LexicoNordica

LexicoNordica

26 · 2019

NORDISKE ORDBØKER
MED TO ELLER FLERE SPRÅK

NORDISK FORENING FOR LEKSIKOGRAFI

LexicoNordica 26 · 2019
Nordiske ordbøker med to eller flere språk

Hovedredaktører
Sturla Berg-Olsen
Emma Sköldberg

Redaksjonskomité
Anna Braasch
Kjetil Gundersen
Annika Karlholm
Mariann Skog-Södersved
Ásta Svavarsdóttir

© 2019 LexicoNordica og forfatterne

Omslag og sats: Laurids Kristian Fahl
Trykk: Tarm Bogtryk a-s, Danmark

LexicoNordica trykkes med støtte fra
Nordplus Nordens Språk

ISSN 0805-2735
ISSN 1891-2206 (online)

Innhold

Sturla Berg-Olsen & Emma Sköldberg
Nordiske ordbøker med to eller flere språk 9

Tematiske bidrag

Sofia Björklöf & Saarni Laitinen
Korpus och enspråkig ordbok som bas för ny estnisk-finsk ordbok 17

Anna Helga Hannesdóttir
Glossarium – Sveo-Gothicum: en underlig ordbok eller ett bidrag till forskningen? 35

Poul Hansen
Statistisk modellering og prognosticering af efterspørgslen efter netordlister 55

Helga Hilmisdóttir
ISLEX på finska – redigering, användning och mottagande 75

Patrick Leroyer & Henrik Køhler Simonsen
Google Translate: trussel eller redning for oversættelsesordbøger? 95

Jacqueline Levin & Frederikke Ulv
"Hvordan siger man ...?" Om direkte brugerkontakt i ordbogsarbejdet 117

Morgan Nilsson
Somaliska tvåspråkiga ordböcker i Norden:
utbud och kvalitet 135

Margunn Rauset
Bokmålsordboka og *Nynorskordboka* – einegga, toegga
eller siamesiske tvillingar? 155

Trond Trosterud
Kva bruker vi minoritetsspråksordbøker til? Ein studie av
brukarloggane for tolv tospråklege ordbøker 177

Ikke-tematisk bidrag

Kristian Blensenius
Revision av konstruktionsuppgifter i *Svensk ordbok*
utgiven av Svenska Akademien 203

Anmeldelser

Ebba Hjorth
To store leksikografer – om Svend Eegholm-Pedersens
Mothstudier 227

Tor Erik Jenstad
Norsk banneordbok – ei omfattande samling med eit
noko vidt banneomgrep 239

Lennart Larsson
En nyliberal ordlista 251

Geir Lima
NE:s stora italienska ordbok – ny og utvidet utgave av den største ordboken mellom svensk og italiensk 259

Morgan Nilsson
Kunnskapsforlagets *Stor norsk-polsk ordbok* – ett informationsspäckat monumentalverk 275

Carina Nilstun
Samvittighetsfullt revidert *Norsk ordbok* 295

Mihkel Tõnnov
Eesti-soome sõnaraamat – en lovande ny estnisk-finsk ordbok på nätet 311

Meddelelser

Lotte Follin
Mindeord for Jens Axelsen 329

Pär Nilsson
Rapport från styrelsen för Nordiska föreningen för lexikografi 333

Redaksjonelt 339

Nordiske ordbøker med to eller flere språk

Sturla Berg-Olsen & Emma Sköldberg

Med dette legger vi fram det 26. nummeret av *LexicoNordica*, som utgis av Nordisk forening for leksikografi (NFL). Nummeret inneholder tradisjonen tro en tematisk del med artikler basert på innlegg holdt under *LexicoNordica*-symposiet, som ble avholdt på Lysebu i Oslo 10.–12. januar 2019. I tillegg til ni tematiske artikler består nummeret av ett ikke-tematisk bidrag, sju anmeldelser og et par meddelelser av forskjellig karakter. Bakerst i boka finner leseren opplysninger om tidsskriftet og instruksjoner til bidragsforfattere.

Årets tema er altså nordiske ordbøker med to eller flere språk. For mange ordbokbrukere er det nettopp dette som *er* ordbøker, for to- og flerspråklige ordbøker spiller en viktig rolle som hjelpemiddel i mange ulike situasjoner der det er snakk om kommunikasjon over språkgrenser. Også historisk er det den to- og flerspråklige leksikografien som var den primære, mens enspråklige ordbøker kom til først seinere. I en del tilfeller lar det seg ikke uten videre slå fast om et ordbokverk er enspråklig, tospråklig eller flerspråklig. Dette er noe av det Anna Helga Hannesdóttir tar opp i sin artikkel om Haquin Spegels *Glossarium* fra 1712, for dette verket har trekk som kjennetegner hver av de tre typene. Margunn Rauset er inne på noe av det samme når hun reiser spørsmålet om hvorvidt *Bokmålsordboka* og *Nynorskordboka* bør regnes som to ordbøker for hvert sitt språk, eller som én ressurs med to målformer.

Felles for mange av artiklene er at de på ulike vis tar opp forholdet mellom dem som redigerer ordbøkene, og dem som bruker dem, og ikke minst hvordan viten om brukernes behov kan samles inn og brukes til å forme ordbøker og andre leksikografiske ressurser.

Poul Hansen viser at den statistiske metoden tidsserieanalyse kan avdekke informasjon om brukernes atferd som ikke nødvendigvis trer klart fram ved bruk av mer tradisjonelle metoder. I sin artikkel analyserer Trond Trosterud brukerlogger for tolv ulike elektroniske ordbøker til og fra minoritetsspråk i Norden. Dataene viser blant annet at bruksmønstrene varierer fra språksamfunn til språksamfunn, noe ordbokredaktørene åpenbart bør være klar over og ta hensyn til. Jacqueline Levin & Frederikke Ulv gjør på sin side rede for hvordan man driver direkte brukerkontakt i Ordbogen A/S, og hvordan kontakten med brukerne brukes til å bygge ut og forbedre de tospråklige ordbøkene på nettstedet ordbogen.com. I sitt bidrag viser Patrick Leroyer & Henrik Køhler Simonsen hvordan Google Translate i dag er et nyttig og selvsagt verktøy for oversettere av fagtekster. Gitt dette mener de at leksikografiske ressurser i større grad bør skreddersys til målretta etterredigering av maskinoversatte tekster.

ISLEX er en spesiell ordbok fordi den er flerspråklig og omfatter alle de statsbærende språkene i Norden. I sin artikkel drøfter Helga Hilmisdóttir noen av utfordringene med redigeringa av den finske delen av ISLEX. Hun presenterer data fra en brukerundersøkelse som blant annet viser at det gir brukerne en merverdi at finsk integreres i det fellesnordiske og flerspråklige ISLEX. Sofia Björklöf & Saarni Laitinen skriver om sine erfaringer fra arbeidet med den nylig lanserte estisk-finske ordboka; den er basert på en enspråklig korpusbasert estisk ordbok, og det har både sine fordeler og sine ulemper. Morgan Nilsson gjør opp status når det gjelder nordiske tospråklige ordbøker til og fra somalisk. Somalisk er et av de største minoritetsspråkene i Norden, men de ordbøkene som finnes, er ikke alltid av så god kvalitet som man skulle ønske.

Årets eneste ikke-tematiske artikkel er skrevet av Kristian Blensenius, som viser hvordan man arbeider for å revidere føringa av konstruksjonsopplysninger i den kommende utgaven av *Svensk ordbok utgiven av Svenska Akademien*.

Som vanlig inneholder nummeret en rekke anmeldelser av nylig utkomne ordbøker og leksikografiske ressurser av ulik karakter – interessant nok handler seks av de sju anmeldelsene om utgivelser som foreligger i papirformat. Gitt årets tema er vi glade for å ha med tre anmeldelser av tospråklige ordbøker. Mihkel Tõnnov skriver om den nye estisk-finske nettordboka som ble publisert i januar 2019, altså samme ordbok som omhandles i Björklöf & Laitinens tematiske bidrag. Morgan Nilsson har tatt et dypdykk i den svært omfangsrike *Stor norsk-polsk ordbok*, som er resultatet av over tjue års arbeid. Geir Lima tar for seg den nye, oppdaterte og utvidede utgaven av *NE:s stora italienska ordbok*, som er publisert både på papir og elektronisk.

Tor Erik Jenstad gransker Ruth Vatvedt Fjelds *Norsk banneordbok*, en utgivelse som på en systematisk måte kartlegger et område av ordforrådet som preges av humor og stor kreativitet. Carina Nilstun anmelder en enspråklig bokmålsordbok som kom ut i 2018 og primært retter seg mot elever og studenter.

Utgangspunktet for Ebba Hjorths bidrag er Svend Eegholm-Pedersens storverk om kildegrunnlaget for Moths ordbok fra omkring år 1700. I tillegg til en presentasjon og vurdering av verket får vi et interessant innblikk i Hjorths egne erfaringer med Moths verk. Lennart Larsson anmelder *Nyliberal ordlista*, en bok som på overflaten ser ut som en ordbok, men som etter alt å dømme primært har andre formål enn de reint leksikografiske.

Under rubrikken Meddelelser står det et minneord over Jens Axelsen og en rapport fra styret i NFL, bl.a. med informasjon om den 15. konferansen om leksikografi i Norden, som fant sted i Helsingfors i juni 2019.

Den avsluttende delen av nummeret handler om *LexicoNordica* og tidsskriftets redaksjon. Hovedredaktører for dette nummeret er Emma Sköldberg (Sverige) og Sturla Berg-Olsen (Norge). Som tidligere er det fem landsredaktører som representerer hvert sitt land: Anna Braasch (Danmark), Kjetil Gundersen (Norge), An-

nika Karlholm (Sverige), Mariann Skog-Södersved (Finland) og Ásta Svavarsdóttir (Island). På NFLs generalforsamling i juni 2019 ble Anna Helga Hannesdóttir valgt som ny hovedredaktør etter at Emma Sköldbergs periode gikk ut. Lennart Larsson ble valgt som ny landsredaktør for Sverige og Harry Lönnroth som ny landsredaktør for Finland.

Redaksjonen har bestemt at de kommende symposiene skal behandle disse temaene:

2020: kontroversielle ord i nordiske ordbøker
2021: nordisk leksikografi – nå og i framtida

Redaktørene tar gjerne imot forslag til foredragsholdere i 2021 samt forslag til framtidige *LexicoNordica*-temaer.

Vi vil takke alle landsredaktørene for innsatsen og for et godt og velsmurt samarbeid på tvers av Norden. Takk til Rikke E. Hauge for all hjelp med forberedelser til og avvikling av symposiet på Lysebu. Vi er også takknemlige for det arbeidet Laurids Kristian Fahl har lagt ned i settinga av tidsskriftet. Nok en gang har Fondet for dansk-norsk samarbeid gjort det mulig å avholde symposiet til en gunstig pris, og Nordplus Nordens Språk har bevilga støtte til symposiet og til trykking av tidsskriftet. Takk til begge disse. Til slutt vil vi rette en stor takk til NFLs styre og i særdeleshet til Pär Nilsson for godt samarbeid i året som har gått.

Sturla Berg-Olsen
seniorrådgiver, dr.art.
Språkrådet
Postboks 1573 Vika
NO-0118 Oslo
sturla.berg-olsen@sprakradet.no

Emma Sköldberg
universitetslektor, docent
Institutionen för svenska språket
Göteborgs universitet
Box 200
SE-405 30 Göteborg
emma.skoeldberg@svenska.gu.se

TEMATISKE BIDRAG

Korpus och enspråkig ordbok som bas för ny estnisk-finsk ordbok

Sofia Björklöf & Saarni Laitinen

The new Estonian-Finnish online dictionary, *Viro–suomi-sanakirja. Eesti-soome sõnaraamat*, published in January 2019, has been compiled using the new corpus-based Estonian monolingual dictionary. This article examines how editing the dictionary is made easier due to it being based on the corpus-based monolingual dictionary, while also discussing how this editing process can result in some problems. Technology helps in compiling bilingual dictionaries, but do the advantages of using a monolingual dictionary as a basis for a bilingual one outweigh the resulting problems? As there are quite a few emerging issues, it is important that sufficient time and labour resources are given to projects of this sort.

1. Inledning

Den nya ordboken *Viro–suomi-sanakirja. Eesti-soome sõnaraamat* är den första estnisk-finska ordboken som publicerats enbart elektroniskt. Ordboken baserar sig på den nya estniska definitionsordboken (*Eesti keele sõnaraamat* 2019) som för sin del är korpusbaserad. Därmed är den estnisk-finska ordboken också den första i sitt slag som baserar sig på korpus. Syftet med denna artikel är att behandla och diskutera hur detta å ena sidan underlättar redigeringsarbetet och å andra sidan leder till olika problem. Författarna av artikeln har varit redaktörer för den estnisk-finska ordboken.[1]

1 Huvudredaktörer Heikki Hurtta, Margit Langemets, redaktörer Sofia Björklöf, Anu Haak, Külli Kuusk, Saarni Laitinen, Ilona Paajanen, Ulla Vanhatalo.

2. Bakgrund

Initiativtagaren till skapandet av en ny estnisk-finsk ordbok var Institutet för det estniska språket (Eesti Keele Instituut) i Tallinn. Ordboken redigerades i samarbete med Institutet för de inhemska språken i Helsingfors. Projektet påbörjades i september 2016 och avslutades i december 2018. Ordboken publicerades den 9 januari 2019. Att den publicerades just då beror delvis på att Finland fyllde hundra år 2017 och Estland 2018 och att de båda instituten ville uppmärksamma de här jubileumsåren med det nya verket.

Ordboken har redigerats med redigeringsprogrammet EELex, som skapades vid Institutet för det estniska språket för ungefär tio år sedan. Den tekniska sidan har överlag skötts av esterna. I redaktionen har både estnisk- och finskspråkiga lexikografer ingått, de flesta med målspråket finska som modersmål, vilket gör ordboken till den första estnisk-finska ordboken med en tvåspråkig redaktion.

Ett viktigt skäl till att den aktuella ordboken utarbetades var att de tidigare ordböckerna var föråldrade. Den första estnisk-finska ordboken, *Virolais-suomalainen sanakirja*, sammanställdes för hundra år sedan av en finsk språkvetare, Lauri Kettunen (1917). Den följande, *Eesti-soome sõnaraamat* med 36 000 ord, publicerades i Estland 1972. Det gavs ut ett par nya upplagor av den och år 2015 publicerades en reviderad upplaga med 9 000 nya uppslagsord (*Eesti-soome sõnaraamat*, sammanlagt 45 000 ord). Den nyaste upplagan är trots allt lite föråldrad och också den innehåller bara få exempel, fraser och sammansatta verb. Redaktionen var dessutom enbart estnisk. Det behövdes trots den reviderade upplagan därför en ny modern ordbok vars redaktörer har målspråket finska som modersmål och som skulle innehålla fraser och sammansatta verb som lemman samt rikligt med exempel.

3. Tänkta användare

Traditionella och rentav självklara ordboksanvändare är översättare och studerande, vilket gäller också denna ordbok. Dessutom pendlas det mycket mellan Finland och Estland. Förutom de som arbetar i grannlandet har även turister och andra resande nytta av ordboken. Finskan och estniskan är nära besläktade östersjöfinska språk, som har rikligt med gemensamma ord men samtidigt gott om falska vänner. Både språkens och ländernas närhet bidrar till att det finns många som vill lära sig grannspråket.

Också de som bor mer eller mindre permanent i grannlandet är potentiella användare. I slutet av 2018 bodde det ca 49 700 estnisskspråkiga i Finland och estniskan är därmed det näst största invandrarspråket i Finland efter ryskan som talas av 79 200 personer (Statistikcentralen 2019). Det finns en estnisk skola i Helsingfors och några skolor med estniska klasser. År 2011 bodde det ca 2 600 finskspråkiga i Estland och ca 167 000 estländare kunde finska på någon nivå (Statistikaamet 2011a; 2011b). I universitetsstaden Tartu finns det också en finsk skola.

4. Elektronisk publicering

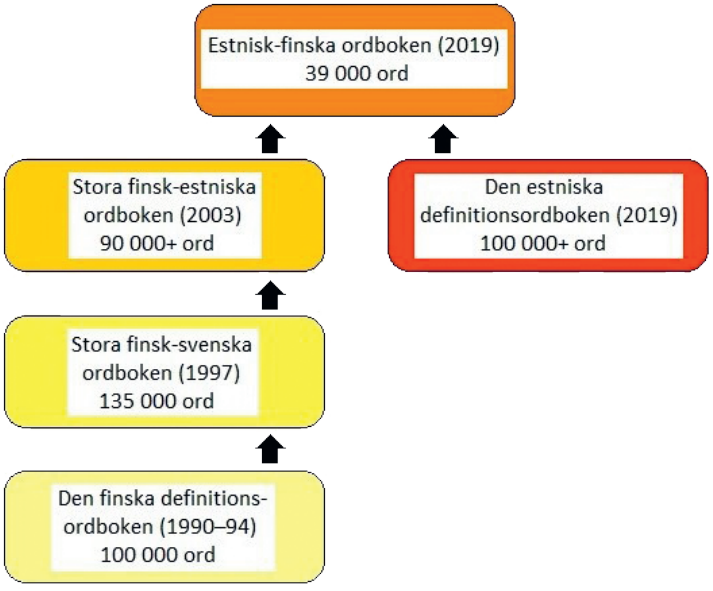
Att den estnisk-finska ordboken publicerades enbart elektroniskt gav en hel del möjligheter. Man kunde ge många exempel eftersom man inte behövde oroa sig för brist på utrymme. Den elektroniska ordboken kunde också genast göras fritt tillgänglig för användarna. En av den elektroniska publiceringens avigsidor är att det var möjligt att ge ut ordboken med bara ca 29 000 kontrollästa artiklar samt ca 10 000 artiklar som inte ännu blivit kontrollästa. Detta tycker vi är mycket problematiskt bl.a. med tanke på ordbokens pålitlighet trots att de okontrollerade artiklarna är tydligt utmärkta. För tillfället uppdateras ordboken fortlöpande men tyvärr tar det en lång tid att kontrolläsa alla artiklar.

