsträckning upprättar databaser över de ingående språken, börjar skillnaden mellan käll- och målspråk i ett lexikografiskt hänseende att suddas ut. Anna Helga Hannesdóttir behandlar utvecklingen av tvåspråkig lexikografi i sin artikel om Islex, den elektroniska flerspråkiga ordbok som upprättats med isländska som källspråk och danska, norskt

bokmål, nynorska och svenska som målspråk. Ordboken är tänkt att vara en produktionsordbok för islänningar och en avkodningsordbok för de övriga språken. Genom att det digitala formatet tillåter att man också har något av de andra språken som ingång i sökningarna kan man dock säga att distinktionen mellan källspråk och målspråk försvunnit. Hannesdóttir beskriver i sin artikel hur man kan använda materialet i Islex för kontrastiva studier av enskilda ord och uttryck, och på det viset utnyttja materialet i ordboken som en parallellkorpus där man kan få fram även annan information än lexikalisk. Kanske kan framtidens tvåspråkiga lexikografi genom den digitala tekniken på detta sätt utvecklas på ett nytt sätt som bättre tillvaratar de möjligheter som ryms i jämförelsen mellan flera språk.

Ordboken och datorn

På internet finns mycket lexikaliskt material tillgängligt, både gratis och som betaltjänst. Kvaliteten kan vara mycket skiftande och det framgår inte alltid vad lexikonen bygger på. Google tillhandahåller tjänster som bygger på statistik framräknad på deras mycket stora korpusar. Inte minst deras översättningstjänst är flitigt anlitad, på gott och ont. Även här kan man se att det uppstår problem när man ska översätta mellan språkpar där det ena språket inte har så stora material tillgängliga för analys.

Svenskan är ett språk som är väl beskrivet både lexikaliskt och grammatiskt, och vi har stora datamängder tillgängliga för maskinell lexikologi, något som Lars Borin belyser i sin artikel. Människor och datorer har emellertid olika förutsättningar när det gäller att processa språkliga yttranden. Detta medför i sin tur att mänskliga och maskinella lexikon ser olika ut och bygger på delvis olika information. Datorernas lexikon är inte alltid ämnade för människor utan de behövs primärt för att kunna generera mer information som datorn kan fortsätta att bygga analyser, översättningar och andra applikationer på. Artikeln belyser på flera sätt den problematik som måste behandlas utifrån aspekten människa – maskin inte minst vad gäller urval av information; vad behöver en människa respektive en dator ha för information om ett ord för att kunna förstå det på samma sätt? Människans kunskap om världen är svår att bibringa en dator, och tänker vi egentligen ens på vilken kontext vi använder för att tolka ett yttrande? Och ser kontexten lika ut för olika människor?

För en människa är det också enklare att hantera den suddiga gränsen mellan lexikon och syntax men för en dator måste de reglerna beskrivas mer stringent. Också betydelseutveckling får effekt på gränsdragningen mellan olika kategorier i ett lexikon, för att inte tala om flerordsuttryck. I mångt och mycket kan man säga att datamaskinell lexikologi upphäver

många begränsningar som sätts av den gängse, pappersbaserade tryckta ordboken, men att den i stället kräver exakt formulerad information om hur ord ska förstås och tolkas.

Att beskriva språk över tid

SAOB, eller *Ordbok över svenska språket, utgiven av Svenska Akademien*, som den egentligen heter, är den stora, historiska ordbok som skildrar utvecklingen av svenskt ordförråd från 1500-talet och framåt. Ordboken har producerats sedan slutet på 1800-talet och den beräknas vara klar 2018. Trots denna oerhört långa utgivningstid är ordboken förvånansvärt enhetlig i sin beskrivning av materialet. För att snabba upp den långsamma lexikografiska processen har man genom åren beslutat om vissa inskränkningar i antalet ord och i utformningen av artiklarna, men i princip bygger artiklarna i dag på samma metod som när första häftet kom ut 1893.

Pär Nilssons bidrag handlar om betydelsebeskrivningen i SAOB. Han beskriver hur betydelsebeskrivningen kan sägas vara baserad på hur människan uppfattar och kategoriserar betydelser utifrån sin sociokulturella bakgrund och situation, något som också utgör fundamentet för prototyp teorin. I artikeln beskrivs hur prototyp teorin kan läggas till grund för lexikografiskt arbete och hur man kan formulera betydelsebeskrivningar i definitionsformat utifrån de olika grader av ”normalitet” som ett ords betydelse kan anses ha i relation till någon slags prototypuppfattning. Artikeln är intressant inte minst som kontrast till den diskussion som förs av Borin om de precisa beskrivningar som ett maskinellt lexikon behöver ha för att kunna tolka text.

Sprungnen ur den nationalromantiska tankevärld som genomsyrade Norden mot slutet av 1800-talet är SAOB det absolut största ordboksprojekt som Sverige skådat. Eftersom man inte uppdaterat de tidigare delarna säger det sig självt att man måste påbörja en uppdatering ganska omgående efter att första upplagan blivit klar, eftersom ordförrådet ändrats kraftigt sedan slutet på 1800-talet. Att man då kan konstatera att det definitionsformat som används i SAOB till stor del grundats på en slags prototyp teoretisk tanke långt innan den formulerades av någon semantiker, och att det har hållits i princip konstant över så lång tid, borde vara en bra utgångspunkt för uppdateringen. Om man därtill kan skapa hjälpmedel att maskinellt kunna få fram prototypiska betydelser ur äldre språkmaterial för att underlätta uppdateringsprocessen skulle man skapat en skön förening mellan historisk och framtida lexikografi.

Svenskan i Finland

Ytterligare en ordbok som emanerar ur den nationalromantiska synen på den egna nationen och dess språk som ett utslag av nationalidentiteten är *Ordbok över Finlands svenska folkmål*. Bakgrunden till denna, och andra finlandssvenska ordboksprojekt, tecknas av Caroline Sandström. Hon ger en fyllig beskrivning av en del av svensk lexikografi som inte alltid uppmärksammas i Sverige, nämligen den finlandssvenska. Sandström beskriver också hur den finlandssvenska lexikografiska traditionen relateras till den sverigesvenska. I artikeln diskuteras också hur man hanterar strävan att försöka återge så många belägg och varianter som möjligt av svenska dialekter i Finland. I detta stora ordboksprojekt har man dessutom behövt ta sig an just den problematik som uppstår när man mitt under pågående ordboksprojekt beslutar sig för att ge ut materialet digitalt, vilket ytterligare exemplifierar den situation som lexikografin idag befinner sig i, i en övergångsperiod mellan traditionell lexikografi och elektroniskt baserade ordböcker.

Framtidens lexikografi

Ordböcker utgör en viktig del i beskrivningen av svenskans ordförråd ur såväl ett språkvetenskapligt som kulturellt perspektiv. Genom lexikografihistoriska undersökningar kan man kartlägga och förstå hur synen på svenskan som språk utvecklats genom tiderna och hur såväl språknormen som det lexikografiska arbetssättet kommit att stabiliseras. Utvecklingen av nya tekniker visar samtidigt att man ständigt måste förfina de språkliga analyserna för att kunna etablera nya arbetssätt, inte minst datamaskinella, och att lexikografi långt ifrån är ett uttömt ämne.

Tvärtom visar artiklarna i denna bok att svensk lexikografi idag står inför en stor utmaning som både består i att förena ny teknik med ett etablerat arbetssätt, och att samtidigt bibehålla en god kvalitet på det som produceras. Ordböcker kan idag distribueras via appar som gör lexikaliskt material lättillgängligt. Detta ställer stora krav på såväl avvägningar vad gäller språknormering som användarvänlighet.

Därmed är lexikografi idag kanske viktigare än någonsin genom att tillgängligheten för användaren ökar samtidigt som i princip alla kan vara lexikografer. Ordboken ligger idag mer än någonsin i användarens hand och att dagens lexikografiska verksamhet via internet också engagerar andra än språkvetare i lexikografins tjänst ger självfallet ökade möjligheter till ytterligare beskrivningar av språket. Samtidigt innebär det att många lexikografiska internetprojekt påbörjas men överges utan att någon an-

svarar för kontinuiteten. Dagens lexikografi står därför inför många stora utmaningar som rätt använda borde kunna bidra till att stärka ordbokens roll i framtiden.

Lena Rogström
prefekt vid Institutionen för svenska språket

Lexikografi för maskiner och lexikografi för människor¹

Lars Borin

The author has long experience in developing lexical resources for use in language technology applications, as well as a strong and active interest in typological linguistics, including lexical typology. Against this background the article discusses some important differences between machine lexicons and lexicons for human use, as well as some methodological considerations which arise when the contents of lexical products intended for human consumption are reused and adapted in order to develop machine lexicons.

1. Bakgrund

Följande betraktelse är skriven ur ett visst perspektiv, nämligen av någon som

- är professionell utvecklare av lexikon för maskiner (som i princip inte utesluter människor som användare)
- är flitig användare av lexikon för människor
- har en akademisk bakgrund inom
 - datalingvistik/språkteknologi
 - slavisk och finsk-ugrisk språkvetenskap
 - storskalig jämförande språkforskning:
 - språktypologi (inklusive morfologisk och lexikalisk typologi)
 - arealingvistik
 - historisk lingvistik
 - men *inte* nordiska språk eller svensk lexikografi.

¹ Det arbete som beskrivs här har möjliggjorts genom finansiering från flera håll: Vetenskapsrådet (projektet *Svenskt FrasNät++*; VR dnr 2010-6013), EU (projektet *META-NORD*; ICT PSP, kontrakt nr 270899), samt Göteborgs universitet genom dess stöd till Språkbanken och styrkeområdet språkteknologi (*CLT*).

Denna bakgrund ger intressanta alternativa infallsvinklar på lexikons uppbyggnad och innehåll.² Språkteknologin ställer framförallt kravet att lexikon måste vara fullständiga och totalt explicita, medan min bakgrund i typologisk språkvetenskap bidrar med teoretiska insikter med konsekvenser för vilka sorts lexikonenheter man kan och bör räkna med. Nedan kommer jag att kontrastera vissa aspekter av maskinell och mänsklig lexikografi, med utgångspunkten i det maskinella och i viss mån i det språktypologiska.

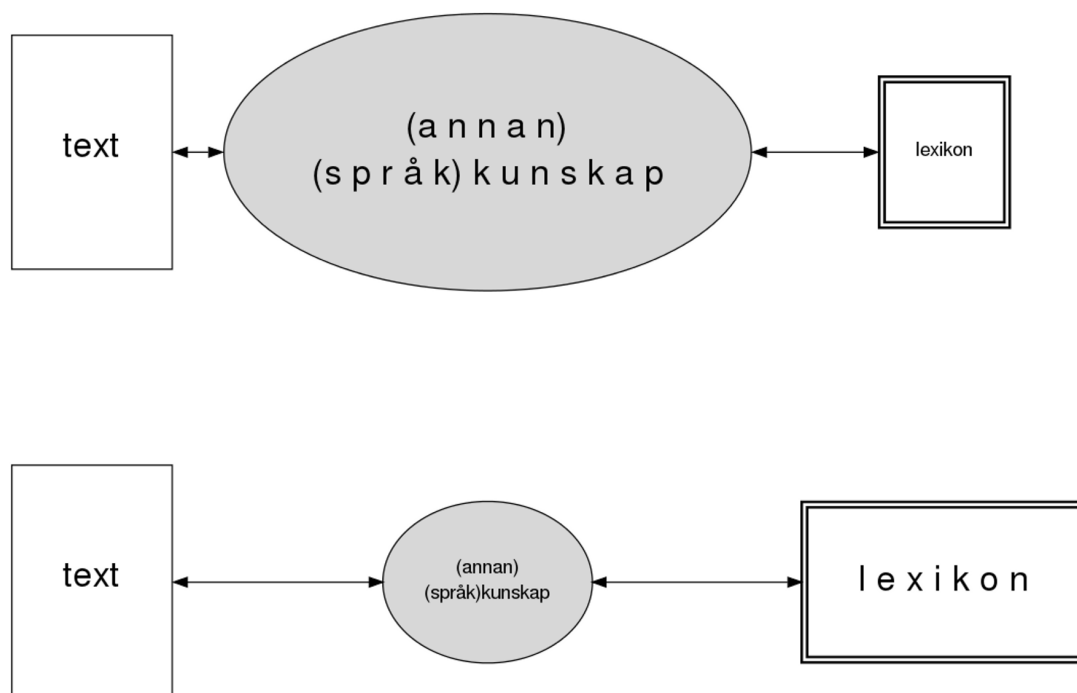
2. Lexikon för människor och lexikon för maskiner

Det är faktiskt förvånande att lexikon för människor och lexikon för maskiner ibland kan se ganska lika ut, eftersom de i själva verket har mycket olika förutsättningar och ändamål. Maskinlexikon måste i någon mening ”bära med sig” all den språkkunskap som behövs för att en språkanalysuppgift ska kunna lösas. Jfr. Jonathan Swifts akademi i Lagado i *Gullivers resor*:

The other project was, a scheme for entirely abolishing all words whatsoever; and this was urged as a great advantage in point of health, as well as brevity. For it is plain, that every word we speak is, in some degree, a diminution of our lunge by corrosion, and, consequently, contributes to the shortening of our lives. An expedient was therefore offered, 'that since words are only names for things, it would be more convenient for all men to carry about them such things as were necessary to express a particular business they are to discourse on.' [...] many of the most learned and wise adhere to the new scheme of expressing themselves by things; which has only this inconvenience attending it, that if a man's business be very great, and of various kinds, he must be obliged, in proportion, to carry a greater bundle of things upon his back, unless he can afford one or two strong servants to attend him. (Jonathan Swift: *Gulliver's travels*, Part 3, Ch. 5)

² Med ”lexikon” menar jag här genomgående en artefakt, alltså en ordbok avsedd för människor (som kan vara digital) eller en databas – en lexikonresurs – för någon språkteknologisk tillämpning. Det handlar här alltså inte om lexikonet i betydelsen ’mentalt ordförråd’. De språkteknologiska tillämpningar som diskuteras är dessutom huvudsakligen inriktade mot det skrivna språket (men inte med nödvändighet det standardiserade skriftspråket). Lexikon för talspråk ställer eventuellt andra krav än lexikon för skriftspråk. Detta är ytterst en empirisk fråga, som jag inte har möjlighet att belysa närmare här.

Lexikon för människor kan förlita sig på att användaren redan besitter det mesta av den kunskapen; se figur 1. Vi människor kan förlita oss på en stor mängd kunskap om språket och världen när vi använder oss av ett lexikon för att klarlägga någon aspekt av en text, där vi dessutom normalt redan förstår det mesta.³



FIGUR 1. *Textförståelse för människor (upptill) och maskiner (nedtill).*

Positivt när man står inför uppgiften att bygga lexikon för maskiner är att man tvingas att tänka på många metodfrågor i språkbeskrivningen som man lätt kan sopa under mattan om man kan förutsätta en mänsklig språkande agent (se avsnitt 4 nedan). Å andra sidan uttrycks svaren på metodfrågorna inte alltid nödvändigtvis bäst i termer som en språkvetare skulle känna igen eller ens acceptera. Det har nämligen visat sig genom åren att de system som byggs genom att man strikt håller sig till vedertagna språkvetenskapliga lösningar inte gör särskilt väl ifrån sig vid analys av godtycklig text eller tal, en omständighet som ibland ironiskt sammanfattas genom att man anför ett citat som tillskrivs en av talteknologins förgrundsgestalter,

³ Här finns en uppenbar skillnad mellan L1- (modersmåls-) användare och L2- (andra- eller främmandespråks-) användare. Även om jag i denna framställning i första hand har L1-fallet i åtanke, kan det vara värt att komma ihåg att även L2-lexikonanvändningen bygger på att vi redan kan ett eller flera språk (många L2-lexikon är ju dessutom tvåspråkiga), på att vi har en stor kunskap om världen samt på vår kunskap om hur lexikon brukar vara organiserade. Dessutom är ju som framgått ovan min egen akademiska språkvetenskapliga bakgrund i hög grad präglad av L2-studier.

Frederick Jelinek: ”Anytime a linguist leaves the group the recognition rate goes up.” (Jurafsky & Martin 2009:83)

Den för en mänsklig betraktare kanske mest främmande typen av maskinella lexikala representationer kommer ur statistisk språkteknologi och bygger på ords distribution i stora mängder text som kan vara obearbetad eller försedd med lingvistisk information. I det senare fallet talar man ofta om *annoterad* text.

lexikon..nn

[4.5, -4.5, 5, -10, -39, 24.1667, 0, 15, 6.5, -5, 16.5, 3, -1, 2.5, -2, 8, 9.5, 1, 0, 5.5, 9.5, 21, 14, 97, -7.5, 2, 1.5, -5.5, 30.5, -11.6667, 6.5, -1, 0, 22.3333, -2, -31, 1, 14, -3.5, -4, 0.5, -2.83333, 1.5, 3, 6, -1, 19, 38.5, -3, -0.5, -3, 1.66667, 8.66667, 25.5, 24, 6, 1, 62, 5, -9, 8, 13, -1, 10.5, -10.5, -2.5, 5, -6, 14.3333, 5, 1.66667, 7.16667, 4, 6.5, 10, -6.5, -5, -1.5, 3, 3.5, 1, 1, 2, 251, 16, 4, 115.5, -6, 1.5, 0, -4.5, -36.5, -9, -11.6667, 7, 9.33333, 1, 0, 5, 0, -1, 145.5, -9, 6, 0, -11, 109, -2, 2, 1, 0, 1.5, -96, 6, -2.5, -2.5, 6, 0, -4.5, 30, 4.33333, 104, -9.5, 5, 10, 5.5, -1, -9.5, -3, -1.5, -5.5, -3, 3.5, -1, -3.5, 30, 2, 1, -1, -8.5, 7.5, 2, 6, -25, -4.5, 2.66667, -7.66667, 3, -7.5, 109.833, 0, -2, -10.5, -7, -10, -2.5, -1, -3, 5.5, 2.33333, 3, 9, 1.83333, -17, 5.5, 0, 23.5, -11.5, -2.5, 31.5, -2.5, 119, -6, 6, 8.5, 26, 3.5, -25.8333, 20.5, 3.66667, 9.5, -4.16667, 18, 5, 5.5, 35.5, -91, 27.5, 0, 4, -0.5, 9.33333, -9, 1.5, 2, 8.5, -2.33333, -7, 37.6667, -20, 0, -34.5, 0, 13, -37.5, -5, -3, -2.5, 8, -37, -6, -5.33333, -8.5, 5.83333, -20, -4.5, -16.5, 0, 4.33333, -3, 1.5, 14, 6.5, 2, -16, 0, -3, 28, 12, -1.5, 1, -21, 9, -9.5, -11.5, -0.5, 2.5, -8.83333, 3, 3.83333, 3, -7, -48, -3.66667, 5, 4.5, 4, 1.83333, -26, 26, 19, 3.5, -35, 3.5, 8, -7, -4.5, 2, -7.5, -2, -4, -20.5, -53.6667, -7.33333, 0.166667, -1.5, 4, -2.5, -2, -7.5, -1, 17, -5.5, -8, 2.83333, -8.5, -2, 8, -34, 0, -4.5, 11.5, -3.33333, 27.1667, 3, -4, -4.5, -2.5, 36, 2, 4, 4, -29, -27, -111, 0, 1.83333, -6, -78.8333, 21, 2.5, -15.8333, -2, 8, -9.5, 20, -7, -3.5, -13, 0.5, -11, -9.5, -115.667, 18, -24.5, 4, -3, -10, -10.6667, 5, -4, -5.5, -5, -2.5, -27, 1, 0.5, -4, -8, 0.5, -22, -29, -93, -117.5, 13.8333, -17.6667, -71, 4.5, -149, -15, -28, -19, 12.5, 4, -16.5, -1.5, -5.5, 1, -4.66667, 25, 10, -46.8333, 13, -3.16667, 5, 1, -46.5, -8, 3.5, 11, 11.3333, 8, 1.5, 2.5, 146, 2.5, -4, 4.5, -1.5, 84, -25.5, -0.5, 6, 6.66667, -4, 2.5, 21, 1, 10, 2, 0.5, 1.5, -3, -17.5, -8, 26.5, -4, -34, 6.83333, 5.5, 17, -8, -3, 6, 2, 7.5, 0.5, 16.5, 6.5, 4, 15, 0.5, -8, -9, -11, 9.5, 21.5, -1.5, -23, 6.5, 19, 2, 0.5, -147.5, 4.5, -0.166667, -4.5, -40, 43, -7.83333, 10, 2, 1.33333, 1, 24, 6, -0.5, 4.5, 4, -3, 7, -1.5, 16, -22.5, -22, 2, 8, -21, 12.5, -3.33333, 3.5, 33, 7.5, 2, -6.33333, 2.5, 5, 0, 13, -3.5, -7.83333, -2, 5.66667, -5, -2.5, 12, -16, 1, -27.5, -4, -3, 8, 2, -6, -6.5, -14, -6, 9.5, -8, -7, 0, 0, -2.66667, 25, 0.5, -4, -4.5, 5.5, 40, 3.5, -8.33333, 9, -11, -4.5, 16.5, 89.5, 0.833333, -2.33333, 38, 17, -4.83333, -1.16667, 1.5, 0, 8.5, 2.66667, 1, -5, 7.16667, -7, 4.5, -1, 7, 0, 45.5, -2, 12.5, 2, -7, 3, -33.3333, 0, -2.5, -3, 13, -11.5, 10, 3, -7, -6.5, -14.5, 3.5, 4, -2.5, 2.5, -3.5, -30.6667, -25, 6.5, 7, 36, -8.5, 19.5, -11, 2, 51.5, -7.5, -5.5, -10, 57.5, 8, 15, 5, -123, 94.3333, -14, 6, 0.5, -2, 35, -6, -1, -21.5, 10, 30.3333, 6, 43, -14.1667, -3, -0.5, 3.5, 7.5, 11.6667, -3, 2.5, -2.5, 2.5, -7, 0.5, 23, 5.5, 2.33333, -90.5, 5, 33, -7.5, 57, -14.5, 30, 2.66667, 1, -4, -7.66667, -20.5, -7.83333, 14.5, 3, -27, -17, -47, 2.5, 15, 21.3333, 9, -0.5, 5, 5.5, -21, -3, -5.66667, -5, 5, -17.5, -0.166667, 0, -2, 5, 12.6667, -8.33333, 29.5, 0.5, 5, 4, 31, 9.5, 15, 1, -14, 1, -33.8333, 0, 3, -31.5, 2, 33, -3.5, -2, 1, -11, -4.66667, -35.5, -6.5, -5, -9.5, 118, 73.5, -5.5, -8, -6, 12.6667, -1, -1, -1, -10.5, -10, 0, 2, -12, -3, -5.5, -26, -16, 24.5, -19, -5, -13, 12.5, -15, -28, -8.16667, -7.5, -3.5, -8.33333, -52, -46, 1, -45.5, -6.5, -13.5, -29.5, -13.8333, 13, -43.5, -34, -5, -4.5, -33.5, -130, -36, -6.66667, -12, -151.167, -104.5, 23, -42, 39.6667, -1, -5.33333, 7.5, 1, -9.5, 7.33333, -2, 14.5, 2.5, -16.5, 5, -6, 1.5, 2, -36, 6.33333, 2, -5.5, 3, -26, -2.5, 2, -2, 3, -18.5, 47, -21.3333, 21, 13, -3, -4, 31.5, 0, 4, -4, 5.33333, 4, 24, -1, -2, -5.5, -3, -31.5, 51, -10.5, 2.5, 4.5, -1.5, 1, 0, 0.5, -3.33333, -4, 13, -5, -11, -1.5, -3, -32.5, -7.5, -6, 38.8333, 40.5, -2, -1, -5, -40, -2, 3, -9.5, -0.833333, 0, 30.5, 43, 8.5, 0.5, 2.16667, 2.5, -6, 1, -35, 3.5, 14.5, 2.5, -25.5, -0.5, -2.5, -2, 30, -3, 8.5, 8, -6, 3, -2.16667, -2.5, 1, 24, 1, 4.5, 11, 10, 3.5, 10.3333, -1, -46, -6, -6, -15, 21, 32, 20, 0, 139.833, 10, 14.5, -2, 1, 48.5, 19.5, -65.5, 10.5, -3, 2, 31, 2.5, 4, 6.5, -16.5, 4.5, -4, 2, 9, -3, -10, -6.5, 35.5, -10, 1, 3, 5, 17.5, 1.83333, -3.66667, -5, 3.5, 6, 4, 35.5, -12.5, 0, 2.5, 3, -3.5, 7.33333, -23.5, 19.1667, -0.5, 1, 2.5, 2.16667, 7.33333, 116.5, -5, 6, 1, -8, -0.5, 3, -1.5, 5.5, -119.5, 7.83333, -13, -3, -7, -12.1667, 10.5, -8.5, 7, -5.16667, 0.5, 190.5, -1.66667, -3.5, 1, -4.83333, -20, 20, 1, -5, 9, -4.5, -1, 44, -1, -1, 2.5, 1.33333, 5, 7, 1.83333, 3.66667, -9, -3, -104, -7.5, 8, 2.33333, 5, -17, -4, -2, -116, 5, 13, 8, -14, -3.5, 4, 7.5, 6, 1, 8.33333, -29.5, -8.83333, 5, -7.5, -102, -30, 1.5, -5, 0.5, 2.5, 7.33333, -5, 14, -4, 5.5, 8.33333, 2.66667, -1.83333, 1.66667, 1.33333, -0.5, 1, 2, -6, 14, 1, -3, 6, -31, 12, -1.5, -1.5, 16, -3, -13, -16.8333, -153.833, 3, 6, -3.5, 1.5, 17.5, -3.66667, 3, -2, 2, 7, -5, -15, -3.5, -7, -126, 6, 1, 2, -0.5, 1, -12.1667, -1.33333, 5.5, -4, 21, 11, 5, -16.5, 11.5, -19.6667, 5.5, 5, 50, -9.66667, 116.5, 2.5, 9.66667, -4, -11, 153.833, 51.3333, -9.5, -3.5, 1.5, -1, -55, -9.5, 11, -17.5, -31, -5.5, -3, -0.333333, 2.5, 6.16667, 14.5, -4, -24.5, -18, -4, 20.1667, -1, 2.5, -0.333333, -9, -44.6667, -29.5, -27.5, 13, -6, 32, -5, 5, 4, 8, -25.5, -11]

FIGUR 2. En ”lexikoningång” för ordet lexikon i ett slags maskinlexikon.

Den lexikala representationen – lexikoningången – blir i sådana fall oftast en vektor – alltså ett matematiskt objekt – med utgångspunkten i origo i

ett koordinatsystem med ett stort antal dimensioner och ord relateras till varandra genom något avståndsmått för vektorer.

I grova drag får man fram dessa representationer så här:⁴

- (1) Ett urval av ord görs som gemensamt antas vara diagnostiska för att särskilja ordbetydelser i deras omgivning. Det handlar om några tiotal till några hundratal ord. Mycket högfrekventa ord (typiskt funktionsord) fungerar inte bra. Allför lågfrekventa ord kan heller inte användas, eftersom slumpen då ges för stort spelrum. Urvalet kan vara mekaniskt, helt baserat på ordens frekvens och spridning i ett korpusmaterial.
- (2) I en stor korpus samlar man ihop och räknar de diagnosord som står inom ett visst avstånd från varje ord som ska beskrivas, t.ex. inom 5 ords avstånd, i samma mening eller i samma stycke.
- (3) Varje diagnosord definierar en dimension i beskrivningsordens ordvektorer, som man kan tänka på som listor med en siffra för varje diagnosord. Siffran räknas fram på grundval av summan av diagnosordets förekomster i den bestämda kontexten för beskrivningsordet. Anta att diagnosordet *läsa* definierar dimension 27 i ordvektorerna, och att ordet uppträder 23 gånger i kontexten för totalt 100 förekomster av ordet *ordbok* men bara 11 gånger i kontexten för 121 förekomster av ordet *lexikon*. Då ska siffran i dimension 27 i respektive vektor på något vis avspegla detta. Exakt hur det görs varierar. Man kan t.ex. använda något samförekomstmått, och olika skalningsmetoder förekommer. Oavsett hur man gör är det en rent mekanisk operation att få fram siffran, givet ingångsvärdena enligt föregående punkt.
- (4) När man väl har alla vektorer kan man sen jämföra dem rent matematiskt. En möjlighet är att se dem som punkter i ett rum med lika många dimensioner som diagnosord och låta jämförelsemåttet vara det euklidiska avståndet mellan punkterna. Om man normaliserar vektorerna till att ha längden 1 – så att den punkt vektorn anger ligger på avståndet 1 från origo – så kan man även använda något trigonometriskt mått (vanligen kosinus) på vinkeln mellan vektorerna (som ju alla börjar i origo). Ett avståndsmått kan i sin tur lätt konverteras till ett likhetsmått.

⁴ Följande beskrivning är starkt förenklad i syfte att göra framställningen klarare, och sopar därmed en rad metodologiska frågor under mattan.

I figur 2 visas ”lexikoningången” för substantivet *lexikon* (lexikon.nn), framtagen (av Richard Johansson, Språkbanken) på det sätt som beskrivs ovan på basis av ordets distribution i ett stort korpusmaterial (en betydande del av Språkbankens dryga miljarden ord modern svensk text, sökbar online via gränssnittet Korp; se Borin et al. 2012).

För en mänsklig läsare är den här informationen inte särskilt lättsmält. Eftersom den är numerisk, har den som just nämnts den stora fördelen att vara direkt omsättbar i olika slags ordlikhetsmått, genom att man alltså rent matematiskt jämför ords vektorer med varandra. I figur 2 ser vi vilka andra ord som är mest lika *lexikon* respektive *ordbok* med den vektormodell som har valts.

Likhet mäts här på en skala i intervallet $[0,1]$, där 1 betecknar maximal likhet, eller faktiskt identitet, eftersom vi ser att ordens likhet med sig själva är 1, vilket är betryggande. Likhetsmättet är symmetriskt, vilket leder till att medan *lexikon* står först i *ordbok*:s tjugo i topp-lista, så gäller inte det omvända, eftersom det finns minst 19 andra ord som är mer lika *lexikon* än vad *ordbok* är (t.ex. *replik*, *skrift* och *uppslagsbok*) enligt det använda likhetsmättet. Likheten mellan *lexikon* och *verk* är ju 0,960237, medan likheten mellan *lexikon* och *ordbok* bara är 0,943426.

lexikon.nn	ordbok.nn
lexikon.nn: 1	ordbok.nn: 1
replik.nn: 0.967299	lexikon.nn: 0.943426
stil.nn: 0.966401	spalt.nn: 0.936092
skrift.nn: 0.966314	krets.nn: 0.933194
bok.nn: 0.964195	skrift.nn: 0.932667
tidning.nn: 0.963404	term.nn: 0.931388
nummer.nn: 0.962569	verk.nn: 0.930797
häfte.nn: 0.962523	klass.nn: 0.929038
form.nn: 0.962522	tal.nn: 0.928482
värld.nn: 0.962477	stil.nn: 0.9284
katalog.nn: 0.962464	fyrtiotal.nn: 0.928224
återge.vb: 0.962155	intervju.nn: 0.928196
stycke.nn: 0.962144	nummer.nn: 0.927245
namn.nn: 0.962048	termer.nn: 0.927234
uppslagsbok.nn: 0.961056	socken.nn: 0.927132
mönster.nn: 0.960947	sal.nn: 0.926919
översättning.nn: 0.960425	församling.nn: 0.926583
dikt.nn: 0.960397	artikel.nn: 0.926555
kedja.nn: 0.960353	protokoll.nn: 0.926446
verk.nn: 0.960237	arkiv.nn: 0.9264
[...]	
ordbok.nn: 0.943426	

FIGUR 3. Likhet mellan lexikoningångar i maskinlexikonet från figur 2.

Avsikten är att man med hjälp av vektorrepresentationen och ett lämpligt jämförelsemått ska kunna gruppera ord i enlighet med deras semantiska

och syntaktiska egenskaper. I exemplet i figur 3 kan man med lite god vilja se hur orden i toppen av de två listorna verkar bära några slags gemensamma semantiska särdrag. Den informationen kan ligga till grund för vidare mänsklig beskrivning eller också användas direkt i automatiska språkverktyg, t.ex. för att öka träffsäkerheten vid informationssökning genom att systemet kan komplettera sökbegreppen med semantiskt närliggande ord. Den matematiska representation av ordförekomster i en korpus som här använts i illustrationssyfte hör i själva verket till de enklaste som har föreslagits. Det finns även betydligt mer sofistikerade metoder för att räkna fram ordvektorerna, där man t.ex. tar hänsyn till syntaktiska relationer mellan beskrivningsord och diagnosord (se t.ex. Baroni & Lenci 2010 och referenserna där).

Något som man får på köpet med den här typen av representation är att ordbetydelser inte blir skarpt avgränsade. Man kan experimentera med olika tröskelvärden eller grupperingsalgoritmer och därmed få en indelning i ordbetydelser som passar ett visst ändamål, eller helt enkelt – och mer radikalt – aldrig betrakta ordbetydelser som diskreta i den mening som normalt antas i traditionell lexikografi⁵ (jfr Kilgarriff 1997; Hanks 2000, 2013; Reiter & Sripada 2002; Erk 2010). Men det finns gott om mer ”människa” maskinlexikon också, ofta av rent pragmatiska skäl. Även om de storskaliga statistiska ansatserna idag är förhärskande i språkteknologin, har de en relativt kort historia jämfört med fältet som helhet, eftersom det var först någon gång under 1990-talet som utvecklingen av hård- och mjukvara nådde en punkt där de blev praktiskt tillämpbara. Behovet av lexikonresurser för automatisk språkanalys har givetvis funnits längre än så och de första större maskinlexikonerna kom till med utgångspunkt i maskinläsbara versioner – i form av fotosättningsfiler – av vanliga ordböcker för mänskliga brukare, t.ex. LDOCE (Boguraev & Briscoe 1989). Det mest kända och använda maskinlexikonet i världen – Princeton WordNet (Fellbaum 1998) – är i grund och botten en traditionell lexikografisk produkt utförd av traditionellt skolade lexikografer även om den konstruerades som en elektronisk orddatabas redan från början (se figur 4). Detsamma gäller i stora stycken även Språkbankens centrala lexikonresurs SALDO (Borin et al. 2013).

Av ett lexikon som ska kunna användas i språkteknologiska tillämpningar krävs att all information måste vara *explicit* och *uttömmande* – all information som behövs för språkanalys (eller -generering) måste finnas i

⁵ Det är förvisso sant att det vimlar av formuleringar som ”jfr”, ”se även”, etc. i traditionella ordböcker. Sådana formuleringar kan ibland i princip förstås som fingervisningar om mindre diskreta betydelser, men det är helt omöjligt att avgöra i ett enskilt fall om det är det eller något annat som avses.

lexikonet eller i någon annan komponent i systemet (inga ”o.s.v.”, ”o.dyl.”, ”och liknande”, etc. tillåts) – samt *formellt manipulerbar* i teori och praktik. Det sistnämnda innebär att informationen uttrycks i ett formellt språk – alltså avkodbart enbart utifrån en syntaktisk specifikation – med en explicit semantik och tillhörande inferensmekanismer av hanterbar komplexitet.

Det just sagda utesluter inte ett probabilistiskt element. Tvärtom tyder mycket på att vår språkanvändning i grund och botten bäst beskrivs med någon form av sannolikhetsdistribution (Manning 2003; Norvig 2011) – och därmed sällar sig till många andra typer av mänskligt beteende – medan de strikt kategoriska språkbeskrivningar som dominerar lingvistisk lexikografi och grammatik i stor utsträckning ytterst verkar motiveras av en preskriptiv språklig tradition som delvis har sina rötter i den filosofiska och matematiska logiken.⁶

The screenshot shows the Princeton WordNet search interface. At the top, it says "WordNet Search - 3.1" with links to the home page, glossary, and help. Below this is a search bar with the word "lexicon" entered and a "Search WordNet" button. Underneath the search bar are "Display Options" with a dropdown menu set to "(Select option to change)" and a "Change" button. A key explains that "S:" shows synset (semantic) relations and "W:" shows word (lexical) relations. Below the key, it says "Display options for sense: (gloss) 'an example sentence'". The main content is under the heading "Noun" and lists two senses of "lexicon":

- **S: (n) vocabulary, lexicon, mental lexicon** (a language user's knowledge of words)
 - [direct hypernym](#) / [inherited hypernym](#) / [sister term](#)
 - [part holonym](#)
 - [derivationally related form](#)
- **S: (n) dictionary, lexicon** (a reference book containing an alphabetical list of words with information about them)
 - [direct hyponym](#) / [full hyponym](#)
 - [part meronym](#)
 - [direct hypernym](#) / [inherited hypernym](#) / [sister term](#)
 - [derivationally related form](#)

FIGUR 4. Princeton WordNet (från <<http://wordnet.princeton.edu/>>).

⁶ Hur vår språkanvändning och vår språkkunskap hör ihop är en stor fråga, som i lingvistikens hittills huvudsakligen har varit föremål för mer eller mindre välgrundade gissningar snarare än seriöst empiriskt utforskande.

3. Det bästa av två världar?

Befintliga ordböcker för människor innehåller mycket värdefull information, men med en metaforjämförelse lånad från McCawley (1983) är de inte skattkammare fyllda med direkt användbara värdeföremål, utan snarare på sin höjd guldgruvor: Rikedomarna finns där, men det kan inte nog framhållas att det krävs mycket arbete för att man ska kunna göra om en digital ordbok för mänskligt bruk till en lexikal resurs för maskiner. Framställningen i ordboken förlitar sig i hög grad på implicit information, i ett format som inte är entydigt (på formplanet, och ibland inte ens om man förstår innehållet).

Idag finns det alternativ för insamling av den lexikala kunskapen som behövs i språkteknologisystem. Förutom de redan nämnda statistiska ansatserna används även s.k. *crowdsourcing*, alltså storskaliga kollaborativa amatörinitiativ som Wiktionary och Wikipedia (se t.ex. Gurevych & Wolf 2010; Munro et al. 2010). Är det mot den bakgrunden verkligen värt besväret att försöka utgå från befintliga ordböcker för människor istället?

I Språkbanken anser vi att det under vissa förutsättningar trots allt kan vara det. Vi har redan en relativt stor erfarenhet av att använda ordböcker skrivna för människor som utgångspunkt för utveckling av lexikala resurser för språkteknologisystem och vi har i det arbetet sett klara synergieffekter när (även heterogen) information från flera källor kombineras med användning av delvis automatiska metoder. Än så länge är också informationen i många avseenden av högre kvalitet än den som går att åstadkomma med statistiska metoder eller med *crowdsourcing*.

Dessutom är många av våra forskningssamarbeten sådana som involverar lingvistisk forskning och språkvetare med förväntningar om en viss typ av språkbeskrivning i de resurser som Språkbanken utvecklar och tillhandahåller. Som vi sett ovan är detta något som man absolut inte får med de statistiska ansatserna och knappast heller med *crowdsourcing*.

Den absolut grundläggande förutsättningen för att det ska vara värt besväret är att rättighetsfrågorna kan lösas. Vår syn på forskning och forskningsinfrastruktur – som detta i högsta grad handlar om – kräver s.k. *Open Content*, alltså att de resulterande lexikonresurserna görs tillgängliga på följande villkor:

- De får användas till alla slags ändamål (även kommersiella).
- De får modifieras/vidareutvecklas och vidare distribueras fritt.
- Vid användning och distribution ska resursens ursprung anges, helst genom referens till en vetenskaplig publikation.

En licens som motsvarar dessa krav är *Creative Commons erkännande* (CC BY; <<http://creativecommons.org/licenses/by/2.5/se/>>, <<http://creativecommons.org/licenses/by/3.0/>>), vilken följaktligen är Språkbankens normallicens för de resurser vi utvecklar och tillgängliggör.