Ordbokens gränssnitt är responsivt, vilket gör den bekväm att använda med mobiler och surfplattor. Folk är nuförtiden vana vid att ha mobilen med sig hela tiden och söka information på nätet.

Figur 1: Startsidan på datorskärmen.

5. Ordböcker som källor

Att den nya estnisk-finska ordboken är korpusbaserad betyder egentligen att den baserar sig på den nya estniska definitionsordboken, som i sin tur är korpusbaserad. Utöver det utgicks det från gamla ordböcker: finska uppslagsord togs automatiskt från den stora finsk-estniska ordboken och de blev potentiella ekvivalenter i den estnisk-finska ordboken. Den stora finsk-estniska ordboken baserar sig på den stora finsk-svenska ordboken, som i sin tur baserar sig på den finska definitionsordbokens version från 1990-talet (se figur 2). Potentiella ekvivalenter krävde mycket redigeringsarbete, eftersom en stor del av dem var oanvändbara och måste gallras bort. Dessutom måste redaktörerna ofta lägga till ekvivalenter. Alla ekvivalenter kom automatiskt in i den första betydelsegruppen och måste därefter placeras i rätt grupp. (Metoden diskuteras närmare i avsnitt 7.)



Figur 2: Andra ordböcker som källa för den estnisk-finska ordboken.

Den stora finsk-estniska ordboken (*Suomi–viro-suursanakirja*) gavs ut som bok år 2003. På nätet har den funnits gratis från och med november 2017. Den är den nyaste och största finsk-estniska ordboken som finns.

6. Korpus

Korpusen som den nya estnisk-finska ordboken baserar sig på är sammanställd vid Institutet för det estniska språket, vi finskspråkiga redaktörer har inte arbetat på den. Korpusen uppdateras regelbundet; år 2017 bestod den av 1,3 miljarder textord. Korpusarbetet påbörjades på 1980-talet vid Tartu universitet och Tartu-korpusen med ca 250 miljoner textord användes som bas för den som nu sammanställs vid Institutet för det estniska språket. Den nuvarande korpusen baserar sig på många slags estniska texter samlade

från internet, bl.a. tidnings- och tidskriftsspråk samt myndighetsspråk, men också t.ex. kommentarer till tidningsartiklar. Nästan ingen skönlitteratur finns dock med. De flesta texterna är skrivna från och med början av 1990-talet.

Verktyget för korpusarbetet som görs vid Institutet för det estniska språket är Sketch Engine. Allt som är skrivet på estniska på internet laddas ner; det finns dock några filter, t.ex. blogginlägg filtreras bort.

6.1. Fördelar

När det gäller språkparet estniska och finska är denna ordbok som redan konstaterats den första som är korpusbaserad. Vi upplever att användning av en korpus som bas för en tvåspråkig ordbok har flera goda sidor. För det första är korpusen i fråga väldigt stor, med mer än en miljard textord, och för det andra är den aktuell (vilket också skapar problem, se avsnitt 6.2). För det tredje tror vi att det kan vara nyttigt för användaren att exemplen är från levande språkbruk; texterna i korpusen är skrivna av väldigt många olika individer, också helt vanligt folk. För det fjärde möjliggör korpusen en indelning av orden i frekvensgrupper. De mest frekventa 35 000 orden har delats in i sju frekvensgrupper och de första grupperna finns med i sin helhet i ordboken.

Artiklarna **higine** 'svettig' och **nägu** 'ansikte' fungerar som exempel på levande språkbruk. I satsen *ta oli näost täitsa higine* 'han/hon/hen var helt svettig i ansiktet' finns det estniska ordet *täitsa*, som är den talspråkiga varianten för *täiesti* 'helt'. I estniskan kan ordet **nägu** 'ansikte' också betyda 'min'. I satsen *sul on sihuke paipoisi nägu* 'du har/uppvisar en oskyldig uppsyn' har man använt det talspråkliga ordet *sihuke* 'sån'. Vi finskspråkiga redaktörer har översatt exemplen så exakt som möjligt och om exemplet innehåller talspråk har vi försökt använda talspråk också i översättningen.

Att blogginlägg, som ofta utgörs av mer talspråksliknande tex-

ter, filtreras bort leder till att korpusen och därmed också ordboken inte innehåller så många talspråkliga ord och uttryck som den kunde. Det är ändå glädjande att det trots allt finns en del sådana ord och uttryck i ordboken för att de har använts i andra slags texter som ingår i korpusen.

6.2. Problem

Det finns naturligtvis också problem med att använda en korpus. Vi har märkt att vissa ord som är vanliga i talet kan saknas eller se mindre frekventa ut eftersom korpusen enbart innehåller skrivna texter. Också gammalt och litterärt ordförråd kan saknas, eftersom de flesta texterna är från år 1990 och framåt. Myndighetsspråk är däremot överrepresenterat.

I den estniska definitionsordboken och därmed också i vår ordbok har man beslutat ordna betydelsegrupperna enligt frekvens, vilket vi anser är problematiskt. Traditionellt har betydelsegrupperna ordnats så att den konkreta och ursprungliga betydelsen kommer först och de bildliga eller annars yngre betydelserna efter den (Kamppi 2014:152). Enligt Wojciechowska (2012:70) är dock betydelsernas frekvens i korpusen nuförtiden det vanligaste sättet att ordna betydelsegrupperna.

Också de estniska exemplen härstammar från korpusen. De kan vara problematiska med tanke på innehållet, mängden eller placeringen p.g.a. att redaktören som valde dem har förhastat sig. Ett problem för finskspråkiga redaktörer är att exemplen kan ha för knapp kontext, vilket kan försvåra översättandet och t.o.m. leda till felaktiga tolkningar.

6.2.1. Betydelsegrupperna i frekvensordning

När betydelsegrupperna ordnas enligt frekvens i korpusen, kan ordningen ändå bli traditionell tack vare att betydelsernas frekvens i korpusen är sådan. I artikeln **jäine** är ordets första betydelse 'isig'

och andra betydelse 'iskall' (se figur 3). Den bildliga betydelsen av 'iskall' hör således till den andra gruppen och kommer sist. Vi anser att betydelsernas ordning i artikeln **jäine** är bra eftersom det kan vara pedagogiskt att ange betydelserna med den konkreta och ursprungliga först. Problemet är att betydelsegrupperna som sagt alltid ordnas enligt frekvensen, vilket leder till innehållsmässigt olika resultat i olika artiklar. Ordningen ser därför slumpmässig ut och blir inte transparent för användaren. Det finns naturligtvis problem också med andra sätt att ordna betydelsegrupperna, men ofta vore det möjligt att nå ett resultat som är mer koherent och därför också begripligare för användaren (Meer 2002:514).

jäine M ‹adj›
1 *jääst koosnev või jääga kaetud, jäätunud (mitte veekogu kohta)*
jäinen, iljanteinen
▪ öine külm oli maandumisraja jäiseks muutnud halla oli muuttanut kiitoradan jäiseksi / halla oli tehnyt kiitotiestä iljanteisen
2 *külm kui jää, jääkülm*
jäinen, jäätävä, hyinen, hyytävä
▪ kõrvetava päikese käes maitses jäine kokteil lausa taevalikult paahtavassa auringossa jäinen cocktail maistui suorastaan taivaalliselta
▪ jäist tuult trotsides rühiti edasi tarvottiin eteenpäin hyytävää tuulta uhmaten
■ *(tunnete, psüühilise sfääri kohta:) äärmiselt ükskõikne, kalk*
jäätävä KUV., hyytävä KUV.
▪ külaline võeti vastu jäise viisakusega vieras otettiin vastaan jäätävän kohteliaasti

Figur 3: En fungerande ordning på betydelsegrupperna (utdrag ur artikeln **jäine**).

Uppslagsordet **hiilgus** representerar ett motsatt fall. Ordet har ursprungligen betytt 'glans' i konkret betydelse, men denna betydelse har hamnat sist (se figur 4). Den bildliga betydelsen 'prakt' an-

ges först. Särskilt med tanke på att ordboken i fråga är tvåspråkig, skulle det vara användarvänligt att betydelsegrupperna ordnades traditionellt. Att utgå från korpusen är förstås väldigt konsekvent, men användaren har ingen aning om det eftersom man inte heller nämner denna princip i förordet.

Frekvensordningen har motiverats bl.a. med att användarna oftast antas söka just den vanligaste betydelsen, och att de snabbt vill hitta det de söker. Den frekvensbaserade ordningen snabbar visserligen upp sökandet av enskilda betydelser (vilket enligt kritikerna är det enda som denna slags ordning duger till). Däremot säger frekvensen ingenting om semantiska länkar och mönster. Den frekvensbaserade ordningen gör det svårare för användaren att bekanta sig med och lära sig betydelserna och få en sammanhängande uppfattning om betydelsen. Man borde hellre sträva efter att skapa en sådan struktur som hjälper språkinlärare och andra ordboksanvändare att förstå att betydelserna ofta är förknippade med varandra. Forskningen i språkinlärning visar att ordens betydelser är lättare att begripa, lära sig och komma ihåg när de kognitiva sambanden mellan betydelserna framhävs tydligt. (Meer 2000:420, 2002:510–515 och 2004:808; Kamppi 2014:149)

hiilgus M ‹s›
1 *väline sära, toredus, uhkus või kuulsus*
loisto, loistokkuus, hohto, hohdokkuus
kuuluisuus
- linna kunagisest hiilgusest on vaid varemed järel kaupungin entisestä loistosta on vain rauniot jäljellä
- kuulus lauljatar suri oma hiilguse tipul tunnettu laulajatar kuoli ollessaan kuuluisuutensa huipulla

2 *tugev ühtlane helendus, valguskiirgus või sära*
hohde, hohto, loiste, paiste
- keskpäevase päikese pimestav hiilgus keskipäivän auringon sokaiseva loiste

Figur 4: En problematisk ordning på betydelsegrupperna (utdrag ur artikeln **hiilgus**).

6.2.2. Problematiska exempel och exempel med för lite kontext

Vissa exempel kan bli mindre lyckade bland annat på grund av tempus eller brist på kontext. Det sammansatta verbet **alla heitma** betyder bl.a. 'underordna; kuva, förtrycka'. I artikeln fanns det ursprungligen ett exempel som löd 'esterna kuvas genom att röva och beskatta', d.v.s. i presens. Det verkade underligt med tanke på att Estland är ett modernt europeiskt land och att uppslagsordets betydelse är enbart konkret, inte bildlig. Till all lycka gick det att hitta satsen på nätet. Kontexten var en lång text om Estlands historia, som hade skrivits i presens och där den aktuella satsen ingår i ett avsnitt som skildrar 1200-talet. För att göra satsen förståelig ändrades tempuset till imperfekt: 'esterna kuvades genom att röva och beskatta'. Texten hade skrivits 1922 och lagts ut på nätet 2009, vilket visar att också äldre texter kan förekomma sporadiskt i korpusen.

7. Enspråkig ordbok som bas

Den nya estnisk-finska ordboken baserar sig som sagt på den nya estniska definitionsordboken. Definitionsordboken gjordes först – fast den publicerades lite senare, 14 februari 2019 – och i slutfasen redigerades de båda ordböckerna parallelt. Den nya estniska definitionsordboken har publicerats endast som en del i en ny portal som samlar ihop flera resurser. Såvitt vi vet är det sällsynt att redigera en tvåspråkig ordbok med en enspråkig som bas. I vårt fall tog man den enspråkiga ordbokens databas och lade till fält för finska ordförklaringar, ekvivalenter och översättning av exempel. Man tog också bort några fält, t.ex. fältet för det estniska uppslagsordets etymologi.

7.1. Fördelar

Att använda en enspråkig ordbok som bas för en tvåspråkig ordbok har en del goda sidor. Det gör redaktionsarbetet snabbare eftersom basen är färdig och omfattande. Det finns över 100 000 uppslagsord i den estniska definitionsordboken och vi hade i princip möjligheten att göra en lika stor tvåspråkig ordbok. Uppslagsordets betydelsegrupper har var sin ordförklaring, vilket inte är vanligt i tvåspråkiga ordböcker. Ordförklaringarna underlättar redaktionsarbetet och kan också hjälpa ordbokens användare. Förklaringarna var till stor nytta när vi diskuterade några krångliga ord med den estniska redaktionen och tidvis kunde vi även skriva om förklaringarna i samarbete med esterna.

7.2. Problem

Uppslagsordets betydelser har delats i grupper med tanke på den enspråkiga definitionsordbokens behov, och därför kan det finnas antingen för många eller för få betydelsegrupper för den tvåspråkiga ordboken. Exemplen är inte heller valda för en tvåspråkig ordbok. Allt detta kräver uppmärksamhet och bearbetningen tar tid.

7.2.1. Uppdelning

Artikeln **filosoofia** är ett exempel på för fin uppdelning för den tvåspråkiga ordboken (se figur 5). Det finns fyra betydelsegrupper fast det skulle räcka med färre. Till exempel består artikeln ”filosofi” i den gamla estnisk-finska ordboken endast av uppslagsordet **filosoofia** och den finska ekvivalenten *filosofia*. Enligt Svensén (2000:153) ska grupperingen av betydelser i en tvåspråkig ordbok inte återspegla källspråkets betydelsestruktur. På grund av att den estnisk-finska ordboken baserar sig på den enspråkiga ordboken har det varit svårt att följa denna princip.

filosoofia M ‹s›
teadus, mis uurib maailma ja mõtlemise kõige üldisemaid seaduspärasid
filosofia
■ *mingi koolkonna, teatava autori filosoofiline õpetus*
filosofia
▪ Aristotelese filosoofia Aristoteleen filosofia
■ *mingi valdkonna kõige üldisemad seaduspärad, nende uurimine ja rakendamine*
filosofia
▪ juhtimisfilosoofia johtamisfilosofia
▪ haridusfilosoofia koulutusfilosofia; kasvatusfilosofia
■ *(mingi valdkonna tavade ja kommete kohta)*
(*toimintatavoista, perinteistä tms.*)
filosofia
▪ meie veinitegemise filosoofia on täiesti erinev naabrite omast
viininteon filosofiamme on aivan erilainen kuin naapureilla

Figur 5: För fin uppdelning i betydelsegrupper.

Det följande exemplet handlar också om för fin uppdelning, men i det här fallet har artikelns struktur korrigerats (se figur 6). Den övre bilden är från den enspråkiga ordboken och där ser man att det finns två betydelsegrupper. På den nedre bilden – från den tvåspråkiga ordboken – syns det att grupperna har förenats. Det estniska ordet *jalam* betyder ’fot’ av t.ex. ett berg eller torn. Esterna har uppdelat betydelsen så att den första gruppen består av geografiska former och den andra av sådant som människan har byggt. Det kändes vettigt att förena grupperna, eftersom betydelserna ändå är så nära varandra och ekvivalenten är densamma för båda grupperna.

{F2} **jalam** ‹s› ‹-i 2›
KOHT_LOODUS {OSA} GEOGR mäe, künka vms nõlva alumine osa ja selle lähem ümbrus
▫ *Vulkaani jalamile rajati rahvuspark.*
■ KOHT_HOONE {OSA} samba, torni vm ehitise alumine osa või alus
▫ *Tuletorni jalamil peetud pidulik tseremoonia.*

F2 **jalam** ‹s›
ab GEOGR mäe, künka vms nõlva alumine osa ja selle lähem ümbrus; samba, torni vm ehitise alumine osa või alus
jonkin ympäristöstään kohoavan vierusta, lähiympäristö
juuri [▷ *3. vuoren tms.* jalam ◁]
▪ *vulkaani jalamile rajati rahvuspark* tulivuoren juurelle perustettiin kansallispuisto
▪ *tuletorni jalamil peetud pidulik tseremoonia* majakan juurella pidetty juhlallinen seremonia

Figur 6: För fin uppdelning i betydelsegrupper och uppdelningen efter bearbetning (skärmbild från redigeringsprogrammet EELex).

Ibland är uppdelningen däremot inte tillräckligt fin. Uppslagsordet **algdokument** 'originalhandling' har endast en betydelsegrupp i den enspråkiga ordboken trots att den har två ordförklaringar åtskilda av ett semikolon (se figur 7). På finska har vi inget ord som skulle motsvara den första ordförklaringen. När ekvivalenten saknas ges det en ordförklaring på finska och ekvivalenten ersätts med ett understreck. I det här fallet har den ursprungliga betydelsegruppen dessutom delats upp, eftersom det finns en ekvivalent som motsvarar den andra förklaringen.

{F4} **alg+dokument**
ESE_SEMIO majandustehingu kohta koostatud originaaldokument (nt pangaväljavõte, ostu- või müügiarve, töötasu arvestamise dokumendid); muu originaaldokument
▫ *Raamatupidamise algdokumendid.*

F4 **alg+dokument**
majandustehingu kohta koostatud originaaldokument (nt pangaväljavõte, ostu- või müügiarve, töötasu arvestamise dokumendid) /SBjorklof/
rahankäyttöön liittyvä alkuperäinen asiakirja (esim. tilitosite, lasku, palkanmaksuun liittyvät paperit)

_
■ muu originaaldokument /SBjorklof/
alkuperäinen asiakirja [▷ aito [1*F]: algdokument ◁], alkuperäis+asiakirja

Figur 7: Inte tillräckligt fin uppdelning i betydelsegrupper och uppdelningen efter bearbetning (skärmbild från redigeringsprogrammet EELex).