Det är tyvärr ofta omöjligt att använda existerande lexikon på dessa villkor. Ibland beror detta på kommersiella hänsyn, då den som har upphovsrätten till ordboksmaterialet är orolig för att en fritt tillgänglig elektronisk version ska leda till att försäljningen minskar. Detta måste man givetvis ha förståelse för och helt enkelt acceptera att en sån ordbok inte går att använda som underlag för eller komponent i en resurs som ska vara fri för all slags användning. Svårare att svälja är däremot de fall där en elektronisk lexikonresurs har tagits fram med offentliga medel – projektanslag från forskningsråd eller reguljära fakultetsmedel vid ett universitet – men där resursen ändå inte görs tillgänglig under former som tillåter vidareutveckling och fri återpublicering av den.⁷

I de fall där det är möjligt att använda en existerande ordbok eller digitalt lexikon, har man enligt min mening mycket att vinna på det, inte minst därför att man i arbetet med maskinlexikon får nya infallsvinklar på flera av de klassiska lexikaliska frågorna inom språkvetenskapen, infallsvinklar som därmed direkt kan ställas i kontrast till om och i så fall hur samma frågor hanteras av den konventionella lexikografin.

Då man ger sig i kast med att anpassa informationen i en traditionell lexikografisk produkt så att den ska kunna användas i språkteknologiska tillämpningar, så handlar det först och främst om att se till att den resulterande resursen uppfyller de oavvisliga kraven på *total explicitet* (i all information) och *full täckning* (av alla språkliga fenomen i de texter som ska behandlas).

Härvid explicitgörs med nödvändighet implicita antaganden hos ordboksförfattaren om den språkkunskap som ordboksanvändaren förväntas besitta om språket, t.ex. om produktiv ordbildning, på grund av möjliga lexikoningångar som systematiskt saknas i ordboken. Man får även anled-

⁷ Detta leder bland annat till att man löper risken att undersökningsresultat inte kan reproduceras, men även – och betydligt allvarigare – till att forskningsresultat inte kan kumuleras, vilket förorsakar onödigt dubbelarbete och därmed i slutändan slöseri med offentliga medel (Lieberman 2012). Språkbanken utvecklar nu ett eget svenskt ordnät från grunden, trots att sådant med över 20.000 ingångar har byggts med forskningsrådsmedel för ungefär ett decennium sen, samt även en egen moderniserad version av Brings (1930) svenska version av Rogets thesaurus (1852), trots att en sådan modernisering redan påbörjats en gång av anställda forskare vid Stockholms universitet. I båda fallen blir de resulterande nya resurserna från Språkbanken tillgängliga på det sätt som just redogjorts för, med en CC BY-licens.

ning att reflektera över den gränsdragning mellan språklig och encyklopedisk kunskap som ordboken – igen implicit – ger uttryck för.

En i det här sammanhanget högst relevant egenskap hos storskaliga statistiska metoder för utvinnande av språkliga generaliseringar ur stora textkorpusar är att de inte tillhandahåller något bekvämt sätt att på förhand peka ut vilka språkliga drag som bör ligga till grund för beräkningarna. I allmänhet antar man alltså inte en strikt induktiv procedur à la Harris (1946,1955; se även Hammarström & Borin 2011), utan metoderna bygger i stor utsträckning på att man har en hypotes om vilka drag som ska visa sig användbara som grundval för den typ av generalisering man är intresserad av. Här är den lingvistiska ordforskningens – lexikologins, morfologins, den lexikaliska semantikens och den lexikaliska typologins – landvinningar naturligtvis viktiga, liksom de konkreta lexikografiska produkternas explicita och implicita organisation, genom att de kan peka på signifikanta språkliga särdrag som sedan kan testas empiriskt genom att de används som indata för de statistiska metoderna.

4. Några konkreta frågor om metod och teori

Nedan tar jag upp några konkreta metodiska och teoretiska frågor som har uppstått i vårt arbete med anpassning, utvidgning och sammanlänkning av traditionella lexikon till en språkteknologisk lexikalisk makroresurs (Borin et al. 2010). Frågor om vilken typ av språklig information som faller inom lexikonets domvärjo dryftas (eller manifesteras i konkreta beslut) av flera olika grupper av språkvetare, som delvis verkar vara isolerade från varandra. Åtminstone (generativa) grammatiker, (lexikala) typologer och lexikografer/lexikologer har bidragit till denna diskussion. I de följande avsnitten berör jag några av de aktuella frågorna, dock utan att på något sätt försöka vara uttömmande. De kan alla sägas handla om – ur något olika synvinklar – var lexikonets gränser ska anses gå: Var slutar lexikonet (och var börjar det)?

4.1. Vad är lexikalisk kunskap?

Vilken sorts lingvistisk information är att anse som lexikalisk information? Om man tittar i en stor enspråkig referensordbok, finner man typiskt information om lexikoningångarnas stavning, uttal, betydelse(-relationer), omvärldsfakta, bruklighet/stilnivå, domän, ordklass, böjning, valens, korpusfrekvens, intern (t.ex. ordbildnings-) struktur.

I grammatikforskningen har man rört sig mellan två ytterligheter när det gäller den språkliga kunskap som rubriceras som ”lexikal”, alltså antas finnas i det mentala lexikonet. I den klassiska generativa grammatiken är lexikonet en slags lista med relativt informationsfattiga medlemmar (ord eller morfem, beroende på version av formalismen) och huvudansvaret för att bitarna hamnade på rätt plats antogs vila på den syntaktiska komponenten, som åtminstone i tidig generativ grammatik även tänktes inkludera morfologin. Reaktionen mot detta synsätt lanserades under rubriken ”radikal lexikalism” (Karttunen 1986) med LFG (Bresnan 1982) som ett av de tidigaste och mest kända exemplen.

Konstruktionsgrammatiken (Hoffman & Trousdale 2013) kortsluter helt framt denna motsättning genom sin inställning att alla språkliga fenomen bortom fonologin – från morfem till syntaktiska regler – kan ses som ”konstruktioner” av olika generalitet och komplexitet.

Lexikografi för maskiner tenderar här att dras mot den mer lexikaliska polen. Rent praktiskt upptäcker man när man arbetar med konkreta textanalystillämpningar att de generella grammatiska mönstrens förklaringskraft är relativt svag och att mycket av språkbeskrivningen kommer att vila på (små) klasser av lexikonenheter eller enskilda ord. Detta verkar egentligen även – fast ofta inte uttryckt – vara den mänskliga lexikografins ståndpunkt. Däremot blir det naturligt att dra gränsen mellan språklig och icke-språklig (”encyklopedisk”) kunskap annorlunda än den av hävd dras i mänskliga lexikon när målet är automatisk textförståelse, helt i linje med att maskiner måste ”bära med sig” all information som behövs för avkodningen (se avsnitt 2 ovan). I ett lexikaliskt orienterat system för automatisk språkanalys blir den naturliga platsen för den ”icke-språkliga” kunskapen just i anslutning till de relevanta lexikonenheterna. Som ett konkret exempel kan vi anföra egennamnen i SALDO. De finns där delvis därför att de har ett morfologiskt beteende, som lämpligen redovisas i lexikonet, men de är dessutom försedda med en del semantisk information (se figur 5).

BETYDELSE	LEMGRAM	ORDKLASS	PRIMÄR	SEKUNDÄR	BARN (PRIMÄRA)	BARN (SEKUNDÄRA)
⚙️ Paris	Paris (egennamn) ↗️ 560102	egennamn	huvudstad	Frankrike	Lutetia parisisk	Bastiljen Champs-Élysées Eiffeltornet Lidbom Louvren Notre Dame Orly Parisyflygplats Pariskommunen pariserliv parismode färre...

FIGUR 5. Informationen under betydelsen Paris i SALDO.

4.2. Vad är en lexikoningång?

Det ortografiska ordet har en stark ställning som central enhet både i traditionell lexikografi och i språkvetenskap i allmänhet. Det finns en god medvetenhet om att flerordsenheter spelar en viktig roll i språket (Jackendoff 1997; Sag et al. 2002; Baldwin & Kim 2010), men det är ytterst sällan som man i praktiken ser språkbeskrivningar där flerordsenheter behandlas på sina egna villkor, som fullvärdiga medborgare i språket.

I brist på kunskap om hur vårt språkande egentligen fungerar inne i våra huvuden har vi naturligtvis stor frihet i hur vi organiserar våra språkbeskrivningar, om bara slutprodukterna i någon mening innehåller samma information. Alltså är det inget fel i att t.ex. hantera flerordsenheter som syntaktiska fraser som genomgår speciell efterbehandling på grund av sin icke-kompositionella betydelse. Det är heller inte fel att tilldela de ingående orden ”dekompositionell betydelse” (Baldwin et al. 2003), alltså en betydelse som enbart kommer fram i någon eller några få kontexter (se även avsnitt 4.3 nedan).

Därmed torde det även vara fritt fram att se i vilken utsträckning det går att behandla flerordsenheter som ord, vilket är vad som görs i SALDO. På betydelsesidan ställer detta inte till med några större besvär, utan flerordningar beskrivs på samma sätt som alla andra lexikoningångar i SALDO. Logiskt sett kommer vi på det sättet naturligtvis att få färre fall av dekompositionell betydelse för enordningarna; när *av ondo* har en egen lexikoningång, behövs naturligtvis ingen ingång för enbart *ondo*.⁸

På formsidan har flerordningarna i SALDO en grundform, en ordklass och ett böjningsparadigm, precis som de enkla orden. Sålunda är *all världens väg, av födsel och ohejdad vana* och *av och an* samtliga av ordklassen *abm* (flerordsadverb), trots att de som fraser har olika kategorier (substantivfras, prepositionsfras och adverbfras). På samma sätt är *ur gängorna, ute i ogjort väder* och *yr i mössan* av ordklassen *avm* (flerordsadjektiv).

Naturligtvis finns komplikationer på så sätt att många av flerordsuttrycken får ”inre böjning” – *enarmad bandit ~ enarmade banditer* – och i somliga fall (mestadels verb) behöver man räkna med diskontinuerliga former (t.ex. partikelverb) eller rentav att delarna kan byta ordningsföljd (t.ex. stödverbskonstruktioner). Detta för med sig vissa nya krav på de språkanalyskomponenter som ska använda lexikoninformationen, men komplexiteten i analysystemet som helhet blir inte med nödvändighet större, eftersom den syntaktiska analysen blir enklare om den kan erbjudas större färdiganalyserade bitar (Attardi & Dell’Orletta 2008; Kokkinakis

⁸ I en helt digital lexikonresurs finns förstås inga hänvisningsingångar; det är inte svårare att söka på en del av ett uppslag än på hela uppslaget.

2008). Det innebär även införandet av böjningsparadigm som vi inte är vana vid i svensk grammatik, men globalt sett förekommer alla dessa formella mekanismer i ordböjningen i något språk, eventuellt med undantag för regelstyrd ordningsföljdsvariation (Anderson 1985; Aikhenvald 2007).

4.3. Hur många och vilka betydelser finns i lexikonet?

Fenomenet ”dekompositionell betydelse” som berördes i föregående avsnitt kan ses som ett specialfall av ett mer generellt problem, nämligen frågan om hur mycket av en ordanvändningsinstans betydelse som kommer från ordet självt och hur mycket som kommer från kontexten plus generella texttolkningsmekanismer? Beroende på hur man svarar på den frågan sällar man sig till dem som föredrar att laborera med mer detaljerade (”splitters”) eller mer övergripande (”lumpers”) ordbetydelser. Relevant i det här sammanhanget är att det tycks krävas lång lexikografiskolning innan människor klarar av att med god säkerhet och reproducerbarhet tilldela ord i fri text den rätta av många finfördelade betydelser (Kilgarriff 1997; Hanks 2000). Det är välkänt att automatisk lexikal disambiguering brottas med samma problem och empiriska studier tyder på att antagandet av både mer övergripande betydelser och av överlappande betydelser kan vara till hjälp (Erk 2010). Den svenska lexikografiska tradition som vi åtminstone till en del stöder oss på har länge strukturerat ordbetydelser i huvud- och underbetydelser, alltså hierarkiskt (t.ex. *Nationalencyklopedins ordbok* och *Svensk ordbok utgiven av Svenska Akademien*)⁹ och i vårt arbete med ett nytt svenskt ordnät (Borin & Forsberg 2010, 2011, 2014) använder vi oss av graderade synonymiömdömen från en användaruppbyggd lexikonresurs (Synlex; Kann & Rosell 2006). Därmed har vi goda förutsättningar att experimentera med hur användningen av vidare och snävare ordbetydelser inverkar på språkteknologisystems prestanda.

4.4. Hur många lexikon?

Rubrikens fråga hänger ihop med det bara delvis lingvistiska problemet att avgöra var ett språk slutar och nästa börjar. De språkteknologiska tillämp-

⁹ Vid en närmare anblick visar sig hierarkin dock vara skenbar, eftersom ’underbetydelser’ bär många olika slags lexikalisk-semantiska relationer till huvudbetydelserna, introducerade av uttryck som ”spec.” (som antyder hyponymi), ”även allmännare” (som pekar på hyperonymi), ”även bildligt”, etc.

ningar som vi utvecklar ska bland annat kunna hantera texter från många genrer, ämnesområden och tidsperioder med bästa möjliga resultat.

I en ordbok för människor ska ordförrådet vara samtida och i huvudsak allmänspråkligt. För äldre ord eller speciella fackområden får man konsultera äldre ordböcker (eller ordböcker över äldre språkformer) eller facklexikon. Speciellt i det förra fallet kommer det att finnas ett stort överlapp i ordförrådet mellan ordböcker för språkformer från olika tider. Brings *thesaurus* (1930) innehåller totalt omkring 60.000 grundformer fördelade på ett ännu inte närmare bestämt antal betydelser, av vilka uppskattningsvis en fjärdedel till en tredjedel är föråldrade, medan naturligtvis en stor mängd moderna ord saknas. Men det betyder att två tredjedelar av Bring är ”onödiga” i relation till den moderna lexikonresursen SALDO, eller är de det? På vilka grunder säger vi att substantivet *båt* i Bring är detsamma som substantivet *båt* i SALDO? Hur kan och bör vi formellt hantera den typen av utsagor?¹⁰

Eftersom Språkbankens korpusar omfattar text från den fornsvenska perioden och framåt (Borin et al. 2012) och eftersom målet är att automatiskt förse hela textmängden med goda och rika lingvistiska annoteringar, behöver annoteringssystemet tillgång till lexikoninformation för alla tider, genrer och ämnesområden.

Samtidigt vill man å ena sidan undvika redundans i systemet, bland annat eftersom det mycket lätt leder till att motstridig information kommer in. Å andra sidan är som sagt vissa ord – eller värre: vissa användningar av ord – knutna till vissa tider, genrer eller ämnesområden. Det är alltså viktigt att lexikonen inte går i vägen för varandra utan kompletterar varandra på bästa sätt.

Detta är således ett intrikat och intressant metodologiskt problem, eftersom ord eller ordbetydelser inte förändras abrupt över en natt eller ens från ett år till ett annat (Wang 1969). Med andra ord fungerar i praktiken varken en enkel tidsuppmärkning eller etiketter som ”åld.”, utan det blir snarare tal om en sannolikhetsdistribution: Hur troligt är det att ordet *O* uppträder i betydelsen *B* vid tidpunkten *T* i en text av en viss typ?

¹⁰ Ett sätt att se på saken kan förstås vara att följa Leipzigtypologerna och betrakta de språkformer som beskrivs i Bring och SALDO som separata *dokulekter* (se t.ex. Nordhoff & Hammarström 2011). Det för oss dock i sig inte närmare en lösning på problemet med överlappande information.

5. Mot framtiden

Mitt fokus i den här artikeln har legat på lexikon för maskiner, även om jag emellanåt har försökt jämföra dem med lexikon för människor. Som förhoppningsvis har framgått tror jag att de krav som automatisk språkanalys ställer på lexikonresurser ifråga om explicititet, fullständighet, etc., men framförallt deras fundamentalt empiriska natur – de ska ju användas för att analysera faktisk naturlig text – generellt kan föra den lexikografiska metodutvecklingen framåt.

Speciellt passar den moderna språkteknologins fokus på distributionella (statistiska) metoder (Erk 2010) ihop med en ökande medvetenhet om att många – kanske de flesta – språkliga generaliseringar bäst uttrycks i statistiska snarare än kategoriska termer. Detta gäller naturligtvis även generaliseringar på det lexikaliska planet.

Litteratur

- Aikhenvald, Alexandra Y. 2007. Typological distinctions in word-formation. I: Shopen, Timothy (red.), *Language Typology and Syntactic Description (2nd ed.)*. Volume III: *Grammatical Categories and the Lexicon*. Cambridge: Cambridge University Press. S. 1–65.
- Anderson, Stephen R. 1985. Inflectional morphology. I: Shopen, Timothy (red.), *Language Typology and Syntactic Description (1st ed.)*. Volume III: *Grammatical Categories and the Lexicon*. Cambridge: Cambridge University Press. S. 150–201.
- Attardi, Giuseppe & Felice Dell’Orletta 2008. Chunking and dependency parsing. I: *LREC Workshop on Partial Parsing*. Marrakech: ELRA. S. 27–32.
- Baldwin, Timothy, Colin Bannard, Takaaki Tanaka & Dominic Widdows 2003. An empirical model of multiword expression decomposability. I: *Proceedings of the ACL 2003 Workshop on Multiword Expressions: Analysis, Acquisition and Treatment*. Sapporo: ACL. S. 89–96.
- Baldwin, Timothy & Su Nam Kim 2010. Multiword expressions. I: Indurkha, N. & Damerau, F. J. (red.), *Handbook of Natural Language Processing*. Boca Raton: CRC Press. S. 267–292.
- Baroni, Marco & Alessandro Lenci 2010. Distributional memory: A general framework for corpus-based semantics. I: *Computational Linguistics* 36(4). S. 673–721.
- Boguraev, Bran & Ted Briscoe (red.) 1989. *Computational Lexicography for Natural Language Processing*. London: Longman.

- Borin, Lars, Dana Danélls, Markus Forsberg, Dimitrios Kokkinakis & Maria Toporowska Gronostaj 2010. The past meets the present in Swedish FrameNet++. I: *14th EURALEX International Congress*. Leeuwarden: EURALEX. S. 269–281.
- Borin, Lars & Markus Forsberg 2010. From the People's Synonym Dictionary to fuzzy synsets – first steps. I: *Proceedings of the LREC 2010 workshop Semantic relations. Theory and Applications*. Valletta: ELRA. S. 18–25.
- Borin, Lars & Markus Forsberg 2011. Swesaurus – ett svenskt ordnät med fria tyglar. I: *LexicoNordica* 18. S. 17–39.
- Borin, Lars & Markus Forsberg 2014. Swesaurus; or, The Frankenstein Approach to Wordnet Construction. I: *Proceedings of the 7th Global WordNet Conference (GWC 2014)*. Tartu: Global WordNet Association. S. 215–223.
- Borin, Lars, Markus Forsberg & Lennart Lönngren 2013. SALDO: a touch of yin to WordNet's yang. I: *Language Resources and Evaluation* 47(4). S. 1191–1211.
- Borin, Lars, Markus Forsberg & Johan Roxendal 2012. Korp – the corpus infrastructure of Språkbanken. I: *Proceedings of LREC 2012*. Istanbul: ELRA. S. 474–478.
- Bresnan, Joan (red.) 1982. *The Mental Representation of Grammatical Relations*. Cambridge, Massachusetts: MIT Press.
- Bring, Sven Casper 1930. *Svenskt ordförråd ordnat i begreppsklasser*. Stockholm: Hugo Gebers förlag.
- Erk, Katrin 2010. What is word meaning, really? (And how can distributional models help us describe it?). I: *Proceedings of the 2010 Workshop on GEometrical Models of Natural Language Semantics (GEMS 2010)*. Uppsala: ACL. S. 17–26.
- Fellbaum, Christiane (red.) 1998. *WordNet: An Electronic Lexical Database*. Cambridge, Massachusetts: MIT Press.
- Gurevych, Iryna & Elisabeth Wolf 2010. Expert-built and collaboratively constructed lexical semantic resources. I: *Language and Linguistics Compass* 4(11). S. 1074–1090.
- Hammarström, Harald & Lars Borin 2011. Unsupervised learning of morphology. I: *Computational Linguistics* 37(2). S. 309–350.
- Hanks, Patrick 2000. Do word meanings exist? I: *Computers and the Humanities* 34. S. 205–215.
- Hanks, Patrick 2013. *Lexical Analysis: Norms and Exploitations*. Cambridge, Massachusetts: MIT Press.
- Harris, Zellig 1946. From morpheme to utterance. I: *Language* 22(3). S. 161–183.

- Harris, Zellig 1955. From phoneme to morpheme. I: *Language* 31(2). S. 190–222.
- Hoffmann, Thomas & Graeme Trousdale (red.) 2013. *The Oxford Handbook of Construction Grammar*. Oxford: Oxford University Press.
- Jackendoff, Ray 1997. *The Architecture of the Language Faculty*. Cambridge, Massachusetts: MIT Press.
- Jurafsky, Daniel & James H. Martin 2009. *Speech and Language Processing: An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition*. Prentice Hall series in artificial intelligence (2nd ed.). Upper Saddle, New Jersey: Prentice Hall.
- Kann, Viggo & Magnus Rosell 2006. Free construction of a free Swedish dictionary of synonyms. I: *Proceedings of the 15th NODALIDA Conference*. Department of Linguistics, University of Joensuu. S. 105–110.
- Karttunen, L. (1986) Radical lexicalism, CSLI Report No. 86-68, CSLI, Stanford, CA.
- Kilgarriff, Adam 1997. I don't believe in word senses. I: *Computers and the Humanities* 31(2). S. 91–113.
- Kokkinakis, Dimitrios 2008. Semantic pre-processing for complexity reduction in parsing medical texts. I: *Proceedings of MIE 2008*. Amsterdam: IOS Press. S. 217–222.
- Liberman, Mark 2012. Big inaccessible data. Inlägg på Language Log den 4/6 2012. <<http://languagelog.ldc.upenn.edu/nll/?p=3999>>.
- Manning, Christopher D. 2003. Probabilistic syntax. I: Bod, Rens, Jennifer Hay, & Stefanie Jannedy (red.), *Probabilistic Linguistics*. Cambridge, MA: MIT Press. S. 289–341.
- McCawley, James D. 1983. Recension av: *Anna Wierzbicka, Lingua Mentalis: The Semantics of Natural Language*, Sydney & New York, Academic Press, 1980. I: *Language* 59(3). S. 654–659.
- Munro, Robert, Steven Bethard, Victor Kuperman, Vicky Tzuyin Lai, Robin Melnick, Christopher Potts, Tyler Schnoebelen & Harry Tily 2010. Crowdsourcing and language studies: The new generation of linguistic data. I: *Proceedings of the NAACL HLT 2010 Workshop on Creating Speech and Language Data with Amazon's Mechanical Turk*. Los Angeles: ACL. S. 122–130.
- Nordhoff, Sebastian & Harald Hammarström 2011. Glottolog/Langdoc: Defining dialects, languages, and language families as collections of resources. I: *CEUR Workshop Proceedings 783: Proceedings of the First International Workshop on Linked Science 2011*.
- Norvig, Peter 2011. On Chomsky and the two cultures of statistical learning. <<http://norvig.com/chomsky.html>>.
- Reiter, Ehud & Somayajulu Sripada 2002. Human variation and lexical choice. I: *Computational Linguistics* 28(4). S. 545–553.

- Roget, Peter Mark 1852. *Thesaurus of English Words and Phrases*. London: Longman.
- Sag, Ivan, Timothy Baldwin, Francis Bond, Ann Copestake, & Dan Flickinger. 2002. Multi-word expressions: A pain in the neck for NLP. I: *Proceedings of CICLING-2002*. Berlin: Springer. S. 1–15.
- Wang, William S.-Y. 1969. Competing changes as a cause of residue. *Language* 45. S. 9–25.

Tvåspråkig lexikografi – ny teknik, nya funktioner Den digitala ordboken som parallellkorpus

Anna Helga Hannesdóttir

This article illustrates how the bilingual electronic dictionary can function as a parallel corpus and as such provide a basis for various linguistic studies. The dictionary in question is the online multilingual *Islex* dictionary, here used as a bilingual Icelandic-Swedish dictionary. The generous access to the different information categories, encoded in the dictionary, makes the examples just as accessible as the traditional lexical entry, and target language units of different kinds are as easy to reach as the source language units. Here, all the occurrences of the Icelandic idiom *koma í ljós* and their translations are reviewed, and the semantic and syntactic properties of the idiom and their equivalents are briefly described as well as similarities and differences in the information structure in Icelandic vs. Swedish.

1. Inledning

Den tvåspråkiga ordboken har genom den digitala tekniken utökats med nya funktioner. Använd som parallellkorpus kan den läggas till grund för kontrastiva studier av olika egenskaper hos de språk den behandlar, inte minst av drag som ligger på gränsen mellan lexikografi, lexikologi och grammatik.

Här redovisas en begränsad studie, avsedd att illustrera ordbokens möjliga funktion som parallellkorpus. Den ordbok som används som åskådningsexempel är den flerspråkiga nätordboken *Islex*, och de här aktuella språken är isländska och svenska. Inledningsvis ges en kort beskrivning av det teoretiska läget inom lexikografen idag och *Islex* presenteras. Därefter redovisas resultaten av en kontrastiv undersökning med utgångspunkt i det isländska idiomet *koma í ljós* och dess svenska motsvarigheter, följt av en betydligt kortare diskussion av idiomerna *láta í ljós* och *leiða í ljós* och deras motsvarigheter.

2. Bakgrund

En fråga som under de senaste decennierna diskuterats i lexikografiska kretsar har gällt lexikografins status som vetenskaplig disciplin. Grundförutsättningarna är ovedersägliga: verksamheten har uråldriga anor och bygger på starka traditioner utvecklade kring det praktiska lexikografiska hantverket – men hur är det med teorin?

Någon fullständig enighet kan inte sägas råda. Såsom läget vanligen har beskrivits förefaller lägret vara uppdelat i praktiker å ena sidan och teoretiker å den andra. I den nu aktuella lexikografiteoretiska debatten, där funktionalitet och användaranpassning tilldelas en viktig roll, gäller diskussionen främst den enspråkiga lexikografen. Mera sällan framträder den skiljelinje som länge har funnits mellan den enspråkiga och den tvåspråkiga lexikografen. Den här presenterade studien är främst avsedd att illustrera hur den tvåspråkiga ordboken kan användas som utgångspunkt för språkvetenskapliga studier men även att belysa den teoretiska utmaning som den tekniska utvecklingen innebär också för den tvåspråkiga lexikografen.

3. Lexikografins vetenskapliga status

I rapporten från den första konferensen om lexikografi i Norden, anordnad i Oslo 1991, summerar Vatvedt Fjeld situationen för den nordiska lexikografen i allmänhet (Fjeld 1992). Hon konstaterar att lexikografen står inför en brytpunkt mellan traditioner och ackumulerad praktisk kunskap å ena sidan och nya teorier, ny metod och ny teknologi å den andra och att det gäller att förena det gamla och det nya. Praktik och teori måste följas åt: ”det må ikke bli teoriutvikling for teoriens egen skyld, kunnskapen bør føre til at det skapes gode leksikografiske produkter i fremtiden [...]” (Fjeld 1992:6).

Närmare tjugo år senare presenterar Hartmann, vid den tionde konferensen i samma serie, en översikt över lexikografins avancemang inom vetenskapen. Han tar sin utgångspunkt bland annat i den lexikografiska litteraturen, som han menar tenderar att polarisera de båda aspekterna av fältet, den teoretiska och den praktiska (Hartmann 2010). Han låter Atkins & Rundell representera praktikerna och anför, lite tillspetsat, att de i sin handbok i lexikografi ”take a strongly anti-theoretical position” (Hartmann 2010). I själva verket tillmäter Atkins & Rundell teoretiska landvinningar inom framför allt semantik, korpuslingvistik och lexikologi stor betydelse för den praktiska lexikografen (Atkins & Rundell 2008:130ff.). På teoretikernas barrikader ställer Hartmann H.E. Wiegand, som i sina ar-

beten ”denies practical lexicography the status of a discipline” (Hartmann 2010:19).

Att det tycks föreligga en uppdelning av fältet kan således lätt konstateras. En av dem som fört diskussionen vidare är Rundell. I sin genomgång av lexikografiska teorier och deras relevans för den lexikografiska praktiken karakteriserar han Wiegands arbeten som *metalexikografi*, avsedda för andra metalexikografer snarare än för lexikografer. Rundell deklarerar tydligt sin syn på metalexikografins relation till lexikografen: ”... many aspects of metalexigraphy (in particular, the passion for complex typologies, minute categorization, and exhaustive nomenclature) can often seem quite irrelevant” (Rundell 2012:54). Däremot framhåller han än en gång vikten av semantiska, korpuslingvistiska och lexikologiska utgångspunkter för den lexikografiska praktiken. Skiljelinjen tycks idag inte i första hand gå mellan lexikografi och lingvistik utan snarare mellan lexikografi och metalexikografi.

I en artikel från 1989 ger Wiegand sin syn på lexikografins vetenskapliga status, en översikt som sedan dess ofta fått bilda utgångspunkt för lägesbeskrivningar av den typ som återopats ovan. Han är klar över den viktiga roll datorerna och datatekniken måste tilldelas inom lexikografen. Avslutningsvis pekar han mot den framtid vi nu befinner oss i:

Die hier gegebene Statusbestimmung der gegenwärtigen Lexikographie ist so angelegt, daß eine weiter fortschreitende Einbindung des Computers in den lexikographischen Prozeß eine Änderung erst dann erforderlich macht, wenn eine konkrete metalexikographische Utopie zum Alltag werden sollte, nämlich die, daß die allermeisten lexikographischen Prozesse, die neu geplant werden, nicht mehr zu Sprachwörterbüchern in Buchform führen werden (Wiegand 1989:272).

Huruvida Wiegand redan 1989 förutsåg publicering på Internet är tveksamt men han var redan då på det klara med att den digitala ordboken inte skulle begränsas av bokens format.

Nu befinner vi oss i den framtid som Wiegand pekar fram emot. Nätet och elektronisk publicering erbjuder inte bara en ny arena för lexikografisk verksamhet. Att den också har banat väg för nya typer av lexikografiska produkter och för nya aktörer inom lexikografen framgår av exempelvis Sköldberg & Mattsson (i denna volym). De aspekter av elektronisk lexikografi och elektronisk publicering som uppmärksammas mest är dels av språkteknologisk natur, dels gäller de relationen mellan lexikografiska produkter och deras användare. De språkteknologiska aspekterna diskuteras av Borin (i denna volym). Inom den elektroniska lexikografi (elexikografi) som vänder sig direkt till mänskliga användare har tankarna

på anpassning av den lexikografiska produkten till användarens önskemål och behov drivits längst inom den så kallade *funktionsteorin*. Denna presenteras här kort i relation till den tvåspråkiga lexikografen.

4. Den tvåspråkiga lexikografen, användarperspektivet och funktionsteorin

De lexikografiska produkternas anpassning till en tydligt definierad målgrupp problematiserades först inom ramen för den tvåspråkiga lexikografen. Främst var det då just relationen mellan ordbokens båda språk och presentationen av dessa å ena sidan och å den andra användarnas behov som kom att uppmärksammas och diskuteras (Kromann et al. 1984). Utgångspunkten för hur de behandlade språken, vart och ett, skall beskrivas blev dels deras funktion som källspråk respektive målspråk i ordboken, dels huruvida respektive språk utgör modersmål (L1) eller främmande språk (L2) för de tilltänkta användarna. För varje given konstellation av två språk konstaterade man, behövs åtminstone fyra olika ordböcker, utformade med utgångspunkt i den avsedda målgruppens relation till de aktuella språken, vilket av dem som är modersmål och vilket som är det främmande språket. Till detta kommer alltså språkets funktion i ordboken; dvs. om det är källspråk eller målspråk. Detta användarperspektiv på ordböckerna har sedan blivit en självklarhet inom den tvåspråkiga lexikografen, och även enspråkiga ordböcker utformas med hänsyn till den målgrupp de vänder sig till och den funktion de är avsedda att fylla (se exempelvis Malmgren 2009).

Användarperspektivet har kommit att utvecklas ytterligare och då i relation till de nya förutsättningarna och de nya medierna, framför allt inom den så kallade Århuskolan och dess version av funktionsteorin. I denna tappning av teorin betonas en strikt, funktionell anpassning till användarens i princip individuella behov av ett lexikografiskt verktyg, ett verktyg som skall vara avpassat till användarens specifika behov av information. Vidare menar förespråkare för funktionsteorin att lexikografen måste frigöra sig från språkvetenskapen: ”Current lexicographic theory owes much to the efforts by linguists, but this does not imply that current and future lexicographic theory must play a subordinate role in its relation to either linguistics or linguists” (Gouws 2011:22). Istället för tillämpningen av språkvetenskapliga teorier och metoder skall e-lexikografen ta konsekvenserna såväl av informationsteknologins betydelse för disciplinen som av de sociala aspekter av informationssökning och informationshantering som också framhålls inom Århuskolan: ”Electronic dictionaries are typical products of what we call knowledge and information society, which de-

mand a different approach to electronic lexicography from the one which started at the beginning of the electronic-dictionary age” (Fuentes-Olivera & Bergenholtz 2011a:1). Lexikografins samhällsvetenskapliga karaktär betonas och lexikografi skall snarare betraktas som en självständig samhällsvetenskaplig eller informationsvetenskaplig disciplin än som en språkvetenskaplig verksamhet. De lexikografiska produkternas relevans som utgångspunkt för språkvetenskapliga studier av olika slag uppmärksammas däremot inte.

Funktionsteorin såsom den har utvecklats på senare tid har sina förespråkare och sina kritiker. De senare menar att man driver användarperspektivet för långt och att det snarast är frågan om teoriutveckling för teorins egen skull. Det användarperspektiv som utgör kärnan i funktionsteorin är i själva verket väl etablerat i den lexikografiska teorin såväl som i den lexikografiska praktiken. Århuskolan kritiserar dels för att ha drivit kraven på individuell anpassning för långt (se t.ex. Grønvik 2012), dels för att underskatta språkvetenskapens betydelse både för datalingvistik och för lexikografin: ”If we had a database containing all the facts and generalisations about the behaviour of all the words and phrases of the language, optimally structured, then we wouldn’t need linguists. But we don’t” (Kilgarriff 2012:29).

Det är lätt att konstatera att grundvalarna nu skakar för det användarperspektiv som utvecklats inom den tvåspråkiga lexikografin, dvs. språkens funktionsuppdelning i källspråk respektive målspråk. I den elektroniska ordboken kan enheter i båda de behandlade språken på lika villkor utgöra ingångar i ordboken. Konsekvenserna av upplösningen av dessa grundläggande funktioner borde bli att båda språken beskrivs på lika villkor, dvs. både som ”källspråk” och som ”målspråk” och att beskrivningen av dem, vart och ett, i alla avseenden utgår ifrån en kontrastiv analys av de ömsesidiga relationerna dem emellan. Den teoriutveckling som krävs för att beskrivningen och presentationen av ordbokens båda språk skall tillgodose användarnas behov vare sig de är L1- eller L2-talare är ännu bara påbörjad. Samtidigt inbjuder de nya publikationsformerna till helt nya funktioner för den tvåspråkiga lexikografin.

Med utgångspunkt i den flerspråkiga internetordboken *Islex* kommer här vissa av den nya lexikografins möjligheter att belysas. Det som fokuseras är den generösa åtkomststrukturen och de möjligheter den erbjuder till kontrastiva studier av olika slag. Här följer först en kortfattad presentation av Islexprojektet såsom det bedrevs och av den färdiga ordboken.

5. *Islex*-ordboken

Den flerspråkiga ordboken *Islex* har från det att projektet startade varit avsedd för publicering på Internet och det redaktionella arbetet har varit nätbaserat. Ordboken innehåller isländska, danska, norskt bokmål, nynorska och svenska. Arbetet har bedrivits i samarbete mellan akademiska institutioner på Island, i Sverige, Norge och Danmark (se islex.se). Senare har såväl Fróðskaparsetur Føroya som Helsingfors universitet anslutit sig till projektet och nu inkluderas även färöiska och finska i ordboken.

Den isländska redaktionen har ombesörjt utformningen och utvecklingen av den databas som ligger till grund för *Islex* och för urval och bearbetning av det isländska material som utgör ordbokens källspråk. Redaktionen har vidare ansvarat för det redigeringsystem som använts liksom redaktionsgränssnitt och till slut användargränssnittet. De övriga redaktionerna har ansvarat för arbetet med sina respektive språk som målspråk till det isländska källspråket. Redaktionernas arbete har i stort sett helt bekostats med offentliga medel från respektive stat och ordböckerna är kostnadsfritt tillgängliga på Internet. De nås på adressen *islex* följt av respektive lands suffix: <islex.dk>, <islex.is>, <islex.no> och <islex.se>.

Projektet inleddes 2006 och visst arbete pågår fortfarande i Sverige och Norge medan det redaktionella arbetet med danskan har avslutats. Den isländska redaktionen fortsätter sitt arbete med underhåll av databasen och med vissa kompletteringar och korrigeringar av ordmaterialet samtidigt som den bistår den färöiska och den finska redaktionen, där det redaktionella arbetet fortfarande pågår. Olika aspekter av *islex*-projektet har belysts såväl vid Euralex-konferensen (Sigurðardóttir et al. 2008) som vid konferensen om lexikografi i Norden (se exempelvis Rauset et al. 2012 samt Jónsdóttir & Úlfarsdóttir 2012). Ordboken har också nyligen fått en fyllig presentation i *Orð og tunga* 15 (Úlfarsdóttir 2013) och den recenseras i *LexicoNordica* 20 (Sanders 2013).

5.1. Målgruppen för *Islex*

Islex främsta syfte är att fungera som en produktionsordbok för islänningar. Källspråket är deras modersmål och målspråket ett främmande språk: ordboken fungerar då som en L1 > L2-ordbok. Dessutom är den avsedd som en avkodningsordbok för danskar, norrmän och svenskar. Det isländska källspråket är ett främmande språk för dessa användare och deras respektive modersmål utgör ordbokens målspråk. För dem fungerar ordboken då som en L2 > L1-ordbok. I det avseendet är *Islex* en traditionell, tvåspråkig lexikografisk produkt mellan två ”små” språk, där så många funktioner

som möjligt bör fyllas av ett och samma arbete – traditionellt alltså inom ramen för bokformatet. En sådan ordbok är exempelvis *Svensk-isländsk ordbok* (1983); i inledningen anges att den ”är avsedd både för svenska och isländska läsare” (s. IX).

5.2. Innehållet i *Islex*

Ordboken innehåller nästan 50 000 isländska uppslagsord. Förutom uppslagsordet innehåller artiklarna sammanlagt omkring 30 000 språkprov och dessutom rikligt med kollokationer, idiom och andra fraser och flerordsenheter av olika slag. Dessa illustrerar uppslagsordets egenskaper och användning. Härfter används beteckningen *språkprov* för alla dessa typer av illustrerande material. Ordförrådet liksom också språkproven i *Islex* är i stor utsträckning korpusbaserat (Úlfarsdóttir 2013).

Uppslagsorden förses med ordklassmarkering och i substantiv-, adjektiv-, pronomen- och verbartiklarna finns det en länk till den morfologiska fullformsdatabasen *Beygingarlýsing íslensks nútímamáls*, <bin.arnastofnun.is>. Dessutom finns ljudillustrationer; vid samtliga uppslagsord ges inspelat uttal av uppslagsformen, vare sig det rör sig om ettordslemman eller flerordsenheter (se bild 1). Ordboken är också rikt illustrerad med bildmaterial, och även en del rörliga illustrationer förekommer.



BILD 1. *Artikeln stökkönd med svenska som metaspråk och svensk ekvivalent.*

Till allt källspråksmaterial redovisas motsvarigheter på målspråken. Dessutom har de danska och svenska ekvivalenterna försetts med böjningsuppgifter, som framträder när muspekaren placeras på ordet. Möj-

ligheten finns att även foga inspelat uttal till ekvivalenterna. Från artiklarna finns hyperlänkar främst inom ordboken och till den morfologiska databasen. I fråga om utrymme har *Islex* sålunda sprängt de strikta ramar som bokformatet utgör.