7.2.2. Exempel inte valda för en tvåspråkig ordbok

Det framgår av både innehållet och mängden av exempel i den enspråkiga ordboken att de inte är valda för en tvåspråkig ordbok. I vissa fall finns det för många exempel, som inte ger någon ny information och som inte är kontrastivt intressanta (se figur 8). Sådana har vi lämnat utan översättning och de syns inte heller i ordboken. I artikeln **horoskoop** 'horoskop' är det enda exemplet som översatts "enligt horoskopet i tidskriften ligger det svartsjuka i luften".

F2 **horoskoop** ‹s›
astroloogiline tabel või skeem saatuse või sündmuste ennustamiseks taevakehade seisu järgi kellegi sünnihetkel; ennustus selle alusel
horoskooppi [▷ horoskoop ◁]
ab▪ *horoskoobi järgi on tema salapahe laiskus* ##
▪ *ajakirja horoskoop ütleb, et õhus on armukadedust* lehden horoskoopin mukaan ilmassa on mustasukkaisuutta
ab▪ *sünnihoroskoop* (ls) ##
ab▪ *isikuhoroskoop* (ls) ##

Figur 8: För många exempel för en tvåspråkig ordbok (skärmbild från redigeringsprogrammet EELex).

I andra fall är exemplen för få. Till exempel beträffande uppslagsordet **poleks** 'skulle inte vara (eller ngt annat verb)' räcker det med endast två exempel i definitionsordboken (se figur 9). Ordet har dock många finska ekvivalenter p.g.a. att det i finskan finns ett böjligt nekningsverb som saknas i estniskan. Av den anledningen behövs det också flera exempel, helst ett i varje person. Därför sökte vi lämpliga exempel i ordbokens databas och kompletterade artikeln med dem (se figur 10).

poleks
VERB_ABSTR • VERB_GRAMMATILINE ei oleks
▫ *Kellel siis muresid poleks.*
▫ *Poleks kunagi uskunud!*

Figur 9: Artikeln **poleks** i den enspråkiga ordboken: för få exempel för en tvåspråkig ordbok (skärmbild från redigeringsprogrammet EELex).

poleks
ei oleks
en olisi yks.1.p,
et olisi yks.2.p,
ei olisi yks.3.p;
emme olisi mon.1.p,
ette olisi mon.2.p,
eivät olisi mon.3.p
▪ *eladeski poleks arvanud, et sellist tööd tegema hakkan* en olisi ikinä uskonut, että alan tehdä tällaista työtä
▪ *siit ma oleks alla sadanud nagu purujutt, kui sa poleks taibanud mind kinni hoida* olisin pudonnut tästä päätä pahkaa, ellet olisi huomannut pitää minusta kiinni
▪ *kellel siis muresid poleks* kenelläpä ei olisi huolia
▪ *matkajad asusid teele varahommikul, sest päevakuumuses poleks nad seda suutnud* retkeilijät lähtivät matkaan aamuvarhaisella, sillä päivän kuumuudessa he eivät olisi siihen kyenneet
▪ *küllap oleksid Balti riigid tahtnud saada iseseisvaks ka juhul, kui Nõukogude Liit poleks olnud diktatuur* eivätköhän Baltian maat olisi halunneet itsenäistyä myös siinä tapauksessa, että Neuvostoliitto ei olisi ollut diktatuuri
▪ *Poleks kunagi uskunud!* Ei olisi uskonut!; En olisi ikinä uskonut!
▪ *kui tark ta ka poleks, kõike ta ikka ei tea* niin viisas kuin hän onkin, kaikkea hän ei kuitenkaan tiedä

Figur 10: Artikeln **poleks** i den tvåspråkiga ordboken: exempel har lagts till (skärmbild från redigeringsprogrammet EELex).

8. Sammanfattning

I denna artikel har vi behandlat några exempel som illustrerar de fördelar och problem som har uppstått när en korpus och en enspråkig ordbok använts som bas för den nya estnisk-finska ordboken. Till korpusens fördelar hör att den är omfattande, innehåller mycket nutidsspråk och gör det möjligt att dela orden i frekvensgrupper. Det som vi anser är problematiskt med korpusen är att betydelsegrupperna har ordnats enligt betydelsernas frekvens i korpusen. Dessutom har betydelserna delats i grupper med tanke på den enspråkiga ordboken. Att exemplen härstammar från korpusen gör arbetet snabbare men skapar samtidigt svårigheter eftersom exemplen är valda enligt den enspråkiga ordbokens behov. Att arbetet blir snabbare är överlag en av de största fördelarna med att använda korpus och enspråkig ordbok som bas. Å andra sidan måste redaktörerna göra en hel del förändringar som de hela tiden ska vara medvetna om. Bearbetningen tar tid, och vissa förändringar förblir tyvärr ogjorda p.g.a. tidsbrist.

Teknologin gör det lättare att producera tvåspråkiga ordböcker, men är fördelarna med att använda en enspråkig ordbok som utgångspunkt så stora att de väger upp nackdelarna? Med tanke på ovannämnda problem borde man se till att sådana här projekt hade tillräckligt med tid och andra resurser.

Litteratur

Ordböcker

Eesti keele sõnaraamat 2019. Margit Langemets m.fl. (red.). Tallinn: Eesti Keele Instituut. <https://sonaveeb.ee> (januari 2019).

Eesti-soome sõnaraamat. Eestiläis-suomalainen sanakirja (1972). Valdek Pall m.fl. (red.). Tallinn: Valgus.

Eesti-soome sõnaraamat. Viro–suomi-sanakirja (2015). Sirje Laidre m.fl. (red.). Tallinn: Valgus.

Kettunen, Lauri (1917): *Virolais-suomalainen sanakirja. Eesti-soome sõnaraamat*. Helsinki: Otava.

Suomen kielen perussanakirja (1990–1994). Risto Haarala m.fl. (red.). Helsinki: Kotimaisten kielten tutkimuskeskus / Valtion painatuskeskus.

Suomi–viro-suursanakirja. Soome-eesti suursõnaraamat (2003). Valdek Pall m.fl. (red.). Tallinn/Helsinki: Eesti Keele Instituut / Kotimaisten kielten keskus.

Suuri suomi–ruotsi-sanakirja. Stora finsk-svenska ordboken (1997). Birgitta Romppanen m.fl. (red.). Porvoo/Helsinki/Juva: WSOY.

Viro–suomi-sanakirja. Eesti-soome sõnaraamat (2019). Verkkosanakirja [webbordbok]. Heikki Hurtta m.fl. (red.). Tallinn/Helsinki: Eesti Keele Instituut / Kotimaisten kielten keskus. <www.eki.ee/dict/efi/> (januari 2019).

Annan litteratur

Kamppi, Nina (2014): Metonyymiset merkitykset sanakirjoissa [recension av Wojciechowska (2012)]. I: *Virittäjä* 118, 148–153.

Meer, Geart van der (2000): Core, subsense and the *New Oxford Dictionary of English* (NODE). On how meanings hang together, and not separately. I: *Euralex 2000 Proceedings*. Stuttgart: Institut für Maschinelle Sprachverarbeitung, Universität Stuttgart, 419–431. <www.euralex.org/elx_proceedings/Euralex2000/> (januari 2019).

Meer, Geart van der (2002): Dictionary entry and access trying to see trees *and* woods. I: *Euralex 2002 Proceedings*. Copenhagen: Center for Sprogteknologi, 509–519. <www.euralex.org/elx_proceedings/Euralex2002/> (januari 2019).

Meer, Geart van der (2004): On defining: polysemy, core meanings and 'great simplicity'. I: *Euralex 2004 Proceedings*. Lorient: Uni-

versité de Bretagne Sud, 807–816. <www.euralex.org/elx_proceedings/Euralex2004/> (januari 2019).

Statistikaamet (2011a): Statistika andmebaas. RL0431: Rahvastik emakeele ja soo järgi, 31. detsember 2011. Tallinn: Statistikaamet. <http://andmebaas.stat.ee/Index.aspx?lang=et&DataSetCode=RL415> (mars 2019).

Statistikaamet (2011b): Statistika andmebaas. RL0438: Rahvastik võõrkeelte oskuse, vanuserühma ja elukoha järgi, 31. detsember 2011. Tallinn: Statistikaamet. <http://andmebaas.stat.ee/Index.aspx?lang=et&DataSetCode=RL415> (mars 2019).

Statistikcentralen (2019): Finlands officiella statistik (FOS): Befolkning efter språk 31.12.[2018]. Helsingfors: Statistikcentralen. <https://www.tilastokeskus.fi/tup/suoluk/suoluk_vaesto_sv.html#befolkningefterspråk31.12> (mars 2019).

Svensén, Bo (2000): *Handbok i lexikografi*. Stockholm: Norstedts Ordbok.

Wojciechowska, Sylwia (2012): *Conceptual metonymy and lexicographic representation*. Frankfurt am Main: Peter Lang.

Sofia Björklöf
doktorand, FM
Helsingfors universitet
PL 24, Unionsgatan 40
FI-00014 Helsingfors universitet
sofia.bjorklof@helsinki.fi

Saarni Laitinen
forskningsassistent, HuK
Helsingfors universitet
PL 24, Unionsgatan 40
FI-00014 Helsingfors universitet
saarni.laitinen@helsinki.fi

Glossarium – Sveo-Gothicum: en underlig ordbok eller ett bidrag till forskningen?

Anna Helga Hannesdóttir

Haquin Spegel's *Glossarium – Sveo-Gothicum* (1712) marks an important step in the history of Swedish lexicography. Liberating his dictionary from the Latin-based tradition, the author himself compiled the lemma list. However, the information provided in the entries is not typical of any specific type of dictionary. In this article, the mono-, bi- and multilingual features of the *Glossarium* are described. This survey leads to the issue of Spegel's aim with his effort. In addition to producing a Swedish dictionary, he also intended to further the knowledge of Swedish in the international science community and to contribute to an ongoing work on a polyglot dictionary: a lexicon harmonicum.

1. Inledning

År 1712 utkom Haquin Spegels *Glossarium – Sveo-Gothicum eller Swensk Ordabook/ Inrättat them til en wällmeent anledning/ som om thet härliga språket willia begynna någon kunskap inhämta*. Den har ömsom ansetts utgöra ett viktigt steg i utvecklingen mot en svensk lexikografi, ömsom betraktats som en underlig och föga användbar ordbok. Det som främst har kritiserats är det egensinniga ordförrådet och de disparata informationskategorier som i olika konstellationer beskriver uppslagsordets egenskaper (se Hannesdóttir 2014 och där anförda referenser).

Här nedan kommer *Glossarium* att beskrivas med avseende på artiklarnas innehåll och struktur, med fokus på drag som illustrerar dess status som enspråkig såväl som två- och även flerspråkig ordbok. Särskilt uppmärksammas presentationen av ekvivalenter

och de språk som dessa i de enskilda fallen hämtats ur. Avslutningsvis relateras urvalet av ekvivalenter till ett av de många olika syften som presenteras i ordbokens förord. Bland annat ville Spegel bidra till den internationella etymologiska språkforskningen genom att lägga fram en beskrivning av det svenska språket med fokus på det samtida och utökade svenska ordförrådet och dess relationer till andra språk (se Hannesdóttir 2007:134–137). Möjligen såg han också sitt inlägg som ett bidrag till det stora, flerspråkiga ordboksprojekt som Olof Rudbeck d.y. då arbetade med: ett polyglottlexikon med fokus på språkens släktskap, ett s.k. lexicon harmonicum.

Här närmast följer en kort översikt över den äldre svenska lexikografin och de tidiga ordböckerna; dels för att ge en – visserligen avlägsen – bakgrund till senare tiders två- och flerspråkiga lexikografi, dels för att sätta in *Glossarium* i ett lexikografihistoriskt sammanhang.

2. Den tidiga lexikografin och föregångare till *Glossarium*

Det kan vara på sin plats att precisera vad som här räknas som ordbok eller ordboksarbete. Att såväl lexikografisk teori som lexikografisk praxis kan förankras i aristotelisk kunskapsteori är välkänt (Grønvik 2016:31 och där anförd litteratur). Det är inte därmed sagt att de tidiga ordboksarbetena utmärks av en kontinuerlig utveckling sedan antiken. Snarare krävs det en både generös och pragmatisk syn på den lexikografiska genren för att också de tidiga verken ska kunna betraktas som ordböcker. En i sammanhanget optimal definition har formulerats av Einar Haugen:

> A dictionary may be broadly defined as *a systematic list of lexemes, i.e. words or wordlike units, plus whatever informa-*

> *tion the author chooses to encode with each lexeme*. (Haugen 1984:1, hans kursiv; se även Hannesdóttir 1998:8)

Spegel var inte den förste som åstadkom en svensk ordbok. Med Haugens vida definition som utgångspunkt kan antalet föregångare till *Glossarium* räknas till drygt 30 ordboksarbeten av högst varierande slag.

Kännetecknande för den tidiga svenska lexikografin är att den är starkt präglad av den kontinentala, latinbaserade lexikografin. Några av föregångarna till *Glossarium* är tvåspråkiga och innehåller bara svenska och latin. De flesta kan dock betraktas som flerspråkiga eftersom de förutom dessa två även innehåller ytterligare något språk: grekiska, finska, tyska eller franska. I bara fyra av föregångarna saknas latinet helt.

Så gott som alla ordböcker som gavs ut under 1500- och 1600-talet bygger direkt eller indirekt på kontinentala förlagor. Relationerna mellan de svenska ordböckerna och deras respektive förlagor och förebilder är nu relativt väl kartlagda (se exempelvis Johansson 1997). Detta gäller bl.a. den stora, alfabetiskt ordnade *Lexicon Lincopense* (1640) av Jonas Petri Gothus. I ordbokens huvuddel förses de latinska uppslagsorden med svenska och tyska motsvarigheter. Denna del följs först av ett svenskt–latinskt och därefter ett tyskt–latinskt index. Medan den latinska källspråksdelen har kontinentala föregångare (Johansson 1997:121–122) är lemmalistan i det svensk–latinska indexet av allt att döma sammanställd av Gothus själv (Johansson 1997:56). Den är dock sekundär i förhållande till det latinska ordförrådet i den latin–svenska delen på så vis att den bygger på de ekvivalenter som anges till de latinska uppslagsorden. Det svenska indexet kom att bilda utgångspunkt för lemmalistan i flera av de svenska arbeten som utkom senare under 1600-talet.

Sådant var läget 1712, när *Glossarium – Sveo-Gothicum eller Swensk Ordabook* publicerades.

3. Innehållet i Spegels *Glossarium*

Det som främst har anförts som Spegels bidrag till lexikografihistorien är att hans *Glossarium* bryter med den ditintills latindominerade lexikografin. Framför allt utgör ordbokens ordförråd en viktig lexikografisk insats. Till skillnad från förhållandet i tidigare ordböcker där ordförrådet vanligen, direkt eller indirekt, har utgått ifrån kontinentala föregångare, är ordförrådet i *Glossarium* sammansatt av Spegel själv.

Glossarium innehåller en opaginerad inledning,"Företaal", på drygt 30 sidor, skriven på omväxlande svenska och latin. Ordboksdelen består av ca 5 000 artiklar. Förutom den artikelinledande enheten innehåller den ytterligare drygt 7 000 ord och fraser som behandlas. Makrostrukturen är (någorlunda) alfabetisk, men flertalet artiklar består av nästen som innehåller ord som i sin tur också kommenteras: sammansättningar där uppslagsordet utgör efterled, prefixavledningar, närsynonymer och fraser som i många fall bryter makrostrukturens alfabetiska ordning (se exempel 4 och 7 nedan).

Redan ordbokens titel, *Glossarium – Sveo-Gothicum*, ger en fingervisning om det presenterade ordförrådets ramar. Spegel avser att behandla det "svensk-götiska språket", dvs. den samtida svenskan förstärkt med ord från det medeltida språket och också från den av skandinaverna då relativt nyligen upptäckta isländskan. Ordförrådet kommenterar han i inledningen och han redovisar vilka ord han har för avsikt att beskriva, nämligen inte bara "alla Swenska ord/ som antingen nu i bruk äre/ eller wore tå wår Helga Bibel uttolckades och språket icke war så mycket med främmande ord förblandat" utan dessutom "the gamla Göthiska orden/ som finnas i Sagurna eller Sweriges åtskillige Lagböcker och äre med stoor flit sammanhämtade och förklarade af Hr. Verelio uthi hans Lexico Scandico" (Spegel 1712:[xxiv]).

Med den ”Helga Bibel” avser han den översättning av Bibeln som var färdig 1541 och med de gamla götiska orden syftar han på det medeltida nordiska ordförrådet, hämtat dels i de svenska och norska medeltidslagarna, dels i svenska utgåvor av den isländska medeltidslitteraturen.

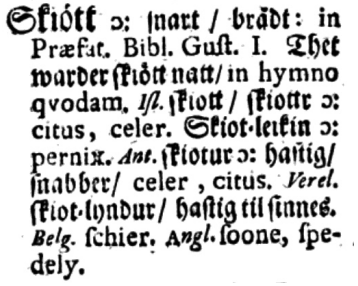
4. Artiklarna i Spegels *Glossarium*

Här nedan illustreras de drag i *Glossarium* som kännetecknar en-, två- respektive flerspråkig lexikografi. De avsnitt som står i fokus utgörs av slutet i bokstavskapitlet *N*, sidorna 320–325 (*niugg–nytta*), och början av *O*, sidorna 325–335 (*o-achtat–öfwer*), samt sidorna 420–435 (*skilling–skuld*), avsnitt som kan anses vara representativa för *Glossarium* beträffande artiklarnas innehåll och struktur. Sammanlagt upptar dessa 30 sidor 222 artikelingångar. De flesta av de enheter som inleder artiklarna är ettordsenheter, ibland kompletterade med en eller flera alternativformer såsom ”Nytr eller nyter”, ”Skipa/skippa vel skipta” eller med en närsynonym, ”Ogäfwer/odugelig”, ”Ordan eller orden”.[1] Även flerordsenheter eller fraser förekommer som artikelingångar: ”Ocker och fördeel”, ”O-uttydeliga tungo”. Presentationsformen är vanligen grundform, men även böjda former förekommer. Det lexikografiska metaspråket är som regel latin: *item* (’även’), *nunc* (’nu’), *vid.* (’se’), *vel* (’eller’) etc.

1 Användningen av typografiska tecken som skiljetecken avviker i de tidiga 1700-talstexterna från det som nu är brukligt. I stället för kommatecken används ofta snedstreck (/) och jämte divis används tankstreck eller likhetstecken utan någon uppenbar systematik. Vidare varierar det huruvida skiljetecknen föregås eller åtföljs av blanksteg. Här har återgivningen av de förekommande skiljetecknen standardiserats till de konventioner som gäller idag.

4.1. En någorlunda typisk artikel i *Glossarium*

Innan ordbokens en-, två- och flerspråkiga drag redovisas, återges här en artikel som får tjäna som illustration till hur en artikel i *Glossarium* kan vara uppbyggd. Även om det är vanskligt att utse en allt igenom representativ artikel i ordboken innehåller artikeln *Skiött* (figur 1) tillräckligt många av de drag som förekommer i ordbokens artiklar för att kunna räknas som en sådan.


Skiött ɔ: ſnart / brådt: in Præfat. Bibl. Guſt. I. Thet warder ſkiött natt/ in hymno qvodam. *Iſl.* ſkiott / ſkiottr ɔ: citus, celer. Skiot-leikin ɔ: pernix. *Ant.* ſkiotur ɔ: haſtig/ ſnabber/ celer, citus. *Verel.* ſkiot-lyndur/ haſtig til ſinnes. *Belg.* ſchier. *Angl.* ſoone, ſpedely.

Skiött ɔ: snart/ brådt: *in* Præfat. Bibl. Gust. I. Thet warder skiött natt/ in hymno qvodam. *Isl.* skiott/ skiottr ɔ: citus, celer. Skiot-leikin ɔ: pernix. *Ant.* skiotur ɔ: hastig/ snabber/ celer, citus. *Verel.* skiot-lyndur/ hastig til sinnes. *Belg.* schier. *Angl.* soone, spedely.

Figur 1: Artikeln *Skiött* i Spegels *Glossarium* (1712). Bilden visar artikeln i det exemplar som finns tillgängligt i digital form i Google Books.

Uppslagsordet är *Skiött*.[2] Det förses med två närsynonymer, *snart* och *brådt*, följda av en källhänvisning: ordet förekommer i inledningen till Gustav Vasas Bibel. Dessutom ges ett belägg: *Thet warder skiött natt*, taget ur en psalm (jfr ”in hymno qvodam”, dvs. ’såsom det står i psalmen’). Därpå följer en ny enhet i nästet inledd med den diaintegrativa markeringen *Isl.* (en uppgift om ordets språktillhörighet; se Svensén 2004:379). Denna följs av två former av en isländsk motsvarighet till uppslagsordet *skiött*: adverbet *skiott* och adjektivet *skiottr*. Dessa förklaras med två latinska ekvivalenter: *citus* och *celer*. I nästet infogas också den isländska adjektiviska sammansättningen *skiot-leikin*, där den nästinledande enheten

2 Som framgår av figur 1 används i ordboken frakturstil för det germanska språkgodset och antikva för övrigt. Denna distinktion upprätthålls inte här.

skiottr ingår som förled. Som ekvivalent till sammansättningen ges latinets *pernix* ('vig'). Därefter följer ytterligare ett näste som inleds med den diakroniska markeringen *Ant.* (förkortning av en form av lat. *antiquus*, 'forn'). Denna avser formen *skiotur*, där den omisskännligt isländska ändelsen *-ur* tydligt utmärker ordet som isländskt snarare än som allmänt ålderdomligt eller föråldrat. Ordet tillfogas svenska och latinska ekvivalenter: *hastig*, *snabber* respektive *celer*, *citus*. Nästet fortsätter med källangivelsen *Verel.* som syftar på Olof Verelius *Index lingvæ veteris schytho-scandicæ* (1690). Därifrån har ytterligare en representant för det aktuella semantiska fältet hämtats: sammansättningen *skiot-lyndur* med sin svenska parafras "hastig til sinnes".

Artikelns avslutande del inleds med den diaintegrativa markeringen *Belg.* Vilken av de föregående enheterna som det markerade *schier* adresserar framgår inte. Därefter följer ytterligare en diaintegrativ markering, *Angl.*, och de två engelska ekvivalenterna *soone* och *spedely*. Inte heller för dessa framgår vad det är som adresseras, *skiött* eller någon av de andra svenska eller "götiska" enheter som artikeln innehåller. De engelska ekvivalenterna faller dock in i det semantiska fält som avspeglas i artikeln.

Denna artikel uppvisar drag som kännetecknar den enspråkiga lexikografin såväl som sådana som är karakteristiska för den kontrastiva två- eller flerspråkiga lexikografin.

4.2. *Glossarium* – enspråkiga drag

Det ymniga inslaget av förklaringar på svenska i artiklarna i *Glossarium* hör principiellt till den enspråkiga lexikografin. För att disambiguera homonymer användes dock redan i de tidiga två- och flerspråkiga ordböckerna närsynonymer till uppslagsordet eller parafraser av dess betydelse. Så är exempelvis fallet i Lincopensens svenska index, där omkring 8 % av de svenska uppslagsorden förses med "en synonym eller ett ekvivalent parallellord" (Johansson

1997:59). Enligt en tidigare uppskattning förses omkring 40 % av orden i *Glossarium* med något slags information på svenska (Hannesdóttir 1998:160). Denna information kan bestå av en eller flera närsynonymer, diverse språklig eller encyklopedisk information, en definition, parafras eller språkprov. Som framgår av artikeln *Skiött* (figur 1) kan exempelvis ord som hämtats ur Bibeln även förses med hänvisning till bibelstället och information om den där aktuella förekomstens betydelse (ex. 1–2).[3]

(1) **Skåda**. al. skåda ɔ: see. Jag skodde/ Prov. 24, 33. nunc skådde ɔ: jag beskådade. Men Skodde/ Eph. 6, 15. vid. Sko.[4]

(2) **Sköfla** ɔ: plundra/ borttaga/ utblotta. vid. Skiöfla. Til sköfwels/ Esa. 42, 24. al. til spillo ɔ: til förderf.

Exempel (1) inleds med den allmänna betydelsen av *skåda*, dvs. 'se', illustrerat med ett exempel ur Salomos ordspråk 24:33. Preteritumformen *skodde*, som förekommer i Efesierbrevet 6:15, är dock den homografa preteritumformen till verbet *sko*. För att få reda på mer om denna form hänvisas användaren till artikeln *Sko* där såväl formen *skodde* som källhänvisningen återfinns. Exempel (2), *Sköfla*, inleds med närsynonymer till uppslagsordets allmänspråkliga betydelse och en hänvisning till artikeln *Skiöfla*. Därefter kompletteras beskrivningen med frasen *til sköfwels*, belagd i Jesaia 42:24. Artikeln *Skiöfla* i sin tur innehåller förutom närsynonymen *plundra* och frasen *til sköfwels* ytterligare information, såväl ekvivalenter som flera belägg med hänvisningar (se exempel 26 nedan).

Det är inte bara ord som tillhör de kategorier Spegel särskilt lyfter fram i sin inledning, dvs. de bibliska och ålderdomliga, som beskrivs på svenska. Också ord som han markerar som dialektala

3 Artiklarna återges efter Kungliga bibliotekets exemplar av *Glossarium* som finns tillgängligt i digital form i Litteraturbanken.

4 Artikeln föregås av *Skock* och följs av *Skof*. Avsikten har troligen varit att inleda den med "**Skoda**. al. skåda" istället för "**Skåda**. al. skåda".

eller folkliga – eller inte alls – kan förklaras med hjälp av en encyklopedisk utläggning (ex. 3), parafraser och/eller närsynonym(er) (ex. 4–5).

(3) **Orf.** *OstGoth*. kallas thet trä eller skaft/ på hwilket lian blifwer bunden. al. Lie-skaft.
(4) **Skiöfwe**/ *O-Goth*. al. skiäfwe ɔ: skiul.
Wagnskiöfwe ɔ: skiul til at sättia wagnar uti: al. wagnslider.
(5) **Ny-torfwa** ɔ: nyttia/ tarfwa.

Vid polysema ord och homonymer disambigueras de olika betydelserna med svenska definitioner eller parafraser (ex. 6):

(6) **Skrubba** betyder at giöra thet någorlunda slätt/ som ojämt är. *Belg*. schrobben. *Angl*. to rubb, to scrape. Ther af kallas en Skrubhöfwel. *Lat*. radula, runcina.
Ellies brukas thet ord i then mening/ som banna/ hårdt tiltala. *Lat*. increpare, objurgare. Sic dicitur: tu får skrubb.
Skrubba betyder och bergskrefwa/ ther man kan dölja sig uti/ item groop. [...]
Skrubber ɔ: plump/ ohöflig sälle. *Belg*. een schrobber. *Angl*. a scurvie fellow.

Av exempel (4) framgår att även sammansättningar som redovisas i artiklarnas nästen kan förses med definitioner, parafraser och/eller svenska närsynonymer. Detta gäller också de avledningar, fraser och språkprov som inlemmas i nästena (ex. 7–11).

(7) **Nööt** ɔ: bos, jumentum. *Dan*. Nød. *Ant*. Naut; qvæ vox etiam significat rem pretiosam: a niota ɔ: niuta/ frui.
Ök-nööt / *Smol*. ɔ: nööt som kan gå under ook.

Nythia-nöth kallades en oxe/ som hafwer dragit trädeståck i 3 år. item koo/ som hafwer buret 3 kalfwar. **Nööt-achtig** / bovinus.

(8) **Odal** [...] Odalboren ɔ: then som är född til sine fäders fasta egendom eller frelsegods. [...]

(9) **Of** [...] Ofgiri ɔ: åtrå/ stor begärelse. [...]

(10) **Sko.** [...] Kip-skodd ɔ: när skorna äro icke igenbundne eller spände/ utan sittia lösa på foten. [...]

(11) **Skolka** [...] Skolka ur scholan ɔ: göma sig undan/ tå man bör gå i scholan.

Den variation som gäller både för vilka informationskategorier som ingår och för vilken konkret information som ges inom ramen för varje kategori bidrar till intrycket av godtycklig struktur och innehåll. Intrycket förstärks av den avsaknad av konstanta informationskategorier som bygger upp den enskilda artikeln. Den enda konstanta informationskategorin är artikelingången, medan de enheter som uppträder i den kategorin inte utgör en konsistent lexikalisk kategori.

4.3. *Glossarium* – tvåspråkiga drag

Att *Glossarium* innehåller inslag som är karaktäristiska för den enspråkiga ordboken räcker inte för att den ska kvalificera sig som en enspråkig ordbok. Lika lite är den en tvåspråkig ordbok, även om den också innehåller renodlade sådana artiklar. Dessa är i de flesta fall svensk–latinska (ex. 12–16).

(12) **Niupon**. *Lat*. morum rubi.

(13) **Nosa** ɔ: olfacere, indagare, naribus subodorari, a näsa vel noos ɔ: nasus.
Niosna ɔ: inqvirere, olfacere.
Niosnar men ɔ: speculatores.
Niosn ɔ: rumor: item speculatio, proditio.

(14) **Ogannur**/ *ant*. ɔ: inæqualis; alias ojämn.
(15) **Skip**. *Isl.* simil. ɔ: navis, navigium. *G. A.* vid. Skep.
(16) **Skof** ɔ: intervallum.
Tu skof ɔ: twå gånger. *Lat*. bis.
Skofwetals ɔ: interrupte, per intervalla, sensim.

Källspråksenheterna varierar även i denna typ av artiklar. De kan tillhöra någon av de grupper som Spegel särskilt lyfter fram i sin inledning, såsom det ålderdomliga *ogannur* och det angivet isländska *skip*, men även ord som torde ha varit välbekanta på hans tid inkluderas, såsom *niupon* och *nosa*. Att inte heller latinet är konstant som det enda ekvivalentspråket framgår av ex. 17, där franskan får fullgöra den funktionen – om än kompletterad med uppgiften om ordets ursprung i tyskan.

(17) **O-achtat**/ är utan twifwel taget af tyska ordet/ ungeacht.
Gall. Non obstant.

4.4. *Glossarium* – flerspråkiga drag

Också de artiklar som innehåller ekvivalenter på flera språk varierar när det gäller källspråksenhetens art och utformning liksom vilken typ av information som ges om den (ex. 18–23).

(18) **Niugg**. *Dan*. nygg/ niding. *Angl*. niggard. *Gall*. taquin, chiche, sordide, vilain; *ant*. chichemaile. *Lat*. parcus, restrictus, avarus, tenax. *Ital*. scarso.
Niugghet. *Angl*. niggardliness. *Gall*. taquinerie, villaine.
(19) **Niura.** *Dan*. nyre. *Germ*. nieren. *Belg*. de nieren. *Græc*. νεφρός. *Lat*. ren, renes, plur. *Angl*. reins, kidneys. *Gall*. rognons. *Ital*. rognoni. *Hisp*. ronnones.
Niure-siuka. *Lat*. dolor renum. *Græc*. νεφριτις.
Niura-talg; alvina, circa renes.

(20) **Offra**. *Dan*. ofre. *Germ*. opffern. *Angl*. to offer, or sacrifice. *Belg*. offerhandt, ofte sacrificie doen. *Lat*. offerre, sacrificare.
Offer. *Dan*. simil. *Germ*. opfer. *Angl*. offering. *Lat*. oblatio, sacrificium, hostia. [...]

(21) **Ogräs.** *Germ*. unkraut/ lulch / trespen *Dan*. Klinte. *Belg*. harick. *Angl*. tares *Lat*. lolium. *Gall*. juraye; *ant*. leville *Ital*. loglio.

(22) **Sköllja**. *Dan*. skølle. *German*. ausspühlen. *Lat*. eluere.

(23) **Skjna**. *Dan*. skine. *Germ*. scheinen. *Belg*. schijnen. *Angl*. shine. *Lat*. splendere, lucere, coruscare. A.G. scimo ɔ: splendor. *A. S.* Sciman ɔ: fulgere.

Det som är påfallande i ekvivalenträckorna i artiklarna ovan är att de representerade språken varierar. Av de tio ord som beskrivs i dessa sex artiklar ges latinska motsvarigheter till alla utom ett, *niugghet* i artikeln *Niugg* (ex. 18), där ekvivalenterna hör hemma i engelska och franska. Såväl informationskategorierna som ekvivalentspråken och deras inbördes ordning i de enskilda artiklarna varierar utan någon uppenbar systematik. I dessa artiklar hämtas ekvivalenterna i åtta olika språk: belgiska, danska, engelska (*Angl.*), franska (*Gall.*), grekiska, italienska, tyska (*Germ.*) och spanska (*Hisp.*). I ordbokens hela artikelbestånd förekommer också ekvivalenter från arabiska, finska, hebreiska, holländska, isländska, longobardiska, norska, persiska, polska, pommerska, ryska, saxiska, syriska, turkiska, ungerska och möjligen ytterligare andra, samt ord och ordformer som uppges vara fornspråkliga, folkliga och troligen andra varianter av flera av dessa språk. Exakt vilka språk eller varieteter som avses med vissa av beteckningarna är något oklart. Exempelvis förekommer som framgår ovan belgiska, pommerska och longobardiska. Möjligen återgår beteckningarna på bruket i den källa där Spegel har funnit det enskilda belägget.

Artiklarna i *Glossarium* uppvisar således en stor mångfald beträffande format, struktur och innehåll. Också i mikrostrukturen är relationen mellan de enheter som presenteras svår att urskilja: det är i vissa fall vanskligt att avgöra vart den information som ges hör (se genomgången av artikeln *Skiött* ovan).

5. Ekvivalenter och kognater – falska och äkta

Av genomgången ovan förefaller Spegels ordbok inte helt ligga i linje med den utveckling som lexikografin kom att genomgå och som manifesteras i dagens kontrastiva lexikografi.

I *Nordisk lexikografisk ordbok* definieras *ekvivalent* som ”leksikalsk enhet som viser ekvivalens med en bestemt leksikalsk enhet i et annet språk” (NLO 1997:107). Definitionen av *ekvivalens* lyder i sin tur: ”semantisk og funksjonsmessig overensstemmelse mellom ord eller uttrykk i to eller flere språk” (NLO 1997:106). Spegel inskränker inte sin beskrivning av de ord han behandlar till att belysa enbart de semantiska och/eller funktionella relationer som kan etableras mellan det beskrivna ordet och det/de beskrivande. Även ordens formella sida beaktas. Urvalskriterierna för de motsvarigheter som presenteras till varje enskild enhet tyder i många fall även på en strävan efter att redovisa kognater, dvs. formmässiga motsvarigheter i de språk där sådana förefinns. För uppslagsordet *ook* (ex. 24) ges ekvivalenter i de germanska språken samt latin, franska, italienska, spanska och grekiska.

(24) **Ook**. *Germ*. joch. *Belg*. jock. *Angl*. yoak, yoke. *Lat*. jugum. *Gall*. Joug. *Ital*. giogo. *Hisp*. yugo, Græc. ζυγός.

Medan formlikheten mellan de olika målspråksenheterna och det svenska *ook* är tydlig är det för vissa av dem svårt att idag bedö-

ma graden av semantisk och funktionell överensstämmelse med uppslagsordet. Spegel nöjer sig med att rada upp kognater. Också beskrivningen av det dialektala ordet *babbe* (ex. 25, hämtat från en annan del av ordboken) har formen av en räcka av kognater (se även Hannesdóttir 1998:153–154).

(25) **Babbe**/ *Smoland*. & *Bleking*. liten pilt/ barn/ gåße. *Angl*. a babe. *Gall*. poupart, poupon ɔ: petit enfant, mais en burlesque. *Germ*. bub/ bube / betyder på somlige orter en liten gåße/ på andra en skalk eller boof. *Ital*. bambino, bambolo, bambolino: Men Babbo är så mycket som Pappa.