5.3. De olika språken och presentationsformen

I användargränssnittet erbjuds olika visningsalternativ. Standardinställningen för metaspråk följer den valda hemsidan; om ordboken öppnas i *islex.se* är metaspråket svenska. Inne på sidan kan man dock växla till isländska, danska, bokmål eller nynorska.

landsmaður subst m

 **uttale**

 **BØYING**

lands-maður

 indbygger (i et land)

 innbygger (i et land), landsmann, borger

 innbyggjar (i eit land), landsmann, ibuar, borgar

 invånare (i ett land)

 íbúgvi

 asukas

landsmenn gengu til kosninga í vor

 indbyggerne gik til valg i foråret

 innbyggerne gikk til valg i vår

 innbyggjarane gjekk til val i vår

 invånarna gick till val i våras

 íbúgvarnir fóru á val í vár

 asukkaat äänestivät (asiasta) keväällä

BILD 2. Artikeln **landsmaður** med nynorska som metaspråk och ekvivalenter och översättningar på alla målspråken.

Vidare finns det två alternativ beträffande vilken ordbok som tillgängliggörs. Det automatiskt valda är att den tvåspråkiga ordboken visas, det isländska källspråket och ett av målspråken (bild 1). För att få fram den flerspråkiga ordboken, där alla målspråken exponeras samtidigt (bild 2), väljer användaren alternativet *alla*.

Även om varken finska eller färöiska ännu har publicerats med en egen adress och inte är sökbara visas det finska och färöiska material som finns i databasen när visningsalternativet *alla* har valts.

5.4. Sökmöjligheterna

Datatekniken har revolutionerat all lexikografi, den enspråkiga och kanske ännu mer den tvåspråkiga. De förutsättningar som gällde för den tvåspråkiga lexikografin publicerad i bokform är inte längre aktuella. Det som mest skiljer den digitala ordboken från tryckta ordböcker är de alternativa ingångarna i ordbokstexten. Genom de generösa sökmöjligheter som tekniken erbjuder har distinktionen källspråk–målspråk i princip upphävts. Därmed har ordboken också berikats med funktioner som vida överträffar det som tidigare har förknippats med ordböcker och lexikografi. Även om *Islex* är utarbetad med utgångspunkt i den tvåspråkiga ordbokens traditionella funktionsuppdelning mellan källspråk och målspråk, utnyttjas teknikens sökmöjligheter på ett sätt som gör uppslagsenheterna och ekvivalenterna till likvärdiga ingångar i ordboken.

Användargränssnittet erbjuder två huvudalternativ för sökningar: *enkel sökning* respektive *avancerad sökning*. Vid enkel sökning redovisas bara träffar bland de isländska uppslagsorden och de ekvivalenter som fogats till dessa. Alternativet avancerad sökning tillåter dessutom fritextsökning såväl på enstaka ord som på fraser och samförekomster av ord som finns i språkproven, både på isländska och på det aktuella målspråket. En sökning på strängen ”gick till val i våras” i *all artikeltext* (se bild 3 nedan) leder sålunda till den artikel som visas på bild (2), *landsmaður*.

5.5. Språkproven i *Islex* som parallellkorpus

Tillgången till parallellkorpusar har berikat såväl den tvåspråkiga lexikografin som den kontrastiva lingvistik. Att parallellkorpusarna även utgör en stark konkurrent till ordboken som hjälpmedel vid översättning har Gellerstam med all tydlighet visat (Gellerstam 1999). Hans utgångspunkt är en jämförelse mellan en svensk-engelsk parallellkorpus och den tryckta *Stora engelska ordboken* (1989). När det gäller parallellkorpusar mellan isländska och de andra nordiska språken är tillgången inte särskilt stor. En liten isländsk-skandinavisk parallellkorpus på ca 115 000 ord bestående av texter från Nordiska rådet har emellertid skapats av Jansson, som också har demonstrerat dess förtjänster i det lexikografiska arbetet (Jansson 2012).

Som framgått ovan innehåller *Islex* en mängd språkprov av olika slag. Dessa textsträngar och översättningarna av dem kan betraktas som en liten, flerspråkig parallellkorpus. Eftersom vart och ett av språkproven utgör en komplett liten text fungerar de utmärkt som källa för språkvetenskapliga studier av vissa egenskaper i de olika språken. När det gäller språkprov i

ordböcker är det framför allt kriterier för urvalet av dem som har diskuterats och då främst urvalet av språkprov i enspråkiga ordböcker (se exempelvis Kilgarriff et al. 2008). Hur språkproven i tvåspråkiga ordböcker skall hanteras berörs hastigt av Atkins & Rundell (2008:506ff.), som betonar vikten av det kontrastiva perspektivet och att språkproven väljs för att illustrera relationen mellan de aktuella språken. Eftersom det isländska ordförrådet i *Islex* redan från början har varit avsett att utgöra källspråk för flera målspråk har de valts enbart med utgångspunkt i det isländska uppslagsordets egenskaper. Hur väl de illustrerar relationen mellan isländskan och det enskilda målspråket är därför i princip slumpmässigt. De utgör emellertid såväl formellt som innehållsligt möjliga isländska fraser som översätts till ekvivalenta, idiomatiska motsvarigheter på respektive målspråk. Det bör dock påpekas att dessa översättningar, liksom översättning i allmänhet, alltid har tillkommit som en respons på ett semantiskt innehåll som redan fått en språklig form. Eftersom språkprovets semantiska, pragmatiska och formella egenskaper inte alltid är möjliga att fånga i en översättning, är det inte givet att det alternativ som lexikografen valt är det naturligaste sättet att uttrycka språkprovets innehåll sett ur ett enspråkigt perspektiv. Där-
emot bildar både källtexten och översättningen ett komplett sammanhang, utan syftning och referenser utanför den korta text som utgör språkprovet.

Nedan redovisas resultatet av en mindre, kontrastiv studie där tre isländska konstruktioner och deras realisation på svenska granskats. Förutom det rent lexikaliska perspektivet på de aktuella ekvivalensrelationerna framträder även intressanta drag i språkens syntax och informationsstruktur. Observera att syftet med den analys som här presenteras inte är att i detalj kartlägga och beskriva de skillnader mellan isländska och svenska som kan observeras vid jämförelsen. Det är snarare frågan om att illustrera hur materialet i *Islex* kan utnyttjas för kontrastiva studier av olika slag.

6. Tre isländska idiom

De konstruktioner som undersökts utgår från prepositionsfrasen *í ljós* 'i ljus' som tillsammans med verben *koma* 'komma', *láta* 'låta' och *leiða* 'leda' bildar idiomerna *koma í ljós*, *láta í ljós* och *leiða í ljós*. Verbet kan uppträda i sina olika former, medan prepositionsfrasen däremot är konstant. Av de tre idiomerna är *koma í ljós* det som förekommer oftast i *Islex* och det är också det idiom som uppvisar störst variation när det gäller konstruktionssättet. Därför ägnas det här störst uppmärksamhet. *Láta í ljós* och *leiða í ljós* behandlas endast kortfattat.

6.1. *Idiomens presentation i sina artiklar*

Alla tre idiomerna redovisas som just idiom försedda med komplement i artikeln *ljós* (ex. 1). Både *koma í ljós* och *leiða í ljós* ingår i levande språkprov som har formen av en fullständig sats, medan *láta í ljós* illustreras med hjälp av ett s.k. dödt språkprov, dvs. en formulering utan subjekt och med verbet i infinitiv. Som allt annat material i *Islex* förses alla slag av illustrativt språkmaterial, dvs. bruksexempel, kollokationer, idiom och andra flerordsförbindelser, med motsvarigheter på målspråken. Här redovisas bara isländskan och svenskan. Idiomerna och de svenska enheter som uppstår ekvivalensrelationen kursiveras. I exempel (1) förses de isländska språkproven också med lexikalisk svensk glossning. Vinkelparentes sätts i ordboken runt material som illustrerar antingen en syntaktisk struktur som är representativ för den enhet som är i fokus, eller den typ av semantiskt komplement som vanligen uppträder i den aktuella konstruktionen. I bästa fall kombineras båda dessa funktioner så att såväl den syntaktiska valensen som de semantiska selektionsrestriktionerna illustreras.

- | | | |
|-----|------------|---|
| (1) | källspråk: | <i>það kom í ljós að</i> <ég hafði misskilið þetta> |
| | glossning: | [det kom i ljus att jag hade missförstått detta] |
| | målspråk: | det <i>visade sig</i> att <jag hade missförstått detta> |
| | källspråk: | <i>láta í ljós</i> <álit sitt> |
| | glossning: | [låta i ljus åsikt sin] |
| | målspråk: | <i>uttrycka</i> <sin åsikt>, <i>säga</i> <sin mening> |
| | källspråk: | <rannsóknin> <i>leiðir í ljós að</i> <hann er smitaður> |
| | glossning: | [undersökningen leder i ljus att han är smittad] |
| | målspråk: | <undersökningen> <i>visar</i> att <han har smittats> |

Idiomerna *koma í ljós* konstrueras här med formellt subjekt *það* och ett egentligt subjekt i form av motsvarigheten till en att-sats: *að ég hafði misskilið þetta*. Också *leiða í ljós* presenteras i en fullständig sats med subjektet *rannsóknin* och ett objekt som utgörs av en att-sats: *að hann er smitaður*. Idiomerna *láta í ljós*, som försetts med två ekvivalenta svenska uttryck, redovisas med ett objekt som enda komplement: *álit sitt*. Syftet med döda språkprov är som sagt att visa uppslagsordets generella egenskaper. Fördelarna med döda språkprov i tvåspråkig lexikografi är att de bättre än levande språkprov illustrerar just den enskilda enhetens lexikaliska egenskaper, utan distraktorer och begränsningar av det slag som även den kortaste kontext kan innebära.

Idiomen *láta í ljós* och *leiða í ljós* redovisas också i samband med att respektive verb behandlas, dvs. i artiklarna *láta* och *leiða* (exempel 2); båda har formen av döda språkprov och båda har försetts med två ekvivalenta svenska uttryck. Här och i det följande ges den svenska motsvarigheten för det mesta utan glossning.

(2) *láta* <skoðun sína> *í ljós*
vísa <vad man tycker>, *säga* <sin mening>

leiða <þetta> *í ljós*
påvísa <detta>, *avslöja* <det här>

Idiomet *koma í ljós* redovisas inte i den omfattningsrika artikeln *koma* och inget av idiomerna förekommer i den artikel som behandlar prepositionen *í*.

6.2. Fritextsökningen: idiomerna i all artikeltext

En sökning på strängen *í ljós* i all artikeltext (bild 3) i *Islex* ger sammanlagt 49 träffar. Det är inte möjligt att koncentrera sökningen enbart på formen *ljós*; fritextsökningen ger därför även former av *ljós* som inte är relevanta här. I sökningen kan visserligen ordslut markeras med blanksteg, men eftersom blanksteg inte finns inlagt finalt i språkproven utesluter en sådan sökning träffar där *ljós* förekommer finalt (som i exempel 2 ovan). Sökningar på *í ljós* i all artikeltext ger därför 11 förekomster som inte hör till något av de tre idiomerna. Dessa träffar kan bestå i olika böjningsformer av adjektivet och substantivet (*ljósum*, *ljósi*), av substantivet *ljós* som led i sammansättningar (*í ljósmyndinni*; 'på fotografiet') eller av prepositionsfrasen i andra sammanhang än de idiom som det rör sig om här. De tre idiomerna förekommer i sammanlagt 38 språkprov varav 22 innehåller *koma í ljós*, 9 *láta í ljós* och 7 *leiða í ljós*.

The screenshot shows the search interface in Islex. At the top, there is a navigation bar with letters: á ð é í ó ú ý þ æ ö ä å ø and a link for 'Enkel sökning'. Below this is a search input field containing 'í ljós' and a 'Sök' button. Under the search field, there are two columns of checkboxes for filtering the results:

- uppslagsord
- ekvivalent
- all artikeltext
- varianttålig
- danska
- bokmål
- nynorska
- svenska

At the bottom left of the filter area, the word 'sökning' is displayed.

BILD 3. Avancerad sökning på *í ljós* i "all artikeltext" i *Islex*.

7. Idiomet *koma í ljós*

Den vanligaste fasta konstruktionen med *í ljós* är, som framgår ovan, tillsammans med verbet *koma*: *koma í ljós*. I ett kontrastivt isländsk-svenskt perspektiv är det ganska lätt att upprätta ekvivalens mellan det isländska idiomet och en lexikalisk enhet i svenskan. Här nedan kommer först ekvivalensen såsom den upprättas mellan det isländska *koma í ljós* och de svenska översättningarna att diskuteras (7.1–7.3). Därefter redovisas hur isländskans syntaktiska egenskaper realiserar i de svenska översättningarna (7.4) och slutligen de skillnader i de båda språkens informationsstruktur som kan urskiljas med utgångspunkt i de parallella textsträngarna (7.5).

7.1. Ekvivalentparet *koma í ljós/visa sig*

Det uttryck som oftast förekommer som ekvivalent till *koma í ljós* är det reflexiva *visa sig*, nämligen i 12 av de 22 förekomsterna. Subjektet är i 11 av dessa 12 fall abstrakt medan det i ett fall är konkret. De abstrakta subjekten består i sin tur i sju fall av en bisats varav sex utgörs av en att-sats. I två av dessa språkprov kompletteras att-satsen av ett spetsställt formellt subjekt *það* (ex. 3 och 4). Även översättningarna konstrueras på ett parallellt sätt med det formella subjektet *det*. De enskilda språkproven kommer i det följande att upprepas allteftersom nya perspektiv anläggs. Som exempel återges de då vanligen med ny numrering.

- (3) *það kom í ljós að* <ég hafði misskilið þetta>
det visade sig att <jag hade missförstått detta>
- (4) *það kom í ljós að* hún var prýðilega pennafær
det visade sig att hon var en utmärkt skribent

Av de övriga fyra språkproven där subjektet utgörs av att-bisats inleds tre av ett adverbial bestående av prepositionsfraser (ex. 5–7), medan ett av dem (ex. 8) illustrerar isländskans möjlighet att spetsställa ett infinit huvudverb.

- (5) við nánari aðgæslu *kom í ljós að* bíllinn var bensínlaus
 när vi tittade närmare efter *visade det sig att* vi hade fått bensinstopp
- (6) við nánari athugun *kom í ljós að* við erum skyldar
 vid närmare undersökning *visade det sig att* vi är släkt

- (7) til að kóróna allt saman *kom í ljós* að bíllinn var stolinn
till rága þá allt *visade* det *sig* att bilen var stulen
- (8) *komið* hefur *í ljós* að börn sjúga oft á sér hendurnar í móðurkviði
det har *visat sig* att barn ofta suger þá sína händur redan i moderlivet

Förutom de sex att-bisatserna förkommer en interrogativ bisats (ex. 9) som subjekt.

- (9) í síðari hálfleik *kom í ljós* hve mannválið var mikið í liðinu
i andra halvlek *visade* det *sig* vilka utmärkta spelare laget består av

I samtliga översättningar kompletteras subjektbisatsen med ett formellt subjekt, *det*, medan subjektfunktionen i de isländska källtexterna endast upprätthålls av bisatsen. Denna skillnad är dock ett utslag av svenskans generella subjektstvång och inte av konstruktionsrestriktioner specifika för just *visa sig*.

I exempel (5) som upprepas nedan, har den svenska lexikografen fört in en agent i översättningen av det isländska språkprovet vilket lett till en mera aktiv motsvarighet än vad som är nödvändigt med utgångspunkt i den isländska förlagan. Den isländska prepositionsfrasen *við nánari aðgæslu* har ersatts av en temporal bisats med personligt pronomen, *vi*, som subjekt och ett finit verb *tittade*. Matrissatsen har utökats med det formella subjektet *det* och det personliga pronomenet återkommer som subjekt i att-satsen.

- | | | | |
|-----|--|----------------|-------------------------------------|
| (5) | við nánari aðgæslu | kom í ljós | að bíllinn var bensínlaus |
| | [vid närmare granskning | kom i ljus | att bilen var utan bensin] |
| | när <i>vi</i> <i>tittade</i> närmare efter | visade det sig | att <i>vi</i> hade fått bensinstopp |

Denna modifikation kan möjligen tillskrivas semantiska faktorer. I svenskan motsvaras frasen *vera/verða bensínlaus* snarare av *fá bensinstopp* än av *vara/bli utan bensin* och rollen PATIENT till frasen *fá bensinstopp* innehas oftare än i isländskan av ett mänskligt subjekt (som refererar till bilförare, passagerare etc.) än av ett inanimat (t.ex. ett fordon). Den metonymiska relation som råder mellan bilen och dess förare och passagerare är dock möjlig att uttrycka på båda dessa sätt i båda språken: *bíllinn var/við vorum bensínlaus* respektive *bílin/vi fick bensinstopp*.

Övriga språkprov som försetts med ekvivalenten *visa sig* inleds i den isländska versionen med subjektet. I exempel (10) utgör hela satsen *þetta kemur í ljós* ett typiskt, fakultativt komplement till ett annat idiom: *þegar fram líða stundir*. I exemplen (11) och (12) består subjektet av tunga nominalfraser medan subjektet i exempel (13) och (14) består av ett enkelt substantiv i bestämd form.

- (10) <þetta kemur í ljós> þegar fram líða stundir
<det visar sig> med tiden
- (11) miklir listamannshæfileikar hans komu snemma í ljós
hans stora konsthæfileikar hans visade sig tidigt
- (12) stórvægilegir gallar á húsinu hafa komið í ljós
det har visat sig att huset har allvarliga fel
- (13) gallinn kom síðar í ljós
felet visade sig senare
- (14) <ég tók mynd> um leið og <fuglinn kom í ljós>
<jag tog ett kort> så fort <fågeln visade sig>

I samtliga fall är ekvivalensen mellan *koma í ljós* och *visa sig* oproblematiske. Varken det isländska idiommet eller den svenska reflexivkonstruktionen *visa sig* tycks vara underkastade särskilt stränga semantiska restriktioner när det gäller subjektets egenskaper. Det isländska subjektet tycks dock främst vara inanimat och även om de flesta subjekten här är abstrakta illustrerar *fuglinn* i exempel (14) att de även kan vara animata och konkreta.

7.1.1. Det omvända perspektivet: ekvivalenter till *visa sig* i *Islex*

Här ovan har utgångspunkten varit de isländska språkprov där *koma í ljós* ingår och som har fått en översättning som innehåller *visa sig*. I det elektroniska formatet fungerar som sagt den tvåspråkiga ordbokens båda språk som ingångar i ordbokstexten. Lika enkelt som det är att få fram alla förekomster av *koma í ljós* i *Islex* är det därför att få fram alla förekomster av *visa sig*.

En fritextsökning på de olika formerna av *visa sig* ger sammanlagt 79 förekomster, dvs. betydligt fler än de 12 som redovisats ovan. I dessa fall, som ekvivalent till *koma í ljós*, kan noteras den relativt sparsamma förekomsten av enkla, spetsställda subjekt av den typ som finns i exempel (10) *þetta*, (13) *gallinn* och (14) *fuglinn* ovan. Framför allt är det personbetecknande subjekt som i denna kontext lyser med sin frånvaro. En genomgång av de sammanlagt 79 förekomsterna av *visa sig* i översättningar av språkproven och en granskning av de isländska formuleringar som ger upphov till dessa avslöjar en syntaktisk asymmetri mellan *koma í ljós* och *visa sig* när det gäller konstruktionernas valens. *Koma í ljós* kan, till skillnad från *visa sig*, inte konstrueras med nexusingativ. Så framgår det att exempelvis

den isländska utgångspunkten för översättningen ”professorn visade sig *vara en bra kock*” inte innehåller idiomet *koma í ljós* utan dess partiella närsynonym *reynast*: ”prófessorinn reyndist *vera góður kokkur*”. Av de två närsynonymerna *reynast* och *koma í ljós* förekommer *reynast* mer än dubbelt så ofta som *koma í ljós*, dvs. i 49 språkprov, ofta just i strukturer med nexusinfinitiv. *Reynast* i språkproven har i 35 fall fått motsvarigheten *visa sig* i den svenska översättningen.

7.1.2. Kontrastivt perspektiv i den enspråkiga beskrivningen

Det kontrastiva perspektiv som är kärnan i den tvåspråkiga lexikografin kan bidra till nya insikter som också gagnar den språkinterna beskrivningen. En distinktiv semantisk beskrivning kan belysa fina betydelskillnader mellan närsynonymer som exempelvis *koma í ljós* och *reynast*. Vidare illustrerar diskussionen om ekvivalensrelationerna mellan *koma í ljós* och *reynast* å ena sidan och *visa sig* å den andra behovet av en kontrastiv lexikalisk teori och modeller för hur partiell ekvivalens kan hanteras i den digitala, tvåspråkiga ordboken.

7.2. Det isländska idiomet *koma í ljós* och det svenska komma i dagen

Som framgått ovan är den vanligaste ekvivalenten till idiomet *koma í ljós* inte ett idiom utan ett reflexivt verb. I svenskan finns emellertid ett idiom som kan etableras som ekvivalent till det isländska idiomet: *komma i dagen*. Det som kan förefalla anmärkningsvärt är att detta idiom, som uppvisar stora likheter såväl semantiskt som syntaktiskt med *koma í ljós*, bara i två fall, exempel (15) och (16), ges som översättningsekvivalent.

- (15) ýmsar fornminjar *komu í ljós* við uppgröftinn
vid utgrävningen *kom* diverse fornlämningar *i dagen*
- (16) við raskið í túninu *komu í ljós* merkar fornminjar
när man grävde i slåttervallen *kom* intressanta fornminnen *i dagen*

En förklaring till det asymmetriska förhållandet kan naturligtvis vara de enskilda lexikografernas preferenser, men skillnaden kan också vara pragmatiskt betingad. Sålunda förefaller *koma í ljós* vara stilistiskt neutralt och frekvent förekommande i såväl tal som skrift medan ett nedslag i ett urval av materialen i Språkbanken ger vid handen att *komma i dagen*

har en något skriftspråklig karaktär och att det inte förekommer lika ofta i vardaglig, informell svenska <Språkbanken.gu.se>.¹

Även om båda idiomerna kan konstrueras med flera olika slags subjekt tycks selektionsrestriktionerna för det svenska idiomets konkreta betydelse vara betydligt striktare än för *koma í ljós*. Av förekomsterna i de material i Språkbanken som redovisas ovan kan utläsas, att de konkreta subjekt som uppträder tillsammans med *komma i dagen* i första hand betecknar föremål eller formationer som från att ha varit dolda i marken eller i landskapet blir synliga. Den svenska motsvarigheten till de isländska språkproven i exempel (15) och (16) ovan, där det är frågan om fornminnen som grävs fram, illustrerar tydligt denna användning av det svenska idiomerna. Däremot redovisas i *Islex* inget exempel på den överförda betydelsen hos *komma i dagen* i en konstruktion med ett abstrakt subjekt av typen ”Vid inspektioner ombord har allvarliga brister kommit i dagen” (GP 1994, <<http://spraakbanken.gu.se/korp/>>). Korpussökningen visar att den användningen trots allt är betydligt vanligare än den konkreta.

I exempel (16) har de prepositionsfraser som bildar det inledande adverbialiet i det isländska språkprovet återgivits i form av en temporal bisats med utsatt mänskligt subjekt, *man*. I stället för verbalsubstantivet, *raskið* ’rubbingen’, har översättningen kompletterats med ett finit verb, *grävde*.

7.3. Andra ekvivalenter

Två av förekomsterna av *koma í ljós* förses med den ekvivalenta frasen *bli synlig*, exempel (17) och (18). I båda fallen är det fråga om konkreta subjekt, *stjörnurnar* respektive *þúst*, som rent konkret blir synliga för ögat.

(17) himinninn dökknar og stjörnurnar *koma í ljós*
himlen mörknar och stjärnorna *blir synliga*

(18) *í ljós kom* svört þúst á sjónum
på havet *blev* en svart silhuett *synlig*

¹ Sökningen i Språkbanken gjordes med hjälp av korpussökverket Korp i två ungefär lika stora textmaterial. Det ena materialet består av korpusen Bloggmix, som innehåller informella, talspråksnära texter med drygt 439 000 löpord. Det andra materialet består av fem korpusar innehållande texter av skriftspråklig karaktär, sammanlagt nästan 475 000 löpord. I Bloggmix förekommer idiomerna 65 gånger mot 333 gånger i de skriftspråkliga texterna. Det skriftspråkliga materialet består av korpusarna Akademiska texter, Skönlitteratur, Tidningstexter, Tidskrifter och Svenska Wikipedia <<http://spraakbanken.gu.se/korp/>>. Sökningen är gjord i september 2013.

De övriga sex ekvivalenter till *koma í ljós* som finns i materialet förekommer bara en gång var: *dyka upp*, *märkas*, *upptäcka*, *uppstå*, *visa* och *framgå*. I ett av de sex språkproven är subjektet konkret (exempel 19), medan de övriga fem är abstrakta (exempel 20–24).

- (19) hvalur *kom í ljós* aftur undan skipinu
det *dök upp* en val bakom fartyget
- (20) það *kom* strax *í ljós* að tækið var gallað
det *märktes* med en gång att apparaten var defekt
- (21) leyndur galli á bílnum *kom í ljós*
man *upptäckte* ett dolt fel på bilen
- (22) fljótlega *komu í ljós* þverbrestir í hjónabandi þeirra
det *uppstod* snart sprickor i deras äktenskap
- (23) hans innri maður *kom í ljós* í erfiðleikunum
han *visade* sitt rätta jag när det blev problem
- (24) hún er góð söngkona eins og berlega *kom í ljós* á tónleikunum
hon är en bra sångerska såsom tydligt *framgick* på konserten

I några av dessa översättningar har ekvivalenten en annorlunda uppsättning semantiska aktanter än konstruktionen *koma í ljós*. I exempel (20) förutsätter den svenska översättningen, trots den passiva konstruktionen, en aktör som noterar apparatens brister, och i (21) införs en sådan aktör genom ett indefinit pronomen; *man* upptäckte de fel som utgör subjekt i den isländska satsen. Liknande perspektivskifte har genomförts i exempel (22) och (23), medan ekvivalenten *framgå* i exempel (24) har en semantisk struktur som är parallell med *koma í ljós*.

Översättningen i exempel (22) ovan, där *koma í ljós* återges med *uppstå*, har även ett något annorlunda fokus än den isländska förlagan. Det isländska språkprovet och den svenska översättningen skiljer sig åt med avseende på verbets aktionsart. Betydelsen hos *komu í ljós* i den isländska versionen är neutral när det gäller uppkomsten av det som subjektet betecknar; det enda som sägs är att sprickorna visar sig, blir synliga. *Uppstå* i den svenska versionen fokuserar istället uppkomsten av det som subjektet betecknar: sprickorna *uppstår*. Underförstått kan dock utläsas att de för att bli synliga redan måste ha uppstått. Exemplet illustrerar emellertid väl hur konnotativa och pragmatiska ekvivalensrelationer ibland prioriteras vid översättningen av språkproven före en strikt denotativ eller formell ekvivalens.

7.4. Syntaktiska drag

Av de 22 isländska satser som innehåller idiomet *koma í ljós* består 18 av enkla huvudsatser. Av dessa inleds nio av ett egentligt subjekt, tre av ett formellt subjekt *það* och sex av ett adverbial. I två av språkproven ovan förekommer idiomet i en adverbial, efterställd bisats som i exempel (14) och (24), här upprepade som (25) och (26). I ett fall spetsställs ett huvudverb, *komið hefur í ljós*, exempel (8), här upprepat som (27), och i exempel (18) spetsställs ett prepositionsled, *í ljós kom*, här upprepat som (28).

- (25) <ég tók mynd> *um leið og <fuglinn kom í ljós>*
<jag tog ett kort> *så fort <fågeln visade sig>*
- (26) *hún er góð söngkona eins og berlega kom í ljós á tónleikunum*
hon är en bra sångerska såsom tydligt framgick på konserten
- (27) *komið hefur í ljós að börn sjúga oft á sér hendurnar í móðurkviði*
det har visat sig att barn ofta suger på sina händer redan i moderlivet
- (28) *í ljós kom svört þúst á sjónum*
på havet blev en svart silhuett synlig

Bland de subjekt som inleder språkproven finns som sagt några mycket enkla sådana bestående av ett enkelt nomen (*stjörnurnar, hvalur, þetta, gallinn*). Subjekten i exemplen (12) och (21) ovan, här upprepade som (29) och (30), är däremot både semantiskt och syntaktiskt mer utbyggda i form av nominalfraser med såväl framförställda som efterställda attribut.

- (29) *stórvægilegir gallar á húsinu hafa komið í ljós*
[allvarliga fel på huset har kommit i ljus]
det har visat sig att huset har allvarliga fel
- (30) *leyndur galli á bílnum kom í ljós*
[dolt fel på bilen kom i ljus]
man upptäckte ett dolt fel på bilen

Översättningarna till båda dessa språkprov har en annorlunda syntaktisk struktur än förlagan, de har byggts ut med ett spetsställt, semantiskt lätt subjekt. I exempel (29) inleds den svenska översättningen med ett formellt subjekt och isländskans subjekt har gjorts om till en att-sats: ”det har visat sig att huset har allvarliga fel”. I exempel (30) uppbär det indefinita pronomenet *man* subjektsfunktionen och isländskans subjekt blir då objekt: ”man upptäckte ett dolt fel på bilen”. Övriga spetsställda subjekt,

formella såväl som egentliga, illustrerar en hög grad av överensstämmelse mellan isländskans och svenskans syntaktiska krav på subjektet.

De strukturer som inte har överförts till en direkt parallell syntaktisk struktur i svenskan är dels spetsställningen av det infinita verbet (ex. 8 och 27), här igen upprepat som (31), dels idiomets spetsställda prepositionsfras (ex. 18), här upprepat som (32).

(31) komið hefur í ljós að börn sjúga *oft* á sér hendurnar í móðurkviði
[kommit har í ljús att barn suger ofta på sig händerna i moderlivet]
det har visat sig att barn *ofta* suger på sina händer redan i moderlivet

(32) í ljós kom svört þúst á sjónum
[i ljús kom svart silhuett på havet]
på havet blev en svart silhuett synlig

Språkprovet i exempel (31) är konstruerat med en typ av spetsställning som är fullt möjlig och ganska vanlig i isländskan men som inte är neutral (om ens möjlig) i svenskan. Dessutom framgår avvikelser i de båda språkens bisatsordföljd. I den svenska översättningen infogas ett formellt subjekt *det* och ledföljden i bisatsen har ändrats: ”*det* har visat sig att barn *ofta* suger på sina händer redan i moderlivet”. För övrigt förefaller språkens olika preferenser när det gäller den inbördes ordningen mellan olika typer av innehållsadverbial vara värda att studera närmare. I exempel (32) har det i isländskan finala adverbialet placerats i fundamentet i den svenska översättningen: ”*på havet* blev en svart silhuett synlig”.

Såväl isländskan som svenskan är V2-språk och de uppvisar stora likheter beträffande de lösningar som tillgrips för att det finita verbet skall placeras på andra plats i satsen. Av exemplen (31) och (32) framgår emellertid att isländskan tillåter vissa strategier för spetsställning som knappast utnyttjas i svenskan.

7.5. Informationsstruktur

I det lilla material som här är aktuellt påträffas vissa avvikelser som kan tillskrivas skillnader i isländskans och svenskans informationsstruktur. Språkproven i exempel (33) och (34), även redovisade som exempel (12) och (15) ovan, inleds med tunga, indefinita subjekt.

(33) *stórvægilegir* gallar á húsinu hafa komið í ljós
[omfattande fel på huset har kommit i ljús]
det har visat sig att *huset* har *allvarliga fel*

- (34) *ymsar fornminjar* komu í ljós við uppgröftinn
 [diverse fornminnen kom i ljus vid utgrävningen]
 vid utgrävningen kom *diverse fornlämningar* i dagen

Även om isländskan strävar efter att placera det kända först och ny information sent i satsen (Þráinsson 2005:506ff.), och på så vis tillämpar samma principer för informationens distribution i satsen som svenskan, tyder exempel (33) och (34) på större flexibilitet i detta avseende i isländskan än i svenskan. I exempel (33) kan *stórvægilegir gallar* antas vara ny information. För att översättningen skall följa svenskans tema-remapricip byggs den ut med ett formellt subjekt och informationen i den isländska nominalfrasen görs om till ett egentligt subjekt i form av en att-sats. I exempel (34) framstår *uppgröftinn* i sin definitiva form som redan känd bakgrundsinformation medan resultatet av den, i form av ett semantiskt indefinit nominal, *ymsar fornminjar*, utgör den nya informationen i meningen. Genom att spetsställa adverbiallet *vid utgrävningen* får subjektet, som bärare av den nya informationen, en position som bättre stämmer överens med svenskans tema-remapreferenser.

I exempel (35) och (36), även redovisade som (16) och (11) ovan, sammanfaller isländskans ledföljd med svenskans krav på informationens organisering. I (35) markerar den bestämda nominalfrasen, *raskið í túninu*, känd information, medan den obestämda nominalfras som utgör subjektet, *merkar fornminjar*, är rema och placeras som sådant sist i satsen.

- (35) við raskið í túninu komu í ljós *merkar fornminjar*
 [vid rubbningarna i slåttervallen kom i ljus intressanta fornminnen]
 när man grävde i slåttervallen kom *intressanta fornminnen* i dagen

- (36) miklir listamannshæfileikar hans komu snemma í ljós
 hans stora konstnärliga talang *visade sig* tidigt


Också subjektet i exempel (36) följer svenskans informationsstruktur; bestämda nominalfraser signalerar att innehållet är känt och att de kan utgöra satsens tema.


Det material som här har undersökts är naturligtvis för litet för att några långtgående slutsatser skall kunna dras om skillnader i de båda språkens informationsstrukturella preferenser. Den skillnad som emellertid har observerats tycks även vara relevant i relationen mellan isländskan å ena sidan och danska och norska å den andra.


Som framgår av bild (4) modifieras översättningen av språkprovet så att den nya informationen också i danska och norska placeras i slutet av satsen. Den danska och de båda norska översättningarna ligger närmare den isländska källtextens struktur än den svenska. I stället för att som i

svenskan införa en agent, *man*, behålls den opersonliga strukturen med hjälp av ett formellt subjekt *der* respektive *det*. Den nya informationen redovisas då som egentligt subjekt i form av en att-bisats.

leyndur galli á bílnum kom í ljós

 der viste sig at være en skjult fejl på bilen

 det viste seg at det var en skjult feil på bilen

 det synte seg at det var ein løynd feil på bilen


 man upptäckte ett dolt fel på bilen

BILD 4. *Tungt, indefinit nominal i fundamentet i isländskan och översättningarna till danska, norskt bokmål, nynorska och svenska.*

Den svenska översättning som här har valts är naturligtvis inte den enda gångbara. En motsvarighet till danskans översättning är fullt möjlig också i svenskan, *det visade sig vara ett dolt fel på bilen*, liksom en motsvarighet till de norska översättningarna, *det visade sig att det var ett dolt fel på bilen* samt ytterligare några innehållsmässigt fullt ekvivalenta alternativ. Vid översättningen av språkproven, liksom vid översättning i allmänhet, finns det som regel en rad alternativa lösningar som alla framstår som goda översättningar. Förutom de lexikala skillnaderna mellan språken är det viktigt att också redovisa övriga skillnader som avslöjas vid den kontrastiva analysen av källspråket och målspråket. I det här fallet är det skillnaden i isländskans och de skandinaviska språkens restriktioner när det gäller placeringen av ny information i satsen som illustreras. Dessa restriktioner omöjliggör en direkt överföring av isländskans struktur: *ett dolt fel på bilen visade sig*.

8. Idiomen *láta í ljós* och *leiða í ljós*

Som framgår ovan förekommer idiomerna *láta í ljós* och *leiða í ljós* inte lika ofta i *Islex* som *koma í ljós*: sammanlagt inte mer än 16 gånger. Till skillnad från *koma í ljós* är dessa idiom relativt enkla såväl semantiskt som syntaktiskt. Idiomens valens begränsas av att både *láta* och *leiða* är transitiva, och även deras selektionsrestriktioner är stränga när det gäller såväl subjektets och objektets semantiska egenskaper. De båda idiomerna beskrivs därför betydligt mera översiktligt än *koma í ljós* ovan.

8.1. Láta í ljós

Islex-databasen innehåller nio förekomster av idiomet *láta í ljós* (ex. 37–45). I fyra av de nio språkproven ges verbet *visa* som ekvivalent till *láta í ljós*, i ett fall, i exempel (40), kompletterad med den alternativa ekvivalenten *säga*.

- (37) krakkarnir létu óspart í ljós gleði sína
ungarna *visade* tydligt sin glädje
- (38) enginn dirfðist að láta óánægju sína í ljós
ingen dristade sig att *visa* sitt missnöje
- (39) *láta í ljós* hrifningu <sína>
visa <sin> förtjusning
- (40) *láta* <skoðun sína> í ljós
visa <vad man tycker>, *säga* <sin mening>

Säga förekommer dessutom i ytterligare tre språkprov, (41)–(43), i (43) kompletterat med ekvivalenten *uttrycka* som även ges som ekvivalent i exempel (44).

- (41) *láta í ljós* skoðun sína
säga sin mening
- (42) hann lét skoðanir sínar hispurslaust í ljós
han *sa* rakt ut vad han tyckte
- (43) *láta í ljós* <álit sitt>
uttrycka <sin åsikt>, *säga* <sin mening>
- (44) ég vil *láta í ljós* þakklæti mitt fyrir alla hjálpina
jag vill *uttrycka* min tacksamhet för all hjälp jag fått

Slutligen har *láta í ljós* i ett fall, i exempel (45), försetts med ekvivalenten *kommunicera*.

- (45) ung börn *láta í ljós* vanlíðan með gráti
smábarn *kommunicerar* fysiskt obehag med gråt

I fyra av de nio språkproven, (39)–(41) och (43), illustreras idiomets användning med döda språkprov, dvs. ofullständiga satser utan subjekt. De övriga fem språkproven illustrerar idiomets krav på en mänsklig agent som subjekt: *krakki* 'barn, unge', *enginn*, *hann*, *ég* och *barn*. Även objektet

är semantiskt begränsat till att uttrycka någon typ av känsla, inställning eller mentalt tillstånd, vilket framgår av de objekt som redovisas i språkproven: *gleði* 'glädje', *óánægja* 'missnöje', *hrifning* 'förtjusning', *skoðun* 'mening', *álit* 'åsikt', *þakklæti* 'tacksamhet' och *vanlíðan* 'fysiskt obehag'.

Det svenska verbet *uttrycka* skulle kunna etableras som en denotativ ekvivalent till *láta í ljós* i de flesta av fallen i *Islex*. Det förekommer två gånger i exemplen ovan. Men i stället för att i alla sammanhang återge *láta í ljós* med *uttrycka* konstrueras översättningarna oftare med någon av två hyponymer till *uttrycka*: *visa* respektive *säga*. I exempel (37) skulle visserligen den alternativa översättningen *ungarna uttryckte tydligt sin glädje* vara denotativt ekvivalent med språkprovet men samtidigt framstå som pragmatiskt avvikande och ett mindre lyckat sätt att återge språkprovets konnotativa betydelse.

De alternativa ekvivalenterna *visa* och *säga* är endast i liten utsträckning utbytbara mot varandra. Båda är polysema och den betydelse hos var och en av dem som här aktualiseras är mer specifik än det isländska idiomets betydelse. Medan föremålet för den aktion som uttrycks i *láta í ljós* kan utgöras av nästan vilken känsla eller uppfattning som helst, kan de lösningar som här har valts tolkas som att svenskan tenderar att skilja mellan verbal aktion och mentalt tillstånd. I de fall där aktionen är mental och objektet betecknar en känsla som inte i första hand tar sig verbala uttryck, som i exempel (37) *glädje*, (38) *missnöje*, (39) *förtjusning*, redovisas aktionen *láta í ljós* genom ekvivalenten *visa*. Att en inställning eller ett förhållningssätt även kan uttryckas verbalt illustreras av exempel (40), där alternativa formuleringar presenteras. Där den handling som uttrycks i språkprovet har verbal karaktär återges den med ekvivalenten *säga* såsom i fallen där det isländska objektet översätts med *sin mening* (ex. 40, 41 och 43).