Här tycks den formella likheten vara den primära. Spegel redovisar noggrant vilka såväl semantiska som funktionella inskränkningar som föreligger mellan den svenska enheten och några av de enskilda språkens motsvarigheter. Vid den franska kognaten kompletterar han den diaintegrativa markeringen *Gall.* med en diafasisk precisering: den är burlesk eller folklig. Tyskans kognat är i vissa dialekter en äkta kognat, 'en liten gosse', medan den i andra tyska dialekter snarare är en falsk sådan eftersom motsvarigheten där endast är formell och inte semantisk. De italienska kognaterna *bambino*, *bambolo* och *bambolino* lämnas utan kommentar, medan det formmässigt snarlika *babbo* redovisas som falsk kognat; det betyder inte 'en liten gosse' utan 'pappa'.

I artikeln *Skiöfla* (ex. 26) ges först tyska och latinska ekvivalenter till uppslagsordet, men beskrivningen av formen *sköfwels* mynnar ut i en förklaring av vad danskans falska kognat "*skiøling* eller *skyfling*" betyder.

(26) **Skiöfla** ɔ: plundra. *German*. plündern. *Lat*. deprædari, med fientligit wåld borttaga och intet lemna.
Sköfwels. Til sköfwels/ Esa. 42, 24. Jer. 17, 3 ɔ: til rofs.

> *Luth*. in die rappuse. *V.V*. in direptionem. *Angl*. robbing. *Dan*. Skiøling eller skyfling kallas/ när en mans egendom giörs ofri och gifwes til prjs. vid. Gloss. *Veil*.

Den språkliga mångfald som präglar många av artiklarna i *Glossarium* och som ger karaktär åt ordboken bidrar onekligen till att dess funktion som hjälpmedel för att främja kommunikation kan ifrågasättas. Troligen vägde andra syften minst lika tungt.

6. *Glossarium* – ett bidrag till den internationella forskningen

Som framgår av de artiklar som presenteras ovan utgör *Glossarium* en ypperlig illustration till den mångfald som ryms inom Haugens definition av en ordbok: ”[...] *a systematic list of lexemes, i.e. words or wordlike units, plus whatever information the author chooses to encode with each lexeme*” (se ovan). Den som ger sig i kast med ett så omfattande arbete som det att författa en ordbok kan antas ha en avsikt med arbetet. I den innehållsrika inledningen till *Glossarium* redogör Spegel för ett antal olika syften. Bland annat vill han förklara och återuppväcka bibliska och andra föråldrade ord i svenskan och göra dem lättare att förstå för invånare i de tidigare danska, då nyförvärvade svenska landskapen (Företaal:[xxvi]). Dessa syften uppfylls emellertid inte annat än på ett tämligen ytligt plan (Hannesdóttir 2007, 2014).

En stor del av inledningen ägnar Spegel åt att ge en översikt över den samtida internationella språkforskningen. Hans överblick över forskningsläget imponerar. Sammanlagt hänvisar han till över 200 olika källor, både personer och skrifter. Dessa skrifter och de många skiftande synpunkter som där framförs kommenterar han i vissa fall ingående och ofta kritiskt, huvudsakligen på latin. Frågan om urspråket och den roll som syndafloden

och byggandet av Babels torn spelade i den språksplittring som konstaterades råda i världen hade länge satt ramarna för språkforskningen. Jämförelser av enskilda språks ordförråd i syfte att kartlägga språkens släktskap, dels sinsemellan, dels (och främst) med det hebreiska urspråket, diskuterades ingående i den språkfilosofiskt och etymologiskt dominerade litteraturen. Spegel förhåller sig skeptisk till att lexikaliska likheter mellan språken tillmäts så stor betydelse i frågan om språksläktskap (Hannesdóttir 2007 och där anförd litteratur). Genom att redovisa inte bara det svenska ordförrådets ekvivalenter i en rad språk utan såväl äkta som falska kognater understryker han den arbiträra relationen mellan uttryck och innehåll:

> [...] thet är oß nu aldeles okunnigt/ hwarföre somblige ting heta så eller så; altså är oß och jämwäl omöjeligt at finna rätta roten til hwart och it ord/ fast man än wiste alla Språk i Werldena. (Företaal:[viii])

Förutom att bidra till den etymologiska diskussionen vill Spegel också främja det internationella lärda samhällets kännedom om det svenska språket. Avslutningsvis i sin inledning åberopar han en icke namngiven ”fremmad Scribent” som hävdat att vissa språk skulle vara särskilt lämpade för vissa användningsområden. Hebreiska, grekiska och latin, påstår denne, krävs för att kunna följa Guds och människors lag; spanska, franska, arabiska och ”sclavoniska” behövs för resor och handel samt franska, castilianska och florentinska för välljud och sköna konsters skull (Företaal:[xxix–xxx]). Spegel påpekar att

> [...] om then[ne] Italianske Scribent hade haft kunskap om thet rätta Götha språket/ och förstådt huru gammalt/ reent/ ymnogt/ prydeligit/ Liufligit/ ärbart thet är/ tå hade han utan twifwel icke utelemnat thet samma/ utan snarare satt thet framom the andra. (Företaal:[xxx])

Hade skribenten varit förtrogen med svenskan, eller det svensk-götiska språket, hade han, menar Spegel, inte bara tagit med det i sin kategorisering av språken utan rangordnat det högst av dem alla. Sådan okunnighet har han med sitt *Glossarium* velat avhjälpa. Han fortsätter:

> At jag nu så giordt hafwer/ thet hoppas jag ingen rättsinnig kunna misstyckia/ omskönt jag bekänner gierna/ thet jag icke hafwer så fulkomlig kunskap til at giöra et Lexicon harmonicum [...]. (Företaal:[xxx])

Spegel har satt ”thet rätta Götha språket [...] framom the andra” och genom att anföra motsvarigheter från en lång rad andra språk tillgängliggjort dess ordskatt och redovisat dess relationer till andra språk. Han åberopar den tyske forskaren Daniel Georg Morhof, som i sin översikt över det tyska språket har efterlyst just ett ”Lexicon harmonicum”, dvs. ett regelrätt polyglottlexikon. För att ett sådant ska kunna åstadkommas hävdar Morhof – enligt Spegel – nödvändigheten att

> [...] utsöka gamla Vocabula Gothica, Runica, Anglo-Saxonica, Cimbrica, Gallica, Hispanica, Islandica och jämföra them med thet nu brukeliga Tyska språket/ samt thes Dialecter, jämwäl och med thet Swenska/ Danska/ Norska/ Holländska och Ängelska/ betygandes alle plåga sig med een fåfäng möda/ som wilia utleta språkens ursprung och lijkhet och icke grundeligen förstå alla thessa. (Företaal:[xxx])

Att åstadkomma ett lexicon harmonicum, dvs. ett mångspråkigt, jämförande lexikon med fokus på språkens släktskap, låg under 1600-talet i den internationella språkforskningens frontlinje. I Sverige arbetade Olof Rudbeck d.y. vid 1700-talets början just med

ett sådant lexikon. Spegel ser sitt *Glossarium* som ett bidrag till Rudbecks arbete, som syftade till att påvisa likheten

> [...] icke allenast emellan thet Finska och Lapska språket/ utan och emellan thet Swenska/ Isländska/ Norska/ Danska/ Anglo-Saxiska/ Tyska/ Holländska/ Fransöska/ Italianska/ Spanska/ Sclavoniska/ Grækiska/ Latinska och Ebraiska. &c. (Företaal:[xxx])

Av Rudbecks stora arbete publicerades endast en liten del som provtryck. Resten blev, trots Spegels insats, aldrig utgivet.

7. Avslutning

Av eftervärlden har Spegels *Glossarium* blivit betraktat och bedömt som en ordbok – och då uppfattats som en underlig sådan. Hans avsikt var troligen inte bara – och kanske inte i första hand – att åstadkomma en ordbok, vare sig enspråkig eller två- eller flerspråkig. Bland de många syften som han presenterar för sitt *Glossarium* återfinns visserligen främjandet av kommunikation (främst mellan danskar och svenskar), men också berikandet av det svenska ordförrådet (genom att återinföra gamla och bortglömda ord) och upplysande av det internationella forskarsamhället om det svensk-götiska språket. I stället för praktisk nytta för studier och kommunikation bereds likheter och skillnader i språkens ordförråd en framträdande plats i Spegels arbete.

Inledningen är avsedd som ett bidrag till den internationella, lärda diskussionen om språkens ursprung och släktskap. I första hand verkar Spegel ha vänt sig till svenska användare, då den beskrivande delen av ordbokens titel och långa stycken i inledningen är på svenska. Men med de latinska passagerna och mångfalden i fråga om antal anförda språk i ordboksdelen utgör den även ett in-

lägg i den internationella debatten. Spegel ville sprida kunskapen om svensk-götiskans status som rikt och viktigt språk, väl i paritet med de stora europeiska språken.

Den undersökning som här redovisas syftar i första hand till att beskriva hur Spegels *Glossarium* utgör en enspråkig, tvåspråkig och flerspråkig ordbok. Dessutom har urvalet av ekvivalenter i de två- och flerspråkiga artiklarna relaterats till den diskussion om språkens släktskap som Spegel för i inledningen till ordboken. Fortfarande är dock de obesvarade frågorna kring *Glossarium* många. För troligen var Haquin Spegels projekt mycket mera avancerat än senare tiders forskare utgått ifrån när de har beskrivit *Glossarium – Sveo-Gothicum* som en ordbok bland andra.

Litteratur

Ordböcker

Lexicon Lincopense = Jonas Petri Gothus (1640): *Dictionarium Latino–Sveco–Germanicum*. Lincopiæ: Ch. Güntherus.

NLO = Henning Bergenholtz, Ilse Cantell, Ruth Vatvedt Fjeld, Dag Gundersen, Jón Hilmar Jónsson & Bo Svensén (1997): *Nordisk leksikografisk ordbok*. Oslo: Universitetsforlaget.

Spegel, Haquin (1712): *Glossarium – Sveo-Gothicum eller Swensk Ordabook/ Inrättat them til en wällmeent anledning/ som om thet härliga språket willia begynna någon kunskap inhämta.* Lund: Abrah. Habereger. <https://litteraturbanken.se/forfattare/SpegelH/titlar/GlossariumSveoGothicum/> (september 2019), <https://books.google.se/books?id=u48NAAAAQAAJ&printsec=frontcover&hl=sv&source=gbs_ge_summary_r&cad=0#v=onepage&q&f=false> (september 2019).

Annan litteratur

Grønvik, Oddrun (2016): Vitskaplegheit og samfunnsrelevans for store ordboksverk. I: Asgerd Gudiksen & Henrik Hovmark (red.): *Nordiske studier i leksikografi* 14. København: Nordisk forening for leksikografi, 27–61.

Hannesdóttir, Anna Helga (1998): *Lexikografihistorisk spegel. Den enspråkiga svenska lexikografins utveckling ur den tvåspråkiga*. (Meijerbergs arkiv för svensk ordforskning 23.) Göteborg.

Hannesdóttir, Anna Helga (2007): Först och sist. Haqvin Spegel och ordförrådets förnyelse. I: *LexicoNordica* 14, 125–139.

Hannesdóttir, Anna Helga (2014): Haquin Spegels arbete med orden. I: Johnny Hagberg (red.): *Haquin Spegel – en antologi*. Skara: Skara stiftshistoriska sällskap, 149–178.

Haugen, Einar (1984): Introduction. I: Eva L. Haugen: *A bibliography of Scandinavian dictionaries. With an introduction by Einar Haugen*. White Plains, New York: Kraus International Publications.

Johansson, Monica (1997): *Lexicon Lincopense. En studie i lexikografisk tradition och svenskt språk vid 1600-talets mitt*. (Meijerbergs arkiv för svensk ordforskning 21.) Göteborg.

Litteraturbanken. <https://litteraturbanken.se> (mars 2019).

Svensén, Bo (2004): *Handbok i lexikografi. Ordböcker och ordboksarbete i teori och praktik*. Stockholm: Norstedts Akademiska Förlag.

Anna Helga Hannesdóttir
professor
Institutionen för svenska språket
Göteborgs universitet
Box 200
SE-405 30 Göteborg
anna.hannesdottir@svenska.gu.se

Statistisk modellering og prognosticering af efterspørgslen efter netordlister

Poul Hansen

In this paper, human behaviour in interaction with three specialized bilingual dictionaries between Danish and Swedish on the Web is studied with focus on the recorded activity from 2006 to 2018. A statistical seasonal-trend model is applied on the time series from log files, and challenges and applicability of the method are tested and discussed. The results indicate that seasonality is important to consider and could be useful in forecasting.

1. Indledning

For en ordbogsudvikler er det vigtigt, at en ordbog svarer til brugernes behov og interesser. Tidligere var det salget af ordbøger, som gav en klar indikation om dette. Men med ordbøger liggende ude på nettet er det blevet betydeligt vanskeligere at danne sig et overblik over den aktivitet, der finder sted. Samtidig giver en ordbog på nettet en central og enestående kilde til data, der kan fortælle om, hvordan brugerne anvender ordbogen, f.eks. via besøgstrafik og logfiler. At trænge helt ned i datamængden og forstå baggrunden for tallene er dog langt fra ligetil – og vejen er belagt med faldgruber.

Der er mange gode grunde til, at ordbogsudviklere alligevel bør interessere sig nærmere for den information, som er tilgængelig om aktiviteten på nettet. Ikke mindst fordi det kan give gode værktøjer til at følge strømninger og forandringer relateret til ordbøgernes anvendelse, så man kan blive bedre til at modificere, opdatere og tilføje elementer i henhold til brugernes interesser og behov. Og det vil man naturligvis hellere gøre på basis af objektive

analyser og prognoser end på subjektive vurderinger. Dette vil forudsætte, at ordbogssiden måles og analyseres løbende – og at man benytter sig af mere differentierede parametre end blot det ofte brugte "antal besøg" eller "bruger-sessioner" pr. uge, måned eller år.

I dette arbejde gøres brug af en statistisk metode, *tidsserieanalyse*, til at få en dybere forståelse og mere formaliseret beskrivelse af brugernes adfærd, når de besøger ordbogssider. Interessen er rettet mod dels at forbedre tolkninger af de tidsforandringer og fluktuationer, der ses i besøgstrafikken, dels at udarbejde modeller, som kan beskrive den fremtidige udvikling ved hjælp af allerede tilgængelig information.

1.1. Relateret forskning

En systematisk litteratursøgning efter studier relateret til anvendelse af statistisk tidsserieanalyse af logfilsdata gav relativt få hits, og der blev ikke fundet et eneste eksempel på anvendelse af metoden på trafik til netordbøger. Dog er metoden blevet anvendt sporadisk inden for andre områder til at få en bedre forståelse af de faktorer, der forårsager tidsvariationer i besøgstrafikken til et websted over et længere tidsforløb. Et relativt ofte citeret logfilsstudie, som tester og anvender tidsserieanalyse som prognoseinstrument, er Zhang, Jansen & Spink (2009).

Eksempler på deskriptive fremstillinger af flerårige tidsserier med data fra logfiler (dvs. uden forsøg på at modellere udviklingen med statistisk tidsserieanalyse) er imidlertid lettere at finde i litteraturen, f.eks. Serrano-Cobos (2009), Martinez-Gil & Aldana-Montes (2013), Generous et al. (2014), Park & Lee (2016) og He & Tian (2017).

Der er også eksempler på logfilsstudier, hvor der peges på mere avancerede tidsserieanalyser som en metode, f.eks. Kramár & Bieliková (2014), som argumenterer for, at der bør tages hensyn til sæsonvariationen i analyser af logfilsdata.

1.2. Undersøgelsens formål

Formålet med undersøgelsen er:

1. Verifikation af, at tidsserieanalyse baseret på logfilsdata kan benyttes til at undersøge dynamisk adfærd hos ordsøgere eller brugere af netordbøger. Hvis vi kan identificere en sammenhæng mellem teori og praksis vedrørende udnyttelsen af en ordbog i et tidsperspektiv, kan det give os bedre forståelse af brugernes interaktioner med ordbogen og bedre værktøjer til videreudvikling. I første omgang drejer det sig om at få afklaret, i hvilken udstrækning tidsserieanalyse er en relevant metode i denne sammenhæng.
2. Identifikation af tendenser i den tidsmæssige udvikling over lang tid. Udskillelse af karakteristiske mønstre kan give værdifuld information om tilbagevendende brugeradfærd, der kan være værd at kende til, hvis man f.eks. vil optimere annoncerings- eller informationskampagner.
3. Prognosticering af brugeradfærden ud fra karakteristiske mønstre i tidsserieudviklingen. Med tidsseriemodeller tages skridtet fra en rent deskriptiv beskrivelse af logfilsdata til en modelbaseret beskrivelse. Det giver en mulighed for at beskrive udviklingen bagud i tiden på en formaliseret måde og at se fremad i tiden ved hjælp af beregnede prognosemodeller.

1.3. Hvad er statistisk tidsserieanalyse?

Tidsserieanalyse bruges til at undersøge om tilsyneladende helt vilkårlige fluktuationer kan tilskrives samspil eller interaktion mellem to eller flere bagvedliggende faktorer eller må konstateres at være 100 % uforklarlige og vilkårlige. Dette kan vi nemlig ikke afgøre ud fra, hvad øjet kan se. Vi må ty til sofistikerede beregnin-

ger, hvis vi vil vide svaret. Metoden er især brugt inden for økonomisk videnskab til sæsonkorrigering af data, dvs. når man ønsker at få fjernet sæsonvariation fra en måleserie, så en eventuel trend bedre kan identificeres og studeres. Men metoden har også en anden interessant anvendelse. Hvis det lykkes os at udskille tidsvariationen i separate komponenter, vil det også være muligt at bruge disse parametre til at forudsige det videre forløb.

Normalt gennemføres tidsanalyse med avancerede og ikke særligt lettilgængelige metoder (for en oversigt, se Zhang & Qi 2005). Her er det dog tanken at bruge en forenklet metode, som kan implementeres i Excel, hvilket vil gøre det lettere for de læsere, som kunne have lyst til at teste på eget materiale.

Først og fremmest vil det være nødvendigt, at man sætter sig ind i lidt formel teori, der danner grundlag for og forklarer, hvorfor og hvordan det er muligt at opløse en tilsyneladende vilkårligt varierende tidslinje i helt separate komponenter.

Det centrale i teorien er, at en tidsserie, som består af observerede værdier y_t, har to komponenter, sæsonkomponenten S_t og trenden T_t, som giver de observerede værdier i tidsvariationen, hvis de multipliceres med hinanden, hvilket vi kan udtrykke i denne enkle formel:

$$(1) \qquad y_t = S_t \cdot T_t$$

Tidsserieanalyse går ud på at bruge de observerede værdier y_t til at beregne de ukendte værdier S_t og T_t. Umiddelbart kan det måske lyde som en helt umulig opgave, men det er muligt, fordi den ene parameter, sæsonvariationen, har et tilbagevendende mønster hvert år og derfor følger en typisk kurve, som vi kan tilnærme os via gennemsnitsberegninger. Og når man først har fået fastlagt sæsonvariationen, er det nemt at beregne den anden parameter, trendkomponenten, ved hjælp af formlen.

Det er muligt at gå et skridt videre. Man kan nemlig nedbryde

trendkomponenten i yderligere to komponenter; en forudsigelig komponent formuleret som et matematisk udtryk (bestemt ud fra den gennemsnitlige forandringstendens) og en uforudsigelig komponent (al restvariation, som vi ikke kan forklare, hverken som en sæsonvariation eller en konstant forandringstendens).

Restvariationen er et vigtigt element i analysen, fordi den vil være lille, hvis der er stor forudsigelighed i variationen, og stor, hvis der ikke er noget tydeligt mønster, som karakteriserer variationen. Ved at dividere den såkaldte forklarede (mønsterrelaterede) variation med den totale variation får vi et direkte udtryk for, hvor relevante og tilpasningsbare vores beregnede modeller er for de foreliggende data. Dette kaldes determinationskoefficienten eller forklaringsgraden, R^2, og udtrykkes i procent. I denne undersøgelse bruges denne værdi til at validere og sammenligne forskellige modeller.

Formlen (1) er klassisk inden for tidsserieanalyse og er særligt velegnet til anvendelse på tidsforløb med stigende eller faldende tendenser, hvilket er tilfældet med de data som undersøges i dette arbejde.

For den, der har lyst til selv at prøve metoden i Excel på egne data, anbefales den pædagogiske YouTube-video "Excel – Time Series Forecasting" (Jalayer Academy 2013).

1.4. Datamaterialet

De analyserede data er hentet fra operatøren One.com, som står for webhosting af det websted, hvor det analyserede ordbogsmateriale ligger, og består af de registrerede månedlige aktivitetsdata for tre ordlister i en 13-års periode fra og med januar 2006 til og med oktober 2018 (se afsnit 1.5). Til analysen valgtes parameteren *antal eksponeringer* (antal gange, der er blevet vist et link til hjemmesiden på skærmen i et søgeresultat). En alternativ parameter er *antal besøg* (antal gange, der er klikket ind på hjemmesiden),

som især er interessant for den, som har økonomiske interesser som udgangspunkt. I denne artikel blev antal eksponeringer valgt ud fra det synspunkt, at denne parameter numerisk ligger meget højere (op til omkring 6.000 om måneden). Cifrene for antal klik er betydeligt lavere og kan i højere grad være influeret af bevidst overaktivitet fra enkelte brugere, hvilket kan give et falsk billede af det reelle antal brugere.

Datamaterialet består af webhosterens egen bearbejdning af de såkaldte rå logfiler i form af de sammenstillinger af trafikdata, som webhosteren gør tilgængelig for alle kunder, der har hjemmesider hos dem. Hermed inkluderes også aktivitetsdata fra andre browsere og adgangsveje end f.eks. Google.

I 2014 indførte webhosteren et nyt og mere moderne program til at udføre datasammenstillingerne, hvilket medførte, at data før og efter 2014 ikke blev sammenstillet på eksakt samme måde som tidligere og derfor adskiller sig lidt. Men der er en etårig periode, hvor begge programmer kørte samtidigt, hvilket muliggør kontrol og sammenligning, så tidskurverne for perioden før og efter 2014 evt. kan kobles sammen og overlappe hinanden i overgangsåret i det samme diagram.

Et andet problem er, at der har været havarier hos webhosteren, hvilket har medført uigenkaldelige tab af data. Det er imidlertid sjældent forekommende, ofte enkelte dage. En enkelt gang skete et større havari, hvorved der forsvandt 13 måneders data for den svenske hjemmeside.

1.5. Ordlistematerialet

De tre analyserede ordlister er udviklet af artikelforfatteren med støtte fra Nordplus og er blevet undersøgt og rapporteret i et tidligere arbejde (Hansen 2016). De indgår i et større materiale og foreligger i dansk-svensk version på en svensk hjemmeside (www.danska-svenska.se) og i svensk-dansk version på en dansk hjem-

meside (www.dansk-og-svensk.dk). Man finder normalt specialordlisterne ved at angive søgeord i browserens søgefelt. Ordlisterne er søgemaskineoptimerede hver for sig (nøgleord: svensk + dansk + "specialord") og har normalt høje placeringer i browsersøgninger på grund af optimeringen. Nedenfor følger beskrivelse af de tre ordlister i analysen.

1.5.1. Beskrivelse af ordlisterne

Svensk-dansk bandeordliste omfatter ca. 200 lemmaer og er skabt med henblik på at opdele bandeord i tematiske grupper for at karakterisere bandeords "alvorlighedsgrad". Den var tiltænkt en anvendelse som støttemateriale ved tolkning og oversættelse mellem dansk og svensk. Bandeordene er indsamlet ved "wild-card"-søgninger i chatrooms og lignende fora på Google samt ved gennemgang af relevante artikler i nyhedsmedier og tidsskrifter. Ordlisten er placeret på det danske website (dansk-og-svensk.dk). I ordlisten indgår også supplerende excerpter fra Ljung (1984), Kotsinas (1998), Gänger (2003) og Doggelito & Kotsinas (2004).

Svensk-dansk socialordliste omfatter ca. 750 lemmaer og er en grundordliste, der hovedsageligt er udarbejdet på basis af myndighedstekster, personudredninger, spørgeskemaer og excerperet materiale i forbindelse med et større oversættelsesprojekt inden for socialvidenskab. Ordlisten er placeret på det danske website (dansk-og-svensk.dk). I ordlisten indgår også excerpter fra Kolding & Reintoft (1992).

Dansk-svensk hverdagsordliste omfatter ca. 7.000 lemmaer og er en fortegnelse bestående af en grundordliste, som oprindeligt er konstrueret på basis af bl.a. fortegnelsen i *Svensken i Danmark* (Malmström 1975; med indhentet tilladelse fra udgiveren) samt gennemgang af ordudvalget i ældre fortegnelser, f.eks. over turistord og flere nyere udgivne ordbøger. Dette er suppleret med almensproglige ord noteret i det daglige oversættelsesarbejde samt analyser af logfiler med hyppige søgeord og indkomne spørgsmål

fra hjemmesidens brugere. Ordlisten er placeret på det svenske website (danska-svenska.se).

1.6. Undersøgelsens design og gennemførelse

Undersøgelsen blev gennemført i tre momenter: 1) udarbejdelse af diagrammer til sammenligning af de beregnede kurver for tidsudviklingen med de observerede data (afgrænset til perioden 2006-2014), 2) validering af sæson-trend-modellen, dvs. hvor godt den kombinerede sæson- og trendmodel forklarer tidsvariationen i forhold til to alternative enklere trendmodeller (her benyttes forklaringsgraden R^2 som kriterie), og 3) undersøgelse af prognosticerbarheden, dvs. hvor gode sæson- og trendmodellerne fra 2006-2014 er til at prognosticere den fremtidige udvikling (repræsenteret som observationsperioden 2014-2018).

2. Resultat

I det følgende præsenteres resultatet af undersøgelsens tre momenter: sammenligning af beregnede og observerede tidsserier, validering af den sammensatte model og prognosticering af den fremtidige udvikling ved hjælp af modellen.

2.1. Beregnede og observerede tidsserier 2006-2014

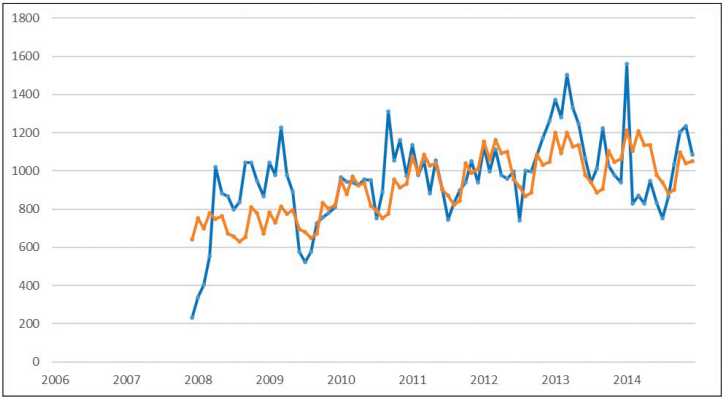
I figur 1-3 præsenteres diagrammer for relationen mellem de oprindelige tidsserier for ordlisterne og de tidsserier, som estimeres af sæson-trend-modellerne. I diagrammerne er de oprindelige tidsserier angivet med blåt, og modellens beregnede værdier angivet med orange.

For svensk-dansk socialordliste kunne modellen ikke beregnes efter midten af 2011 på grund af manglende data. Svensk-dansk

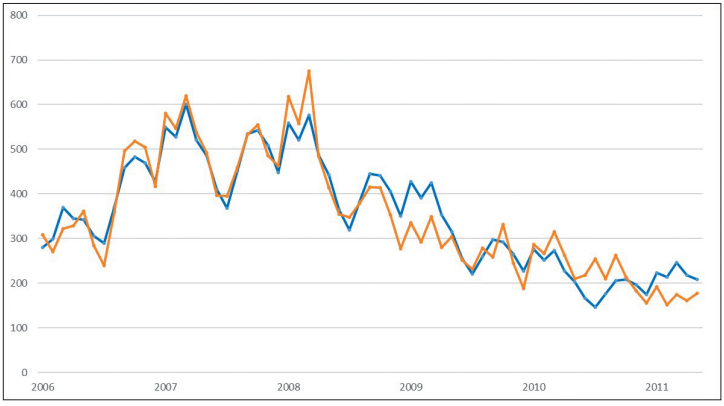
bandeordliste blev lagt ud på nettet senere end de andre (i 2007), så begyndelsen af kurven repræsenterer opstartsforløbet. De andre to ordlister blev etableret allerede i 2005.

Afstanden mellem kurverne i diagrammerne repræsenterer den del af variationen, som modellen ikke forklarer. Det fremgår tydeligt, at der er relativt stor overensstemmelse mellem model og data i alle tre diagrammer. Mindst afvigelse mellem kurverne ses for svensk-dansk socialordliste (figur 2) og dansk-svensk hverdagsordliste (figur 3). Svensk-dansk bandeordliste (figur 1) har lidt større afvigelser.

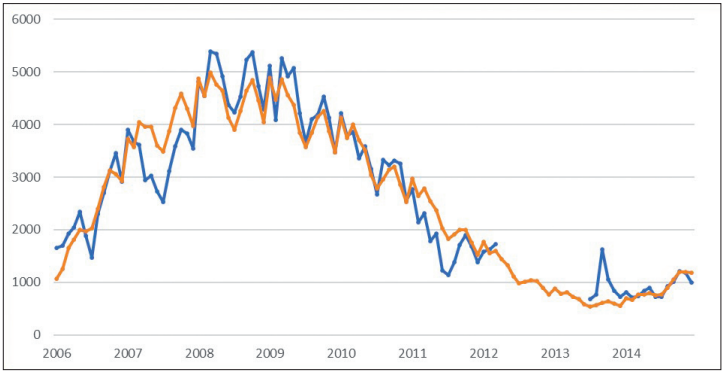
Trendkomponenten i modellerne beskrevs bedst med et såkaldt tredjegradspolynomium, dvs. en model, hvori aktiviteten beskrives med transformeringer af tidsvariablen, dels førstegrad (tidsvariablen uden transformation), dels andengrad (tidsvariablen multipliceret med sig selv to gange) og tredjegrad (tidsvariablen multipliceret med sig selv tre gange). En sådan model gav mindst restvariation i statistiske test (den eksakte formel blev estimeret ved hjælp af statistikmodulet i Excel).



Figur 1: Resultatet for svensk-dansk bandeordliste. Blå linje er tidsserien for antal eksponeringer, orange linje er sæson-trend-modellen, baseret på sæsonvariationen og en polynomisk trend ($y = 0{,}0004x^3 - 0{,}1586x^2 + 15{,}268x + 625{,}15$). Korrelationen mellem linjerne er statistisk signifikant, $R^2 = 34$ %.



Figur 2: Resultatet for svensk-dansk socialordliste. Blå linje er tidsserien for antal eksponeringer, orange linje er sæson-trend-modellen, baseret på sæsonvariationen og en polynomisk trend ($y = 0{,}01x^3 - 1{,}1727x^2 + 34{,}045x + 210.5$). Bemærk, at perioden kun strækker sig til midten af 2011. Korrelationen mellem linjerne er statistisk signifikant, $R^2 = 90$ %.



Figur 3: Resultatet for dansk-svensk hverdagsordliste. Blå linje er tidsserien for antal eksponeringer, orange linje er sæson-trend-modellen, baseret på sæsonvariationen og en polynomisk trend ($y = 0{,}0317x^3 - 5{,}9277x^2 + 276{,}06x + 694{,}06$). Korrelationen mellem linjerne er statistisk signifikant, $R^2 = 88$ %.

2.2. Validering af de beregnede sæson-trend-modeller 2006-2014

Tabel 1 viser en sammenligning af, hvor effektiv sæson-trend-modellen er i forhold til alternative modeller, som ikke tager hensyn til sæsonvariation. Som alternative modeller blev to modeller testet; en af typen lineær model og en af typen tredjegrads polynomisk model.

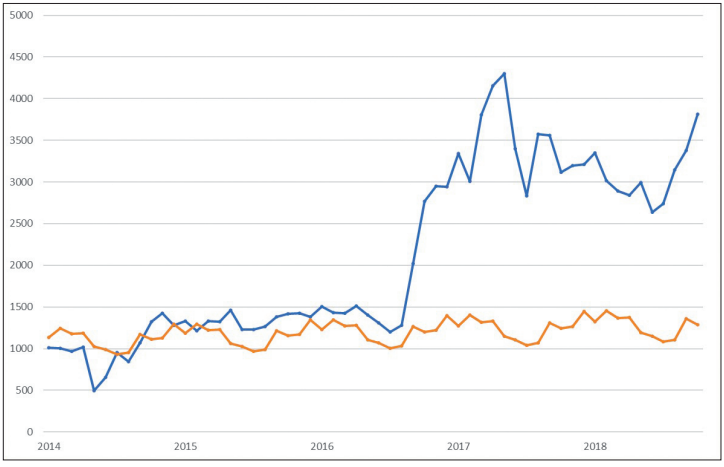
Resultatet viser en klar forbedring af forklaringsgraden ved at inkludere sæsonvariationen. Det er især bemærkelsesværdigt, at forklaringsgraden kan blive så høj som 88-90 %.

Forklaringsgrad R2	Lineær model	Tredjegrads polynomisk model	Sæson-trend-model
Svensk-dansk bandeordliste	22 %	29 %	34 %
Svensk-dansk socialordliste	39 %	76 %	90 %
Dansk-svensk hverdagsordliste	5 %	80 %	88 %

Tabel 1: Sammenligning af sæson-trend-modellen med to modeller uden sæsonvariation, en ren lineær model og en ren tredjegrads polynomisk model, ud fra forklaringsgraden R^2, dvs. hvor gode modellerne er til at reducere den ikke-forklarede variation.

2.3. Prognoser for perioden 2014-2018

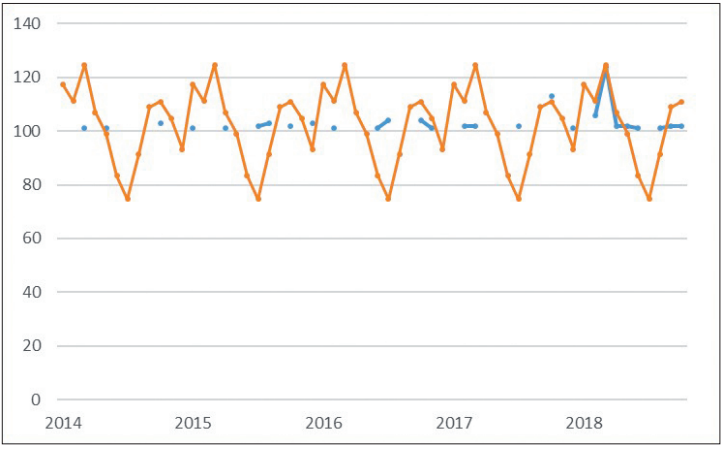
I figur 4-6 testes prognosemodellernes evne til at udføre forudsigelser uden for det tidsinterval, de er definerede i.



Figur 4: Resultatet for svensk-dansk bandeordliste. Blå linje er tidsserien for antal eksponeringer for 2014-2018, orange linje er prognosemodellens forudsigelse fremad i tiden, baseret på data fra 2006-2014.