Fraserna *visa sin glädje*, *säga sin mening* och *uttrycka sin åsikt* respektive *sin tacksamhet* är snarast att betrakta som kollokationer där kollokatorerna *visa*, *säga* och *uttrycka* inte är helt godtyckligt utbytbara mot närsynonyma verb. De svenska nominalfrasernas konstruktionsmönster framstår dock som allt annat än enkla. Kompletterade med en prepositionsfras kan preferenserna beträffande verbet ändras så att exempelvis *glädje över ngt* konstrueras med *uttrycka* snarare än *visa*. En något friare konstruktion är den som redovisas i exempel (45). Trots att subjektet även där är mänskligt, *småbarn*, är den verbala förmågan inte given.

När det gäller idiomets syntaktiska egenskaper kan det konstateras att mellan verbet *láta* och prepositionsfrasen *í ljós* kan såväl ett adverbial av något slag (ex. 37) som objektet (ex. 38 och 40) placeras, liksom också båda dessa led (ex. 42).

8.2. Leiða í ljós

Idiomet *leiða í ljós* förekommer i sju språkprov i *Islex*. Av dessa består ett (ex. 46) av ett dött språkprov i form av en ofullständig sats, med det för idiomet obligatoriska objektet illustrerat av ett demonstrativt pronomen utan korrelat: *þetta*. De övriga sex, exempel (47)–(52), består av levande språkprov. I likhet med *láta í ljós* är de semantiska ramarna för *leiða í ljós* mycket strikta. Samtliga subjekt betecknar någon typ av undersökning eller granskning och objektet betecknar resultatet av undersökningen eller granskningen.

- (46) *leiða* <þetta> *í ljós*
pávisa <detta>, *avslöja* <det här>
- (47) greining á sýninu *leiddi í ljós* krabbamein
analysen av provet *pávisade* cancer
- (48) <rannsóknin> *leiðir í ljós* að <hann er smitaður>
<undersökningen> *visar* att <han har smittats>
- (49) yfirferð á reikningum fyrirtækisins *leiddi ýmislegt í ljós*
en granskning av företags bokföring *avslöjade* ett och annat
- (50) læknisrannsóknin *leiddi í ljós* hjartagalla
läkarundersökningen *avslöjade* ett hjärtfel
- (51) hlustun *leiddi í ljós* lungnabólgu
vid auskultationen *konstaterades* lunginflammation
- (52) könnun á hafsbottinum hefur *leitt ýmislegt í ljós*
undersökningar av havsbotten har *gett* en hel del ny kunskap

De redovisade ekvivalenterna till *leiða í ljós* i exempel (46), *pávisa* och *avslöja*, upprepas båda i språkproven, *pávisa* i exempel (47) och *avslöja* i (49) och (50). Dessutom förekommer *visa*, som precis som i förlagan följs av ett objekt i form av en att-sats (ex. 48). I ingen av översättningarna redovisas det svenska idiomet *lägga i dagen* – eller något annat idiom – som ekvivalent till det isländska *leiða í ljós*.

I exempel (49) och (52) har ett semantiskt lätt objekt i form av det substantiverade adjektivet, *ýmislegt*, placerats mellan verbet och prepositionsfrasen. Något exempel på att objekt av annat slag kan placeras så förekommer inte i detta begränsade material. Det går dock inte att dra några generella slutsatser om eventuella restriktioner beträffande led som kan placeras mellan verbet och prepositionsfrasen; för det är materialet

alltför litet. Med utgångspunkt i språkproven i *Islex* framstår användningen av idiomet *leiða í ljós* som mycket regelbunden såväl semantiskt som syntaktiskt.

9. Sammanfattning

De sökningar som kan göras i en nätordbok som *Islex* upphäver i princip funktionsuppdelningen mellan källspråk och målspråk. Målspråket är lika åtkomligt som källspråket i en konventionell ordbok är ensamt om att vara, och fritextsökningarna öppnar möjligheten att undersöka drag som inte bara – eller inte ens – i första hand är av lexikografisk natur.

Den mycket begränsade undersökning som här har redovisats har på intet sätt syftat till att ge en uttömmande beskrivning av de tre idiomens alla egenskaper utan enbart till att i grova drag illustrera de möjligheter som *Islex* erbjuder som parallellkorpus. Fokus har legat på idiomens semantiska och syntaktiska egenskaper. Vissa syntaktiska skillnader mellan isländska och svenska har därvid kunnat dels bekräftas, dels anas, och de båda språkens informationsstruktur förefaller vara värd att studera närmare. Det måste dock påpekas att innehållet i de isländska språkproven naturligtvis i samtliga fall hade kunnat återges på annat sätt än det som lexikografen slutligen valt. Så är det med all översättning.

Sin primära funktion har den tvåspråkiga lexikografen haft som hjälpmedel vid översättning i vid mening. Medan det nu höjs röster för att frigöra lexikografen från språkvetenskapen har syftet här snarare varit det omvända, nämligen att illustrera hur den tvåspråkiga lexikografen kan utgöra underlag för språkvetenskapliga studier. Med sina små, välavgränsade och kompletta lexikaliska parallellkorpusar och de generösa sökmöjligheter som publikationsformen tillåter kan ordböcker som *Islex* sålunda användas i syften som tidigare inte har gått att förena med de tvåspråkiga ordböckerna.

De möjligheter som här har antytts måste för *Islex* del i första hand tillskrivas mediet och dess egenskaper. När det gäller den lexikografiska strukturen i *Islex* har den inte fullt ut lösgjort sig från de ramar som lexikografen traditionellt har arbetat inom. Den teoretiska utgångspunkten, där det centrala är just funktionsuppdelningen mellan ett källspråk som den enda ingången i ordboken å ena sidan och ett underordnat målspråk å den andra, gäller fortfarande för *Islex*. På så sätt har denna nätordbok inte riktigt nått den framtid som Wiegand (1989) förutspådde, där lexikografen och lexikografen verkligen frigjort sig från bokens begränsande format.

Litteratur

- Atkins, B.T. Sue & Michael Rundell 2008. *The Oxford Guide to Practical Lexicography*. Oxford: Oxford University Press.
- Beygingarlýsing íslensks nútímamáls*. Stofnun Árna Magnússonar í íslenskum fræðum: <<http://bin.arnastofnun.is>>.
- Borin, Lars 2016. Lexikografi för människor och lexikografi för maskiner. I: Hannesdóttir, Anna Helga (red.), *Framtidens lexikografi. Rapport från ett symposium i Göteborg 5 oktober 2012*. (Meijerbergs arkiv för svensk ordforskning 42.) Göteborg: Göteborgs universitet. S. 9–27.
- Eaker, Birgit, Lennart Larsson & Anki Mattisson (red.) 2012. *Nordiska studier i lexikografi* 11. Lund: Svenska Akademiens ordboksredaktion.
- Fjeld, Ruth Vatvedt 1992. Innledning. I: Fjeld, Ruth Vatvedt (red.), *Nordiske studier i leksikografi*. Oslo. S. 1–6.
- Fuertes-Olivera, Pedro A. & Henning Bergenholtz 2011a. Introduction: The Construction of Internet Dictionaries. I: Pedro Fuertes-Olivera & Henning Bergenholtz 2011b. S. 1–16.
- Fuertes-Olivera, Pedro A. & Henning Bergenholtz 2011b (eds.). *e-Lexicography. The Internet, Digital Initiatives and Lexicography*. London: Continuum.
- Gellerstam, Martin 1999. Översättaren och ordboken. I: Slotte, Peter, Pia Westerberg & Eva Orava (red.), *Nordiska studier i lexikografi* 4. Helsingfors: Forskningscentralen för de inhemska språken. S.121–129.
- Gouws, Rufus H. 2011. Learning, Unlearning and Innovation in the Planning of Electronic Dictionaries. I: Pedro Fuertes-Olivera & Henning Bergenholtz 2011b. S. 17–29.
- Grønvik, Oddrun 2012. Digital lexikografi og funksjonsteoretiske briller. (Recension av) Pedro A. Fuertes-Olivera and Henning Bergenholtz (eds.). *e-Lexicography. The Internet, Digital Initiatives and Lexicography*. I: *LexicoNordica*19. S. 239–254.
- Hartmann, Reinhard 2010. Has Lexicography Arrived as an Academic Discipline? Reviewing Progress in Dictionary Research during the Last Three Decades. I: Lönnroth, Harry & Kristina Nikula (red.), *Nordiska studier i lexikografi* 10. Tammerfors: Tammerfors universitet. S. 11–35.
- Islex*. <<http://islex.is>>, <<http://islex.dk>>, <<http://islex.no>>, <<http://islex.se>> (september 2013).
- Jansson, Håkan 2012. Parallellkorpuser som resurs i lexikografiskt arbete. I: Eaker, Birgit, Lennart Larsson & Anki Mattisson (red.). S. 328–339.
- Jónsdóttir, Halldóra & Þórdís Úlfarsdóttir 2012. ISLEX – en flersproget nordisk ordbog. I: Eaker, Birgit, Lennart Larsson & Anki Mattisson (red.). S. 353–366.

- Kilgarriff, Adam 2012. (Review of) Pedro A. Fuertes-Olivera and Henning Bergenholtz (eds.). *e-Lexicography: The Internet, Digital Initiatives and Lexicography*. I: *Kernerman Dictionary News*. July 2012. Tel Aviv: K Dictionaries LTD. S. 26–29.
- Kilgarriff, Adam, Miloš Husák, Katy McAdam, Michel Rundell & Pavel Rychlý (2008). GDEX: Automatically Finding Good Dictionary Examples in a Corpus. I: Bernal, Elisenda & Janet DeCesaris (eds.), *Proceedings of the XIII Euralex International Congress*. Barcelona 2007. S. 425–432.
- Kromann, Hans-Peder, Theis Riiber & Poul Rosbach 1984. Überlegungen zu Grundfragen der zweisprachigen Lexikographie. I: H.E. Wiegand (Hrsg.), *Studien zur neuhochdeutschen Lexikographie* V. Hildesheim/New York: Georg Olms. S. 159–238.
- Malmgren, Sven-Göran 2009. On production-oriented information in Swedish monolingual defining dictionaries. I: Nielsen, Sandro & Sven Tarp (eds.), *Lexicography in the 21st Century: In honour of Henning Bergenholtz*. Amsterdam/Philadelphia: John Benjamins Publishing Company. S. 93–102.
- Rauset, Margunn, Anna Helga Hannesdóttir & Aldís Sigurðardóttir 2012. Ein-, to- eller fleirspråkleg ordbok? I: Eaker, Birgit, Lennart Larsson & Anki Mattisson (red.). S. 512–523.
- Rundell, Michael 2012. 'It works in practice but will it work in theory?' The uneasy relationship between lexicography and matters theoretical. I: Fjeld, Ruth Vatvedt & Julie Matilde Torjusen (eds.), *Proceedings of the 15th EURALEX International Congress*. Oslo: University of Oslo. S. 47–92.
- Sanders, Christopher 2013. *Islex* foråret 2013. I: *LexicoNordica* 20. S. 259–277.
- Sigurðardóttir, Aldís, Anna Helga Hannesdóttir, Halldóra Jónsdóttir, Hákan Jansson, Lars Trap-Jensen & Þórdís Úlfarsdóttir 2008. ISLEX. An Icelandic-Scandinavian Multilingual Online Dictionary. I: Bernal, Elisenda and Janet DeCesaris (eds.), *Proceedings of the XIII Euralex International Congress*. Barcelona 2007. S. 779–789.
- Sköldberg, Emma & Christian Mattsson 2016. Ful- och finlexikografi? Om ordboksverksamhet i Sverige i dag och i morgon. I: Hannesdóttir, Anna Helga (red.), *Framtidens lexikografi. Rapport från ett symposium i Göteborg 5 oktober 2012*. (Meijerbergs arkiv för svensk ordforskning 42.) Göteborg: Göteborgs universitet. S. 111–129.
- Språkbanken*. <<http://spraakbanken.gu.se>> (september 2013).
- Svensk-isländsk ordbok* 1983. Lund: Walter Ekstrands Bokförlag & Reykjavík: Almenna Bókafélagið.

- Úlfarsdóttir, Þórdís 2013. ISLEX – norræn margmála orðabók. I: *Orð og tunga* 15. Reykjavík: Stofnun Árna Magnússonar í íslenskum fræðum. S. 41–71.
- Wiegand, Herbert Ernst 1989. Der gegenwärtige Status der Lexicographie und ihr Verhältnis zu anderen Disziplinen. I: Franz Josef Hausmann et al. (Hrsg.), *Wörterbücher. Dictionaries, Dictionnaires. Ein internationales Handbuch zur Lexikographie*. Erster Teilband. Berlin & New York: de Gruyter. S. 246–280.
- Þráinsson, Höskuldur 2005. *Setningar. Handbók um setningafræði. Íslensk tunga* III. bindi. Reykjavík: Almenna bókafélagið.

Teori i praktiken

Om prototyp teorin och SAOB

Pär Nilsson

Prototype theory is a central element of semantic theory, in particular within cognitive linguistics, and the theory is often suggested as a starting point for semantic categorization within lexicographic theory.

It is sometimes claimed that there is a gap between academic semantics and dictionary semantics; that the former is being too theoretical and difficult to apply, and the latter too practice-oriented.

The aim of this paper is to analyze the methods used in the Swedish Academy Dictionary (SAOB) in the light of prototype theory and with main focus on the graphic representation of polysemous words and the formation of definitions. Are the methods theoretically motivated and is the theory applicable in the lexicographic practice? The conclusion is that there is a foundation in the approach of the dictionary that seems to follow the ideals of prototype theory.

1. Inledning

Svenska Akademiens ordbok (SAOB) är en historisk ordbok som beskriver det svenska språket från 1520-talet fram till det datum då aktuellt ordbokshäfte trycktes. Det första bandet gavs ut 1898 och band 36, som sträcker sig t.o.m. ordet UTSUDDA, kom 2012. Arbetet på redaktionen sker inte utifrån någon fastlagd och explicit teori. Även om redaktörer och ordbokschef under tidens gång i olika utsträckning satt sin prägel på arbetet utifrån sina respektive kunskaper och teoretiska ideal och samtidens övergripande språksyn, är utgångspunkten snarare en mer implicit semantisk modell som tar sikte på betydelseutveckling. Det har hävdats och hävdas fortfarande emellanåt att det finns en klyfta mellan universitetssemantiken å den ena sidan, och ordbokssemantiken å den andra. Att den ena är för teoretisk och svår att tillämpa, den andra för praktisk och inte tillräckligt systematisk. Prototyp teorin har åberopats inom den teoretiska semantiken för att förklara hur begrepp förhåller sig till varandra.

I det här bidraget kommer jag att titta lite närmre på prototypeteorins huvudidéer och deras tillämpning, och sedan undersöka hur SAOB står sig i förhållande till dem. Min utgångspunkt är Hovmark (2012) och Atkins & Rundell (2008), två texter där praktisk tillämpning av prototyp teorin inom lexikografen diskuteras.

2. Teoretisk bakgrund

Prototyp teorin har sina rötter i den kognitiva psykologin. Teorin utvecklades på 1970-talet av den amerikanska psykologen Eleanor Rosch (se t.ex. Rosch 1978), och den försöker förklara hur vi människor kategoriserar – dvs. hur vi uppfattar vår omgivning och grupperar och strukturerar information om den. Den traditionella modellen för kategorisering (som kan härledas till Aristoteles logik) baseras på formeln *nödvändiga* och *tillräckliga kriterier*. Ett antal kriterier måste med nödvändighet uppfyllas för att något ska tillhöra en viss kategori, och motsatt: om något uppfyller dessa kriterier är detta tillräckliga villkor för att tillhöra den aktuella kategorin. Antingen hör en individ eller enhet till en viss kategori eller också inte. Säger vi t.ex. att ett barn är en människa under skedet mellan födelsen och vuxen ålder¹, är kriterierna för kategorin barn: *människa* och *mellan födelsen* och *vuxen ålder*.

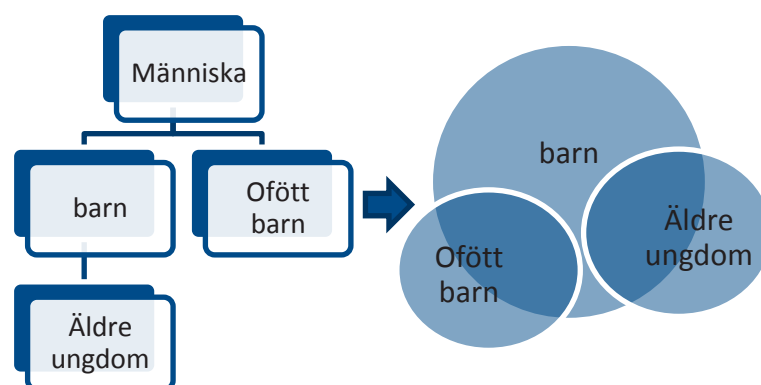
Ett problem med detta synsätt som prototyp teorin tar fasta på är att verkligheten sällan eller aldrig är så svart eller vit. Vi stöter ofta på företeelser i gråzonen mellan två kategorier som är svåra att föra till den ena eller andra. Är t. ex. en person som befinner sig i senare delen av ungdomsåren, och som ännu inte hunnit bli vuxen, fortfarande ett barn? Eller är det här en person i skedet någonstans mellan barn och vuxen? Enligt de uppsatta kriterierna ovan har vi fortfarande att göra med ett barn. Vi kan också vända på resonemanget och fundera över begreppet foster (jfr sådana uttryck som ”ofött barn” och ”(ha ett) barn i magen” osv.). Kan detta begrepp föras till kategorin barn? Enligt samma kriterier är det inte fråga om ett barn.

Rosch (1978:36) menar utifrån sådana företeelser i gränlandet mellan kategorier, att en kategoris gränser inte är absoluta, utan istället ska uppfattas som just oskarpa (se figur 1).

Ur ett annat perspektiv uppmärksammar prototyp teorin svårigheten i att hitta den uppsättning kriterier som alla individer i en kategori måste uppfylla för att räknas som medlemmar. Trots att man i verkligheten kan känna sig säker på att en kategori omfattar vissa enheter eller individer är det svårt att sätta fingret på just vilka kriterier det är som uppfylls. Den här tanken

¹ NE:s (1990) definition av barn.

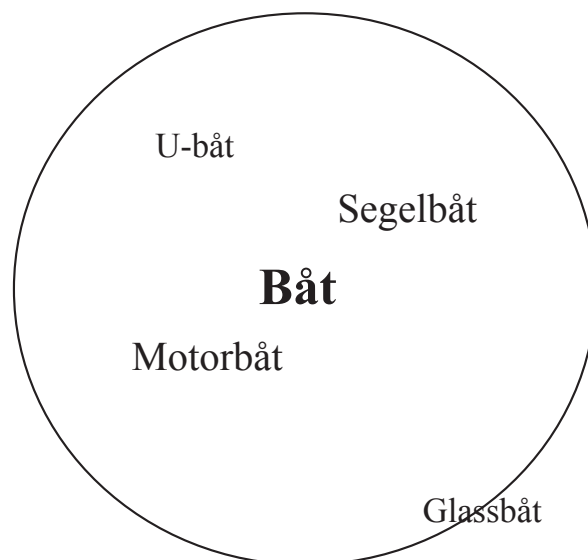
lutar sig mot Wittgensteins (berömda) resonemang om familjelikhet: det är ofta möjligt att ringa in centrala drag eller attribut som förekommer hos de olika enheterna, men att välja ut en specifik uppsättning attribut som ska gälla alla enheter går inte. I själva verket är kategorimedlemmarna besläktade i varierande grad och medlemskapet eller kategoritillhörigheten framhäver olika särdrag hos medlemmarna beroende på den aktuella kontexten (se t. ex. Wittgenstein 1978:58f.). Kategoriernas struktur ska därför uppfattas som ett överlappande nätverk snarare än ett strikt system av klart avgränsade kategorier.



FIGUR 1. *Kategoristruktur, från traditionell kategoriseringsmodell till prototypeteori.*

Centralt i prototypeteorin är också den empiriska grunden. De slutsatser som dragits är underbyggda av omfattande undersökningar och experiment på försökspersoner. Det har i studier visat sig att folk i regel är överens om en kategoris centrala innehåll, trots att man är oense om gränserna (se Rosch 1978:36f.). Ett trivialt, men relevant påpekande är att vi kan bedöma de tydliga fallen av en viss kategori, utan att ens reflektera över kategorins gränser. Prototypeteorin fokuserar just på de här klara fallen, typfallen.

Prototyperna, eller de prototypiska kategorimedlemmarna, är de som uppfattas såsom mest representativa för kategorin, de som innehåller de mest representativa attributen för alla enheter inom gruppen. Dessa prototyper tenderar att användas vid referens till hela kategorin. När man således talar om en viss kategori av föremål eller företeelser, låt säga en båt eller stol, har man en viss, typisk, individ i tanken som är representativ för hela gruppen. En mindre typisk båt osv. ska då uppfattas i relationen till den mest typiska (se figur 2).



FIGUR 2. *Prototyper av begreppet båt.*

En kategorimedlem kan alltså vara mer eller mindre prototypisk, och man talar om graderat kategorimedlemskap. Ett i sammanhanget ofta anförda exempel härpå gäller rangordning av prototypiska frukter (jfr Hovmark 2012:61). Ber man sålunda 1000 svenskar lista 5 typiska frukter, kan man förvänta sig att hitta äpple och päron på listan, och möjligen apelsin, men mindre sannolikt avokado eller granatäpple. Ett äpple eller ett päron är så att säga mer frukt än en avokado i det här avseendet.

Ju mindre prototypisk en företeelse är, desto längre bort befinner den sig från det centrala innehållet i kategorin — dvs. desto färre drag har den gemensamt med detta innehåll. Det här gör kategoriernas gränser oskarpa. Enligt klassisk kategoriseringsmodell är alla som uppfyller tillräckliga kriterier likvärdiga medlemmar i kategorin, men i verkliga livet har de alltså olika värde, något som redan möjligheten att uttryckligen ange något som ”i praktiken en frukt”, eller ”så gott som en” eller ”tekniskt sett ett” osv. antyder.²

Exemplet med frukter antyder också att kontexten är avgörande. Vilka kategorimedlemmar, eller vilka av deras attribut som uppfattas som centrala, är avhängigt det sociokulturella sammanhanget. Tillfrågas på samma sätt en sydamerikan eller afrikan att göra en fruktlista, kommer den sannolikt att se annorlunda ut än den svenska eller europeiska, även om själva metoden är den samma.

² Jfr Roschs (1978:39) diskussion om ”hedges”.

3. Prototypeteorin inom det språkvetenskapliga området

Prototypeteorin kan även tillämpas på det språkvetenskapliga området, och teorin har diskuterats inom semantiken, lexikologin och inte minst lexikografen. Även inom dessa områden har den traditionella utgångspunkten varit logikens nödvändiga och tillräckliga drag och tankar om hierarkiska system som strikt delar in begreppen i över- och underordnade begrepp. Det som tilltalar med prototypeteorin är att den istället utgår ifrån människans sätt att ordna begreppen. Man är inte längre på jakt efter ett perfekt system (som vid det praktiska arbetet ofta måste modifieras med ad hoc-antaganden). Det är ett annat och mer pragmatiskt sätt att närma sig uppgiften och försöka förklara hur mänsklig kategorisering inom den språkliga sfären går till.

3.1. Inom lexikografen

Prototypeteorin är relativt etablerad idag som utgångspunkt vid semantisk kategorisering, men har kommit i nytt ljus på senare tid, bl.a. genom försök till mer praktisk tillämpning i det lexikografiska arbetet. Detta är den övergripande semantiska teori som Atkins & Rundell (2008) förankrar sin lexikografiska handbok i. Även Hovmark (2012), som tar sin utgångspunkt i danskan och ODS, anlägger detta perspektiv. Vi kommer att återkomma mer i detalj till dessa texter senare, men generellt gäller för båda att fokus nu framförallt är dels den grafiska dispositionen, dvs. själva presentationen av betydelseerna i förhållande till den mentala strukturen, dels betydelsebeskrivningen och utformningen av definitioner. Först tittar vi på den grafiska strukturen.

Prototypeteorin tillbakavisar alltså tanken på att all information, i det här fallet språkliga betydelser, är strikt hierarkiskt ordnad i ett system av kategorier med givet innehåll, med sido- och underordnade avdelningar.³ Utgångspunkten är i stället kontexten och särskilda användningar av en betydelse i förhållande till den mest prototypiska. Följden blir då, om man vill beskriva eller återge systemet, att den traditionella metoden med betydelser och betydelsemoment i linjär eller hierarkisk ordning ifrågasätts. Därför har man föreslagit (och ibland tillämpat) andra varianter för att ge en mer rättvisande bild. Man har bl.a. diskuterat nätverksmodeller där

³ Jag vill här betona ett strikt ordnat system, dvs. tanken på ett komplett system av alla existerande betydelser, med likvärdiga sidobetydelser, och där vi tar oss från det stora till det lilla med bevarade proportioner. Prototypeteorin erkänner givetvis att en viss hierarki råder: en pudel är en hund som är ett däggdjur osv., det förnekas inte, men man anser inte att hela det semantiska arrangemanget är konstruerat på samma vis.

olika delbetydelser hänger samman i korsande och överlappande system (se t.ex. Geeraerts 1990:202). En sådan modell skulle sedan framförallt kunna få representation i digitala ordböcker. En annan variant är att helt slopa momentindelning med siffror och bokstäver och istället presentera ett polysemt ords delbetydelser ihop, med semantiskt närstående betydelser nära varandra. Dessa förslag ska inte recenseras här, men det bör understrykas att ordbokens syfte och målgrupp här givetvis är avgörande. Beskriver man det samtida språket? Vill man beskriva betydelseutveckling eller just det semantiska nätverket?

Vad gäller betydelsebeskrivningen har man traditionellt sett alltså haft stort fokus på betydelseernas avgränsningar mot andra betydelser. Det s.k. polysemiproblemet är välkänt. Vad utgör en egen betydelse? Vad avgränsar den till andra? Hur många finns det, och hur ska man hantera dem i ordböcker? Med den strukturalistiska semantiken kom termer som *fältstruktur* och *arkilexem*, och man talade om åtskillnad genom *minimal opposition*.⁴ Detta gav vatten på kvarnen för tanken om ett ”fyrkantigt” hierarkiskt system. Ett uttalat mål var även s.k. negativa definitioner: för att ge en så exakt bild som möjligt, och för att uttrycka de unika och nödvändiga särdragen hos en betydelse, försöker man så att säga avgöra och beskriva vad en betydelse inte är. En tidigare SAOB-redaktör beskriver en strukturell analys på det här sättet:

Att analysera och beskriva (definiera) en betydelse [...] på strukturell grundval, innebär att grunda analysen och beskrivningen (definitionen) på dels motsättningarna mellan betydelserna inom den (de) betydelsekategori(er), vari ifrågavarande betydelse ingår, dels motsättningarna mellan denna (dessa) kategori(er) och varje annan i systemet motställd kategori. Detta innebär, att beskrivningen (definitionen) med nödvändighet blir negativ i den meningen, att den grundar sig på registrering av distinktiva betydelser (A= icke B, B = icke A). (Eriksson 1975:17)

Den här tanken utgår ifrån en önskan att hitta och beskriva ett fullständigt och objektivt sant system av betydelser som existerar oberoende av människan och tillsynes lever sitt eget liv. Prototypeteorins synsätt är i mycket det motsatta. Som nämndes ovan har vi ett ändrat fokus från gränserna till centrum. Istället för att undersöka vad som nödvändigtvis är fallet, frågar

⁴ Se t.ex. Jonsson (1983:141f.) där semantisk analys och ordbetydelser diskuteras utifrån grammatikens paradigmatiska och formella kriterier, och Eugenio Coserius termer för att förklara ordförrådets struktur presenteras.

man sig: vad är normalt fallet? Greppet blir mindre anspråksfullt, men däremot realistiskt.

För att bli lite mer konkret: både Atkins & Rundell (2008:414ff.) och Hovmark (2012:63ff.), utgår i sina respektive diskussioner om prototypiska definitioner ifrån den s.k. genus proximum-metoden, en metod som varit allmänt vedertagen inom semantisk teori, och som utgör ett sätt att ange nödvändiga/tillräckliga drag. Även Svensén omtalar denna metod i sin *Handbok i lexikografi* (2004) som ett sätt att konstruera en intensionell definition, dvs. den (klassiska) typ av definition där begreppets innehåll (i motsats till dess omfång) fokuseras (s. 274f.). Metoden bygger på tanken om ett hierarkiskt ordnat system, och innebär att man vid definitionen anger det närmast överordnade begreppet för en ordbetydelse + minst ett särskiljande drag. Om man accepterar att begreppet *nåla* ingår i den överordnade kategorin *fästa*, kan en genus proximum-definition utformas som i figur 3.

Nåla	
fästa	med nålar
<i>genus proximum</i>	<i>differentia specifica</i>
Närmast överordnade begrepp	särskiljande drag

FIGUR 3. *Genus proximum-analys av begreppet nåla.*

I fall med uppenbara över- och underordningar och med tydliga kategoriska gränser har den här metoden förtjänster. Det kan då gälla benämningar på biologiska arter eller andra av människan fastställda klassindelningar. Men i många fall är den svårare att tillämpa eftersom den verklighet man vill beskriva är mer komplex. Detta leder vanligen till att definitionerna antingen utformas för snävt: definitionen blir för specifik för att täcka alla användningar som förekommer, eller också motsatsen: definitionen blir så vid att den täcker i princip vad som helst. Därför föreslås att metoden (som alltså trots allt har förtjänster) behöver modifieras lite. Istället för att ange vad som är nödvändigtvis sant, bör man inledningsvis i definitionen ringa in betydelsens centrala kärna. Och istället för de särskiljande dragen, anger man typiska och frekventa fall. Atkins & Rundell (2008:430ff.) ger förslag på hur uppgiften kan lösas: ”A good strategy here is to use a basic definition to describe what is invariably true, then add a clause beginning ‘typically’ or ‘especially’ to describe the prototypical case”. En välfungerande definition i denna anda exemplifieras med ODE-2:s BUS: ”A large motor vehicle carrying passengers by road, typically one serving

the public on a fixed route for a fare” (2008:418). Här anges först det centrala innehållet och efter ”typically” beskrivs en typisk buss. Resultatet för själva definitionen är snarlik genus proximum-metodens, men man har ändrat fokus från att undersöka ett givet strukturellt språkligt system till ett verkligt språkbruk hos människan.

För att som lexikograf komma åt de centrala dragen krävs enligt Hovmark ett stort språkligt material, statistiska verktyg och avancerade korpusökningar (2012:65f.). Detta för att upptäcka särskilda samförekomster av ord eller betydelser och syntaktiska mönster osv. och för att härigenom kunna avgöra vad som är marginellt och vad som är typiskt.

En ytterst intressant poäng, som författarna till båda texterna gör, är att ett tillvägagångssätt som det nyss beskrivna, och själva tanken om betydelsemässig centralitet och svävande gränser, inte har kommit till lexikografen genom prototyp teorin. De har förekommit långt tidigare, men då utan den teoretiska förankringen. Det är alltså inte nytt i sig, det som är nytt är ett samlat teoretiskt helhetsgrepp som verifieras genom empiriska undersökningar.

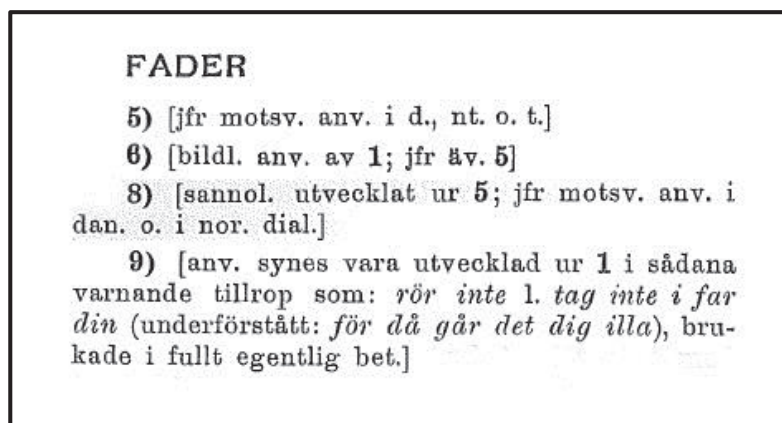
Detta för oss slutligen in på hur SAOB:s metod står sig i förhållande till prototyp teorin. Och liksom i det föregående, tittar vi lite närmre dels på den grafiska dispositionen, dels på utformningen av definitioner.

4. Prototyp teorin och SAOB

Sett till det rent grafiska är SAOB:s struktur hierarkisk. Arrangemanget med huvudmoment och undermoment i olika nivåer känns igen från de andra stora germanska ordböckerna. Strukturen kan göra att en fälthierarkisk tolkning av betydelse- eller begreppssystemet ligger nära till hands; tolkningen att alla betydelser förhåller sig till varandra på samma sätt, med sido- och undermoment med så att säga lika värde.⁵ På ett övergripande plan verkar alltså inte SAOB:s struktur vara förenlig med prototyp teorins ideal. Men som sades tidigare måste man ha ordbokens syfte i tankarna här och SAOB:s fokus är det historiska perspektivet. Formen, etablerad hos de tyska, engelska och nederländska förebilderna, försöker fånga och förklara ordens betydelseutveckling och metoden som används är kanske den enda praktiskt möjliga (och, som påpekades tidigare, viss hierarki förekommer också). Men vad som är intressant ur prototyp teorins synvinkel, trots de strikt ordnade och klart avgränsade avdelningarna, är de upp-

⁵ Jag vill också hävda att det finns tendenser som tyder just på att redaktörer utnyttjat den hierarkiska strukturen djupare och i större utsträckning under perioder då strukturalistiska teorier diskuterats på SAOB-redaktionen (se Nilsson 2012).

lysningar och etymologiska hänvisningar mellan de olika momenten hos polysema ord, som ofta påträffas i ordboken. Det handlar då om sådana uppgifter som återges i figur 4, vilka så att säga korsar varandra linjärt och förklarar icke stringent betydelseutveckling.



FIGUR 4. Uppgifter om betydelseutveckling hos ordet FADER i SAOB.

Här anger upplysningen inom den (etymologiska) hakparentesen den aktuella betydelsens relation till andra betydelser då det linjära systemet (1 → 2 → 3 osv.) inte räcker till. Den här typen av hänvisningar (som förekommer redan i de tidiga banden) understryker prototypeteorins tes att betydelsers inbördes relation ska uppfattas som ett korsande nätverk med kontaktytor åt alla håll.

SAOB har även en särskild modell för att ange att en betydelse eller en del av en betydelse innehåller drag gemensamma med en betydelse som beskrivs i ett annat moment, och således befinner sig i gråzonen mellan två eller flera betydelsemoment. På en nivå längre ner i strukturen, nämligen som ett led (ofta ett avslutande led) i definitionerna förekommer ibland sådana upplysningar som exempelvis: *med bibegrepp av bet. x, med kvardröjande bet. av y, närmande sig l. övergående i bet. z osv.*, eller sådana mer eller mindre ursäktande uppgifter som: *stundom svårt l. omöjligt att skilja från bet x, i ä. språkprov möjl. att hänföra till y osv.* (se exemplet MALÖRT).

MALÖRT

1) växten *Artemisia absinthium* Lin., använd inom medicinen o. (i sht förr) ss. medel mot mal; äv. om drogen, bestående av växtens blad o. blommande grenspetsar (med starkt bitter smak) ... jfr **2**.

2) (numera bl. mera tillf.) dekokt på l. extrakt l. destillat av malört; malörtsdroppar; malörtsdryck; *stundom svårt att skilja från 1*.

SAOB M153 (förf.:s kursiv)

De här upplysningarna visar tydligt svårigheterna i att dra skarpa gränser, och också det faktum att betydelser ofta överlappar varandra. Det ska sägas att anledningen till uppgifterna av det här slaget beror på att SAOB beskriver språket utifrån ett insamlat material som citeras, och varje betydelsemoment beläggs med ett första belägg, och vid inskränkt bruklighet också sista belägg. Och ofta är det betydelserna hos just dessa tidiga belägg som är svåra att bedöma p.g.a. för nutida läsare avvikande syntax eller oklara hänsyftningar och otillräcklig kännedom om kontexten. För att inte riskera att föra ett belägg till fel moment, gör man helt enkelt läsaren uppmärksam på oklarheterna. Men oaktat bakomliggande orsak speglar det människans sätt att kategorisera begrepp genom oskarpa gränser och överlappande struktur.

Om vi istället tittar på betydelsebeskrivningen i SAOB, så sades det tidigare att genus proximum-definitioner ofta kan fungera, men dock i begränsad utsträckning. Problemen som uppstår när metoden inte räcker till är att definitionerna blir för vaga och allmängiltiga, eller också för snäva och specifika. I SAOB kan man stöta på just sådana definitioner, där viljan att täcka alla möjliga situationer står i vägen för läsaren. Nedan exemplifieras ett sådant fall genom SAOB:s definition av SPION.

SPION

1) person som bor l. uppehåller sig i ett land o. där (mot betalning) i smyg l. gm otillåten underrättelseverksamhet införskaffar o. till en (utländsk) uppdragsgivare vidarebefordrar (hemlighållna) uppgifter l. (hemlighållet) material som rör ifrågavarande lands militära förhållanden l. säkerhet o. d.; särsk. dels om förklädd o. med falsk identitet utrustad soldat som i krig har till uppgift att ta sig in i l. bakom fiendens linjer l. in i hans land för att där göra sig underrättad om hans förehavanden l. planer l. militära styrka l. stridsmoral l. dyl. o. rapportera härom till den egna sidan (förr möjl. äv. om soldat som i krig sändes ut för att, under undvikande av kontakt med motsidan, spana l. speja l. rekognosera, spanare, spejare, rekognosör, o. i så fall anträffat bl. i anv. där denna bet. sammanfaller med den ovannämnda), dels om person som (äv. i fredstid) av underrättelsetjänsten i ett land skickas till ett annat land för att där bedriva militär underrättelseverksamhet, (hemlig) agent; äv. övergående i allmännare bet.: person som (mot betalning) spionerar (se **SPIONERA** i allmännare bet.) o. skaffar fram information l. material (om ngn l. ngt); äv. (jfr **2)** oeg. l. mer l. mindre bildl. (särsk. om forskare som gräver l. snokar i källor för att komma ngt på spåren l. klarlägga ngt); särsk. liktydigt med: kunskapare.

En välfungerande prototypisk definition bör i enlighet med Atkins & Rundells resonemang ovan i stället ha sin utgångspunkt i en kärna av centralt innehåll + typiska attribut eller typiska användningar, och jag vill påstå att den här metoden är den som SAOB främst använder sig av för att lösa problem med för stort omfång av betydelseinnehåll. Dvs. vid de fall då den semantiska verklighet som ska beskrivas är mer abstrakt och komplex, och de kategoriska gränserna mer flytande än den strikt hierarkiska genus proximum-metoden tillåter.

Definitionsspråket i SAOB är ytterst formaliserat. En redaktör har i princip ett antal byggstenar att konstruera en definition av. Förkortningen *särsk.* (högfrekvent i definitionstexten) utgör en sådan byggsten. *Särsk.* (utläst *särskilt*), används med speciell innebörd i olika sammanhang, men framförallt används den i definitioner för att markera att det som följer efteråt är ett särskilt fall, ett exempel, eller en speciell variant av det som sades före i direkt anslutning.⁶ En definition av det här slaget är mycket vanlig i SAOB, som i fallen EXEGES och SLÅ UPP:

EXEGES

utläggning, tolkning, förklaring (av ord, mening, skrift); särsk. *teol.* om tolkning o. utläggning av den heliga skrift l. enskilt bibelställe; bibeltolkning.