Prognosemodellen for svensk-dansk bandeordliste blev baseret på den lineære trendmodel ($y = 4{,}3414x + 770{,}55$) og den beregnede gennemsnitlige sæsonvariation. Modellen (figur 4) havde en svagt stigende trend i 2007-2014, og i den prognosticerede periode 2014-2018 var modellens prædiktioner gode frem til midten af 2016, hvor der imidlertid sker en afgørende helt uforudset vækst i efterspørgslen, som får kurven til at lægge sig på et nyt og højere niveau end tidligere. Der ses dog en tendens til, at den nye aktivitetslinje også har en sæsonvariation, som matcher modellens.

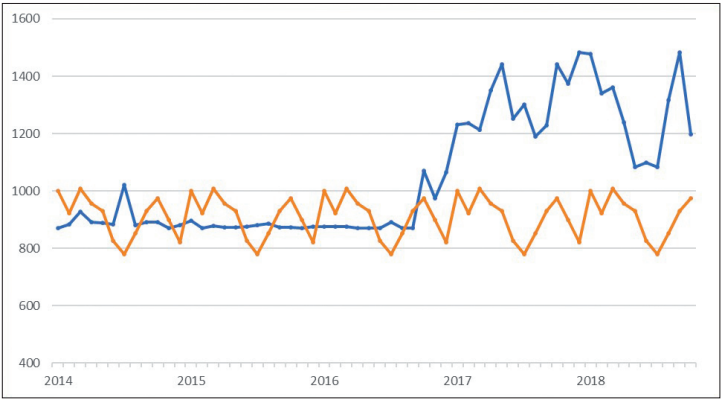
Beregningen af prognosen for svensk-dansk socialordliste blev baseret på gennemsnitstendensen i trenden lige før prognoseperioden og sæsonmodellen (den beregnede gennemsnitlige sæsonvariation). Prognosen i figur 5 forudsiger ganske kraftige fluktuationer i forhold til den registrerede aktivitet. Afvigelserne bør tolkes som en synliggørelse af den beregnede udeblevne brugeraktivitet, dvs. aktiviteten fra brugere er egentlig nede på nul og det, vi har



Figur 5: Resultatet for svensk-dansk socialordliste. Blå punkter/linje er tidsserien for antal eksponeringer for 2014-2018, orange linje er prognosemodellens forudsigelse, baseret på data fra 2006-2014.

registreret, er kun den regelmæssige søgerobotaktivitet. Men mod slutningen af perioden ses en tendens til en svagt stigende aktivitet med en hyppigere forekomst af måneder med registrerede værdier. I denne del af forløbet ses også en vis overensstemmelse med prognosemodellens prædiktioner af høje eller lave værdier.

For svensk-dansk hverdagsordliste (figur 6 på næste side) er prognosemodellen også baseret på gennemsnitstendensen i trenden lige før prognoseperioden og sæsonmodellen (den beregnede gennemsnitlige sæsonvariation). I begyndelsen af prognoseperioden ses tydelige afvigelser mellem prognosen og den registrerede tidsserie, som havde relativt konstante værdier uden større fluktuationer. Efter midten af 2016 ses en vækst i efterspørgslen, som får kurven til at stabilisere sig på et højere niveau i resten af perioden. Samtidig ses en tidsvariation, som i en vis udstrækning følger modellens prædiktion.



Figur 6: Resultatet for dansk-svensk hverdagsordliste. Blå linje er antal eksponeringer for 2014-2018, orange linje er prognosemodellens forudsigelse, baseret på data fra 2006-2014.

Det er naturligvis bemærkelsesværdigt, at der ses så store afvigelser mellem prognoserne og de aktuelle observationer. Men prognosemodellerne spiller her den rolle, at de skaber baggrund for videre analyser. Prognoserne viser f.eks. hvornår og i hvilken udstrækning en ny trend afviger fra tidligere mønstre, og de kan synliggøre graden af udebleven brugeraktivitet, hvis der tidligere var højere aktivitet.

3. Diskussion

Et formål med dette studie var at undersøge, om det er relevant at bruge tidsserieanalyse til at beskrive adfærdsmønstre og efterspørgsel efter ordbøger på nettet. Det kortfattede svar er, at det i det foreliggende materiale ikke blot var forsvarligt at bruge metoden, men den gav også mulighed for at få et mere nuanceret billede af den aktivitet, der kan knyttes til en netordbog, over et længere tidsforløb.

For en ordbogsudvikler eller en ansvarlig for en ordbogsre-

surse på nettet kan tidsserieanalyse være et nyttigt værktøj til at undersøge systematikken og forudsigeligheden i de aktivitetsdata, der er knyttet til en ordliste. Hvad øjet ser som en irregulær varierende linje kan i virkeligheden være et samspil mellem to identificerbare parametre, sæsonvariation og trendvariation. I dette studie viste analysen, at kombinationen af sæson og trend kunne forklare op til 90 % af den samlede variation. Dette giver således et stærkt argument for ikke blot at stole på sit øje ved analyse af tidskurver, men at man altid bør spørge sig selv, i hvilken grad en observeret værdi forklares af tidspunktet på året og den generelle langtidstrend. Hvis man kun analyserer aktiviteten på en ordliste ud fra faste værdier som årsgennemsnit og lignende, går man glip af selve dynamikken i de observerede data.

At det i det hele taget er muligt at adskille sæsonvariation og langtidstrend skyldes, at årstidsvariationen har et regelmæssigt tilbagevendende mønster, som vi kan identificere via gennemsnitsberegninger, når vi har en observationsserie over flere år. Sæsonvariationen er en interessant parameter i sig selv. Man forestiller sig umiddelbart, at kurven skal have en pæn og udglattet form, som svinger op og ned efter årstiderne. Men i de foreliggende data har kurven foruden det forventede hovedforløb også været karakteriseret af typiske afvigelser, der gentages systematisk og tilsyneladende har et relativt stabilt og matchende udseende på tværs af materialet. Det kan også konstateres, at dette formodentlig ikke er unikt for det undersøgte materiale. Lignende modificerede mønstre i sæsonvariation ses også, når man finaflæser de tidsseriekurver af logfilsdata, som er publiceret i litteraturen, f.eks. Serrano-Cobos (2009). Sandsynligvis er det en afspejling af forandret brugeradfærd ved kortere ferier, f.eks. påske og efterårsferie.

Et interessant aspekt vedrørende sæsonvariationen er, at den giver en stærk repræsentation af de aktive brugere. En del af den aktivitet, der registreres for en ordliste, er jo fra robotter, dvs. soft-

ware som automatisk afsøger nettet, og vi kan ikke umiddelbart se på tidslinjen, hvor stor del af trafikken, som skyldes robotter. Men disse robotter søger tilsyneladende helt systematisk uafhængigt af årstiden, hvilket vi bl.a. kan se i figurerne 5 og 6, hvor der forekommer sekvenser af konstante værdier uden årstidsvariation. At separere årstidsvariationen vil i praksis indebære, at robotaktiviteten filtreres bort.

Langtidstrenden repræsenterer, hvordan en ordbogs popularitet eller konkurrenceevne udvikler sig med tiden. I modsætning til sæsonparameteren er den ikke baseret på et systematisk tilbagevendende mønster af delvis kendte årsager (f.eks. klima), men på en generalisering af den konstaterede forandringstendens, hvis mønster vi normalt ikke kender til i forvejen. Parameteren fremtræder mest tydeligt, når der råder samme forudsætninger for søgninger på ethvert tidspunkt, og når der er en generel entydig tendens med hensyn til brugernes antal og aktivitetsniveau. Ændres søgeforudsætningerne, f.eks. ved anvendelse af mere effektive søgeplatforme, påvirker det parameteren. Beregningen af parameteren kan bruges til at indikere forandringer i forhold til tidligere forløb.

Hvis man skal sammenfatte prognoseresultatet, kan det konstateres, at de faktiske data ofte afveg fra det, der kunne forudsiges, men modellerne gav os et vigtigt værktøj til at analysere opkomsten af en ny trend og synliggøre forskelle.

Vi kan bl.a. se, at den udvikling, der er sket fra midten af 2016 til 2018 adskiller sig helt afgørende fra det mønster, som har været kendetegnede for den tidligere udvikling. Det kan hævdes med rimelig sikkerhed, at der er sket noget nyt, som ikke er sket før. Et kig på operatørens statistik over browsingen til ordlisterne viser, at mobiltelefoner blev en mere populær søgeplatform end stationære computere netop i denne periode. Måske har overgangen til søgning via mobiltelefonen medført, at der er kommet mere aktivitet til ordlisterne pga. dels deres oplæg med enkle datastrukturer og

dels ordsøgernes behov for lynhurtige svar og uvillighed til at søge bredere og dybere.

Men det er næppe hele forklaringen. Forandringen har ikke været langsomt glidende, men er sket i løbet af en forholdsvis kort periode, så der bør være noget, som har forstærket udviklingstendensen. Det kan spille rolle, at de nye platforme er mere brugervenlige, mere intelligente og i højere grad tilpasser sig brugernes behov end tidligere samt at brugerne er blevet dygtigere til at finde de rette søgeordskombinationer for at gå direkte til svaret.

4. Sammenfatning

Denne artikel argumenterer for, at man benytter tidsserieanalyse til at undersøge den netaktivitet, der er knyttet til en ordbog i et længere tidsperspektiv. En model af sæsonvariation og generel trend kan forklare en stor del af den tidsvariation, der er knyttet til ordlisterne, og metoden giver os mulighed for at gå langt mere objektivt og rationelt til værks i analyser end ved visuel bedømmelse af tilsyneladende irregulære tidsforløb.

Prognosticering ved hjælp af en model af sæsonvariation og trend er relevant, men forudsætter stabilitet i brugervilkårene. Prognosemodellerne i foreliggende undersøgelse har primært haft den værdi, at de synliggør graden af udebleven brugeraktivitet samt viser hvornår og i hvilken grad en ny trend er opstået.

Litteratur

Ordbøger

Doggelito, Dogge & Ulla-Britt Kotsinas (2004): *Förortsslang*. Stockholm: Norstedts ordbok.

Gänger, Hasse (2003): *Hajaru klyket? Stockholmsk-svensk, svensk-stockholmsk*. Västerås: Sportförlaget i Europa.

Kolding, Hans-Egon & Hanne Reintoft (1992): *Social ABC 1992-93*. København: Aktuelle Bøger.

Kotsinas, Ulla-Britt (1998): *Norstedts svenska slangordbok*. Stockholm: Norstedts ordbok.

Ljung, Magnus (1984): *Om svordomar i svenskan, engelskan och arton andra språk*. Stockholm: Akademilitteratur.

Malmström, Sten (1975): *Svensken i Danmark: en språklig hjälpreda*. 6:e omarb. uppl. Stockholm: Föreningarna Nordens Förbund.

Anden litteratur

Generous, Nicholas, Geoffrey Fairchild, Alina Deshpande, Sara Y Del Valle & Reid Priedhorsky (2014): Global Disease Monitoring and Forecasting with Wikipedia. I: *PLoS Comput Biol 10(11)*. <https://doi:10.1371/journal.pcbi.1003892> (august 2019).

Hansen, Poul (2016): Hvordan har netbrugernes præference for forskellige emnespecifikke ordbøger ændret sig i perioden 2006-2014? I: Asgerd Gudiksen & Henrik Hovmark (red.): *Nordiske studier i leksikografi* 13. København: Nordisk forening for leksikografi, 131-141.

He, Wu & Xin Tian (2017): A Longitudinal Study of User Queries and Browsing Requests in a Case-Based Reasoning Retrieval System. I: *Journal of the Association for Information Science and Technology*, 68/5, 1124-1136.

Jalayer Academy (2013): *Excel - Time Series Forecasting (1-3)*. <https://www.youtube.com/watch?v=gHdYEZA50KE> (august 2019).

Kramár, Tomáš & Mária Bieliková (2014): Context of Seasonality in Web Search. I: *Proceedings of European Conference on Information Retrieval (ECIR 14), LNCS Vol. 8416*, Springer, 644-649.

Martinez-Gil, Jorge & José Aldana-Montes (2013): Semantic similarity measurement using historical google search patterns. I: *Information Systems Frontiers,* 15/3, 399-410.

Park, Minsoo & Tae-Seok Lee (2016): A longitudinal study of information needs and search behaviors in science and technology: A query analysis. I: *The Electronic Library*, 34/1, 83-98. <https://doi.org/10.1108/EL-04-2014-0058> (august 2019).

Serrano-Cobos, Jorge (2009): Combinación de logs internos y externos en la predicción de estacionalidad de búsquedas para el rediseño de webs. I: *El profesional de la información*, 18/1, 11-19.

Zhang, Guoqiang Peter & Min Qi (2005): Neural network forecasting for seasonal and trend time series. I: *European Journal of Operational Research,* 160/2, 501–514.

Zhang, Ying, Bernard Jansen & Amanda Spink (2009): Time series analysis of a Web search engine transaction log. I: *Information Processing & Management*, 45/2, 230-45.

Poul Hansen
fil. dr., translatør
Box 4158
SE-227 22 Lund
poul@danska-svenska.se

ISLEX på finska – redigering, användning och mottagande

Helga Hilmisdóttir

The electronic dictionary ISLEX is a collection of six bilingual dictionaries between Icelandic and other languages in the Nordic region: Swedish, Norwegian Bokmål, Norwegian Nynorsk, Danish, Faroese and Finnish. The main focus of this paper is the Icelandic-Finnish dictionary and its place in the Nordic context. The aim is twofold. First, I discuss some of the challenges of the editing process. As the only non-Indo-European language of ISLEX, Finnish has some syntactic features that call for language-specific solutions. Second, I present the results of a user survey aimed specifically at Finnish speakers.

1. Inledning

ISLEX är en elektronisk ordbok mellan isländska och de övriga statsbärande språken i Norden: svenska, norskt bokmål, nynorska, danska, färöiska och finska. Ordboken består av ca 50 000 isländska uppslagsord, 10 000 idiom och kollokationer, samt 32 000 språkprov som översätts till respektive språk (se Jónsdóttir & Úlfarsdóttir 2011). ISLEX finns gratis på nätet och enligt Google Analytics hade ordboken kring 128 000 användare år 2018 varav närmare 3 000 befann sig i Finland.

ISLEX har från början varit ett internordiskt projekt (se Ljunggren 2011 og Úlfarsdóttir 2013). När den första versionen lanserades i Göteborg 2011 hade ordboken fyra målspråk, svenska, danska, norskt bokmål och nynorska, och den färöiska versionen var redan under arbete. Arbetet med den finska ordboken inleddes sedan i januari 2013 och i mars 2018, när kring 80 % av uppslags-

orden hade översatts till finska, öppnades ISLEX finska version formellt.[1]

Syftet med denna artikel är att diskutera hur väl den isländsk-finska ordboken passar in i den flerspråkiga och nordiska nätordboken ISLEX, dels med tanke på finskans strukturella särdrag som ett icke-indoeuropeiskt språk, dels utifrån användarnas beteende och behov. Jag kommer att dryfta följande frågor: Går det att göra en finsk version av en nordisk ordbok som redan har en fast lexemindelning och färdigvalda exempel? Vilka systematiska problem stöter de finska översättarna på och hur löser de dem? Vilka är det som använder finska ISLEX och hur använder de ordboken? Vad har det för betydelse för användarna att finskan placeras in i ISLEX istället för att skapa en egen nätordbok mellan isländska och finska, utan koppling till de övriga språken i Norden?

Artikeln har följande disposition: I kapitel 2 kommer först en teoretisk diskussion om ISLEX som en digital, flerspråkig ordbok. I kapitel 3 dryftar jag de strukturella problem som den finska gruppen har stött på och i kapitel 4 presenteras en enkätundersökning som fokuserar på användarnas mottagande. Avslutningsvis kommer sedan en kort sammanfattning och diskussion.

2. ISLEX som en digital, flerspråkig ordbok

I en artikel om den norska, svenska och danska versionen av ISLEX konstaterar Rauset, Hannesdóttir & Sigurðardóttir (2011:514) att ordboken inte enbart kan ses som en samling tvåspråkiga och enkelriktade ordböcker. I och med att ISLEX är en digital ordbok uppstår en del nya sökmöjligheter som kan ändra användarens syn på ordboken (se t.ex. Hannesdóttir 2012 och 2015). En ISLEX-an-

1 Tidigare var Järvelä (2008) den enda ordboken som fanns mellan isländska och finska.

vändare kan söka bland de isländska uppslagsorden så som redaktörerna ursprungligen har tänkt (källspråk > målspråk). När användarna matar in sökord på den finska webbsidan får de automatiskt fram målspråket finska men de kan också välja de andra språken genom att klicka på en av flaggorna eller välja ”alla”.

Förutom att gå från källspråk till målspråk kan användarna också gå den andra vägen och söka bland de övriga nordiska språken för att komma åt det isländska uppslagsordet (målspråk > källspråk). De kan till och med söka ett ord på målspråket finska för att via källspråket isländska hitta en ekvivalent på målspråken nynorska eller färöiska (målspråk > källspråk > målspråk). Oavsett ordboksredaktörernas ursprungliga intentioner innebär detta att brukaren har möjlighet att använda ISLEX t.ex. som en finsk-isländsk, finsk-nynorsk eller finsk-färöisk ordbok. Svaren begränsas dock vid de ord som anges som ekvivalenter för de isländska uppslagsorden (se bl.a. diskussion i Hilmisdóttir & Matthíasson 2018).

I detta sammanhang uppstår då frågan om man borde strukturera digitala ordböcker på ett annat sätt än traditionella, tryckta ordböcker som inte tillåter sökningar åt ”fel håll”? Som Hannesdóttir (2015) påpekat leder de digitala ordböckerna till att man inom lexikografin har börjat ifrågasätta grundläggande termer som *källspråk* och *målspråk*. Istället föreslår hon att man talar om *överordnat* och *underordnat språk* och då är det det överordnade språket som avgör vilka ord, idiom, kollokationer och språkprov som finns i ordboken.