SLÅ UPP

29) om ngt abstrakt: uppkomma o. d.; särsk. dels om känsla: flamma upp l. blossa upp ..., dels om sjukdom: blossa upp (i en trakt o. d.), dels om tanke o. d.: dyka upp, bli medveten.

SAOB E814 resp. S7120

Definitionen inleds här i båda fallen med betydelsens kärninnehåll som fångar de olika användningarnas centrala drag. Detta innehåll anges inte alldeles precist, dvs. det är inte nödvändigtvis så att alla delar av det gäller samtliga användningar av den aktuella betydelsen. Denna betydelsemässiga vaghet antyds här i definitionen genom förkortningen *o. d.* (utläst *och dylikt*) och genom fakultativa definitionsled inom parentes. De speciella fall som redaktören presenterar efter *särsk.*, specificerar egentligen inte den övergripande betydelsen eller avgränsar den mot andra betydelser, utan ges snarare som exempel på typisk användning. SAOB-redaktören har även andra möjligheter att presentera sådana typiska och frekventa användningar i definitionen på liknande sätt. Andra termer som förekom-

⁶ För vidare upplysning om användningen av förkortningen *särsk.* i SAOB, se Lundblad (1992).

mer med snarlik funktion är t.ex. *vanl.* (utläst *vanligen*), *ofta*, *stundom* och *i sht* (utläst *i synnerhet*).

Definitionsled inom parentes såsom ett fakultativt tillägg till en definition (som vid exemplet EXEGES ovan) förekommer ofta i SAOB för att ange typiska fall eller attribut för betydelsekategorin i fråga. Som framgår av exemplet TIDNING kan sådana led i princip förekomma på vilken position som helst i definitionen.

TIDNING

2) regelbundet (*numera ofta dagligen*) utkommande tryckalster (*i större format*) med allmänt (*aktuellt*) innehåll (*förr i sht bestående av officiella kungörelser l. meddelanden o. d.*)

SAOB T1068 (förf.:s kursiv)

I definitioner av det här slaget är det möjligt att tolka parentesernas innehåll såsom exempel på prototypisk användning, vid sidan av betydelsens övergripande och mer eller mindre konstanta kärninnehåll.

Metoden att peka på det centrala eller att snäva in sfären och sedan ange typfall är egentligen en pragmatisk lösning. Den beror helt enkelt på att man inte kan definiera exakt vid alla tillfällen, och att, som vi såg tidigare, en mer specifik definition inte automatiskt innebär en bättre definition. Men metoden är, enligt min mening, förenlig med tanken om prototyper och benägenheten att kategorisera prototypiskt. Jag ska också tillägga att metoden givetvis inte är unik för SAOB. Så här går man tillväga också vid andra lexikografiska arbeten och har gjort sedan lång tid tillbaka. Men man har ofta gjort det utan försök att förankra sin metod och sitt praktiska arbete i vetenskapliga teorier.

För att kunna avgöra vilka användningar som är prototypiska, vilka drag eller attribut som är centrala och vad som är mer marginellt, bör en redaktör (i enlighet med Hovmark) arbeta med ett rikligt språkligt material, med stora sökbara korpusar. Ett problem för SAOB är att dess material utgörs av en excerptsamling, ett insamlat material i lappformat, med utdrag ur utvalda texter.⁷

Ett välexcerperat ord med många belägg, i praktiken alltså många lappar, kan vara tillräckligt för att skapa sig en rättvis bild av användningarna

⁷ Se Rosqvist (2011) och Svensson (1992) för en närmare analys av SAOB:s excerptsamling. Rosqvist (2011:16f.) diskuterar frågan om representativiteten hos SAOB:s material, i språket i stort och i förhållande till ett korpusmaterial. I sin undersökning av kollokationer visar hon att SAOB:s täckningsgrad för hennes utvalda flerordsuttryck är hög.

av ordet och för att upptäcka och förstå frekventa uttryck och snävare specialanvändningar.

Men ofta har excerptisterna antecknat väldigt speciella och märkliga ordanvändningar, sådana som sticker ut och som man lätt reagerar på när man läser en text, och som i sig kanske inte är särdeles representativa för eller frekventa i språket i stort. Redaktörens material bli då skevt och kan leda tolkningen åt fel håll. Dessutom finns det ord och användningar som är sparsamt belagda med få excerpter och som av den anledningen blir svårbedömda. SAOB:s material begränsar i det avseendet möjligheten att göra en mer objektiv bedömning.

För inte så länge sedan innebar det här att en SAOB-redaktör vid dylika tillfällen endast hade sin egen (och sina kollegors) språkkänsla att lita till, och andra ordböcker och liknande. Men med dagens tekniska utveckling ser verkligheten annorlunda ut och man har nu tillgång till korpusar och textsökningsverktyg. Uppgiften är fortfarande att beskriva språket utifrån excerpterna i materialet och att ge citat ur de särskilt angivna källorna. Men redaktörernas analys består nu i hög grad av korpusbaserad sökning som referens och för att undersöka kontexter, bruklighet, frekvens och speciellare användningar. Detta faktum, och det att också allt fler av de källor som SAOB citerar blir tillgängliga och sökbara i digital form, gör det lättare att avgöra vad som är centralt och vad som är perifert. Arbetet och analysen förankras därmed i högre grad i empiri, i enlighet med prototypeteorins strävan.

Att SAOB:s huvudsakliga material utgörs av en excerptsamling innebär att redaktörerna inte har samma möjlighet att jämföra frekvenser och dra statistiska slutsatser som om materialet baserats på korpusar. Det ska dock framhållas att det kan finnas fördelar med ett sådant material i förhållande till ett korpusbaserat, särskilt när det gäller historiska ordböcker som beskriver diakrona betydelseförändringar. Lundbladh (1996:96f.) framhäver t. ex. nyttan av att kunna välja ut kontextuellt relevanta textavsnitt ur de excerpterade källorna. Svensén (2004:56) som jämför beläggmetoden med korpusmetoden föreslår för historiska ordböcker att excerpter och korpusmaterial kan kombineras, dels för att kunna analysera betydelseförändringar på djupet, dels eftersom äldre källor inte finns att tillgå digitalt i tillräcklig grad. I takt med att källorna som SAOB citerar alltså digitaliseras i större utsträckning balanseras också SAOB:s kombination av de båda metoderna upp.

5. Sammanfattning

Efter en genomgång av SAOB:s metod utifrån grafisk uppställning, utformning av definitioner och delar av den språkliga analysen vill jag hävda att det finns en bas i ordboken som tangerar prototypeteorins ideal. Den här basen, eller de tankegångar den representerar, har inte tillfogats SAOB efter det att prototyp teorin kommit till och utvecklats, utan har funnits där redan från arbetets början. Av detta vill jag dra två slutsatser. För det första: prototyp teorin kan bekräfta SAOB:s metod. Att i teorin få verifierat att arbetssättet speglar hur människan strukturerar begrepp innebär legitimitet för metoden. Och att knyta det praktiska arbetet till en teori skulle kunna innebära förtjänster. Redaktionen kan ta ett samlat grepp, och arbeta mot ett gemensamt mål såväl i övergripande frågor som på detaljplanet. I praktiken kan det fungera som en ingång och korta ner startsträckan för nya redaktörer, och för själva texten kan det ge upphov till smidigare lösningar och relevantare definitioner. För SAOB, som ju ska beskriva ett enormt ordförråd och flera språkskeden bör en tillämpbar teori – just som prototyp teorin – vara bred och övergripande, och täcka olika typer av betydelsesammanhang och tillåta olika lösningar i olika fall. Att utgå ifrån människans sätt att kategorisera underlättar för redaktören, men också för läsaren som ju så att säga har metoden tillgänglig själv.

För det andra (och omvänt): SAOB:s metod kan bekräfta teorin. Från två olika utgångspunkter (den teoretiska och den praktiska) har man nått samma resultat. SAOB:s ordboksartiklar bygger på faktiskt språkbruk och ger ytterligare empiriskt stöd för prototyp teorins tolkning av begreppsstrukturen. I förlängningen kanske detta sammantaget kan bidra till att minska avståndet mellan den akademiska semantiska teorin och den praktiska lexikografin.

Källor och litteratur

- Atkins, B. T. Sue & Michael Rundell 2008. *The Oxford Guide to Practical Lexicography*. Oxford: Oxford University Press.
- Eriksson: Ulrik 1975. Om betydelse, betydelseanalys och betydelsebeskrivning. I: *Kring en ordbok. Festskrift till Sven Ekbo*. Stockholm: Norstedts. S. 1–45.
- Geeraerts, Dirk 1990. The lexicographical treatment of prototypical polysemy. I: Tsohatzidis, Savas. L. (red.), *Meanings and Prototypes. Studies in linguistic categorization*. London: Routledge. S. 195–210.
- Hovmark, Henrik 2012. Betydningsbeskrivelse og prototypeteori. I: *LexicoNordica* 19. S. 59–78.

- Jonsson, Hans 1983. Om historisk lexikografi av typ SAOB. I: *Nysvenska studier* 63. Stockholm. S. 127–158.
- Lundbladh, Carl-Erik 1992. *Handledning till Svenska Akademiens ordbok*. Lund: Svenska Akademien & Norstedts.
- Lundbladh, Carl-Erik 1996. Värdering av SAOB:s korpus och språkprovssamling. I: *LexicoNordica* 3. S. 91–103.
- NE (1989–1996) = *Nationalencyklopedin. Ett uppslagsverk på vetenskaplig grund utarbetat på initiativ av Statens kulturråd*. 1–20. Höganäs: Bra Böcker.
- Nilsson, Pär 2012. Språkteoriens avtryck i SAOB. I: Eaker, Birgit, Lenhart Larsson & Anki Mattisson (red.), *Nordiska studier i lexikografi* 11. *Rapport från Konferensen om lexikografi i Norden*. Lund. S. 479–487.
- ODE-2 (2003) = *Oxford Dictionary of English*, 2 uppl. Oxford: Oxford University Press.
- ODS (1919–56) = *Ordbog over det danske Sprog* 1–28. Köpenhamn: Det Danske Sprog- og Litteraturselskab & Gyldendal.
- Rosch, Eleanor 1978. Principles of Categorization. I: Rosch, Eleanor & Barbara B. Lloyd (red.), *Cognition and Categorization*. Hillsdale, New Jersey: Lawrence Erlbaum Associates. S. 27–48.
- Rosqvist, Bodil 2011. Strängt upptagen och fast besluten. *Beskrivningen av kollokationer i SAOB*. (Meddelanden från Institutionen för svenska språket 64.) Göteborg.
- SAOB (1898–) = *Ordbok över svenska språket utgiven av Svenska Akademien*. 1–. Lund.
- Svensén, Bo 2004. *Handbok i lexikografi. Ordböcker och ordboksarbete i teori och praktik*. Stockholm: Norstedts Akademiska förlag.
- Svensson, Lars 1992. Om SAOB:s material och en bibliografisk databas. I: Fjeld, Rut Vatvedt (red.), *Nordiske studier i leksikografi. Rapport fra Konferense om leksikografi i Norden*. Oslo S. 369–379.
- Wittgenstein, Ludvig 1978. *Filosofiska undersökningar*. Stockholm: Bonniers.

Perspektiv på svensk lexikografi i Finland med *Ordbok över Finlands svenska folkmål* och *Finlandssvensk ordbok* som exempel

Caroline Sandström

In this article I focus on two Swedish dictionaries which are compiled at The Institute for the Languages of Finland in Helsinki. The two dictionaries are *Ordbok över Finlands svenska folkmål* (The dictionary of Swedish dialects in Finland) (1976–) and *Finlandssvensk ordbok* (The Dictionary of Finland-Swedish) (2000; 2008). These projects are discussed in a Scandinavian context. These two dictionaries are investigated in a historical perspective, e.g. they have common roots in the Swedish Nationalistic movement. They are described in detail according to their structure and content with examples of the entries from both dictionaries. Moreover, future possibilities that e.g. electronic lexicography can offer are considered.

1. Inledning

Mitt bidrag fokuserar på svensk lexikografi i Finland. Jag anlägger dels ett tillbakablickande perspektiv, dels blickar jag framåt. I fokus står arbetet med två enspråkigt svenska ordböcker: den finlandssvenska dialektordboken *Ordbok över Finlands svenska folkmål* (1976–) och *Finlandssvensk ordbok* (2000; 2008), som utarbetats på Institutet för de inhemska språken i Helsingfors.

De kan sägas ha uppkommit genom praktiskt arbete inom var sin ordbokstradition efter mönster från tidigare ordböcker i Finland och Sverige, men utan en medvetet kodifierad lexikografisk teori (se t.ex. Fjeld 1992:1f., Ralph 1992:29ff. och 2000:6f.).

Institutet för de inhemska språken är den största lexikografiska institutionen i Finland, och vid sidan av språkvård hör lexikografi till institutets kärnverksamhet (se www.sprakinstitutet.fi/sv/om_oss/). Institutet inrättades 1976, då bl.a. finsk och finlandssvensk språkvård, flera pågående ordboksprojekt, onomastiska arkiv och arkiv för finska bandinspelningar

sammanfördes till Forskningscentralen för de inhemska språken.¹ De aktuella ordboksprojekten var då *Ordbok över Finlands svenska folkmål*, den finska dialektordboken (fi. *Suomen murteiden sanakirja*), ordboken över äldre finskt skriftspråk (fi. *Vanhan kirjasuomen sanakirja*) och ordböcker över de finska släktspråken karelska (fi. *Karjalan kielen sanakirja* 'ordboken över karelska'), mordvinska, mansiska (voguliska) och mariska (tjeremissiska).

När det gäller normeringen av svenskan i Finland är målet för skriftspråket att följa den allmänsvenska normen för stavning, böjning, ordförråd och grammatik (Reuter 2006:42f., af Hällström-Reijonen & Reuter 2008:5). Det utarbetas inte separata nuspråkliga ordböcker för finlandssvenskan, utan finlandssvenskan följer de normer som anges i allmänsvenska ordböcker producerade i Sverige. I verk som *Svenska Akademiens ordbok* (SAOB), *Svenska Akademiens ordlista* och *Svenska Akademiens grammatik* beaktas dock finlandssvenska särdrag i viss mån. Det huvudsakliga syftet med *Finlandssvensk ordbok* är att ge en normerande förteckning över finlandssvenska särdrag på lexikal nivå (af Hällström-Reijonen & Reuter 2008:3).

Dispositionen för mitt bidrag, där *Ordbok över Finlands svenska folkmål* och *Finlandssvensk ordbok* läggs in i ett bredare lexikografihistoriskt sammanhang, ser ut på följande sätt.² Först ges en bakgrund till svensk lexikografi i Finland dels i relation till tvåspråkighetslexikografi i avsnitten 2 och 3, dels i ett nordiskt perspektiv i avsnitt 4. Ordböcker som nationella projekt diskuteras i avsnitt 5. De två ordböckerna definieras typologiskt i avsnitt 6, innan de i avsnitten 7 och 8 avhandlas var för sig och belyses med utdrag. I den avslutande diskussionen dryftas nya utmaningar, som bl.a. publicering på nätet innebär för ordböckerna.

2. Svensk lexikografisk tradition med rötter i tvåspråkighetslexikografi

Lexikografisk verksamhet har funnits i Sverige i över femhundra år. Som Hannesdóttir (1998:1f., 494) framhåller inleddes och utvecklades den svenska lexikografiska traditionen inom den tidiga tvåspråkighetslexikografien. Hannesdóttir (1998) undersöker framväxten av en svensk lexiko-

¹ *Institutet för de inhemska språken* är underställt finska Utbildnings- och kulturministeriet. Fram till 31.12.2011 verkade det under namnet *Forskningscentralen för de inhemska språken*. I fråga om verksamhet som entydigt hänför sig till perioden 1976–2011 använder jag i fortsättningen den etablerade förkortningen *Focis*.

² I fortsättningen används förkortningarna FO för *Ordbok över Finlands svenska folkmål*, FSOB för *Finlandssvensk ordbok*.

grafisk tradition, alltså ”utvecklingen av en adekvat lexikalisk beskrivning av det svenska språket” (s. 494). Hon har i sin undersökning urskilt tre perioder i denna utvecklingsprocess. Under den första perioden fram till 1600-talets slut etableras ordboksproduktionen i Sverige. Den domineras av latinets starka ställning och behovet av att tillägna sig och kunna använda det latinska språket. Under den andra perioden från 1700-talets början till 1800-talets mitt utvecklas en svensk lexikografisk tradition. Den kännetecknas av ett ökat behov av att bredda svenskans funktioner i alla sammanhang i det svenska samhället, men också av krav på större kunskaper i andra europeiska folkspråk. Den tredje perioden innebär en seger för den enspråkiga svenska ordboken; en milstolpe är A.F. Dalins *Ordbok öfver svenska språket* som utkom vid mitten av 1800-talet. Hannesdóttir (1998:495) sammanfattar:

Med Dalins ordbok fastställs den enspråkiga svenska lexikografin, samtidigt som den framträder som fullt utvecklad. Att arbetet till stor del bygger på den redan etablerade tvåspråkiga lexikografin är uppenbart. Både ordbokens omfång och dess innehåll tyder på detta.

3. Tvåspråkighetslexikografi i Finland under 1900-talet

Finland är av hävd ett flerspråkigt land. Allt sedan regeringsformen 1919 och språklagen 1922 trädde i kraft har Finland haft två nationalspråk, finska och svenska (Reuter 2006:29ff.). Därtill finns minoritetsspråken samiska, romani och teckenspråk, och andra språk som under olika tidsperioder har talats eller talas av minoriteter i Finland, t.ex. ryska, tyska, estniska. Tvåspråkighetslexikografin, med en omfattande produktion av tvåspråkiga ordböcker, har därför av praktiska orsaker varit en vital och viktig uppgift för finländska lexikografer (Romppanen 2000, 2001).

Romppanen (2001) studerar finsk-svenska ordböcker med fokus på övergången från receptionsordböcker till produktionsordböcker. Med början från Cannelins *Finskt-svenskt lexikon/Suomalais-ruotsalainen sanakirja* (1908) förändras strukturen i de finsk-svenska ordböckerna under 1900-talets lopp från att primärt ha varit receptionsordböcker för finskan till en kombination av receptionsordböcker för finskan och produktionsordböcker för svenskan. Genom Cannelins *Finskt-svenskt lexikon* (1908) ”kunde den finskspråkige användaren börja använda den finsk-svenska ordboken för produktion av texter på svenska” (Romppanen 2001:245).³ En

³ Knut Cannelin gav fr.o.m. slutet av 1880-talet ut flera finsk-svenska hand- och storordböcker. Hans arbete fortsattes av Aulis Cannelin, Lauri Hirvensalo och Nils Hed-

stor andel av de tvåspråkiga ordböcker som gavs ut i Finland under perioden 1870–1998 hade svenska och finska som målspråk och/eller källspråk (Romppanen 2001:198, 245f., 298ff.).

Romppanen (2001:198ff.) understryker att det i förlagens utgivning av finsk-svenska och svensk-finska ordböcker fanns flera toppar. Det gäller dels utgivningen under 1870- och 1890-talet, dels under perioden 1930–98, och i synnerhet utgivning av ordböcker mellan finska och svenska under 1970- och 1980-talet. Romppanen påpekar att finsk-svenska ordböcker publicerades i många och tämligen stora upplagor. Hon förklarar dessa utgivningstoppar med att det under 1900-talet skedde stora samhälleliga och tekniska omvälvningar i Finland. Faktorer som påverkade efterfrågan och utgivningen av finsk-svenska ordböcker var till exempel att undervisningen i svenska genom grundskolereformen 1968 blev obligatorisk i den finska grundskolan.⁴ Dessutom ökade resandet och migrationen mellan Finland och Sverige kraftigt under 1960-talet. Från 1980-talet och framåt kan modern marknadsföring och distribution av ordböcker genom bokklubbar, varuhus och stormarknader ha haft betydelse för det stora antalet upplagor.

Såväl Hannesdóttir (1998:65, 70ff.) som Romppanen (2001:12ff.) framhåller att den tvåspråkiga lexikografin hamnade i skuggan, då de stora enspråkiga ordboksprojekten etablerades fr.o.m. senare hälften av 1800-talet. Att tvåspråkighetslexikografin i Sverige och Finland kom i skymundan syntes bl.a. vid inriktningen för akademisk forskning och i utdelningen av statliga anslag. Tvåspråkighetslexikografin sågs allt mer som en förlagsbaserad kommersiell verksamhet, medan den enspråkiga lexikografin fick en självklar hemorts rätt inom akademisk forskning (Hannesdóttir 1998:102). Samtidigt bör man hålla i minnet att lexikografisk verksamhet fram till slutet av 1900-talet inte hade en självfallen plats inom lingvistik; ordböcker uppfattades i bl.a. Noreens (1903:209) anda av många lingvister som praktiska hjälpredor (se t.ex. Ralph 2000:5f.).

4. Lexikografisk metod- och teoridiskussion uppstår genom nordiskt samarbete

Under 1980- och 1990-talet inleddes lexikografisk forskning på bredare front i Norden; de första handböckerna sammanställdes (Svensén 1987,

lund. Den mest omfattande av dessa ordböcker, *Finskt-svenskt lexikon*, gavs ut i flera upplagor under perioden 1908–76 (Romppanen 2001:29, 202ff., 245f., 350ff.).

⁴ I praktiken infördes grundskolan i Finland stegvis under 1970-talet. Tidigare hade undervisningen i svenska huvudsakligen omfattat bara de finska elever som gick i läroverk.

Bergenholtz & Tarp 1994) och lexikografen kom in som ett undervisningsämne på universiteten. Den lexikografiska metod- och teoridiskussionen i Norden och framväxten av lexikografen som forskningsgren var ett led i en internationell utvecklingsprocess (se t.ex. Ralph 1992:29ff., 37f., Svensén 2004:7f., 13f., Bergenholtz & Malmgren 2004:2ff., NLO 1997:7). Vid samma tidpunkt utgavs de första översiktsverken där svensk lexikografihistoria beaktades, och forskningsprojektet ”Lexikografisk tradition i Sverige” inleddes. Projektet lade grunden för flera arbeten om lexikografins historia i Sverige; bl.a. de omfattande avhandlingarna (Johansson 1997, Rogström 1998, Hannesdóttir 1998) och många konferensföredrag och artiklar (Ralph 2000).

Under åren 1990–97 utarbetades *Stora finsk-svenska ordboken* som ett samarbete mellan förlaget Werner Söderström Oyj (WSOY) och Focis. När det gäller forskningen i tvåspråkighetslexikografi med anknytning till svenskan i Finland kom projektet att fungera som ett startskott (Romppanen 2001:13). Redan under början av 1990-talet fick redaktörerna på *Stora finsk-svenska ordboken* via det påbörjade redigeringsarbetet tillfälle att inleda samarbete och meningsutbyte med nordiska kolleger inom ramen för Nordisk förening för lexikografi (Romppanen 2001:12ff.). Ordbokens redaktörer Ilse Cantell och Nina Martola kom att få synliga poster i föreningen genom att de aktivt deltog i de konferenser och symposier som arrangerades från den första Konferensen om lexikografi i Norden, som hölls i Oslo 1991 (se t.ex. Martola 1992, 1995 och 1996; Cantell 1995). Martola var finländsk landsredaktör för tidskriften *LexicoNordica* från starten 1994 fram till nummer 18 år 2011. Cantell medverkade som en av redaktörerna i arbetet med *Nordisk lexikografisk ordbok* (NLO) som publicerades 1997. Både Cantell och Martola har varit medlemmar i styrelsen för Nordisk förening för lexikografi.

Utifrån mitt perspektiv – som knuten till det enspråkiga ordboksprojektet FO sedan år 1990 – var kollegerna Cantell och Martola, som arbetade med tvåspråkighetslexikografi, mer etablerade och aktiva inom den nordiska teori- och metoddiskussionen än medarbetarna i mitt eget projekt. Detta kan bero på att den finlandssvenska dialektordboken, liksom andra samtida nordiska dialektordböcker, var förankrad i de (dominerande) akademiska forskningsgrenarna språkhistoria och dialektologi, och därför inte i samma utsträckning hade ett behov att skapa kontakter med lexikografkolleger i Norden. Under 1990-talet och början av 2000-talet var arbetet på FO inriktat på redigering enligt fastlagda principer. De tekniska ramarna för projektet gav inte anledning till nytänkande eller revidering av det redaktionella arbetet, samtidigt som tidtabellen för publicering inte tillät extra arbetsuppgifter i form av lexikografisk forskning eller nordiskt lexikografiskt samarbete (jfr diskussionen i Fjeld 1992:1ff. och Ralph 1992:28ff.).

I liten skala inleddes dock lexikografiskt samarbete mellan de pågående ordboksprojekten på Focis i form av gemensamma arbetsmöten och en lexikografisk studiecirkel. Då Fjärde konferensen om lexikografi i Norden hölls i Esbo i Finland 1997 deltog Focis lexikografer i konferensen – och flera av dem publicerade bidrag i konferensrapporten (Haarala 1999, Kulonen 1999, Lehtinen 1999, Ruppel 1999, Vilppula 1999). Huvudredaktören för FO ledde dessutom redigeringen av rapporten från denna konferens (Slotte, Westerberg & Orava 1999).

Till skillnad från den finlandssvenska dialektordboken insåg redaktionerna för andra nordiska dialektordböcker och historiska ordböcker som *Jysk Ordbog*, *Norsk Ordbok* och SAOB från början betydelsen av att delta i det nordiska lexikografiska samarbetet (se t.ex. Hageberg 1992, Jonsson 1992, Lundbladh 1996, Svensson 1992, Sørensen 1992, Vikør 1995). För de nordiska dialektordböckernas vidkommande kan tilläggas att samarbete och utbyte av erfarenheter mellan dem som arbetar med de stora nordiska dialektordböckerna fortlöpande upprätthållits genom de vart fjärde år återkommande nordiska dialektologkonferenserna, som inleddes 1978 i Göteborg, med de stora dialektordböckerna som ett av konferensens huvudteman (Ahlbäck et al. 1978, Benson 1978).

5. Ordböcker som nationella projekt

I verket *Imagined communities* har Anderson ([1991] 2003) lyft fram betydelsen av ett kodifierat, skrivet standardspråk. Han framhäver att verksamheten att skriva och redigera – och i synnerhet att låta trycka och sprida – ordböcker har varit ett viktigt instrument för att bygga upp och skapa en nationell samhörighetskänsla i många europeiska länder (Anderson ([1991] 2003:3ff., 37ff.).

Svensén (2004:1f.) poängterar å sin sida att ordboken som kulturell företeelse utgör en produkt av det samhälle där den tillkommit. Samtidigt är ordboken en viktig faktor i utvecklingen av denna kultur. Som pendang till *brukslexikografi*, med uppgift att understödja kommunikationen och inläringen av språk, nämner Svensén *dokumentationslexikografi*, vars syfte är ”att tillgodose ett nationellt, politiskt, vetenskapligt etc. kollektivs behov av att bevara information för framtiden” (Svensén 2004:1). (Se även NLO 1997:15f., 169.)

Considine (2008:2ff.) diskuterar också ordboksarbete och lexikografi som bärare av kulturarv och som en form av heroiska narrativer. Han understryker att kulturella företeelser genom lexikografisk verksamhet överförs från generation till generation och nämner som exempel ordboks-

verket *Canadian Oxford Dictionary*. Considine (2008:9) framhåller att ordböcker även är bärare av identiteter:

Dictionaries have sometimes been records of the emerging sense of identity of particular speech-communities or reading communities, groups defined by use of a common spoken language or by commitment to the study of a common body of texts. These communities are often ill-defined, and an individual may belong to more than one. By no means are they always nations, and this is not a book about nationalism: the members of a speech-community may constitute a nation, but they may also spread over several nations, or by a minority in one nation, or both.

Citatet från Considine kan tillämpas på de två aktuella ordböckerna FO och FSOB. Båda behandlar svenskan i Finland, men de kan samtidigt ses som nationsöverskridande. De dokumenterar ord och uttryck som används i Finland, men det finns ett samband med och en koppling till svenska dialekter och varieteter i Sverige. I synnerhet gäller detta FSOB som har som syfte att ge sverigesvenska motsvarigheter och betydelser för ord och uttryck som enbart eller huvudsakligen används av finlandssvenskar.

I fråga om de svenska dialekterna i Finland är det historiska bandet till Sverige starkt, eftersom svenska bosättare från Sverige tog med sig de svenska dialekterna till Finland (Ahlbäck [1956] 1971, Tarkiainen 2008:49ff., 106ff., 218ff.). Genom processer som dialektdivergens och -konvergens har svenska dialekter i Sverige och Finland under århundradenas lopp utvecklats och förändrats efter såväl likartade som olika mönster (se t.ex. Thelander 1996:164ff. och Sandström 2010:43ff.). Finland hör till ett område där språkkontakt förekommit i alla tider; språkkontakt med finska är därför en viktigare faktor att räkna med i Finland än i stora delar av Sverige (Thors 1981, Sandström 2011).

Enspråkiga historiska ordböcker och dialektordböcker har uppfattats som en nationell angelägenhet med uppgift att bevara ett kulturarv. Detta har i Finland varit påtagligt för dokumentationen av och forskningen i de svenska dialekterna, men också för den svenska språkplaneringen. En av förgrundsgestalterna för forskningen i svenska språket i Finland var Axel Olof Freudenthal (1836–1911). Han införde dialekt- och ortnamnsforskning på svenskt håll i Finland. Han var också en av de tidiga aktörerna inom svensk språkvård i Finland och samtidigt en portalfigur inom den svensknationella samlingsrörelsen. Freudenthal blev den förste finländske professorn i svenska språket. I början av 1860-talet studerade han under ett år i Uppsala, där han tog intryck av den filologiskt inriktade dialektforskningen, som han senare förde vidare till sina studenter vid universitetet i Helsingfors (Huldén 1990:45f., Loman 1985:15ff., 25ff. och 1991:45f.,

Slotte 1992:56f., af Hällström-Reijonen 2012:45ff., Thylin-Klaus 2012:60ff., 84ff., 104f.).

Rötterna för dialektologin finns i nationalromantiken och folkrörelserna. Dialektologin uppkom under senare hälften av 1800-talet som en gren av den språkhistoriska forskningen med inriktningar som textfilologi och historisk-jämförande språkvetenskap. Den nationalromantiska och arkaiserande kopplingen framträder bl.a. i synen på dialekterna som forskningsobjekt, där man sökte de äldsta, mest ursprungliga dialekterna. Denna koppling syns också i termer som *folkmål*, *bygdemål* och *landsmål*, som användes av en del dialektforskare så sent som under 1950- och 1960-talet. Folkmål uppträder t.ex. i namnet på den finlandssvenska dialektordboken *Ordbok över Finlands svenska folkmål* (Sandström 2000 och 2013).

Både den svenska dialektforskningen och språkvårdsverksamheten i Finland initierades alltså av Freudenthal under 1870- och 80-talen. Efter mönster från Uppsala grundades 1874 Svenska landmålsföreningen i Finland. Under Freudenthals ledning började föreningen samla in såväl dialektord som finlandismer (af Hällström-Reijonen 2012:50ff., Thylin-Klaus 2012:9f., 80ff.). Finlandismer definierades 1892 av Svenska landmålsföreningen som ”på svensk språkgrund uppvuxna och till stor del af våra folkmål närda egenheter i det svenska språkbruket i Finland” (Vendell 1894:305).

6. Olika typer av ordböcker

I Svensén (2004:15ff.) och i inledningen till NLO § 29–30 diskuteras typologier av ordböcker. Utgående från dem kan de två ordböckerna FO och FSOB som står i fokus i denna artikel placeras in bland enspråkigt svenska specialordböcker. De kan vidare enligt Svensén (2004:43ff.) hänföras till gruppen *avgränsade ordböcker*, som begränsat sitt urval av lexikaliska enheter efter vissa kriterier. I NLO § 30 behandlas avgränsade ordböcker över vissa språkvarieteter, vilka kan delas in enligt tid: *historiska ordböcker* och rum: *dialektordböcker* och *regionalordböcker*.

Svensén (2004:43, 45f., 28) tar upp avgränsade ordböcker som avser vissa lemmatyper, informationstyper, språkvarieteter eller texter. Han presenterar som grupp (c) *Ordböcker över vissa språkvarieteter*. De kan indelas i flera undergrupper efter språkvarieteternas art, t.ex. *historiska ordböcker* ”över tidsmässigt avgränsade språkvarieteter” eller *dialektordböcker* ”över geografiskt avgränsade, icke riksspråkliga språkvarieteter”. En dialektordbok kan beskriva dialekten i ett mindre område, t.ex. i en stad eller i en socken, i en landskapsdel eller i ett helt landskap. Både Svensén

(2004:46) och NLO (1997:98, *dialektordbok*) nämner att det även finns rikstäckande dialektordböcker, t.ex. *Ordbok över Sveriges dialekter*.

NLO definierar en *dialektordbok* som en ”ordbok som behandlar en geografisk avgrenset, ikke riksspråklig språkvarietet”. Beroende på geografisk täckningsgrad kan man för det första tala om en *territorialordbok*, som t.ex. kan omfatta det svenska landskapet Småland, för det andra en *regionalordbok*, som kan omfatta dialekten i ett mindre dialektområde i en del av ett landskap, t.ex. syds måländska götamål, och för det tredje en *lokalordbok*, som omfattar språket i en socken eller en by. Förutom dessa geografiskt avgränsade ordböcker över regionala språkvarieteter tar NLO upp termen *regionalismordbok*.⁵ En *regionalismordbok* behandlar regionalt förankrade varianter (synonymer eller dubletter) som är i stort sett likvärdiga och existerar vid sidan av varandra. I normala fall täcker en regionalismordbok bara en enda region, ytterst sällan ett helt språkområde. I NLO jämförs denna typ av ordbok med de diatopiska markörerna för provinsialt och regionalt, *prov.* och *region.*, som kan förekomma för vissa ord och uttryck i allmänspråkliga ordböcker. En *regionalismordbok* behandlar således regionalt förankrade, i stort sett likvärdiga, varianter som existerar jämsides med varianter i standardspråket. Upplägget och avgränsningen i FSOB gör att den närmast kan betecknas som en *regionalismordbok*.

7. Ordbok över Finlands svenska folkmål (FO)

Enligt den typologiska indelning av ordböcker som ges i NLO och Svensén (2004) är FO en typisk *dialektordbok*, som behandlar de geografiskt avgränsade svenska varieteterna i Finland. Utifrån syftet kan den efter NLO och Svensén (2004) därtill definieras som en *dokumentationsordbok*; se diskussionerna ovan i avsnitt 5 och 6.

7.1. Tidigare dialektordböcker över svenska dialekter i Finland

De svenska dialekterna i Finland dokumenterades tidigt av lexikografer. Ordbelägg från svenska dialekter i Finland ingick redan i Rietz *Svenskt dialektlexikon* (1862–67). Mindre, regionalt begränsade ordböcker, bl.a. över allmogemålet i Nyland (Vendell 1884) och Vöråmålet (Freudenthal 1889), utgavs under 1800-talets sista decennier. I början av 1900-talet sammanställde och publicerade Vendell *Ordbok över de östsvenska dialek-*

⁵ Termen *regionalismordbok* definieras och exemplifieras både under lemmat *dialektordbok* och *regionalismeordbok* i NLO.

terna (1904–07). Wessman supplerade med *Samling av ord ur östsvenska folkmål* (1925–32). Båda verken täcker hela det svenska språkområdet i Finland. Dessa två ordböcker behandlar dock bara specifika dialektord, dvs. det ordförråd som skiljer dialekterna från det svenska standardspråket; se t.ex. diskussion av Reinhammar (1977:140).

Under 1800-talets sista och 1900-talets första decennier dominerades de vetenskapliga publikationerna i svensk dialektologi i Finland dels av dialektordböcker, dels av junggrammatiska monografier (Loman 1991:45ff. och 1993:13ff., 20ff.; Slotte 1992:57f.). Ahlbäck (1946:1) sammanfattar forskningsläget inom dialektologin på 1940-talet:

De finlandssvenska dialekterna ha länge varit jämförelsevis rätt väl undersökta. Det är framför allt O. F. Hultmans avhandling *De östsvenska dialekterna* (1894) och Herman Vendells och V. E. V. Wessmans stora ordböcker som lagt grunden till den goda kännedom om dem som vi äga i detta nu. Många äldre och nyare dialektavhandlingar ha ökat våra kunskaper om dem. Om vi bortse från lexikografien är det i första rummet ljudhistorien som fångat forskarnas intresse.

Ahlbäck poängterar här att lexikografien och ljudhistorien hade ägnats mest uppmärksamhet. Framför allt lyfter han fram Vendells och Wessmans ordböcker. Sammanfattningsvis kan konstateras att den dokumentationslexikografiska aspekten är stark för dialektordböckerna, och att den svensknationella kopplingen varit en av drivkrafterna bakom den livliga utgivningen av svenska dialektordböcker i Finland; jfr avsnitt 5 ovan.

7.2. Projektet *Ordbok över Finlands svenska folkmål*

FO är en dialektordbok över de svenska dialekter som talas och har talats längs kusten i Österbotten, i Satakunta, på Åland, i Åboland och i Nyland. Den bygger på en indelning i drygt 80 dialektsocknar i dessa områden. Ordboksprojektet har sin bakgrund i 1920- och 1930-talet. År 1927 inrättades Kommissionen för undersökning av svenska folkmål och svensk folkkultur i Finland (Folkmålskommissionen), och tanken på en intensifierad insamling av svenskt dialektmaterial med sikte på en ny heltäckande dialektordbok väcktes. I samarbete med Svenska litteratursällskapet i Finland inleddes en systematisk, omfattande insamling, och under slutet av 1940-talet påbörjades arbetet på ett ordregister över det insamlade materialet kompletterat med andra källor. (Se Ahlbäck 1976:VIII–XIX, Slotte 1981:45ff., Cantell & Sandström 2012:157ff.)

FO ger en heltäckande beskrivning av ordförrådet. I ordboken ges uttalsuppgifter, grammatiska uppgifter om ordklass och böjning, betydelseuppgifter och rikligt med exempel som belyser syntaxen i dialekterna. Ofta anges på sockennivå var ett visst uttal eller en viss betydelse har dokumenterats. Ordboken, som är en materialredovisning av allt tillgängligt material som kommit till redaktionens kännedom, bygger på ett ordregister som i dag omfattar ca 1 miljon excerpter. Tyngdpunkten ligger på att beskriva dialektalt språkbruk och samhällsföreteelser i det agrara samhället fram till omkring 1930. FO innehåller etnologisk och folkloristisk detaljkunskap om arbetet och livet förr, livets och årets högtider, folktro och religiösa föreställningar. Dessutom ingår en del fackterminologi, särskilt inom områden som lantbruk, boskapsskötsel, husbygge, sjöfart, fiske, skomakeri, vävning, spinning och mjölkhushållning. Också svordomar, könsord och pejorativa ord och uttryck ingår i materialet.

Till skillnad från Vendells och Wessmans ordböcker beskriver FO det dialektala ordförrådet i hela sin bredd. FO tar upp både specifika dialektord och ord som är gemensamma för det svenska standardspråket och dialekterna. Inte bara särskilda dialektala ord och uttryck, som *bundon* 'sädeskärve', *dåka* 'kasta, slänga; dåna, dundra', *lådskan* 'modlös; vissen, slokande' eller *nyssja* 'pyssla, stöka', redovisas, utan i FO ingår också frekventa substantiv och verb som *bytta*, *hand*, *hus*, *skog*, *göra*, *klappa*, *komma* eller *springa* och adverb och prepositioner som *efter*, *här* eller *in*, som är gemensamma för dialekterna och det allmänsvenska ordförrådet.

Reinhammar (1977:140ff.) framhåller att ordförrådet i FO redovisas utförligt, både när det gäller uttalsformer, betydelser och ordens utbredning. Arbetet uppfyller enligt Reinhammar "högt ställda vetenskapliga krav både i fråga om uppläggning och genomförande" (1977:144). Han understryker att det är motiverat att utarbeta en ny ordbok över de svenska dialekterna i Finland enligt modifierade principer för redigering och avgränsning, därför att materialet om dialekterna vuxit under decennierna efter 1930-talet. Genom att utnyttja det nya insamlade materialet och källkritiskt granska äldre dialektmaterial kompletterar FO de tidigare ordböckerna över svenska dialekter i Finland.

I FO redovisas lånord från framför allt finskan (t.ex. *ajka* 'duktig; ganska'), men också från bl.a. amerikaengelska (t.ex. *moonshine* 'hembränt brännvin'), ryska och franska. Karlholm (2008:240f.) har anfört det stora antalet lånord, som i vissa fall bygger på få belegg från en enda ort, som en svaghet i urvals- och redigeringsprinciperna för ordboken.

Redigeringsarbetet på FO inleddes 1959 av Olav Ahlbäck, som var den förste huvudredaktören (1960–82). Under 1960-talet slogs redigeringsprinciperna fast, strukturen skapades efter mönster från SAOB och redigeringen fortgick. Första häftet på 180 sidor gavs ut 1976. Under perioden

1982–2007 publicerades avsnittet *A–Och* i fyra band på ca 620 sidor per band, vilket omfattar omkring 65 000 lemman. Redaktionen för FO har sedan 1960- och 70-talen bestått av 3–4 redaktörer.⁶ Ahlbäck efterträddes som huvudredaktör av Peter Slotte (1982–2009), och våren 2009 tog jag över huvudredaktörskapet. Sedan fjärde bandet är Svenska litteratursällskapet i Finland förläggare för ordboken. I framtiden kommer nytt material primärt att publiceras på nätet, men enligt överenskommelse med förlaget Svenska litteratursällskapet i Finland utkommer FO också som tryckt ordbok. Den färdiga ordboken beräknas bestå av omkring 125 000 artiklar fördelade på sju band.

7.3. Exemplifiering av artiklarna i FO

I det följande kommer jag med hjälp av några exempelartiklar att belysa hur den tryckta FO är uppbyggd. Vidare görs jämförelser med motsvarande artiklar i Vendells och Wessmans ordböcker för att visa hur de tre dialektordböckerna förhåller sig till varandra.

Den första bilden (1) visar början av artikeln **höst** i FO. *Höst* behandlas uttömligt i FO.

Artikeln **höst** är ett typfall av ord som är gemensamt för dialekterna och allmänspråket. FO ger noggranna uppgifter om varierande uttal av ordet. På de flesta håll i Österbotten, Satakunta, Åboland och Nyland är vokalen *ö* lång och också diftongerade former, *-uä-*, *-yö-*, finns på många håll, särskilt i södra Österbotten, Åboland och delar av Nyland. I det åländska materialet finns mest uttalsformer med kort, öppet *ö*. Genus för detta ord varierar; det kan vara både maskulinum, femininum och neutrum och i Karleby och Nedervetil har det förenklad genusböjning enligt första deklinationen. Alla betydelser hör samman med årstiden höst. I FO upptar **höst** drygt en spalt och användningen av ordet belyses med många exempel på tidsuttryck som *i höst*, *om hösten*, *om höstarna*, *på hösten*, *på höstet*, *i höstas* eller *i sista höstas* etc. Men framför allt finns det under huvudartikeln en lång räkka med drygt 60 sammansättningar vilka upptar nästan fyra spalter och enligt modell från SAOB indelats i A-sammansättningar med förleden *höst-* och B-sammansättningar med förleden *höstas-*.

⁶ Mångåriga medarbetare i redaktionen har bl.a. varit redaktörerna Susanne Bergström-Söderlund, Ann-Marie Ivars, Per-Henrik Solstrand och Carola Åkerlund.

höst 1 *hø:st* nÖB mÖB Ma Sb SK Su Fö Iö Ho Ko Na Pa Hi Fb vNL Si Pä Li Mö Lt St Py, *huäst* Kv, *hyø:st* Kv Ma Pl Nä Sb Hi Fb Br Sj Py, *hyøst* Pl Kn Nä Tk Lf Sb Ki Fb Py, *höst* Ec Ge Vå Lu Kō Ku Bö Lt, *høst* Su Br, *hö:st* Vå Fö So Ku Bö, *ö:st* Ka Iå, *he:st* Bo, *hyäst* el. *hyø:st* Py d.I Kb Nv, n.1 Pe-Pu Mu Or Vö Mm Iå, n. med pl. -ar Tj Kr Öj La Ee Mu Or Vö, f.1 Pe-Pu, m.1 Tj Kv Bj Sv sÖB SK ÅB vNL Li Mö o. altern. Si St, m. med pl. -nar ÅL Iö Pä Py o. altern. Si St. årstid mellan sommar o. vinter, tid fr.o.m. september t.o.m. november. A.a.o.; *to var e kald hø:sti* Tj, *om jeitingbo:en je byggd i jo:len va:l e torran hyøst* Tk, *vi va redd ti ga: hande kveldin, he va hø:stin o mørkt* Sb, *nø min's ja hø:stnar när ä a frussä bittnarä än nu:* Ku, *de e hø:stn nu* Ko, *såm e va hø:stn så sku årmana i vintärlä:gär* Na, *när e blir hø:stn* Sn, *he a:r vari ein ka:lär* (:kall) *hø:st tessn hø:stn* St; folktro FD VII.4, sakreg.; *varinda hø:st sko a u:t o jaga* Tj, *hyø:stin ba:ket to läid di brist* Nä, *ja ha havi son bro:ta* (:förkylning) *he:la hø:stin* Vå, så va e in *ö:st* millan elgån (:helgonmässan) å första advente (som det hände) Iå, *tom fo:r all hø:stnar opp ti Heinola* Pä, *hø:stn fylld ja tretton* Py; i uttr. i höst, under den kommande hösten (Kv Pl Sb Ku Br) el. under den innevarande hösten (Kv Pl Sb Vå Br); i uttr. om hösten Bj Iö Br Sn Li el. om höstarna allm. el. på höstet Nv Ee el. på hösten Lf Te el. på höstarna Kb Tj Kr Ee Bj Ec Fö Bö, i fråga om ngt som sker varje höst el. ofta sker under hösten, höstetid; i uttr. om höstet Tj Pu el. om hösten Lf Sb Iö el. på höstet Kb Öj Ee Pu el. på hösten Sn i fråga om ngt som sker en viss

BILD 1. Början av artikeln **höst** i FO.

[**höst** *hüst* stn. (obs.!) Ingå (Bsd) Vö Pu, delar av Ped o. Esse, Kr Lmo. — fvs. **haust** n. *hestast* Estl, *hestas* Ru, i *hüstars* Pg, i *hyøstars* Kim, i *hyøstars* Hi, i *hüstast* Np adv. (tlls.) Sistlidne höst. — msv. i **høstres**, i **høsther**; Vl. i **høsters** l. i **høstras**. *hesto* adv. Nästa höst, till hösten. Ru — pl. *høstnar* l. *høstnar* Ahd.]

BILD 2. Artikeln **höst** i Vendells ordbok.

Bild (2) ur Vendells ordbok visar att *höst* tagits med i ordboken främst därför att det i vissa dialekter har genus neutrum. Paralleller dras till fornsvenskan och i Vendells ordbok anges några varierande tidsuttryck som har avvikande form i jämförelse med allmänspråket. Bild (3) ur Wessmans ordbok belyser slutligen hur Wessman kompletterar de uppgifter som finns i Vendells ordbok med två orter där *höst* också har genus neutrum.

[*höst høst* stn. Or Mm. — VII 390 a.

BILD 3. *Artikeln höst i Wessmans ordbok.*

Det franska lånordet *estimer* tas upp i Vendells ordbok, se bild (4). Han anför varierande uttalsformer med *-st-* och *-ks-* från Österbotten (socknarna Kronoby, Korsholm [= Must i Vendells ordbok] och Kvevlax), från Åland, Åboland (socknarna Nagu, Pargas, Kimito och Hitis) och Nyland (allmänt i Nyland och specifikt i socknarna Sjundea och Pyttis). Det är ett verb av första konjugationen. De betydelser som ges är 'värdera, akta' respektive 'hörsamma, lyda'. En notis om att verbet förekommer i några landskap i Sverige finns också. Däremot saknas helt språkexempel som kunde belysa syntaxen för detta verb.

[*estimer* fra. *ästumera* Nyl Pg Kr, *ästumær* Must Kv (mfl.), *äkstumér(a)* l. *äkstomer(a)* Nyl Pg Na, *ästumeär(a)* Sj Py, *äkstumeära* Sj, *äkstomeär(a)* Kim, *äkstomeära* Hi, *äksmæra* l. *äksmæra* Åhd sv. vb. tr. 1: 1) Värdera, akta. 2) Hörsamma, lyda. — *-ks-*former förekomma även i Uppl. Gd. VI. Bhl.

BILD 4. *Artikeln estimer i Vendells ordbok.*

Eftersom Wessmans ordbok är ett supplement till Vendells ordbok tar Wessman inte upp ord och uttryck som förekommer i Vendells ordbok, utom i de fall då han funnit nytt material. I detta fall tycks det inte ha kommit in ytterligare material om *estimera/ekstumera* under 1910–20-talen och därför finns ordet inte heller med i Wessmans ordbok.

I FO finns verbet *estimera* däremot med i lemmalistan, se bilderna (5) och (6) nedan. I jämförelse med Vendells ordbok har materialet kompletterats betydligt, både när det gäller uttalsvarianterna och uppgifter om betydelserna. Av utdragen ur de två ordböckerna, i bilderna (4) och (5), framträder också att lemmat har olika form. Vendell har valt den franska

formen *estimer* som lemma och anger samtidigt ordets etymologi, medan lemmat i FO har en försvenskad form i enlighet med principerna för uppslagsord i FO (Ahlbäck 1976:XIV). Som framgår av bild (5) har det i materialet för FO tillkommit ett substantiv *estimera*. Det har dokumenterats i uttrycket *vara ur estimeran* 'vara ur räkningen' i Vörå i Österbotten.

estimera s. i uttr. *va yor ekstomä:ron*, *vara ur räkningen*. Vö.
estimera v. *estome:r* Tj Kr, *ekstome:r* Pe-Pu, *ekstomär* Mu Or, *ekstomä:r* Vö, *estomä:r* Kv Kh Sv, *estume:r* Pl, *istim'e:r* Pö, *estimie:r* Nä, *istume:r* Lf, *estime:r* Sb, *ästmeir* Vb (1911), *eksime:ra* Sa, *eksm'e:ra* Vå vÅL, *ekstumä:ra* Kö, *eksmä:ra* Fö, *äkst(i)me:ra* Ku, *ikstem'e:r* Ho, *ekstume:r* Na, *estom'e:r(a)* Pa, *ekstom'e:r(a)* Pa Te Sn Ka Iå Sj, *ekstomier(a)* Ki, *ekstom'ä:r(a)* Hi Br, *ekstime:r(a)* Ky, *ekstome:r(a)* He St, *estum'e:r(a)* Si Bo Pä Li Lt, *estumie:r* el. *-me:r* Py v.I. - 1) sätta värde på el. uppskatta el.

BILD 5. Utdrag ur FO som visar substantivet *estimera* och inledningen av verbet *estimera*.

hysa aktning för (ngn el. ngt), ha respekt för (ngn); vanl. negerat. Tj Kr Pe-Pu Mu Or Vö Kh Pl Pö Nä Vb vÅL Kö Fö Ku Na Pa Ki Te Ka Iå Db Ky He Si Bo Li St Py; fålke bör (:börjar) int *ekstome:r* gu:d meir Pa, int *ekstome:rar* dom tjörkkan ti nogo me:ra Te, han *estume:ra* iŋga onon Bo. - 2) bry sig om; hörsamma, lyda; vanl. negerat. vÅL Kö Fö Pa Te Sn Ka Sj Si Li; han *ekstomä:rar* int fast ja sä:r åt han Te, vill do int *ekstome:ra* mej! Sj. - 3) negerat: icke tåla. Ki Hi Ky He; han *eina vild* int *ekstomä:ra* han *andär* Hi, *brömsen* *ekstime:ra* han (:hästen) *inga* Ky, *inga* *ekstome:rar* vi tåkå här patarask He. - 4) behaga el. gitta el. värdigas (göra ngt), icke försmå (ngt); vanl. negerat. Or Vö Sv Lf Sb vÅL Ku Ho Ki Hi Br Iå Bo Pä Li; ja *ekstomä:ra* int ga: Or, han *estime:rar* it ti kom ti å:s Sb, ja *äkstme:rar* ra:kt in't gå: di:t Ku, va sko än *iksteme:r* me: ä Ho, *iŋgen* *ekstomä:rar* bœ:ndsmö:r nu: Hi, va sko dom *ekstomä:r* kom o hels po me:g Hi, ja a int *estume:ra* fro:ga Bo; särsk.: (icke) räkna på el. ifrågasätta; han *estume:ra* int ti llo ein ru:to til sondär när an hadd pengar. Bo; refl.: gitta, bry sig om; (det är inte många) såm sko i:s *estom:er* se ti böri bärg sjö:fäodre. Pa.

BILD 6. Betydelserna 1–4 hos verbet *estimera* i FO.

I jämförelse med Vendells ordbok behandlas verbet *estimera* utförligt med varierande uttalsformer och fyra olika betydelsemoment i FO. Dels finns de samma två första betydelsemomenten 'sätta värde på el. uppskatta el. hysa aktning för (ngn el. ngt), ha respekt för (ngn)' och 'bry sig om; hörsamma, lyda'. Men i FO anförs ytterligare betydelsemomenten 'icke tåla' och 'behaga el. gitta el. värdigas (göra ngt)'. Vid redigeringen av artikeln har fler betydelsemoment kunnat urskiljas i FO tack vare det utvidgade material som FO:s redaktörer haft tillgång till. Användningen av verbet *estimera* exemplifieras med många exempel för de olika betydelsemomenten, och det framgår också att flera av betydelsemomenten främst används i negerande uttryck.

*jötas *joutas* Ee, *joutas* Or Mu
Kv. sv. vb. dep. 1. 1) Vara sparsam, girig. Ee Kv. 2) Nännas; vanl. använt i nekande satser. Ee Kv. 3) Hava lust. Ee. 4) Nödgas. Ee. 5) Mäktä, förmä, idas. Or Mu. — Jfr forngottl. *liuta* 'fä', no. *ljota* 'nödgas till att'. Karsten Finsk Tidskrift 1920: 222. Sammanställes med fi. *joutaa* 'hava tid, vara ledig; hinna, medhinna, räcka till, kunna, böra'. Saxén 132.

BILD 7. Artikeln **jötas** i Wessmans ordbok.

jötas *joutis* Kb-Nv Tj Kr Je, *joutas* Kb Nv Tj
Kr Öj La Ee Pe-Pu Nb Je Mu Or Kv Bj,
jöytas Je, *jöötas* Vö, *jeotas* Vö (1861) Re,
jeutas Bj, *jöutas* Re Kn, *jöitas* el. *jöitas* Kh,
jöitas Sv dep.I. – 1) i neg. uttr.: nännas (göra ngt), förmä sig (till att göra ngt). Kb-Nv Tj Kr Öj Ee Pe Pu Nb Je Mu Vö Kv; *ha jöutast int legg ny: tsepkona opo si to e va regnvedri* Tj, *kakon va so grann so an jöutas it eta opp en* Je; i neg. uttr.: anse sig ha råd (att göra ngt). Kb Nv Tj Kr Öj Ee Pe Pu Nb Je Mu; *han var holl (:gnidig) o to jöutist int an tsø:p Vi:borskringlona för di var so dy:r Kb, va sko hate: snå:lindzi jöutas!* Tj; äv.: gnida, vara överdrivet sparsam. Ee Kv. – 2) böra (göra ngt). Tj Ee Or Vö; *noo jöutas ni få skod e Or, han jöotas noo fo e de: Vö, an jöotas no fol fo na li:ti betä po gambädagan* (:borde väl få det litet bättre på gamla dagar) Vö, *nåov jöotas he* (:het pron.

BILD 8. Början av artikeln **jötas** i tryckta FO.

Exemplet i bild (7) återger artikeln **jötas** ur Wessmans ordbok. Eftersom detta verb härletts som ett finskt lån saknas det i Vendells ordbok. Men Wessman tar i sin ordbok upp finska lån som finns belagda i Saxén (1895).

I Wessmans ordbok finns verbet *jötas* med två uttalsvarianter från Esse respektive Oravais, Munsala och Kvelax. Han urskiljer fem olika betydelser för verbet, som antas ha ursprung i såväl fornsvenska som i finska förebilder. Av artikeln framgår att *jötas* diskuterats av tidigare forskare som Karsten år 1920 och Saxen (1895); (jfr t.ex. Sandström 2011 och 2013:94ff.) Wessman ger däremot inte några exempel eller språkprov på hur ordet används i dialekterna, utan nöjer sig med ange betydelseerna.

I bilderna (8) och (9) belyses huvudet – med lemma, uttalsvarianter och grammatisk angivelse av verbkonjugation samt deras utbredning – i artikeln **jötas** dels i tryckta FO, dels i nätpublikationen av FO. Av utdragen framgår att antalet beläggorter och uttalsvarianter har utökats i jämförelse med Wessmans ordbok; se bild (7) ovan. Påfallande är däremot att detta verb enbart används i dialekterna i Österbotten. I de österbottniska dialekterna är det dessutom begränsat till norra och mellersta svenska Österbotten bortsett från ett belägg från Korsnäs i södra Österbotten.

ORDBOK ÖVER FINLANDS SVENSKA FOLKMÅL

Österbotten Åland Åboland Nyland

Startsida Sökord... börjar med Slå upp Avancerad sökning Hjälp Om ordboken Karta Respons

A B C D E F G H I
J K L M N O P Q R S
 T U V W X Y Z Å Ä Ö

jöra
 jöral-gast
 jörog
 jörs
 jörsa
 jös
 jösa s.
 jösa v.
 jöss
 jösse
jötas
 jöten
 jötter
 jöva
 kaa
 kaback
 kabacka
 kabb
 kabba
 kabbasbyckel
 kabb-don

jötas
jötis Kb-Nv Tj Kr Je, *jötus* Kb Nv Tj Kr Öj La Ee Pe-Pu Nb Je Mu Or Kv Bj,
jötys Je, *jötös* Vö, *jeotas* Vö (1861) Re, *jeutas* Bj, *jötus* Re Kn, *jötis*
 el. *jötis* Kh, *jötis* Sv dep.I.

1 i neg. uttr.: nännas (göra ngt), förmå sig (till att göra ngt). Kb-Nv Tj
 Kr Öj Ee Pe Pu Nb Je Mu Vö Kv;
 • *ha jötast int legg ny: tsepkona opo si to e va regnvedri* Tj,
 • *kakon va so grann so an jötus it eta opp en Je*;
 i neg. uttr.: anse sig ha råd (att göra ngt). Kb Nv Tj Kr Öj Ee Pe Pu
 Nb Je Mu;
 • *han var häll (:gnidig) o to jötist int an tse:p Vi:borskringlona för di
 var so dy:r* Kb,
 • *va skø hate: snå:lindsi jötus!* Tj;
 äv.: gnida, vara överdrivet sparsam. Ee Kv.

2 böra (göra ngt). Tj Ee Or Vö;
 • *noø jötus ni få skod e. Or*,
 • *han jötus noø fo e. de: Vö*,
 • *an jötus noø fol fo na li:ti betä pø gambädagan (:borde väl få det
 litet bättre på gamla dagar) Vö*,
 • *nåov jötus he (:het pron. o. k., 5) hon å få si: va tä jä Vö*;
 nödgas, få lov (att göra ngt). Ee (SO) Vö;
 • *han jeotas fol tag a. Vö (c. 1860)*;
 impers.: fordras, vara tvunget. Bj Re Kh Sv Kn;
 • *e. jötus noø ti kje:d posse bra: i tolonje tjö:lð* Bj,
 • *e. jötis noø för an ti föje:n fö di har femm ba:n* Kh.

BILD 9. Början av artikeln **jötas** i nätpublikationen av FO: uttalsvarianterna och betydelsemomenten 1–2.

När det gäller innehållet i ordboksartiklarna har det inte gjorts några ändringar eller tillägg i artiklarna vid publiceringen på nätet. Men en jämförelse mellan bilderna (8) och (9) visar att artikelstrukturen i FO på nätet inte är lika koncentrerad och sammanpressad som i den tryckta FO med komprimerad text som ombrutits i tätt satta spalter för att spara utrymme. Särskilt exemplen i betydelsemomenten, som belyser syntaxen och åskådliggör hur ordet används i löpande tal, har getts större synlighet i nätversionen av FO genom att varje exempel har skilts ut, placerats på ny rad och markerats med en liten boll framför exemplet.

jöss	
jösse	
jötas	
jöten	
jötter	
jöva	
kaa	
kaback	
kabacka	
kabb	
kabba	
kabbasbyckel	
kabb-don	

3 i neg. uttr.: mäktä, förmå; idas. Mu Or Vö (SO, Wessman FmS 17–18: 25); ss. omskrivning för imper. (jfr idas 2):

- *jöutas int tsinn (:kåbbla) imø:t na:!* Pe-Pu.

4 impers.: gå an (äv. om det inte riktigt passar sig el. är att rekommendera), vara möjligt (med visst förbehåll), vara tillbörligt. Tj Nb Je Mu Vö;

- *nø: jöutas tu freist Tj,*
- *nø jöutis ho siti tä o skoda Tj,*
- *vil do ha de ein su:p?, (svar:) takk såm frå:gar, nåo jöutas he Nb,*
- *fæør ja koma me:?, (svar:) noø jöutas he (:för all del) Je,*
- *noø jöutas he skoda Mu;*

med personsobj.:

- *noø jöutas ni få skod e. Or.*

5 i fråga om (tungt) arbete, (krävande) uppgift o.d., impers.: fresta på, vara krävande; ha sina sidor, vara otacksamt. Kb Nv Tj Kr Öj La Ee Pe Nb Je Mu Or Vö Re;

- *nø: jöutas he Kb Tj Kr Ee Pe Nb,*
- *he jöutas nø: ti vara tä(r) ti pi:gø Tj (Nb),*
- *noø jöutas he. ti ha tannværtjün Mu,*
- *e jöøtas noø ti va bønd Vö;*

med personsobj.: utsättas för påfrestningar, ha det svårt el. besvärligt;

- *å han såm ha:r te ti dra:g i: (:att dras med), han jöøtas nåo han Vö,*
- *dø jöøtas noø om dø ska mott va iu:t i de:hä ve:dri Vö.*

Väl av fi. joutaa v., bl.a. hinna, förslå, passa, kunna o. fi. joutua v., bl.a. hamna, nödgas, hinna, skynda, lyckas under påverkan av dial. pret.
*jöut av ljuta v., få på sin lott, få, nödgas.

BILD 10. Utdrag ur artikeln **jötas** i nätpublikationen av FO: betydelsemomenten 3–5 och etymologisk anmärkning.

I fråga om indelningen i betydelser kan sammanfattningsvis konstateras att FO indelar verbet *jötas* i fem olika betydelsemoment; se bilderna (9) och (10). Men i jämförelse med Wessman (se bild (7) ovan) har det nyare materialet gett möjlighet till en mer nyanserad indelning i FO. En del av de betydelser som Wessman skilt ut till nya betydelsemoment har

sammanslagits i FO; se betydelsemoment 1 och 2 i bild (7), vilka i FO sammanförts inom det första betydelsemomentet; se bild (8). På liknande sätt motsvaras betydelsemomenten 3 och 5 i bild (7) enligt Wessmans indelning av moment 3 i FO; se bild (10). Det utvidgade materialet som ligger som grund för redigeringen av FO har alltså gjort det möjligt att urskilja och beskriva ytterligare fler och mer detaljerat beskrivna betydelser för artikeln **jötas** i FO i jämförelse med Wessmans ordbok; se bild (10) betydelsemomenten 4 och 5 i FO.

Utdragen i bilderna (1) – (10) ur de tre finlandssvenska dialektordböckerna Vendells (1904–07) ordbok, Wessmans (1925–32) ordbok och FO (1976–) visar att det material som den aktuella pågående ordboken FO haft tillgång till ger möjlighet till mer omfattande och nyanserad beskrivning av de svenska dialekterna i Finland. Materialet har genom ökad insamling under mitten av 1900-talet utökats både kvantitativt och kvalitativt. Men FO har därtill tillämpat avgränsningsprinciperna annorlunda; till skillnad från föregångarna har redaktörerna för FO valt att redovisa dialekternas centrala ordförråd på ett mer heltäckande sätt; se Ahlbäck (1976:XIII) och Reinhammar (1977:140f.).

8. Finlandssvensk ordbok (FSOB), 2000 och 2008

Enligt den typologiska indelning av ordböcker som ges i NLO (1997) och Svensén (2004) hör FSOB till gruppen avgränsade ordböcker, och kan som framgått av avsnitt 6 ovan definieras som en *regionalismordbok*.

8.1. Syftet och förebilderna för FSOB

Det huvudsakliga syftet med FSOB är att ge en normerande förteckning över finlandssvenska särdrag på lexikal nivå för att därigenom hos finlandssvenskar öka medvetenheten om att vissa särdrag kan uppfattas som främmande eller svårbegripliga av sverigesvenskar (af Hällström-Reijonen & Reuter 2008:5ff.). FSOB är den första egentliga *ordboken* över särfinländska ord och uttryck i finlandssvenskt språkbruk, men det finns mönster och förebilder för den i tidigare verk. Rötterna ligger i det arbete med finlandismer som påbörjades inom Svenska landsmålsförningen i Finland under slutet av 1800-talet; se avsnitt 5 ovan. Detta manifesterades i Bergroths handledning *Finlandssvenska*, som publicerades första gången 1917 och fick en utökad andra upplaga 1928 (se t.ex. af Hällström 2002:52).

Att Bergroths *Finlandssvenska* står som förebild framgår tydligt av förordet till FSOB som inleds med ett konstaterande om att det gått över sjuttio år sedan 1928 då andra upplagan kom ut. En kortfattad version, *Högsvenska*, av detta verk publicerades i flera upplagor fram till mitten av 1960-talet. Också Stenmarks *Finlandssvenska ord och uttryck* (tredje upplagan 1983) nämns i förordet, men i motsats till de två förstnämnda hade den ”inte ett normativt utan ett deskriptivt syfte och riktade sig närmast till sverigesvenskar”. FSOB är en språkhandbok ”som tar upp finlandssvenska särdrag på lexikal nivå”. Enligt redaktörerna, språkvårdarna, Charlotta af Hällström och Mikael Reuter fanns det ”ett behov av och en efterfrågan på en bok som ger en heltäckande bild av finländska särdrag i svenskan och samtidigt ger rekommendationer om språkbruket” (af Hällström & Reuter 2000:3).

I förordet framhävs att ”[s]yftet med ordboken är att öka medvetenheten om finlandismer bland finlandssvenskar och på det sättet hindra finlandssvenskan från att fjärma sig från rikssvenskan.” Helt explicit hänvisas vidare till Bergroths (1917) syfte med verket ”Handledning till undvikande av provinsialismer i tal och skrift” som var att ”bidraga till bortarbetandet av de talrika provinsiella egenheter som vidlåda vår finländska svenska”. Redaktörerna för FSOB säger sig omfatta samma syfte (af Hällström & Reuter 2000:3).

8.2. Principer för språkvården och den finlandssvenska normen

I en inledande text till FSOB läggs principer för den samtida finlands-svenska språkvården fram, då af Hällström & Reuter (2000:5) fastslår:

Finlandssvenskan är en varietet av svenskan, inte ett eget språk. Målet för den finlandssvenska språkvården har alltid varit att svenskan i Finland inte ska fjärma sig från sverigesvenskan utan att vi fortsättningsvis ska kunna tala om en enda svenska, trots vissa nationella och regionala skillnader som gör att vi använder det gemensamma språket på lite olika sätt.

Som Nikula (2000:248) i en recension av FSOB påpekar är det på sin plats att klart redogöra för denna ståndpunkt, eftersom finlandssvenskans status under de föregående decennierna hade debatterats i flera omgångar. Målgruppen för FSOB redovisas klart och tydligt i de inledande texterna: den ”riktar sig i första hand till finlandssvenskar och först i andra hand till sverigesvenskar och finnar”, vilket medfört att det inte har ansetts ändamålsenligt med ”utförligare definitioner av uppslagsorden” (af Hällström & Reuter 2000:5). Nikula (2000:248ff., 256f.) problema-

tiserar dock denna brist på definitioner. Utgående från konkreta exempel ur FSOB, som *parcell* 'tomt', *kontakter till* 'kontakter med' och *jynsa* 'putsa, blanka', understryker hon att normen för finlandssvenskan inte är densamma i hela Svenskfinland. Nikula framhäver att den regionala variationen inte beaktats i tillräcklig utsträckning i FSOB. Hon visar att detta begränsade synsätt kan ha haft konsekvenser på flera olika nivåer i ordboken; i urvalet av uppslagsord, i betydelsebeskrivningarna och i utformningen av exemplen. Nikula (2000:252) skriver:

Den som gör en finlandssvensk ordbok befinner sig i flera avseenden i en svår situation därför att, som författarna mycket riktigt konstaterar, finlandssvenskan inte är ett enhetligt fenomen utan inom sig härbärgerar lokala varieteter (s. 6). Detta ger sig till känna i FSOB både i valet av uppslagsord [...] och kanske framför allt i betydelseangivelserna [...] som främst verkar beskriva en av dessa varieteter. Ord och uttryck som ur österbottnisk eller åboländsk synvinkel upplevs som berikande och, med författarnas egna ord, viktiga för 'den finlandssvenska identiteten' (s. 9) är kanske inte speciellt viktiga för en nylänning eller vice versa. Detta talar också för att man kunnat avstå från graderingen av användbarheten till förmån för enbart allmänsvensk(a) motsvarighet(er).

Nikula är som ovan framgått kritisk till att FSOB i första upplagan inte har en klart uttryckt norm för finlandssvenskan, i stället blir de bedömningsgraderingar författarna gör ofta inkonsekventa och subjektiva. Dessutom påtalar hon den oklara definitionen av termen *finlandism* (Nikula 2000:248).

Den definition av *finlandism* som ges i FSOB (2000:8 och 2008:6) lyder "ord eller uttryck som bara eller huvudsakligen används i svenskan i Finland eller som i Finland används i en annan betydelse än i Sverige". Nikula föreslår att man i stället kunde ha betecknat ord och uttryck som används i finlandssvenskan som *finlandssvenska* och ord och uttryck som används i Sverige som *rikssvenska*. Härigenom hade man sluppit den belastning som tenderar att förknippas med termen *finlandism* (Nikula 2000:248).

Också af Hällström-Reijonen, en av författarna till FSOB, har i senare sammanhang fört en kritisk diskussion kring termen *finlandism*. I en undersökning om finlandssvensk språkvårdshistoria definierar af Hällström-Reijonen (2012:79) *finlandism* mer elaborerat och nyanserat som:

ett ord, en fras eller en struktur som bara används i finlandssvenska samt finlandssvenska regionalspråk eller som i dem används i en annan betydelse än i svenskan i Sverige. Även uttryck som är påfallande mer frekventa i finlandssvenska och finlandssvenska regionalspråk än i svenskan i Sverige räknas till *finlandismerna*

Ursprunget till de särfinlandssvenska orden och uttrycken kan vara olika; det kan vara arkaismer, t.ex. *barmficka* 'innerficka, bröstficka', fennicismer, som uppkommit genom direkt lån eller översättningslån från finskan, t.ex. *kumpan* 'kompanjon' (av fi. kumppani) eller *taltur* 'inlägg' (av fi. puheenvuoro). Det kan också vara fråga om nybildningar t.ex. *inkommande* 'kommande' eller *ta länge* 'ta lång tid' eller ord och uttryck med dialektalt ursprung t.ex. *klotta* 'kladda' eller *lorva* 'gå långsamt; söla'. (Se af Hällström-Reijonen & Reuter 2008:6f. samt de nämnda orden i lemmalistan i FSOB.)

8.3. Jämförelser mellan första upplagan av FSOB (2000) och nya upplagan av FSOB (2008)

Första upplagan av FSOB som publicerades på Schildts förlag 2000 omfattar 2 300 lemman. Den nya upplagan som gavs ut 2008 hade utökats och reviderats och omfattar drygt 2 550 lemman; en del ord och uttryck hade strukits och 440 nya lemman hade tillkommit. Den gavs ut av förlaget Schildts i samarbete med Focis. Särskilt hade af Hällström-Reijonen och Reuter förändrat sättet att presentera materialet och rekommendationerna, som i nya upplagan av FSOB (2008) är klart utskrivna med formuleringar som "använd i stället" och "använd inte i betydelsen x" (af Hällström-Reijonen & Reuter 2008:4); se exempel i bild (13) och (14) nedan.

I Cantells (2009) recension av nya reviderade upplagan av FSOB framhävs att boken modifierats både till innehåll och utformning, däremot har utgångspunkterna för normeringen inte ändrats, inte heller det i viss mån problematiska sättet att göra jämförelser med finskan. Jämförelser med finska görs i FSOB för att visa att motsvarande ord förekommer i finskan och därför kan inverka på och förstärka förekomsten av vissa ord i finlandssvenskan. Detta kan i en del fall ge det felaktiga intrycket av att ett visst ord eller uttryck kan härledas från finskan, även om etymologin för det aktuella ordet bör sökas på annat håll, ofta i dialekterna eller fornsvenskan. Som exempel anförs t.ex. *harm* i FSOB (Cantell 2009:287 och Nikula 2000:257f.).

Nedan i bilderna (11) och (12) har jag illustrerat hur materialet presenteras i FSOB (2000). Uppgifter om orden ges i två spalter; i vänsterspalt finns lemmalistan och i högerspalt rekommendationen och betydelsen för ordet. Bild (11) visar artiklarna **kalvo**, **kamurka**, **kanava**, **kanavanål**, **kanister** och **kantin**. Här finns flera exempel på bruket av en pil för de ord bland finlandismerna som man i FSOB avråder från. Pilen har placerats in i högra spalten framför de ord eller uttryck som i FSOB föreslås som sverigesvensk motsvarighet; se t.ex. artiklarna **kalvo** och **kanister** i bild (11). Av

de aktuella exempelorden i bild (11) är det endast artikeln **kamurka** som inte har försetts med en pil. Det anges ha ryskt ursprung och har försetts med bruksmarkören vardagligt, *vard*. I fråga om *kamurka* ges i stället bara förklarande synonymer som kan vara gångbara i sverigesvenskan. För en jämförelse mellan FSOB (2000) och den nya upplagan av FSOB (2008), jfr bilderna (11) och (13). I bild (13) illustreras hur samma ord presenteras i nya upplagan av FSOB.

kalvo, kalva fi.	→ transparang, skrivfilm, stordia, overheadbild Se även: hinna, pläty
kamurka [ka'murkka, ka'morkka] ry., vard.	skrymsle, garderob, krypin, hörn, vrå
kanava [*kanava]	→ stramalj fi. kanavakangas, fra. cannevas
kanavanål [*kanavanå:]	→ tapisserinål, grövre stoppnål fi. kanavaneula
kanister om kärl för flytande varor	→ dunk ; En <i>kanister</i> är ett förvaringskärl för torra varor. <i>oljekanister</i> rsv.: oljedunk; Se även: bensinkanister, spritkanister
kantin	ofta: → kafé särsk. på sjukhus, sjukhuskiosk ; En <i>kantin</i> är ett serveringsställe vid militärförläggning, marketenteri, eller en låda för servis, mat och dryck att användas vid resa e.d. fi. kanttiini

BILD 11. *Utdrag med artiklarna kalvo, kamurka, kanava, kanavanål, kanister och kantin ur FSOB (2000:92).*

Bild (12) nedan belyser likaså artiklar i ett utdrag ur första upplagan av FSOB (2000). I utdraget förekommer först ett uttryck *lyfta katten på bordet*, som är ett översättningslån från finskan. Också *keikka* som förekommer i flera etablerade betydelser i finlandssvenskan är ett finskt lån. När det gäller *ked* tar FSOB upp avvikande form, böjning och användningsområde i jämförelse med sverigesvenskt språkbruk. Verbet *kavera* hör till gruppen vardagliga finlandismer.

I bild (13) nedan belyses artiklarnas utformning och ombrytning i nya upplagan av FSOB (2008). Bilden visar samma artiklar som finns ovan i bild (11) ur FSOB (2000), alltså artiklarna **kalvo, kamurka, kanava, kanavanål, kanister** och **kantin**. Boken har i den reviderade upplagan fått ett behändigare format med hårda pärmar, men också vid ombrytningen. Man har valt att hålla samman de uppgifter som ges om ett visst ord inom en artikel i samma spalt, i stället för att som i den tidigare upplagan splittra upp informationen och placera lemmat i vänsterspalten och förkla-

ringar, rekommendationer och exempel i högerspalten; jfr bild (11) ovan. Av utdraget i bild (13) framgår också att man i stället för den pil som användes i FSOB (2000) i FSOB (2008) övergått till verbala förklaringar av typen ”använd i stället”.

lyfta katten på bordet	Uttrycket är omöjligt att förstå för den som inte kan finska. Använd i stället uttryck som sluta gå som katten kring het gröt, kalla saker vid deras rätta namn, tala rent ut, ta tjuren vid hornen, se sanningen i vitögat . Uttrycket <i>lyfta katten på bordet</i> används vanligen i betydelsen att ta fram något som alla egentligen känner till men ingen vågar tala om. fi. nostaa kissa pöydälle
kavera [ka've:ra] dial., vard.	överdriva; skrodera, prata, snacka; Även substantiven <i>kaver</i> och <i>kavering</i> (prat) förekommer i fisv.
ked, -en, -er	kedja, -an, -or; Ordet <i>ked</i> används mest i poetiska sammanhang i rsv.
keikka fi., vard.	grej, sak, grunka; uppträdande, konsert, föreställning <i>keikkasångare</i> rsv.: sångare som åker på turné, sångare som har spelningar; <i>Vi har tre keikkor i vår.</i> rsv.: Vi har tre konserter i vår.

BILD 12. Utdrag med uttrycket *lyfta katten på bordet* och orden *kavera, ked* och *keikka* ur *FSOB (2000:94)*.

kalvo, kalva fi. Använd i stället transparang, skrivfilm, stordia, overheadbild . Se även: hinna
kamurka [ka'murkka, ka'morkka] ry., vard. skrymsle, garderob, krypin, hörn, vrå
kanava Använd i stället stramalj . fi. kanavakangas, fra. cannevas
kanavanål Använd i stället tapisserinål, grövre stoppnål . fi. kanavaneula
kanister om kärl för flytande varor Använd i stället dunk . En <i>kanister</i> är en typ av bleckask för matvaror och tobak m.m. <i>ojekamister</i> sv: oljedunk

BILD 13. Utdrag med orden *kalvo, kamurka, kanava, kanavanål* och *kanister* ur *FSOB (2008:90)*.

Nedan i bild (14) ur FSOB (2008) visas hur informationen fördelas spaltvis i den reviderade FSOB. Här syns delar av ett helt uppslag från sidan 90. I utdraget belyses tre av de ord och uttryck som återges ur den första upplagan av FSOB (2000) i bild (12) ovan, nämligen *lyfta katten på bordet*, *kavera* och *keikka*. I bild (14) belyses dessutom två uttryck med verbet *kasta*, fiskeredskapet *katscha*, *kemikalieaffär* och verben *kila* och *kina* samt slutligen *kinuski*. Här framträder tydligt att de särfinländska orden som tas upp i FSOB ofta är vardagliga ord och uttryck som *kasta sig* eller *kina*. Andra typiska betydelsegrupper är redskap som *katscha* eller maträtter som *kinsuski*.

<p>kasta sig vard., skämts. gå och bli, bli <i>Han kastade sig religiös.</i> sv.: Han gick och blev religiös.; fi. heittäytyä joksikin; Se även: släppa sig</p>	<p>keikka fi., vard. grej, sak, grunka; uppträdande, konsert, spelning, gig, föreställning <i>keikkasångare</i> sv.: sångare som åker på turné, sångare som har spelningar; <i>Vi har tre keikkor i vår.</i> sv.: Vi har tre konserter i vår.</p>
<p>kasta titlarna vard. lägga bort titlarna, slänga titlarna, bli du fi. heittää tittelet pois</p>	<p>kemikalieaffär, -handel parfymaffär, parfymeri fi. kemikaali-kauppa, -liike; Se även: kosmetikbutik</p>
<p>katscha [ˈkatʃa], katiska [ˈkatiska] katsa, katse, mjärde fi. katiska</p>	<p>kila [ˈkila] vard. tränga sig förbi i en kö <i>Ida inte kila i kön.</i> sv.: Träng dig inte förbi i kön. Håll dig i kön. <i>Kila</i> [ˈtʃi:la] betyder svenskan föra in som en kil och sätta fast med hjälp av en kil eller förflytta sig snabbt och lätt. fi. kiilata</p>
<p>katt; lyfta katten på bordet sluta gå som katten kring het gröt, kalla saker vid deras rätta namn, tala rent ut, ta tjuren vid hornen, se sanningen i vitögat. Uttrycket <i>lyfta katten på bordet</i> används vanligen i betydelsen att ta fram något som alla egentligen känner till men ingen vågar tala om. fi. nostaa kissa pöydälle</p>	<p>kila [ˈkila] barnspr. kissa Se även: kisa</p>
<p>kavera [kaˈve:ra] vard. överdriva; skrodera, prata, snacka Även substantiven <i>kaver</i> och <i>kavering</i> (prat) förekommer i finlandssvenskan.</p>	<p>kina [ˈkina] vard. gräla, munhuggas, kivas, bråka, polemiserä fi. kinata</p>
	<p>kinuski, tjinuski [tʃiˈnoski, tʃiˈnusski] kola; kolasås; kolaglasyr fi. kinuski</p>

BILD 14. Utdrag ur FSOB (2008:92) med ord och uttryck inom det alfabetiska intervallet ka-, ke- och ki-.

8.4. FSOB och FO som dokumentationslexikografi

I fråga om syftet med den lexikografiska verksamheten är FSOB framför allt *brukslexikografi*, som stödjer kommunikation och inläring för att ”imøtekomme individers umiddelbare behov for informasjon” (NLO 1997:169). Men även om det huvudsakliga syftet är brukslexikografiskt och normativt/preskriptivt – FSOB avråder explicit från vissa särfinländska ord och uttryck – kan den i vissa avseenden uppfattas som en form av *dokumentationslexikografi*. Typiskt för dokumentationslexikografi är

en stark kulturell koppling mellan ordboken och det samhälle där den tillkommit och därtill en målsättning att dokumentera och bevara viss information (Svensén 2004:1f.).

Jag har ovan visat att det finns tydliga svensknationella rötter och kopplingar för såväl FSOB som den finlandssvenska dialektordboken FO. Bakgrunden till arbetet med FSOB ligger i en vilja att bevara och dokumentera svenskan i Finland. Projektet FSOB kan enligt min tolkning också betecknas som dokumentationslexikografi, därför att sammanställningen och utgivningen av ordboken i sig inneburit att ordboks författarna har dokumenterat och befäst de särfinländska lexikaliska egenheterna, även om syftet med ordboken primärt var brukslexikografiskt, dvs. att ge rekommendationer om vilka finlandismer som kan anses vara neutrala och acceptabla och vilka finlandismer som bör undvikas i stilistiskt neutrala texter (af Hällström-Reijonen & Reuter 2008:9).

9. Framtiden för svensk lexikografi i Finland: elektronisk lexikografi och ordböcker på nätet

Electronic lexicography opens up all sorts of radical possibilities that were closed to traditional lexicography: new kinds of evidence, new modes of description, new ways of organizing evidence, new possibilities for exploiting database structure and hypertext links, and the need for new theoretical foundations. [- - -] These and other opportunities offered by electronic lexicography raise interesting issues at both a practical and theoretical level.

(Hanks 2012:57)

9.1. Elektronisk lexikografi och nya utmaningar

I dag sker en snabb utveckling inom såväl lexikografi som tillämpad språkforskning. Med hjälp av effektiva IT-resurser och språkteknologisk expertis har lexikografer och lingvister fått möjlighet att skärpa och utveckla sina analysverktyg. Större empiriska material kan behandlas med nya metoder, då allt större sökbara korpusar över både tal och skrift byggs ut (se t.ex. Granger & Paquot (red.) 2012). Inom dialektologin har landvinningar skett i form av mer utvecklade metoder för fonetisk och fonologisk analys (Swedia-2000-projektet, Aho 2010, Leinonen 2010). Datateknikens förnyelse och utveckling har också gett nya möjligheter till forskning i dialektgeografi (Rabanus 2011, Leinonen 2012). Lexikografen har traditionellt haft en framträdande plats inom finländsk dialektologi (se

t.ex. Ahlbäck 1946, Sandström 2013). Också i framtiden verkar dialektordböcker behålla sin position och kunna utvecklas och dra nytta av de språkteknologiska framstegen.

Förlagsbranschen i Finland och på andra håll i Europa har under det senaste decenniet drabbats av nedskärningar, som påverkat och minskat utgivningen av stora referensverk och ordböcker. Dessa tendenser märks också i Institutet för de inhemska språkens utgivningspolitik; de egna ekonomiska resurserna för tryckning av pappersordböcker är begränsade och under de senaste åren har det blivit svårare att få utomstående finansiering för tryckning.

Under åren 2006–09 genomgick de stora pågående ordboksprojekten på Focis flera utvärderingsprocesser. I ett expertuttalande (2010) om framtiden för Focis fastslogs att statusen som en statlig språklig myndighet skulle stärkas. Kärnområdena skulle vara lexikografi och språkvård, och utgivningen av kulturbärande historiska ordböcker skulle fortsätta. Men framöver skulle ordböckerna primärt publiceras på nätet. En ny lag och förordning trädde i kraft 1.1.2012 och verksamheten fortsatte under det nya namnet Institutet för de inhemska språken.

9.2. Digitalisering av FO

FO är en typisk flergenerationsordbok (Vikør 1999) med en lång tillkomst-historia. Varierande redigerings- och trycktekniker har avlöst varandra under årens lopp, vilket har haft konsekvenser för överföringen av FO till en digital databas. Om en helt ny nätordbok hade byggts upp, hade vi börjat med att planera möjliga sökningar och sedan fattat beslut om databasens innehåll och kodning utgående från önskemålen. Men för FO:s del handlar det om att digitalisera en existerande ordbok; vi utgick från det innehåll som fanns, tänkte ut vad som kunde vara relevant att söka i innehållet och försökte åstadkomma en kodning som gör de önskade sökningarna möjliga. Detta är utmanande, eftersom ordboken inte ursprungligen planerades och byggdes upp med tanke på nätpublicering.

När vi i redaktionen under år 2010 började planera digitalisering och nätpublicering av FO var utgångsläget för digitaliseringsprocessen varierande för de olika avsnitten av ordboken. De två första banden av FO (1976–92) hade redigerats manuellt och sattes vid tryckningen med datorstyrd fotosättning. Under åren 1993–2005 redigerades FO med ett ordbehandlingsprogram, och 2006 övergick redaktionen till ett XML-baserat redigeringsprogram (Cantell & Sandström 2012).

Under perioden 2010–13 överfördes de fyra tryckta banden till en XML-kodad databas och gränssnittet för nätordboken planerades och byggdes

upp. Vid överföringen skannades de två första banden och genomgick textigenkänning (OCR) innan de kunde konverteras till XML-kod och genomgå kontroller. Tredje och fjärde banden hade redigerats i ett ordbehandlingsprogram och de konverterades och kontrollerades vid överföringen till XML-kod. Eftersom den pågående redigeringen av nya artiklar sker med ett XML-baserat redigeringsprogram kan de nyaste artiklarna sammanslås med artiklarna i de fyra första banden till en XML-kodad databas som sedan konverteras till det format användargränssnittet på nätet kräver (Cantell & Sandström 2012).

9.3. Publicering av ordböcker på nätet

Överföringen av materialet i FO till en nätordbok är en del av ett större projekt på Institutet för de inhemska språken. Projektet ”Ordböckerna på nätet (fi. Sanakirjat verkkoon)” har som målsättning att alla ordböcker som ges ut på Institutet för de inhemska språken ska publiceras på nätet. I framtiden planeras en gemensam ordboksportal för nätordböckerna. I portalen kommer länkar att skapas mellan de olika ordböckerna och också extern länkning har diskuterats. För FO skulle det ge ett mervärde att t.ex. kunna länka från ett finskt lånord direkt till motsvarande ord i de finska ordböckerna.

När det gäller FSOB, finns det inte i nuläget konkreta planer på en uppdatering, en ny upplaga eller nätpublicering.⁷ Men om det i framtiden byggs upp en gemensam ordboksportal på Institutet för de inhemska språken kunde också materialet i FSOB i någon form ingå som en del av portalen. De svenska språkvårdarna på Institutet för de inhemska språken har också på andra sätt dragit nytta av de nya tekniska möjligheter som webbsidor och sociala medier kan erbjuda, bl.a. med en samling av de mest aktuella språkfrågorna i databasen *Ordförrådet*, se <<http://kaino.kotus.fi/ordforradet>>.

I februari 2012 kom den finska dialektordboken *Suomen murteiden sanakirja* ut med ett pilotavsnitt på nätet: se <<http://kaino.kotus.fi/sms>>. För närvarande finns ett avsnitt omfattande *huka–lasiäijä* (*huka* sv. ’hake, krok; förståelse, sinne’; *lasiäijä* sv. ’kringresande glasförsäljare’) på nätet. I juni 2013 trädde den finlandssvenska dialektordboken in i en ny era då det första avsnittet på 15 500 artiklar omfattande *I–K* publicerades som nätordbok. Avsnittet *A–H* publicerades i november 2014 och därefter i december 2015 avsnittet *L–O*. Mer material kommer fortlöpande att pub-

⁷ För tillfället är det inte aktuellt med nyutgivning eller uppdatering av FSOB, enligt uppgifter jag fått vid personlig kommunikation med redaktören af Hällström-Reijonen 6.9.2013.

liceras på nätet; i bilderna (9) och (10) ovan belyses hur artiklarna visas i nätdordboken; se <<http://kaino.kotus.fi/fo>> och Karlholm (2015) för en diskussion om FO som nätresurs.

För forskare ger nätdordböcker möjlighet till nya ingångar till språkliga data. Publicering på nätet medför också att dessa ordböcker når ut till nya användare. Stora vetenskapliga ordböcker som är fritt tillgängliga på nätet är ett steg i utvecklingen mot ett mer demokratiskt kunskapssamhälle.

Litteratur

- Ahlbäck, Olav 1946. *Studier över substantivböjningen i Finlands svenska folkmål*. (Studier i nordisk filologi 33–34.) Helsingfors: Svenska litteratursällskapet i Finland.
- Ahlbäck, Olav [1956] 1971. *Svenskan i Finland*. (Skrifter utgivna av Nämnden för svensk språkvård 15.) Stockholm: Läromedelsförlagen.
- Ahlbäck, Olav 1976. Inledning. I: *Ordbok över Finlands svenska folkmål*. Forskningscentralen för de inhemska språken Skrifter 1. Helsingfors. S. VII–XXVIII.
- Ahlbäck, Olav, Vidar Reinhammar, Karen Margrethe Pedersen, Finn Køster, Ove Rasmussen, Gunnar Pedersen & Oskar Bandle 1978. Dialektordböcker, syftemål och problem. I: Benson, Sven och Hugo Karlsson m.fl. (red.), *Dialektologkonferens 1978*, rapport. Göteborg. S. 76–111.
- Aho, Eija 2010. *Spontaanin puheen prosodinen jaksottelu*. PIC Monographs 6. Helsinki.
- Anderson, Benedict [1991] 2003. *Imagined Communities. Reflections on the Origin and Spread of Nationalism*. London: Verso.
- Benson, Sven 1978. Undersökningsobjekt och undersökningsmetoder inom dialektologien. I: Benson, Sven och Hugo Karlsson m.fl. (red.), *Dialektologkonferens 1978*, rapport. Göteborg. S. 10–16.
- Bergenholtz, Henning & Sven-Göran Malmgren 2004. Forord. I: *Lexico-Nordica* 11. S. 1–5.
- Bergenholtz, Henning & Sven Tarp 1994. *Manual i fagleksikografi. Udarbejdelse af fagordbøger – problemer og løsningsforslag*. Herning: Systime.
- Bergroth, Hugo 1917. *Finlandssvenska. Handledning till undvikande av provinsialismer i tal och skrift*. Helsingfors: Schildt.
- Bergroth, Hugo 1928. *Finlandssvenska. Handledning till undvikande av provinsialismer i tal och skrift*. (Andra reviderade och tillökade upplagan). Helsingfors: Schildts Förlags Ab.
- Cannelin, Knut 1908. *Finskt-svenskt lexikon/Suomalais-ruotsalainen sanakirja*. Borgå: WSOY.

- Cantell, Ilse 1995. Målspråkets verbkonstruktioner i en tvåspråkig produktionsordbok. I: *LexicoNordica* 2. S. 19–31.
- Cantell, Ilse 2009. Finlandssvensk ordbok reviderad. I: *LexicoNordica* 16. S. 285–292.
- Cantell, Ilse & Caroline Sandström 2012. Hur blir en traditionell, tryckt ordbok en webbordbok? I: *Nordiska studier i lexikografi* 11. *Rapport från Konferensen om lexikografi i Norden Lund 24–27 maj 2011*. (Skrifter utgivna av Nordiska föreningen för lexikografi. Skrift nr 12.) S. 157–168.
- Considine, John 2008. *Dictionaries in Early Modern Europe. Lexicography and the Making of Heritage*. London: Cambridge University Press.
- Finlandssvensk ordbok* 2008. (Red.) af Hällström-Reijonen, Charlotta & Mikael Reuter. Helsingfors: Schildts i samarbete med Forskningscentralen för de inhemska språken.
- Fjeld, Ruth Vatvedt 1992. Innledning. I: *Nordiske studier i leksikografi. Rapport fra Konferanse om leksikografi i Norden 28.–31. mai 1991*. (Skrifter utgitt av Nordisk forening for leksikografi. Skrift nr. 1.) Oslo. S. 1–6.
- FO = *Ordbok över Finlands svenska folkmål* 1976–.
- Freudenthal, A.O. 1889. *Vöråmålet, Ljud- och formlära, ordlista med register*. (Skrifter utgivna av Svenska litteratursällskapet i Finland 12.) Helsingfors: Svenska litteratursällskapet i Finland.
- FSOB = *Finlandssvensk ordbok* 2000; 2008.
- Granger, Sylviane & Magali Paqout (red.) 2012. *Electronic Lexicography*. Oxford: Oxford University Press.
- Haarala, Risto 1999. Vem är de finska ordböckerna gjorda för? I: *Nordiska studier i lexikografi* 4. *Rapport från Konferensen om lexikografi i Norden Esbo 21–24 maj 1997*. (Skrifter utgivna av Nordiska föreningen för lexikografi. Skrift nr 5.) S. 149–158.
- Hageberg, Arnbjörg 1992. Norsk Ordbok – i 1950 eit nybrottsverk, i 1991 ein anakronisme? I: *Nordiske studier i leksikografi. Rapport fra Konferanse om leksikografi i Norden 28.–31. mai 1991*. (Skrifter utgitt av Nordisk forening for leksikografi. Skrift nr. 1.) Oslo. S. 235–243.
- Hanks, Patrick 2012. Corpus evidence and electronic lexicography. I: Granger, Sylviane & Magali Paqout (red.) *Electronic Lexicography*. Oxford: Oxford University Press. S. 57–82.
- Hannesdóttir, Anna Helga 1998. *Lexikografihistorisk spegel. Den enspråkiga svenska leksikografins utveckling ur den tvåspråkiga*. (Meijerbergs arkiv för svensk ordforskning 23.) Göteborg: Göteborgs universitet.

- Huldén, Lars 1990. Något om nordistikens historia vid Helsingfors universitet. I: Ivars, Ann-Marie & Mirja Saari (red.), *Nordistik i 150 år. Undervisning och forskning i nordiska språk och nordisk litteratur vid Helsingfors universitet 1840–1990*. (Meddelanden från institutionen vid nordiska språk och nordisk litteratur vid Helsingfors universitet Serie B:13.) Helsingfors: Helsingfors universitet. S. 43–71.
- af Hällström, Charlotta 2002. Normeringen i *Finlandssvensk ordbok*. I: *LexicoNordica* 9. S. 51–62.
- af Hällström, Charlotta & Mikael Reuter 2000. *Finlandssvensk ordbok*. Esbo: Schildts.
- af Hällström-Reijonen, Charlotta 2012. *Finlandismer och språkvård från 1800-talet till i dag*. (Nordica Helsingiensia 28.) Nordica, Finska finsk-ugriska och nordiska institutionen. Helsingfors universitet.
- af Hällström-Reijonen, Charlotta & Mikael Reuter 2008. *Finlandssvensk ordbok*. Helsingfors: Schildts i samarbete med Forskningscentralen för de inhemska språken.
- Johansson, Monica 1997. *Lexicon lincopense. En studie i lexikografisk tradition och svenskt språk vid 1600-talets mitt*. (Meijerbergs arkiv för svensk ordforskning 21.) Göteborg: Göteborgs universitet.
- Jonsson, Hans 1992. De stora historiska böckernas särart. I: *Nordiske studier i leksikografi. Rapport fra Konferanse om leksikografi i Norden 28.–31. mai 1991*. (Skrifter utgitt av Nordisk forening for leksikografi. Skrift nr. 1.) Oslo. S. 231–234.
- Karlholm, Annika 2008. Kortare presentation av Ordbok över Finlands svenska folk mål. Band 4. I: *LexicoNordica* 15. S. 239–243.
- Kulonen, Ulla-Maija 1999. Vad söker läsaren i en etymologisk ordbok? I: *Nordiska studier i leksikografi 4. Rapport från Konferensen om leksikografi i Norden Esbo 21–24 maj 1997*. (Skrifter utgivna av Nordiska föreningen för leksikografi. Skrift nr 5.) S. 233–237.
- Lehtinen, Marja 1999. Den första användarresponsen på cd-romversionen av den finska basordboken. I: *Nordiska studier i leksikografi 4. Rapport från Konferensen om leksikografi i Norden Esbo 21–24 maj 1997*. (Skrifter utgivna av Nordiska föreningen för leksikografi Skrift nr 5.) S. 255–263.
- Leinonen, Therese 2010. *An Acoustic Analysis of Vowel Pronunciation in Swedish Dialects*. (Groningen Dissertations in Linguistics 83.) Rijksuniversiteit Groningen.
- Leinonen, Therese 2012. Indelning av finlandssvenska dialekter med särskild hänsyn till Åboland. I: *Folkmålsstudier* 50. S.155–173.
- Loman, Bengt 1985. Axel Olof Freudenthal. Finlands första professor i svenska språket och litteraturen. I: *Folkmålsstudier* 29. S. 15–39.

- Loman, Bengt 1991. Uppsala tur och retur 1861–1914. I: *Saga och sed. Kungl. Gustav Adolfs Akademiens årsbok* 1990. Uppsala. S. 45–55.
- Loman, Bengt 1993. En turistresa genom svensk språkhistoria. I: Wollin, Lars (red.), *Studier i svensk språkhistoria* 3. Uppsala: Institutionen för nordiska språk vid Uppsala universitet. S. 13–30.
- Lundbladh, Carl-Erik 1996. Värdering av SAOB:s korpus och språkprovsamling. I: *LexicoNordica* 3. S. 91–103.
- Martola, Nina 1992. Arbetet med Finsk-svensk ordbok: Normeringsproblem visavi finlandssvenskan. I: *Nordiske studier i leksikografi. Rapport fra Konferanse om leksikografi i Norden 28.–31. mai 1991.* (Skrifter utgitt av Nordisk forening for leksikografi. Skrift nr. 1.) Oslo. S. 515–521.
- Martola, Nina 1995. Substantiv avledda av adjektiv och verb – en jämförelse mellan några ordböcker med finska som källspråk. I: *LexicoNordica* 2. S. 89–108.
- Martola, Nina 1996. Finlandssvenska, finsk-svensk tvåspråkighetslexikografi och korpusar. I: *LexicoNordica* 3. S. 105–119.
- Nikula, Kristina 2000. Finlandssvensk ordbok. I: *LexicoNordica* 7. S. 247–261.
- NLO = *Nordisk leksikografisk ordbok* 1997.
- Noreen, Adolf 1903. *Vårt språk. Nysvensk grammatik i utförlig framställning.* Första bandet. Lund: Gleerups.
- Nordisk leksikografisk ordbok* 1997. (Red.) Henning Bergenholtz, Ilse Cantell, Ruth Vatvedt Fjeld, Dag Gundersen, Jón Hilmar Jónsson, Bo Svensén. (Skrifter utgitt av Nordisk forening for leksikografi. Skrift nr. 4.) Oslo: Universitetsforlaget AS.
- Ordbok över Finlands svenska folkmål* 1976–. (Skrifter utgivna av Forskningscentralen för de inhemska språken 1.) Institutet för de inhemska språken. Helsingfors: Svenska litteratursällskapet i Finland.
- Rabanus, Stefan 2011. The State of the Art in Linguistic Cartography. I: Edlund, Lars-Erik, Lennart Elmevik & Maj Reinhammar (red.), *Studier i dialektologi och sociolingvistik.* (Acta Academiae Regiae Gustavi Adolphi CXVI.) Uppsala. S. 31–52.
- Ralph, Bo 1992. Ord på väg. Några reflexioner kring nordisk lexikografi. I: *Nordiske studier i leksikografi. Rapport fra Konferanse om leksikografi i Norden 28.–31. mai 1991.* (Skrifter utgitt av Nordisk forening for leksikografi. Skrift nr. 1.) Oslo. S. 28–39.
- Ralph, Bo 2000. Svensk lexikografitradition. I: *LexicoNordica* 7. S. 5–22.
- Reinhammar, Vidar 1977. Recension av Olav Ahlbäck, *Ordbok över Finlands svenska folkmål.* Häfte 1 (band I: 1) abbal–bister. Helsingfors 1976. I: *Svenska landsmål och svenskt folkliv.* Uppsala. S. 140–144.

- Reuter, Mikael 2006. Svenskan i Finland på 1900-talet. I: Ivars, Ann-Marie et. al. (red.), *Vårt bästa arv. Festskrift till Marika Tandefelt den 21 december 2006*. (Ekonomi och samhälle. Skrifter utgivna av Svenska handelshögskolan.) Helsingfors. S. 29–45.
- Rietz, Johan Ernst [1862–1867] 1962. *Svenskt dialektlexikon. Ordbok öfver svenska allmogespråket*. Lund: C.W.K. Gleerups förlag.
- Rogström, Lena 1998. *Jacob Serenius lexikografiska insats*. (Meijerbergs arkiv för svensk ordforskning 22.) Göteborg: Göteborgs universitet.
- Romppanen, Birgitta 2000. Från ordlista till ordbok. Utvecklingen av den finsk-svenska ordboken från 1600-talet till i dag. I: *LexicoNordica* 7. S.101–117.
- Romppanen, Birgitta 2001. *Från målspråk till källspråk. Utvecklingen av den finsk-svenska och den enspråkiga finska ordboken*. (Acta Wasensia Nr 91, Språkvetenskap 18.) Vasa: Vasa universitet.
- Ruppel, Klaas 1999. Bedeutungen in einem etymogischen Wörterbuch. I: *Nordiska studier i lexikografi* 4. *Rapport från Konferensen om lexikografi i Norden Esbo 21–24 maj 1997*. (Skrifter utgivna av Nordiska föreningen för lexikografi. Skrift nr 5.) S. 333–339.
- Sandström, Caroline 2000. Från ”svensk språkdialect” till ”svenska folk-mål”. I: Fewster, Derek (red), *Folket. Studier i olika vetenskapers syn på begreppet folk*. (Skrifter utgivna av Svenska litteratursällskapet i Finland nr 626.) Helsingfors: Svenska litteratursällskapet i Finland. S. 177–187.
- Sandström, Caroline 2010. *Genus i östra Nyland – från dialektutjämning till dialektmarkör*. (Nordica Helsingiensia 20.) Dialektforskning 2. Finska, finskugriska och nordiska institutionen. Helsingfors universitet.
- Sandström, Caroline 2011. Finska lån och hur de anpassas i de svenska dialekterna i Finland. I: Edlund, Lars-Erik, Lennart Elmevik & Maj Reinhammar (red.), *Studier i dialektologi och sociolingvistik. Föredrag vid Nionde nordiska dialektologkonferensen i Uppsala 18–20 augusti 2010*. (Acta Academiae Regiae Gustavi Adolphi 116.) Uppsala. S. 249–257.
- Sandström, Caroline 2013. Från folkmålsstudier till interaktionell dialektologi. I: *Folkmålsstudier* 51. S. 86–113.
- Saxén, Ralf 1895. *Finska lånord i östsvenska dialekter. Språkhistoriska studier* I och II. Helsingfors.
- Slotte, Peter 1981. Ordbok över Finlands svenska folk-mål – problem och metoder. I: Loman, Bengt (red.), *De finlandssvenska dialekterna i forskning och funktion*. Åbo: Åbo Akademi. S. 45–64.
- Slotte, Peter 1992. Svenska dialekter och svensk dialektforskning i Finland. I: *Selskab for Nordisk Filologi i København. Årsberetning 1990–1991*. S. 52–63.

- Slotte, Peter 2004. Föreningen för nordisk filologi 1932–2003. *Folkmålsstudier* 43. S.151–173.
- Slotte, Peter, Pia Westerberg & Eva Orava (red.) 1999. *Nordiska studier i lexikografi* 4. *Rapport från Konferensen om lexikografi i Norden Esbo 21–24 maj 1997*. (Skrifter utgivna av Nordiska föreningen för lexikografi. Skrift nr 5.)
- Stenmark, Harry 1983. *Finlandssvenska ord och uttryck*. Stockholm: Proprius.
- Svensén, Bo 1987. *Handbok i lexikografi. Principer och metoder i ordboksarbetet*. TNC. Stockholm: Esselte Studium.
- Svensén, Bo 2004. *Handbok i lexikografi. Ordböcker och ordboksarbete i teori och praktik*. Stockholm: Norstedts Akademiska Förlag.
- Svensson, Lars 1992. Om SAOB:s material och en bibliografisk databas. I: *Nordiske studier i leksikografi. Rapport fra Konferanse om leksikografi i Norden 28.–31. mai 1991*. (Skrifter utgitt av Nordisk forening for leksikografi. Skrift nr. 1.) Oslo. S. 369–379.
- Sørensen, Viggo 1992. Jysk Ordbog – en presentation. I: *Nordiske studier i leksikografi. Rapport fra Konferanse om leksikografi i Norden 28.–31. mai 1991*. (Skrifter utgitt av Nordisk forening for leksikografi. Skrift nr. 1.) Oslo. S. 324–338.
- Tarkiainen, Kari 2008. *Sveriges Österland. Från forntiden till Gustav Vasa. Finlands svenska historia* 1. Helsingfors/Stockholm: Svenska litteratursällskapet i Finland/Atlantis.
- Thelander, Mats 1996. Från dialekt till sociolekt. I: Moberg, Lena & Margareta Westman (red.), *Svenskan i tusen år. Glimtar ur svenska språkets utveckling*. (Skrifter utgivna av Svenska språknämnden 81.) Norstedts Förlag AB. S. 163–181.
- Thors, Carl-Eric 1981. Finska påverkningar på de finlandssvenska folkmålen. I: Loman, Bengt (red.), *De finlandssvenska dialekterna i forskning och funktion*. Åbo: Åbo Akademi. S. 95–99.
- Thylin-Klaus, Jennica 2012. *”Den finländska svenskan” 1860–1920. Tidig svensk språkplanering i Finland ur ett idéhistoriskt perspektiv*. Åbo: Åbo Akademi.
- Vendell, Herman 1884. *Samlingar af ord ur nyländska allmogemålet Nyland*. (Samlingar utgifna af Nyländska afdelningen. Häfte I.) Helsingfors.
- Vendell, Herman 1894. Svenska landsmålsföreningen i Finland 1874–1891 samt 1892–1894. I: *Finländska bidrag till svensk språk- och folkklifs-forskning*. Helsingfors. S. 1–51 samt 304–314.
- Vendell, Herman 1904–1907. *Ordbok över de östsvenska dialekterna*. Helsingfors: Svenska litteratursällskapet i Finland.

- Wessman, V.E.V. 1925–1932. *Samling av ord ur östsvenska folkmål. Tilllägg till H. Vendells ordbok över de östsvenska dialekterna*. Helsingfors: Svenska litteratursällskapet i Finland.
- Vikør, Lars S. 1995. Problem kring ordutvalet i Norsk Ordbok. I: *Nordiske studier i leksikografi 3. Rapport fra Konferanse om leksikografi i Norden Reykjavík 7.–10. juni 1995*. (Skrifter utgitt av Nordisk forening for leksikografi. Skrift nr. 3.) S. 403–412.
- Vikør, Lars S. 1999. Fleirgenerasjonsordbøker og tida. I: *Nordiska studier i leksikografi 4. Rapport från Konferensen om leksikografi i Norden Esbo 21–24 maj 1997*. (Skrifter utgivna av Nordiska föreningen för leksikografi. Skrift nr 5.) S. 395–405.
- Vilppula, Matti 1999. Ordbok över finska dialekter och evigheten. I: *Nordiska studier i leksikografi 4. Rapport från Konferensen om leksikografi i Norden Esbo 21–24 maj 1997*. (Skrifter utgivna av Nordiska föreningen för leksikografi. Skrift nr 5.) S. 407–411.

Ful- och finlexikografi? Om ordboksverksamhet i Sverige i dag och i morgon

Emma Sköldberg & Christian Mattsson

This article approaches the field of contemporary Swedish dictionary making. By extension, the principal aim of the article is to discuss the future outlook of dictionaries and of lexicography as a whole in Sweden. Firstly, we describe the current state of affairs of lexicography in Sweden in terms of key participants – the institutional agents and financiers. Secondly, we recognize differences in standing (high/low) between various kinds of lexicography and lexicographical resources. On one hand, we find high status, scholarly and professionally produced monolingual dictionaries. On the other hand, we discern low status unsystematic, in some cases unreliable, web dictionaries that are often heavily dependent on user-generated content. In a concluding note, we dare to foresee some challenges concerning the future of Swedish lexicography.

1. Inledning

I flera internationella studier, t.ex. de Schryver (2012:142), Granger (2012) och Rundell (2012:72ff.) framhålls den brytningstid som råder när det gäller modern ordboksframställning. Situationen i andra länder, med nya ordboksformat, vikande försäljning när det gäller förlagsordböcker etc., är bekant även ur ett svenskt perspektiv. Just nu tycks framtiden för olika typer av ordböcker – och för lexikograferna själva – knappast självklar.

I det här bidraget ger vi vår bild av situationen för ordboksframställning i Sverige i dag. I korthet rör vi oss kring frågeställningar som: Vem utarbetar ordböcker eller olika typer av lexikaliska resurser? Och, kanske ännu viktigare, vem *finansierar* den ordboksverksamhet som just nu pågår i landet? Vidare undrar vi om vissa former av lexikografi eller vissa konkreta lexikografiska verk har högre status än andra. Kan man tala om en fulare respektive finare lexikografi? Vad beror ett sådant synsätt i så fall på? Vi undrar också hur framtiden kan tänkas se ut. Är den ljusare för vissa typer av lexikografisk verksamhet än för andra? Hur kommer finansieringen av

ordboksverksamheten i Sverige att se ut? Behöver förändringar komma till stånd för att vi ska säkra – eller ännu hellre höja – kvaliteten på framtidens svenska lexikografi? Vilka är i sådana fall dessa förändringar? Som synes rör det sig om frågor som knappast har några enkla svar. Vår avsikt med bidraget är också framför allt att väcka tankar och skapa diskussion, snarare än att besvara alla de spörsmål som ställs ovan.

I nästa avsnitt följer en presentation av ett antal lexikografiska aktörer – både sådana som mer konkret arbetar med lexikografi och sådana som finansierar ordboksframställning i Sverige. För att avgränsa oss koncentrerar vi oss på den ordboksproduktion som inbegriper svenska språket, dvs. enspråkiga svenska ordböcker eller tvåspråkiga ordböcker med svenska som käll- eller målspråk. Vi gör alltså inte anspråk på att ge en fullständig bild; i första hand vill vi uppmärksamma den nuvarande situationen.

2. Lexikografiska aktörer

Om man granskar de ordböcker som nu, eller nyligen, har utarbetats i Sverige, ser man att vissa aktörer återkommer. I det följande tar vi upp bokförlagens ordboksverksamhet, lexikografiskt arbete som finansieras med statliga medel, Svenska Akademiens ordboksprojekt och nätordböcker som helt eller delvis tas fram av användarna själva.

2.1. Förlagen

Traditionellt har förlag stått bakom en stor del av de svenska ordböckerna. Norstedts har i särklass haft den största utgivningen, inte minst som utgivare av *Svenska Akademiens ordlista* (se vidare avsnitt 2.3 nedan), men också tack vare utgivning av andra stora och regelbundet uppdaterade verk. Två andra större svenska förlag som också gett ut både en- och tvåspråkiga ordböcker är Bonniers och Natur och Kultur. Förutsättningarna för förlagen har dock förändrats väsentligt; sammanfattande kan man säga att förlagslexikografien gått från hausse till baisse. Under 1990-talet och fram till ungefär mitten av 00-talet var det högkonjunktur för ordböcker. Höga försäljningsintäkter gjorde det möjligt för förlagen att satsa på ordboksframställning, vilket de också gjorde. Det gavs också ut en rad nya eller kraftigt reviderade tvåspråkiga lexikon (tyska, franska, italienska, spanska, engelska) med svenska som käll- och/eller målspråk, men även svenska enspråkiga ordböcker.

Sedan mitten av 00-talet har emellertid försäljningen sjunkit kraftigt och förlagen har slagit av på ordboksutgivningstakten. Natur och Kultur gav ut

två gedigna enspråkiga verk år 2001 och 2006 (*Natur och Kulturs Svenska Ordbok* respektive *Natur och Kulturs Stora Svenska Ordbok*), men förlaget har nu lagt ner sin ordboksutgivning. Nyligen meddelade också Bonniers att de inte kommer att ge ut en ny upplaga av sin *Bonniers svenska ordbok*, som trots allt kommit ut i 10 upplagor sedan starten 1980. Även Norstedts drar ner. Förlaget har visserligen satsat på betaltjänster och ordboksappar, men ekonomiskt sett ser det sämre ut än tidigare. Problemet med den sviktande betalningsviljan hos ordboksanvändarna har beskrivits av Törnqvist (2010b:485f.), som menar att många användare numera väntar sig att få tillgång till information utan kostnad. Förekomsten av gratisprodukter gör det svårt för ordboksutgivare, som inte finansierar sin verksamhet genom ideellt arbete eller genom anslag från myndigheter och forskningsstiftelser, utan som är beroende av intäkter från försäljning av olika former av produkter. Återanvändning av gammalt och delvis undermåligt material är därför, enligt Törnqvist, alltför vanligt (se vidare Hannesdóttir i denna volym).

2.2. Staten

Det bedrivs också lexikografisk verksamhet i statlig regi, bl.a. vid Institutet för språk och folkminnen som sedan länge gett ut dialektordböcker. Vid Dialekt- och folkminnesarkivet i Uppsala pågår t.ex. arbetet med *Svenskt dialektlexikon*. Ordboken planeras omfatta ett band och ska behandla ”ett urval av ord ur det centrala dialektala ordförrådet” <<http://www.sprakochfolkminnen.se/sprak/dialekter/projekt/svenskt-dialektlexikon-sdl.html>>. Lexikonet blir den första översiktliga dialektordboken sedan Rietz *Svenskt dialektlexikon: Ordbok öfver svenska allmogespråket* som kom ut på 1860-talet.

Vid institutets avdelning Språkrådet och inom Lexin-projektet arbetar man framför allt med lexikon på minoritetsspråk – i dag ett tjugotal sådana. Dessa tvåspråkiga lexikon utgår från en och samma svenska ordbas, *Lexins svenska lexikon* (se vidare Malmgren 2012). Den 1 januari 2013 tog Språkrådet också över utgivningen av ordlistor för tolkar, arbeten som tidigare givits ut av Tolks- och översättarinstitutet (TÖI) vid Stockholm universitet. Ett annat specialområde är nyordslistor. Språkrådets föregångare, Svenska språknämnden, stod bakom två nyordsböcker, *Nya ord i svenskan från 40-tal till 80-tal* (1986) respektive *Nyordsboken. Med 2 000 nya ord in i 2000-talet* (2000) och en uppdaterad ordlista över nyligen tillkomna svenska ord är nu aktuell.

Vidare måste Institutionen för svenska språket vid Göteborgs universitet betraktas som central i lexikografiska sammanhang. Till institutionen hör

Lexikaliska institutet vars huvuduppgift är att bedriva forsknings- och utvecklingsarbete inom såväl lexikologi som lexikografi. Det lexikografiska arbetet vid institutionen omfattar dels uppbyggnad av lexikaliska databaser och ordböcker, dels forskning om ordböcker. Ordboksproduktionen sker främst inom ramen för externt finansierade projekt, inte minst genom ett väl etablerat samarbete med Svenska Akademien (se vidare nedan) men också med Språkrådet. Den svenska målspråksredaktionen inom ISLEX-projektet är också förlagd till samma institution. ISLEX är en allmänt tillgänglig flerspråkig nätordbok där isländska utgör källspråk och danska, norskt bokmål, nynorska och svenska utgör målspråk (se bl.a. Rauset et al. 2012, Hannesdóttir i denna volym). Den svenska delen av projektet har finansierats av svenska staten.

Till institutionen hör dessutom Språkbanken vars uppdrag är att samla in, utveckla och tillgängliggöra språkresurser åt forskare och allmänheten. Antalet korpusar i Språkbanken har ökat avsevärt på senare tid. I december 2012 innehöll de 91 korpusar som var fritt tillgängliga sammanlagt över en miljard löpande ord. Korpusarna innehåller en rad olika texttyper/genrer, t.ex. tidningsmaterial, skönlitteratur, artiklar i Wikipedia och parti-program. Antalet korpusar som man kan söka i stiger också hela tiden. Därutöver har det nya korpussökningsverktyget, Korp, många fördelar i förhållande till tidigare sökverktyg. Sökdelen i verktyget är uppdelad i tre olika versioner: *enkel*, *utökad* och *avancerad*, som var och en möjliggör sökningar med olika grad av komplexitet (se vidare Borin et al. 2012b). Med hjälp av Karp, Språkbankens nya lexikala plattform, kan man navigera bland Språkbankens olika lexikala resurser. Bl.a. kan man söka i flera ordböcker samtidigt, t.ex. Dalins enspråkiga *Ordbok öfver svenska språket* från mitten av 1800-talet, *Swesaurus* (ett svenskt ordnät) och *Svenskt frasnät* (en svensk motsvarighet till Berkleys *FrameNet*) (se vidare Borin et al. 2012a). Språkbankens korpusar och lexikonresurser länkas samman via SALDO, ett betydelselexikon vilket designats för att kunna användas i automatisk språkbearbetning (se t.ex. Borin & Forsberg 2011:24ff. samt Borin i denna volym.) Språkbanken tillhandahåller alltså dels textmaterial och verktyg som kan ligga till grund för lexikografiskt arbete, dels mer renodlade lexikon. Dessa är emellertid i första hand utarbetade för språkteknologiska ändamål.

2.3. Svenska Akademien

Den fristående kulturinstitutionen Svenska Akademien är känd för sitt långvariga engagemang inom svensk lexikografi och den finansierar också en betydande del av det lexikografiska arbete som sker i Sverige just

nu. Det största åtagandet gäller framtagningen av *Svenska Akademiens ordbok* (SAOB, 1898–), den historiska ordbok som beskriver svenskt skriftspråk från 1520-talet och framåt. Enligt planen ska ordboken, som utarbetas vid SAOB-redaktionen i Lund, vara färdig inom de närmaste åren och den beräknas då bestå av 38 band. Häfte nummer 36:2, som kom i december 2012, sträcker sig från artikeln *upphäva* till artikeln *utsudda*. Se bl.a. Rosqvist (2011) för en relativt aktuell beskrivning av arbetet med SAOB; se även Pär Nilssons bidrag i denna volym.

Ett annat projekt som inleddes på 1800-talet är *Svenska Akademiens ordlista* (SAOL). Den trettonde upplagan av ordlistan kom ut 2006 och året efter publicerades samma upplaga i digitalt format, *SAOL Plus* (Berg 2009:147ff.; Gellerstam 2009b:78ff.). Arbetet med SAOL sker vid Institutionen för svenska språket/Lexikaliska institutet vid Göteborgs universitet och uppdateringsarbetet inför den fjortonde upplagan beskrivs bl.a. av Berg & Malmgren (2010).

Vidare övertog Svenska Akademien innehållet i *Nationalencyklopedins ordbok* (NEO, 1995–96) från bokförlaget Bra Böcker i slutet av 1990-talet. Innehållet i NEO-databasen har sedan vidareutvecklats av redaktionen vid Lexikaliska institutet och 2009 publicerades den enspråkiga samtidsordboken *Svensk ordbok utgiven av Svenska Akademien* (SO) i två band (se vidare bl.a. Malmgren 2009a, b).

Svenska Akademien har därmed som finansiär ett stort inflytande över de ordböcker som tas fram i Sverige i dag. När man diskuterar innehållet i Akademiens ordböcker bör man dock komma ihåg dess uppdrag. Svenska Akademiens syfte är sedan instiftandet år 1786 bl.a. att ”arbeta uppå Svenska Språkets renhet, styrka och höghet” (se Loman 1986:5) och detta syfte torde ha påverkat *vilket* ordförråd som behandlats i Akademiens verk genom tiderna och *hur* det behandlats. Loman (1986) menar dock att det inte är omedelbart klart vad Gustaf III kan ha åsyftat med de tre honnörorden på 1700-talet. Av Gellerstam (2009a, b) framgår också att synen på vilka uppslagsord som ska tas upp i SAOL varierat genom tiderna. I modern tid skulle de tre substantiven i citatet ovan emellertid kunna motsvaras av *klarhet*, *uttrycksfullhet* och *anseende* (se <www.saob.se/om/historik/>).

2.4. Svenska folket

Numera finns det också ett antal lexikaliska/lexikografiska resurser på internet som i större eller mindre utsträckning produceras eller vidareutvecklas av användarna själva (se vidare Törnqvist 2010b:487 om brukarmedverkan vid ordboksframställning). Exempel på sådana lexikon är

svenska *Wiktionary*, *Folkets lexikon*, *Tyda.se*, *Synonymer.org*, *Luxikon* och *Folkmun.se*. De flesta av dessa resurser är fritt tillgängliga.

Viss forskning har bedrivits om den här typen av lexikon för andra språk. Bland annat redovisar Fuertes-Olivera (2009) en grundlig genomgång av engelska *Wiktionary* (se även Lew 2011:236f. och Rundell 2012:80f.). Däremot har mycket litet skrivits om motsvarande svenska lexikon (jfr dock Törnqvist 2010a).

En relevant fråga är givetvis vad dessa ordböcker faktiskt innehåller – något som ofta kan vara svårt att avgöra. I t.ex. svenska *Wiktionary* står det att ordboken (i december 2012) innehåller närmare 330 000 huvuduppslag. Vidare innehåller *Synonymer.org*, enligt webbplatsen, mer än 30 000 synonymer. I andra nätlexikon anges ingen exakt siffra, men antalet uppslagsord verkar vara långt färre. Exempelvis har *Luxikon* ca 2 100 uppslagsord.

I *Wiktionary* uppmanas användarna att lägga in ord och uttryck som är etablerade – eller som det står ”belagda” – i svenskan. Etablering är ett komplext fenomen och vad kravet att orden ska vara belagda egentligen innebär kan förstås diskuteras (se t.ex. Svanlund 2009). Vad det betyder i just detta sammanhang förtydligas emellertid på sajten:

Att ett uttryck är belagt betyder någotdera av:

1. uppenbarligen spridd användning
2. använt i välkänt verk
3. förekommer i en granskad akademisk tidskrift
4. använt åtminstone 3 gånger oberoende av varandra med åtminstone 1 års spridning för att uttrycka mening i ett permanent medium.

Bland huvuduppslagen på *Wiktionary* finner man mer traditionella uppslagsord som *arrendator*, men också böjningsformerna *arrendatorn* och *arrendatorns*. Vidare finns där flerordsuttryck (som *rinna ut i sanden*) och rena tecken (exempelvis utropstecken och humörsymboler). Så kallade dagsländor och lanserade nyord accepteras alltså inte (jfr diskussionen kring innehållet i *Luxikon* strax nedan).

I *Folkmun.se* finns inga uttryckta restriktioner vad gäller lemmaurval, men där finner man i huvudsak en viss typ av ord och ordförbindelser, närmare bestämt vardagliga sådana och slanguttryck. Några av dessa är klart mer konventionaliserade än andra. Nedan återges några uppslag i *Folkmun.se* med befintliga betydelsebeskrivningar (hämtade i december 2012). Uppslagsorden och betydelseorna återges i oredigerat skick, dvs. precis som de står i lexikonet.

- (1) **giftpinne** cigarett
- (2) **gaddning** tatuering
- (3) **gyllene måsen** Annat namn på MC Donalds, syftar på det faktum att det stora gula/gyllene bågarna påminner om en slarvigt tecknad fiskmåssedd fram eller bakifrån.
- (4) **Gösta Timmersjö** Försvenskat namn på Justin Timberlake
- (5) **gött mos** Göteborgs uttryck för något som är trevligt, skönt

Medan t.ex. slangordet *giftpinne* är belagt redan 1942 och upptaget i många välkända ordböcker (se bl.a. SO 2009), torde *gaddning* vara yngre. Det behandlas heller inte, vad vi kan se, i några tryckta lexikografiska verk. Ordet har emellertid ca 100 belägg i Språkbankens korpusar – de allra flesta träffarna är i bloggmaterial. Men ordet är även belagt en handfull gånger i *Göteborgs-Posten* runt 2005. Uttrycket *gyllene måsen* är enligt lexikonet en mer svenskklingande beteckning på en välkänd snabbmatskedja. Namnet är förvisso belagt på internet, men verkar inte vara allmänt spritt (jfr den mer etablerade slangbeteckningen *Donken*). Detsamma gäller *Gösta Timmersjö*, en försvenskning av den amerikanske popsångaren Justin Timberlakes namn. Det västsvenska uttrycket *gött mos* har, vad vi kan se, inte letat sig in i tryckta svenska ordböcker ännu. En snabb granskning av innehållet i *Folkmun.se* ger alltså vid handen att detta lexikon, precis som många andra resurser som tagits fram av användarna själva, fungerar som ett gott komplement till mer etablerade svenska ordböcker när det gäller en viss typ av ord och uttryck. I betydelseangivelserna m.m. finns det dock ett stort antal skrivfel och brott mot svenskans skrivregler.

Många av de aktuella resurserna präglas av en upproriskhet gentemot etablissemangets (beskrivning av) språk. Detta framträder inte minst på *Folkmun*:s förstasida där följande fråga ställs: ”Är det Svenska Akademien som bestämmer över det svenska språket eller är det du?”. Frågan tyder på en önskan att uppvärdera ”vanliga” användares språk – det är gott nog. Över huvud taget framträder ett intresse för folkligare sätt att uttrycka sig – inte minst i webbplatsens namn. Detta gäller också andra fria webbplatser, t.ex. *Folkets lexikon*. Också namnet på detta lexikon betonar allmänhetens kännedom om och, inte minst, makt över språket. Inom ramen för *Folkets lexikon* har *alla* möjlighet att bidra och resultatet är gemensam egendom. Eller som det står på sidan: ”Detta engelsk-svenska lexikon tillhör folket och utvidgas av oss alla tillsammans.” (Se Kann & Hollman 2011a, b för kortfattade presentationer av lexikonet.)

De personer som tagit initiativ till dessa nätlexikon verkar hysa ett genuint intresse för språk. Innehållet på webbplatserna präglas av en språkglädje och – inte minst – verbal uppfinningsriktighet. Det framgår tydligt av *Luxikons* förstasida, där det står: ”Hej! Luxikon är en webbplats med ord. Dina ord, andras ord, ord som används ofta, sällan och ord som inte fanns förrän nyss.” I samband med detta citat aktualiseras också frågan om vem som står bakom dessa webbplatser. Enligt *Luxikons* förstasida startades webbplatsen av ”Mattias, Karin och Christoffer i juli 2008”. Angivelsen gör både ett anonymt och familjärt intryck och kontrasten gentemot andra ordböcker där upphovsman e.d. lyfts fram desto mer, är påfallande. Vi tänker exempelvis på de välkända randiga pärmarna på Norstedts lexikon, som ska signalera kvalitet, och Svenska Akademiens logotyp på SAOL och SO.

I detta avsnitt har vi uppmärksammat några olika aktörer som svarar för ordboksframställning i Sverige i dag. Det finns dock inga vattentäta skott dem emellan. *Folkets lexikon*, som nu utvecklas av användarna själva, togs ursprungligen fram av Skolverket, dvs. en statlig myndighet. Likaså kan en produkt ges ut av ett förlag men delvis vara finansierad med statliga medel, t.ex. *Norstedts ryska ordbok* (2006).

3. Ful- och finlexikografi

Låt oss nu återknyta till vår tillspetsade rubrik och nybildningarna *ful-* och *finlexikografi*. Givetvis anspelar vi med dessa på att prefixet *ful-* ibland skämtsamt används som motsats till *fin-* (se betydelse 3 hos uppslagsordet **ful** i *Wiktionary*; jfr även motsatsparet *dålig – bra*). En välkänd sammansättning med *ful-* som förled är *fulöl* som enligt samma ordbok betecknar ’billig, hemmagjord eller direktimporterad öl av låg kvalitet’. Andra mer tillfälliga sammansättningar som belyser förledens produktivitet är *fulkultur*, *fulrum*, *fulsprit*, *fulkaffe*, *fulvirke*, *fulnedladdad* och *fulstajla* som alla är hämtade ur *Göteborgs-Posten* (se GP 2011 och 2012 i Korp).

Kan man då tala om fulare respektive finare lexikografi? Jo, det kan man nog, om man med fullexikografi avser vissa (typer av) ordböcker som inte fungerar som en auktoritet, är mindre erkända i vetenskapliga sammanhang och som anses hålla lägre kvalitet. En särskild sorts lexikografi kan givetvis inte klassas som helt igenom ful och en annan som helt igenom fin. Men vi tror att det kan finnas schablonföreställningar om vad som har högre respektive lägre status inom lexikografien och att det går att urskilja vissa faktorer som spelar in.

En faktor att räkna med i detta sammanhang kan vara *graden av vetenskaplighet och systematik*, dvs. i vad mån det lexikografiska arbetet är

teoribaserat, stringent och bygger på vetenskapliga metoder. Å ena sidan har vi enskilda webbinitiativ, som ibland kan framstå som osystematiska och lite urskillningslösa, kanske otillförlitliga och därmed mindre fina. På förlagen är lexikografin visserligen mer systematisk, men den är oftast inte forskningsbaserad på samma sätt som den ordboksverksamhet som bedrivs på universitet m.m. Förlagslexikograferna tar större hänsyn till användarna, dvs. köparna, som finansierar verksamheten, och själva användbarheten är därför primär. Den lexikografiska verksamhet som bedrivs av forskare kan, å andra sidan, anses ha högst status, just tack vare vetenskapligheten, men även för att de ordböcker som tas fram inom akademien ger mer uttömmande och precisa beskrivningar av språket. Vi har alltså fått intryck av att den språkbeskrivning som strävar efter precision och fullständighet av många användare betraktas som finare än den som är pedagogiskt förenklad. Låt oss ge ett exempel. De flesta kan nog skriva under på att betydelsebeskrivningarna i t.ex. redan nämnda *Natur och Kulturs Svenska Ordbok* (2001, NoK), som i första hand vänder sig till inlärare, är enklare att förstå än betydelsebeskrivningarna i SO (2009). Förklaringarna i inlärningsordboken kan samtidigt upplevas som inexakta, förenklade och ofullständiga. För att förtydliga presenteras beskrivningen av ett och samma verb i de två ordböckerna. Nedan återges hur den första betydelsen hos uppslagsordet **mäta** presenteras dels i NoK, dels i SO:

mäta^v (...)

1 ta reda på t.ex. hur långt eller hur stort något är ♦ *Mät bordet först så vet vi om det får plats i köket!* ♦ *Klockan mäter tiden.* ♦ *Termometern mäter temperaturen.*

(...)

mä`ta (...)

1 fastställa utsträckning av ngt som (vanl.) kan anges med siffror: längd, volym, temperatur, tryck etc.; genom jämförelse med ngn standard, ofta med hjälp av instrument: *mätglas; mätinstrument; mätmetod; kontrollmäta; ~ isens tjocklek; ~ tiden med timglas*

o äv. ngt utvidgat: *~ avståndet med ögonmått; hans insats i danshistorien kan inte ~s ännu; man kan inte ~ med ett mått för pressen och med ett annat för radio och tv; förluster som inte kan ~s i pengar (...)*

Utan tvekan är SO:s betydelsebeskrivning mer uttömmande och precis än NoK:s. Samtidigt torde SO:s beskrivning vara mindre begriplig för användare med begränsade kunskaper i svenska, bl.a. beroende på ordval, meningslängd, bruk av förkortningar och inskjutna parenteser.

Uppfattning om vad som är fult och fint kan också ha med *ordbokstyp* att göra. Tvåspråkig lexikografi är mer sällan forskningsbaserad. Hannesdót-

tir (2012:40f.) menar att den tvåspråkiga lexikografin förlorade i prestige i samband med att den enspråkiga lexikografin etablerades. Medan den enspråkiga lexikografin fick akademisk status kom de tvåspråkiga ordböckerna främst att betraktas som pedagogiska redskap i studiet av främmande språk. De övergick också till att i första hand bli en angelägenhet för förlag och särspråksspecialister (se även Hannesdóttir i denna volym).

Även den tänkta *målgruppen* kan inverka; enligt vår erfarenhet har ordböcker för inlärare och för undervisning inte riktigt lika högt anseende som ordböcker för modersmålstalare (jfr diskussionen om skillnaderna mellan NoK och SO strax ovan).

Själva *ordboksinnehållet*, dvs. det språk som beskrivs i en ordbok, kan också ha högre eller lägre status – vilket i sin tur kan påverka verkets ställning i förhållande till andra verk. Lite hårdraget betraktas nog skriftspråk som finare än talspråk av många språkanvändare (se t.ex. Andersson 1985:75ff. för en mer nyanserad bild av förhållandet mellan tal och skrift). Vidare torde många ord som förknippas med högre stilnivå tillskrivas högre status än t.ex. vardagliga ord. Långt ner i hierarkin återfinns många slangord (se vidare t.ex. Kotsinas 1998:IX). Ord som t.ex. *fecke* 'fest' och *miffad* 'dum, korkad, elak, tokig' behandlas heller inte i allmänspråkliga svenska ordböcker utan tas bara upp i just slangordböcker som *Norstedts svenska slangordbok* (1998) och *Förortsslang* (2004). Ett relevant verk i detta sammanhang är också *Stora Fula Ordboken* (Dagrin 2004) med sina 10 000 uppslagsord, däribland drygt 500 svordomar och mer än 2 400 beteckningar på det manliga respektive kvinnliga könsorganet. Dagrin (2004:5) skriver också mycket riktigt i förordet att boken är ett försök att

presentera ofta försummade och baktalade ord, sådana ord som det stått strid om, som omskrivits eller uteslutits, som stämplats som vulgära, obscena och fula och som därför få ordforskare velat befatta sig med.

Ordbokens *förhållande till den språkliga normen* och *syftet* med ordboken är också relevant. Frågan är alltså om den primära avsikten med verket är att redovisa eller att påverka språkbruket (se vidare om ordböcker och normering i Svensén 2004:29f.). Delar av den språkintresserade allmänheten kan tycka att normativa ordböcker är lite finare – i dessa verk får man ju veta vad som är "rätt", vad "det heter på riktigt". Forskare kan ha en annan syn; för dem kan deskriptiva ordböcker vara relevantere (och därmed lite finare) eftersom språket i beskrivande arbeten inte är fullt så tillrättlat utan mer verklighetsnära.

Avslutningsvis vill vi lyfta fram ännu en faktor som kan påverka synen på, i synnerhet, resultatet av det praktiska arbetet, nämligen det som uppfattas som *drivkraften* bakom verket. Är det ett förlag med vinstintresse

som har tagit fram ordboken? Är upphovsmännen forskare vid ett universitet? Eller rör det sig om språkintresserade lekmän? Man kan tänka sig att personer som på lediga stunder lägger in (påhittade) ord på en webbplats för att det är roligt inte tas på allvar av personer som arbetar professionellt med lexikografi. De senare strävar i högre grad efter struktur, systematik och konsekvens vid beskrivningen av ett etablerat ordförråd.

Sammanfattningsvis, om vi utgår från dessa schabloner, kan man nog säga att enspråkiga svenska ordböcker som framför allt behandlar vårdat, offentligt skriftspråk och dessutom är utgivna av Svenska Akademien har hög status. Däremot har tvåspråkiga ordböcker utgivna på förlag lägre status. Webbordböcker som har utarbetats av användarna själva, och där framför allt det mer privata språket inklusive vardagsspråk och slang redovisas, har däremot oklar status, men vi menar att de flesta av dessa inte är särskilt erkända bland yrkesverksamma lexikografer.

4. Tankar om den närmaste framtiden

Hittills har vi gett vår bild av det rådande läget vad gäller ordboksframställning i Sverige. Givetvis kan vi bara spekulera kring framtiden för svensk lexikografi, men den kan se olika ut för enskilda ordboksproducenter och olika typer av ordböcker.

Först och främst är det sannolikt att förlagens ordboksintäkter fortsätter att sjunka och att förlagen har svårt att satsa på nya ordboksprojekt. Möjligen upphör den förlagsfinansierade ordboksproduktionen helt, och förlagen blir endast distributörer av andras material. Men förlagen kan också komma att fortsätta med bara nättjänster och ordboksappar. Användarnas betalningsvilja – och huruvida den kan tänkas öka i framtiden – spelar givetvis en mycket viktig roll i detta sammanhang.

Statens roll vid finansiering av svensk lexikografi är inte lika självklar som den varit tidigare. Institutet för språk och folkminnen (och därmed Språkrådet) styrs liksom alla statliga myndigheter av en *instruktion*, dvs. ett uppdrag från regeringen. I instruktionen står det visserligen att myndigheten bl.a. ska förmedla och sprida kunskap om dialekter och minoritetsspråk, men däremot inget specifikt om att det ska bedrivas lexikografisk verksamhet. Men instruktioner kan ändras och inget tyder i dag på att statliga myndigheter kommer att stärkas ekonomiskt. Snarare är det tvärtom. Verksamheten med t.ex. *Lexin*-ordböckerna kan därför inte tas för given.

Däremot är det, glädjande nog, desto troligare att Svenska Akademien fortsätter att finansiera arbetet med enspråkiga svenska ordböcker. Ett viktigt riktmärke för Akademien är givetvis den tidpunkt då SAOB ska vara

klar. Sannolikt påbörjas då en uppdatering av ordboken. Vid Göteborgs universitet fortsätter dels det redan påbörjade arbetet med en fjortonde upplaga av SAOL vilken ska vara färdig år 2015, dels arbetet med att ta fram en elektronisk version av samtidsordboken SO.

Avslutningsvis tror vi att de ordböcker som helt eller delvis tas fram av användarna själva är här för att stanna (se t.ex. Kann 2010:114). Mycket talar för att de blir fler och fler. Möjligen höjs deras kvalitet och de kan på så vis bli mer användbara, men troligtvis fungerar de bättre vid beskrivningen av vissa delar av ordförrådet än andra. Rundell (2012:80) återger t.ex. bokförlaget Macmillans erfarenheter med deras s.k. *Open Dictionary*. Användarna tycks kunna bidra mest med bl.a. nyord, regionala varieteter och teknisk terminologi. Ur ett språkvårdsperspektiv kan man dock undra om de aktuella webbplatserna kan utgöra ett hot mot god svenska eftersom inte fullt så gott språk också får stort utrymme i dem (jfr språket i definitionerna hämtade ur *Folkmun.se* i avsnitt 2.4 ovan). Törnqvist (2010a) menar också att en bred krets av bidragsgivare kan leda till att ordboks innehållet blir mer folkligt; ordböckerna kan t.ex. få större inslag av ungdomsspråk, modeord och slang – vilket är på gott och ont. ”Mycket av det nya materialet kan tyckas efemärt, men det återspeglar faktiskt användarnas intressen och informationsbehov” (Törnqvist 2010a:389).

5. Inför framtiden: önskvärda förändringar

Sammanfattningsvis kommer det troligen att vara ett begränsat antal professionella lexikografer (med ett fåtal större uppdragsgivare) som, tillsammans med användarna själva, utarbetar framtidens ordböcker. Frågan är om det är så lyckat. Vilka konsekvenser kommer det att ha för ordböckerna – både när det gäller typ och innehåll? Vad behöver göras för att lexikografin ska utvecklas i annan och bättre riktning?

Först och främst måste de tvåspråkiga ordböckerna uppvärderas. Hannesdóttir (2012:40) konstaterar som sagt att den tvåspråkiga lexikografin tillmätts en underordnad betydelse under de senaste 150 åren. Men den tidiga tvåspråkiga lexikografin spelade en viktig roll för kodifieringen och standardiseringen av svenskan och i själva verket bygger den enspråkiga lexikografin på den språkbeskrivning som etablerades inom tvåspråkig lexikografi på 1700-talet och tidigt 1800-tal. Hannesdóttir (2012:56) poängterar också att det kontrastiva perspektivet inom tvåspråkig lexikografi kan berika den enspråkiga och bidra till viktiga kunskaper om svenskan. Den tvåspråkiga lexikografin måste alltså tillbaka in i finrummet igen. Vi vill även se fler finansieringsmöjligheter, i synnerhet när förlagens intäkter sannolikt fortsätter att minska. Tvåspråkig lexikografi borde kunna

stödjas av större statliga finansiärer (som Vetenskapsrådet), men också av stiftelser (som Riksbankens jubileumsfond). Möjligen skulle det kunna vara intressant för Svenska Akademien att gå in som aktör i detta sammanhang. Tvåspråkiga ordböcker består ju till hälften av svenska språket och för många användare är den svenska som redovisas i tvåspråkiga verk en minst lika viktig förebild som den svenska som presenteras i enspråkiga ordböcker. Om framtidens tvåspråkiga lexikografi i första hand bedrivs av användarna själva (t.ex. inom ramen för webbinitiativ som Folkets engelsksvenska lexikon) kan de tvåspråkiga ordböckerna komma att innehålla en svenska som inte är fullt så standardiserad som tidigare.

Det behövs också en annan syn på – och mer forskning om – de lexikografiska resurser som tagits fram av användarna själva. Frågan är hur de utvecklas och hur stora innehållsmässiga likheter/skillnader det finns mellan dem och mer etablerade ordböcker som t.ex. Norstedts större tvåspråkiga ordböcker och SO. En viktig fråga är också vad de håller för kvalitet.

I framtiden måste användarna i högre grad än nu betraktas som en resurs. Ett viktigt skäl att engagera användarna i ordboksarbetet är att de tillsammans, naturligt nog, har ”kännedom om ett mycket större språk användningsområde än några få ordboksredaktörer” (Törnqvist 2010b:384ff.). Dessutom är användarnas arbetsinsats gratis. Tack vare användarnas bidrag öppnas möjligheter till projekt som inte hade varit genomförbara om projekten skulle ha bemannats av anställd personal. Det finns alltså all anledning att bejaka den pågående utvecklingen som kan leda till att framtidens ordböcker blir mer omfattande, detaljerade, aktuella och tillförlitliga – och dessutom billigare (Törnqvist 2010a:389). Törnqvist ställer sig dock frågande till en framtid för svenska ordböcker som enkom byggs upp av användarna själva. För det första är det helt enkelt svårt att formulera korrekta och begripliga ordboksdefinitioner. För det andra är antalet användare för begränsat. Enligt Törnqvist krävs det relativt många medarbetare för att en wiki, dvs. webbsida eller webbportal där texten redigeras av användarna själva, ska bli stor och få hög kvalitet. Wiki-konceptet fungerar därför bättre vid framställning av t.ex. engelska och tyska ordböcker som har många användare. De bästa resultaten kan dock troligen uppnås genom samarbetsprojekt mellan professionella lexikografer och brukare. Exempel på sådana projekt är redan nämnda *Macmillan Open Dictionary*, *Merriam-Webster's Open Dictionary* och *Wordnik*, som enligt Lew (2011:237) utgör en ”interesting blend of online dictionary genres, involving a collaborative community-driven component built around a 'professional' core”. Samarbetsprojekt av detta slag för svenskans del är således mer än välkomna.

6. Hinder som måste undanröjas

Om vår önskan att svensk lexikografi ska utvecklas i positiv riktning ska uppfyllas måste en del hinder undanröjas, i första hand ekonomiska sådana. En möjlighet är att mer ordboksarbete – såväl en- som tvåspråkigt – bedrivs vid universiteten. I dag är det dock inte helt oproblematiskt att bedriva konkret lexikografiskt arbete inom universitetsvärlden. Svensén (2004:313) konstaterar att erfarenheten tyvärr visar att ordboksarbetet i Sverige i regel tillmätts ett ganska blygsamt meriteringsvärde. Enligt samme författare betraktas också ordboksverksamhet åtminstone av somliga som ett slags tillämpad lingvistik (Svensén 2004:4) och detta kan vara en bidragande orsak till att det har varit svårt att få medel för mer renodlade lexikografiska projekt från finansiärer som i allmänhet stöder humanistisk forskning. Det är dock enkelt att lyfta fram projekt med förutsättningar att resultera i välbehövliga nya ordböcker som skulle kunna bidra med ett kontrastivt perspektiv på svenska, projekt som i dag saknar finansiering. Ett exempel är arbetet med en större och modernare dansk-svensk ordbok. Det finns goda utgångsmaterial för framtagning av en sådan ordbok. Vi tänker exempelvis att innehållet i *Svensk-dansk ordbog* (2010), ISLEX och *Den Danske Ordbog* (DDO 2003–2005) kan läggas till grund för arbetet. Många befintliga ordböcker skulle också kunna bli bättre om det bara fanns ordentlig finansiering. Arbetsuppgifter saknas följaktligen inte, och kompetens finns, men nya finansieringsformer måste alltså hittas.

7. Slutord

I detta bidrag har vi fört fram vår syn på läget för svensk lexikografi av i dag. Vi har bl.a. diskuterat vilka det är som utarbetar och finansierar ordböcker just nu och om vissa typer av verk – och ordboks innehåll – har högre status än andra. Men vi har också funderat över hur framtiden för svenska ordböcker kan se ut, vilka problem som finns och några förändringar som borde komma till stånd för att förbättra villkoren för ordboksframställning i Sverige. Otvivelaktigt befinner sig svensk lexikografi i en brytningstid och framtiden för olika typer av ordböcker – och för lexikograferna – är knappast självklar. Men enligt vår mening är lexikografisk mångfald eftersträvansvärd. Det behövs olika typer av ordböcker som beskriver vårt brokiga språk och som fokuserar på skilda delar av svenskans ordförråd. Vidare tror vi att beskrivningen av svenskans ordförråd gynnas av att ordböcker tas fram vid olika redaktioner och av att lexikograferna inte är för få till antalet. Användarna själva utgör också en mycket viktig

resurs som måste tas tillvara i högre grad. I nuläget ser dock framtiden betydligt ljusare ut för finlexikografin än för den som inte har lika hög status. Mest prekär är situationen för tvåspråkig lexikografi – där måste något göras. Vi som arbetar med lexikografi och lexikaliska resurser vid forskningsinstitutioner såväl som vid förlag behöver samarbeta mer och utbyta erfarenheter för att stärka vårt fält och föra det framåt. Vi behöver också arbeta mer aktivt för att höja lexikografins status i Sverige.

Källor och litteratur

- Andersson, Lars-Gunnar 1985. *Fult språk. Svordomar, dialekter och annat ont*. Stockholm: Carlssons.
- Berg, Sture 2009. Om ordböjning och SAOL Plus. I: Gellerstam, Martin (red.), *SAOL och tidens flykt*. Stockholm: Norstedts. S. 139–165.
- Berg, Sture & Sven-Göran Malmgren 2010. Varför travintresse ska ut och fotbollsintresse ska in. I: *Språkbruk* 4/2010. S. 4–11.
- Bonniers svenska ordbok* 2010. Utarbetad av Sten Malmström, Iréne Györki & Peter A. Sjögren. 10 uppl. Stockholm: Bonniers.
- Borin, Lars 2016. Lexikografi för maskiner och lexikografi för människor. I: Hannesdóttir, Anna Helga (red.), *Framtidens lexikografi. Rapport från ett symposium i Göteborg 5 oktober 2012*. (Meijerbergs arkiv för svensk ordforskning 42.) Göteborg: Göteborgs universitet. S. 9–27.
- Borin, Lars & Markus Forsberg 2011. Swesaurus – ett svenskt ordnät med fria tyglar. I: *LexicoNordica* 18. S. 17–39.
- Borin, Lars, Markus Forsberg, Leif-Jöran Olsson & Jonatan Uppström 2012a. The open lexical infrastructure of Språkbanken. I: *Proceedings of LREC 2012*. Istanbul: ELRA. S. 3598–3602.
- Borin, Lars, Markus Forsberg & Johan Roxendal 2012b. Korp – the corpus infrastructure of Språkbanken. *Proceedings of LREC 2012*. Istanbul: ELRA. S. 474–478.
- Dagrin, Bengt G. 2004. *Stora fula ordboken. Baktalade ord på språkets bakgård*. Stockholm: Carlssons.
- Dalin, Anders Fredrik 1850–1853. *Ordbok öfver svenska språket*. I–II. Stockholm: Författarens förlag.
- DDO = *Den Danske Ordbog* 2003–2005. Köpenhamn: Det Danske Sprog- og Litteraturselskab. <<http://ordnet.dk/ddo>> (december 2012).
- Folkets lexikon*. <<http://folkets-lexikon.csc.kth.se/folkets/>> (december 2012).
- Folkmun*. <<http://www.folkmun.se/>> (december 2012).

- Fuertes-Olivera, Pedro A. 2009. The Function theory of lexicography and electronic dictionaries: WIKTIONARY as a prototype of collective free multiple-language Internet dictionary. I: Bergenholtz, Henning, Sandro Nielsen & Sven Tarp (eds.), *Lexicography at a Crossroads. Dictionaries and Encyclopedias Today, Lexicographical Tools Tomorrow*. Bern: Peter Lang. S. 99–134.
- Förortsslang 2004. Utarbetad av Dogge Doggelito och Ulla-Britt Kotsinas. Stockholm: Norstedts ordbok.
- Gellerstam, Martin 2009a. Om lånordens anpassning till svenskan. I: Gellerstam, Martin (red.), *SAOL och tidens flykt*. Stockholm: Norstedts. S. 105–118.
- Gellerstam, Martin 2009b. SAOL i många upplagor. I: Gellerstam, Martin (red.), *SAOL och tidens flykt*. Stockholm: Norstedts. S. 53–83.
- Granger, Sylviane 2012. Introduction: Electronic lexicography – from challenge to opportunity. I: Granger, Sylviane & Magali Paquot (eds.), *Electronic Lexicography*. Oxford: Oxford University Press. S. 1–11.
- Hannedóttir, Anna Helga 2016. Tvåspråkig lexikografi – ny teknik, nya funktioner. Den digitala ordboken som parallellkorpus. I: Hannedóttir, Anna Helga (red.), *Framtidens lexikografi. Rapport från ett symposium i Göteborg 5 oktober 2012*. (Meijerbergs arkiv för svensk ordforskning 42.) Göteborg: Göteborgs universitet. S. 29–57.
- Hannedóttir, Anna Helga 2012. Ekvivalensrelationer i tvåspråkig lexikografi. I: *LexicoNordica* 19. S. 39–58.
- ISLEX. <<http://islex.se>> (december 2012).
- Kann, Viggo 2010. KTHs morfologiska och lexikografiska verktyg och resurser. I: *LexicoNordica* 17. S. 99–117.
- Kann, Viggo & Joachim Hollman 2011a. *Slutrapport för projektet Folkets engelsk-svenska lexikon*. <<http://folkets-lexikon.csc.kth.se/folkets/projekt/slutrapport.pdf>> (december 2012).
- Kann, Viggo & Joachim Hollman 2011b. *Slutrapport för projektet Vidareutveckling av Folkets lexikon*. <<http://folkets-lexikon.csc.kth.se/folkets/projekt/slutrapport2.pdf>> (december 2012).
- Kotsinas, Ulla-Britt 1998. Introduktion. I: *Norstedts svenska slangordbok*. Stockholm: Norstedts Ordbok. S. VI–XIII.
- Lew, Robert 2011. Online Dictionaries of English. I: Fuertes-Olivera, Pedro A. & Henning Bergenholtz (eds.), *E-lexicography. The Internet, Digital Initiatives and Lexicography*. London/New York: Continuum. S. 230–250.
- Lexins svenska lexikon*. 4 uppl. <<http://lexin.nada.kth.se/lexin/>> (december 2012).

- Loman, Bengt 1986. ”En Inrättning, ägnad endast til Språkets förbättring”. I: Allén, Sture, Bengt Loman & Bengt Sigurd (red.), *Svenska Akademien och svenska språket. Tre studier*. Stockholm: Norstedts. S. 1–141.
- Luxikon*. <<http://www.luxikon.se/>> (december 2012).
- Macmillan Open Dictionary*. <<http://www.macmillandictionary.com/open-dictionary/>> (december 2012).
- Malmgren, Sven-Göran 2009a. Från Nationalencyklopedins ordbok (1995–96) till Svensk ordbok utgiven av Svenska Akademien (2009). Med tillbaka- och sidoblickar. I: *LEDA-Nyt* 47. S. 14–20.
- Malmgren, Sven-Göran 2009b. On production-oriented information in Swedish monolingual defining dictionaries. I: Nielsen, Sandro & Sven Tarp (eds.), *Lexicography in the 21st Century: In honour of Henning Bergenholtz*. Amsterdam/Philadelphia: John Benjamins. S. 93–102.
- Malmgren, Sven-Göran 2012. Från Svenska ord (Lexin) 3 till Svenska ord 4. I: Eaker, Birgit, Lennart Larsson & Anki Mattisson (red.), *Nordiska studier i lexikografi* 11. Lund. S. 454–465.
- Merriam-Webster's Open Dictionary*. <<http://nws.merriam-webster.com/opedictionary/>> (december 2012).
- Natur och Kulturs Stora Svenska Ordbok* 2006. Stockholm: Natur och Kultur.
- Natur och Kulturs Svenska Ordbok* 2001. Stockholm: Natur och Kultur.
- NEO = *Nationalencyklopedins ordbok* 1995–1996. Höganäs: Bra Böcker.
- Nilsson, Pär 2016. Teori i praktiken. Om prototyp teorin och SAOB. I: Hannesdóttir, Anna Helga (red.), *Framtidens lexikografi. Rapport från ett symposium i Göteborg 5 oktober 2012*. (Meijerbergs arkiv för svensk ordforskning 42.) Göteborg: Göteborgs universitet. S. 59–73.
- Norstedts svenska slangordbok* 1998. Utarbetad av Ulla-Britt Kotsinas. Stockholm: Norstedts Ordbok.
- Norstedts ryska ordbok* (rysk-svensk/svensk-rysk) 2006. Stockholm: Norstedts.
- Nya ord i svenskan från 40-tal till 80-tal* 1986. Stockholm: Svenska språknämnden och Esselte Studium.
- Nyordsboken. Med 2 000 nya ord in i 2000-talet* 2000. Stockholm: Svenska språknämnden och Norstedts ordbok.
- Rauset, Margunn, Anna Helga Hannesdóttir & Aldís Sigurðardóttir 2012. Ein-, to- eller fleirspråkleg ordbok? I: Eaker, Birgit, Lennart Larsson & Anki Mattisson (red.), *Nordiska studier i lexikografi* 11. Lund. S. 512–523.
- Rietz, Johan Ernst 1862–1867. *Svenskt dialektlexikon: Ordbok öfver svenska allmogespråket*. Lund: Gleerups.

- Rosqvist, Bodil 2011. Strängt upptagen *och* fast besluten. *Beskrivningen av kollokationer i SAOB*. (Meddelanden från Institutionen för svenska språket 64.) Göteborg: Inst. för svenska språket, Göteborgs universitet.
- Rundell, Michael 2012. 'It works in practice but will it work in theory?' The uneasy relationship between lexicography and matters theoretical. I: Vatvedt Fjeld, Ruth & Julie Matilde Torjusen (eds.), *Proceedings of the 15th EURALEX International Congress 7–11 August, 2012, Oslo*. Oslo: University of Oslo. S. 47–92.
- SAOB = *Ordbok över svenska språket utgiven av Svenska Akademien (Svenska Akademiens ordbok)* 1898–. Lund: Gleerups.
- SAOL = *Svenska Akademiens ordlista* 2006. 13 uppl. Stockholm: Norstedts.
- SAOL Plus = *Svenska Akademiens ordlista* 2007. 13 uppl. Cd-rom. Stockholm: Norstedts (i distribution).
- de Schryver, Gilles-Maurice 2012. Lexicography in the Crystal Ball: Facts, Trends and Outlook. I: Vatvedt Fjeld, Ruth & Julie Matilde Torjusen (eds.), *Proceedings of the 15th EURALEX International Congress 7–11 August, 2012, Oslo*. Oslo: University of Oslo. S. 93–164.
- SO = *Svensk ordbok utgiven av Svenska Akademien* 2009. Stockholm: Norstedts.
- Språkbanken*. <<http://spraakbanken.gu.se/>> (december 2012).
- Svanlund, Jan 2009. *Lexikal etablering. En korpusundersökning av hur nya sammansättningar konventionaliseras och får sin betydelse*. (Acta Universitatis Stockholmiensis. Stockholm Studies in Scandinavian Philology, New series 52.) Stockholm: Institutionen för nordiska språk, Stockholms universitet.
- Svensén, Bo 2004. *Handbok i lexikografi. Ordböcker och ordboksarbete i teori och praktik*. Andra, omarbetade och utökade upplagan. Stockholm: Norstedts.
- Svensk-Dansk Ordbog* 2010. Köpenhamn: Det Danske Sprog- og Litteraturselskab och JP/Politikens Förlagshus A/S.
- Synonymer.org*. <<http://synonymer.org/>> (december 2012).
- Tyda.se*. <<http://tyda.se/>> (december 2012).
- Törnqvist, Lars 2010a. Brukarmedverkan i webbordböcker. I: Jóhannesson, Kristinn, Ida Larsson, Erik Magnusson Petzell, Sven-Göran Malmgren, Lena Rogström & Emma Sköldberg (red.), *Bo65. Festskrift till Bo Ralph*. (Meijerbergs arkiv för svensk ordforskning 39.) Göteborg. S. 383–390.
- Törnqvist, Lars 2010b. Ordböcker på Internet och Internet som ordbok. I: Lönnroth, Harry & Kristina Nikula (red.), *Nordiska studier i lexikografi* 10. Tammerfors. S. 484–493.

Wiktionary (svensk version).

<<http://sv.wiktionary.org/wiki/Wiktionary:Huvudsida>> (dec. 2012).

Wiktionary (engelsk version).

<http://en.wiktionary.org/wiki/Wiktionary:Main_Page> (dec. 2012).

Wordnik. <<https://www.wordnik.com/>> (dec. 2012).