I fallet ISLEX har isländskan en tydlig status av ett överordnat språk. Analysen och lexemindelningen av det isländska materialet gjordes från ett enspråkigt, isländskt perspektiv. Tanken var att samma ordboksbas skulle kunna fungera för danska, norskt bokmål, nynorska och svenska, och i framtiden även för andra, icke-nordiska språk (Rauset, Hannesdóttir & Sigurðardóttir 2011:514).

Under den inledande redigeringsprocessen när redaktörerna arbetade med de svenska, norska och danska ordböckerna kunde framställningen av de isländska uppslagsorden ännu justeras. Den semantiska analysen kunde anpassas till målspråkens behov och man kunde byta ut, ändra eller ta bort mindre relevanta exempel. Rauset, Hannesdóttir & Sigurðardóttir (2011:515) beskriver i sin artikel hur detta ibland kunde leda till att den semantiska indelningen av det isländska uppslagsordet blev orimligt detaljerad i norska och danska på grund av semantiska nyanser i svenska. Enligt dem präglas ISLEX just av det att ha uppstått som ett flerspråkigt lexikon.

I och med att ISLEX-ordboken redan hade publicerats och arbetet med de fyra första målspråken var avslutat när den finska gruppen påbörjade sitt arbete var utgångspunkten från första början att ordboksbasen inte skulle justeras och anpassas till det nya målspråket finska (Hilmisdóttir & Martola 2016:151). De finska översättarnas uppgift var att förhålla sig till de uppslagsord (inkl. semantiska indelningar), fraser, fasta uttryck och exempel som redan fanns och detta kunde medföra en del problem som dryftas i följande kapitel.

3. Finska utmaningar inom ISLEX-projektet

Trots att översättningen av ISLEX till finska oftast går smidigt har Hilmisdóttir & Martola (2016) visat att de syntaktiska skillnaderna mellan isländska och finska också kan medföra redaktionella problem. Ibland kan det vara svårt att översätta idiom, kollokationer och språkprov utan att justera eller precisera det semantiska innehållet. I följande tre avsnitt diskuterar jag vanligt förekommande problem där översättarna har varit tvungna att göra vissa kompromisser.

3.1. Personliga pronomen och genus

I nordiska sammanhang är det ett välkänt fenomen att finska språket inte har grammatiskt genus (se Karlsson 2009). Detta gäller även för personliga pronomen där man i skriftspråk och formellt talspråk använder pronomenet *hän* 'hen' för mänskliga referenter. Detta betyder att när det finns tredje personens personliga pronomen i de isländska språkproven försvinner informationen om referentens genus i den finska översättningen: ***henni*** *voru gefnir leðurhanskar* '**hon** fick skinnhandskar' respektive ***hän*** *sai nahkahansikkaat* '**hen** fick skinnhandskar'.

I de flesta fall orsakar avsaknaden av grammatiskt genus i finska inga problem. Det hör till språkets struktur att genus inte alltid är relevant. Ett undantag utgör dock exempel där det förekommer två olika pronomen i samma exempel, ett maskulint och ett feminint: ***hann*** *gaf* ***henni*** *hring* '**han** gav **henne** en ring' som i ISLEX översätt med ***mies*** *antoi* ***naiselle*** *sormuksen* '**mannen** gav **kvinnan** en ring'. I dessa fall har man för att undvika otydligheter p.g.a. upprepning av samma pronomen skrivit ut två substantiv. Pronomenet *hann* 'han' har ersatts med finskans *mies* 'man' och *hún* 'hon' med *nainen* 'kvinna'. Detta betyder att översättarna har valt ett ord med en snävare betydelse. Det isländska pronomenet *hún* har en mycket bredare syftning än substantivet *nainen* 'kvinna' som uttryckligen syftar på en kvinnlig person i en viss ålder.

3.2. Negationer

Som diskuteras i Hilmisdóttir & Martola (2016:149–150) är negationer inom parentes något som kan orsaka problem i de finska översättningarna. I isländska och de övriga språken är det möjligt att ange ett språkprov som innehåller negation och sedan markera negationen med parentes för att visa att den är optionell: *það festir (ekki) snjó* 'snön ligger (inte)'. I finska använder man ett nega-

tionsverb i negativa satser, vilket betyder att man inte kan använda samma enkla framställning som i de övriga språken. I detta exempel har de finska översättarna löst problemet genom att skriva ut två olika översättningar: *lumi* ***pysyy*** *maassa* 'snön ligger' respektive *lumi* ***ei pysy*** *maassa* 'snön ligger inte'.

När översättarna skriver ut två olika satser förväntas användarna förstå vilka exempel har en negativ och positiv betydelse men så enkelt är det tyvärr inte alltid. Det finns också exempel på att negativa uttryck i isländska formuleras utan negation i finska: *fara (ekki) í manngreinarálit* '(inte) göra skillnad på folk och folk'. När frasen har en negativ innebörd i källspråket översätts den med en positiv fras på finska, *pitää kaikkia ihmisiä tasaveroisina*, som ordagrant betyder 'hålla alla människor jämlika'. Med negationsverb i finska får frasen sedan en motsatt betydelse: ***ei pidä*** *kaikkia ihmisiä tasaveroisina* som ordagrant betyder 'inte hålla alla människor jämlika'. För att finska användare ska förstå vilken form de ska använda måste de förstå ordet *manngreinarálit* men i ISLEX erbjuds ingen annan förklaring än just språkprovet som innehåller en optionell negation. Det samma gäller för isländska användare. Eventuellt skulle översättarna kunna skriva in en kommentar på finska och isländska för att förtydliga exemplet för användarna men det är svårt att se hur den skulle kunna formuleras på ett smidigt och entydigt sätt. Enklast hade det varit om ISLEX-redaktionen inte hade valt att skriva in optionella negationer i ordboksbasen.

3.3. Fakultativa led

Som jag nämnde i inledningen innehåller ISLEX kring 10 000 kollokationer och idiom som alla ska översättas till finska. Dessa uttryck innehåller ofta *fakultativa led*, det vill säga led som måste ingå i yttrandet men kan bytas ut mot ett ord eller fras som har en liknande funktion. Led som kan varieras markeras i ordboken med vinkelparentes. Men i och med att finskan har en morfolo-

giskt komplicerad struktur, bl.a. personböjning för verb och kasusböjning för substantiv, kan de fakultativa leden i den finska översättningen påverka även andra delar av yttrandet som inte markeras med vinkelparentes i källspråket.

Frasen *eyrun á <mér>* '<mina> öron' är ett exempel på detta. På svenska och de andra skandinaviska språken översätter man frasen genom att vända på ordföljden: först kommer det possessiva pronomenet *mina* (och markeras med vinkelparentes) och sedan huvudordet *öron*. I den finska översättningen använder man samma ordföljd som i svenska, men man måste tillägga en extra vinkelparentes: den senare parentesen avgränsar det possessiva suffixet *-ni* (1p sg) som förekommer i ordet *korva<ni>* 'öra'. Frasen blir då *<minun> korva<ni>*. I de övriga nordiska språken påverkas huvudorden i genitivkonstruktioner inte av vem det är som är ägaren. Om vi skulle byta ut pronomenet *mina* i det svenska exemplet med t.ex. *dina* skulle det inte påverka huvudordet: *dina öron*. I finskt skriftspråk eller skriftspråksnära talspråk måste *-ni*-ändelsen bytas ut mot *-si:<sinun> korva<si>* 'dina öron'.

I exempel som detta är lösningen enkel: possessivsuffixet markeras med en vinkelparentes precis som alla andra fakultativa led. I vissa fall kan det dock kännas orimligt att markera alla ord och böjningsändelser som påverkas. Detta gäller i synnerhet för uppslagsord som innehåller opersonliga verb (och har subjekt i ackusativ eller dativ). För att illustrera detta kan vi titta närmare på exemplet *<mér> líkar vel við <hana>* 'jag tycker om henne'. Eftersom finska verbet *pitää* 'tycka om' personböjs och därför bör kongruera med subjektet borde alla tre orden markeras med vinkelparentes:[2] **<minä> <pidän> <hänestä>* 'jag tycker om hen'. Men eftersom det rör sig om själva uppslagsordet verkar det mycket underligt att markera verbet som ett fakultativt led.

2 Verbet *pitää* har ett *d* i stammen i 1:a och 2:a person men ett *t* i 3:e person. Därför kan man inte markera bara böjningsändelsen med vinkelparentes.

För en finsk användare är informationen om vilka ord som kongruerar med de fakultativa leden självklar men för en isländning fungerar markeringen som en viktig påminnelse. För att beakta isländska modersmålstalares behov markeras därför possessiva suffix med vinkelparenteser. När det däremot gäller krav på verbkongruens som här ovan anser vi att det blir för komplicerat att avgränsa dessa fall med vinkelparentes, bl.a. på grund av att det ofta rör sig om själva uppslagsorden. Den slutgiltiga versionen av översättningen som exemplifierar bruket av verbet *líka* 'tycka om' blir då som följer: *<minä> pidän <hänestä>* 'jag tycker om hen'.

4. Användning och mottagande av finska ISLEX

Som jag nämnde i inledningen öppnade den finska versionen av ISLEX formellt våren 2018. I slutet av samma år gjordes det en enkätundersökning för att få en bild av vilka som använder ISLEX, hur de använder den och vad de tycker om ordboken.[3] Enkäten var på finska och innehöll tio olika frågor (p.g.a. utrymmesbrist behandlas inte alla frågorna här). Den delades via sociala medier och via länk på ISLEX finska hemsida från slutet av november 2018 till början av februari 2019. Totalt fick vi in 104 svar.

4.1. Informanternas bakgrund

De första frågorna gäller informanternas bakgrund. Vad har de för modersmål och i vilken ålder är de? Tabell 1 visar fördelningen mellan modersmål.

3 Se även Jónsdóttir & Úlfarsdóttirs (2016) enkätundersökning om den danska, färöiska, norska och svenska delen av ISLEX.

språk	N	%
finska	77	74
tvåspråkigt finska	2	2
svenska	16	15
isländska	6	6
övriga språk	3	3
totalt	104	100

Tabell 1: Informanternas modersmål.

Eftersom enkätundersökningen var på finska och låg ute på ISLEX finska webbsida har majoriteten av de som svarar finska som modersmål. Inkluderar man även de som är tvåspråkiga där det ena språket är finska får vi totalt 79 informanter som har modersmålskompetens i finska. Totalt 16 informanter säger att de har svenska som modersmål. Bland informanterna finns sedan 6 personer som har isländska som modersmål och 3 som har en annan språklig bakgrund (ryska och tyska).

Tittar man på ålder ser man att ISLEX används av vuxna människor i alla åldrar. Tabell 2 visar fördelningen mellan åldersgrupper.

åldersgrupp	N	%
yngre än 19	0	0
19–25	46	45
26–41	33	32
42–67	21	20
68–	3	3
totalt	103	100

Tabell 2: Informanternas ålder.

Av de 103 användare som anger sin ålder är det inte en enda som är i gymnasieåldern eller yngre. Nästan hälften av informanterna

anger att de är i åldern 19–25, dvs. i den ålder många avlägger sina universitetsstudier i Finland. Detta stämmer även överens med följande fråga där informanterna fritt kunde formulera en yrkestitel: av de 68 som svarar uppger 31 att de är studerande. Övriga informanter säger att de arbetar bl.a. som lärare på olika nivåer (14), som översättare (5), som kulturarbetare (4) och som diplomater (2). När man sedan frågar varför informanterna använder ISLEX anger de följande svar. Informanterna kunde välja mellan svar som var formulerade på förhand (a–e) eller tillägga en egen formulering (f–g).

	svar	N	%
a)	jobb	13	12,5
b)	studier	64	62
c)	bor på Island	6	6
d)	resor till Island	3	3
e)	hobby	14	13,5
f)	isländska är mitt hemspråk	2	2
g)	av nyfikenhet	1	1
	totalt	103	100

Tabell 3: Orsaker bakom användningen av ISLEX.

En stor del av de användare som svarade på enkäten svarar att orsaken till att de använder ISLEX är att de studerar (64) men andra vanliga svar är att isländskan är ett hobby (14) eller att de behöver ordboken för sitt jobb (13).

Innan jag går vidare till informanternas användning och åsikter om ISLEX kan man fråga sig hur representativa informanterna kan sägas vara med tanke på dem som verkligen använder ordboken och hur generaliserbara resultaten är. Här är det viktigt att komma ihåg att fast isländska inte är ett centralt språk i det finska samhället ingår den som ett valfritt eller obligatoriskt läroämne för dem som läser nordiska språk vid de stora universiteten. Det

är därför inte oväntat att en stor del av informanterna är unga universitetsstuderande som läser nybörjarkurser i isländska. Yrkesgrupper som använder isländska på sitt jobb, t.ex. språklärare (på alla nivåer), översättare, bibliotekarier, diplomater och kulturarbetare, finns också bland informanterna, men av naturliga skäl är dessa kategorier relativt små. I stort sett tycker jag att de 104 informanterna som har svarat på ISLEX-enkäten ganska väl representerar de grupper som använder ISLEX i Finland och bland informanterna hittar man både nybörjare och avancerade språkanvändare.

4.2. Användningen av ISLEX

Syftet med följande tre frågor är att ta reda på hur informanterna använder ISLEX. Hur ofta använder de ordboken och vilka språk är de intresserade av? Den första frågan gäller besökens frekvens och svaren framgår av tabell 4.

hur ofta?	N	%
(nästan) varje dag	11	11
varje vecka	28	27,5
då och då	42	41
ett par gånger per år	21	20,5
totalt	102	100

Tabell 4: Hur ofta informanterna använder ISLEX.

Angående hur ofta man använder ISLEX framgår det av svaren att 39 informanter besöker ordboken varje dag (11 %) eller flera gånger i veckan (27,5 %). Det betyder att en dryg tredjedel av de som svarar är frekventa användare. Den största gruppen, eller 41 % av de som svarar, besöker ordboken ”då och då”, medan 20,5 % svarar att de använder ordboken endast ett par gånger per år (ISLEX lanserades 9 månader tidigare).

Ännu en fråga är *hur* informanterna använder ordboken och vilka språk de matar in i sökrutan. Tabell 5 visar fördelningen mellan olika språk. Informanterna fick ange hur ofta de använde varje språk för sig, dvs. de kunde ge ett svar för varje språk, men de behövde inte ge respons på alla.

	alltid	ofta	ibland	aldrig	svarar inte
isländska	24 23 %	59 57 %	19 18 %	1 1 %	1 1 %
finska	12 11 %	37 36 %	34 33 %	9 9 %	12 11 %
svenska	9 9 %	37 36 %	31 30 %	11 10 %	16 15 %
färöiska	-	-	8 8 %	55 53 %	41 39 %
bokmål	-	1 1 %	19 18 %	44 43 %	40 38 %
nynorska	-	1 1 %	14 13 %	49 47 %	40 38 %
danska	-	2 2 %	18 17 %	46 44 %	38 37 %

Tabell 5: Språken som informanterna matar in i sökrutan.

Som framgår av tabell 5 svarar 80 % av informanterna att de ”alltid” (23 %) eller ”ofta” (57 %) använder isländska som sökspråk medan 47% ”alltid” (11 %) eller ”ofta” (36 %) använder finska. Det betyder att informanterna oftast använder ISLEX för att söka isländska uppslagsord, men en stor andel använder ordboken även åt det andra hållet. Det som kanske förvånar mest är att 45 % av informanterna ”alltid” (9 %) eller ”ofta” (36 %) använder svenska som sökspråk. Här är det värt att komma ihåg att en stor del av informanterna är universitetsstuderanden som läser nordiska språk. Svaren visar dock en tydlig trend: ISLEX-användare ser inte ISLEX som enbart en ordbok från isländska till finska. ISLEX-användarna

söker också mycket gärna på andra språk. En liten del av informanterna anger även att de ”ofta” eller ”ibland” söker ord på färöiska (8 %), norskt bokmål (19 %), nynorska (14 %) och danska (19 %).

När man frågar informanterna på vilket eller vilka språk de vill hitta ord, dvs. vilket är deras målspråk, framträder en delvis annan bild.[4] Tabell 6 visar fördelningen mellan språken.

	alltid	ofta	ibland	aldrig	svarar inte
isländska	45 43 %	39 38 %	13 12 %	1 1 %	6 6 %
finska	35 34 %	40 38 %	16 15 %	4 4 %	9 9 %
svenska	28 27 %	43 41 %	12 12 %	9 9 %	12 11 %
färöiska	1 1 %	2 2 %	11 11 %	55 53 %	35 34 %
bokmål	3 3 %	4 4 %	20 19 %	34 33 %	43 41 %
nynorska	3 3 %	2 2 %	19 18 %	49 47 %	31 30 %
danska	2 2 %	8 8 %	20 19 %	34 33 %	40 38 %
alla	6 6 %	16 15 %	25 24 %	24 23 %	33 32 %

Tabell 6: Informanternas målspråk.

Det som kanske är mest oväntat angående svaren i tabell 6 är att 93 % av informanterna säger att de ”alltid” (43 %), ”ofta” (38 %) eller ”ibland” (12 %) är ute efter en isländsk översättning. Detta betyder att trots att isländskan från en lexikografisk synvinkel inte är ett målspråk i ISLEX så kan man säga att den ur informanternas

4 Den finska formuleringen som används i enkäten kan översättas till svenska på följande sätt: ”När jag använder ISLEX söker jag ... a) översättningar till isländska, b) översättningar till finska” osv.

synvinkel både fungerar som sökspråk och målspråk. Endast en informant säger att hen aldrig söker en isländsk översättning. På den här punkten kan det finnas olika uppfattningar mellan lexikografer och användare. Medan lexikografer gör stor skillnad mellan källspråk och målspråk (eller överordnade och underordnade språk) är det osannolikt att alla användare gör samma distinktion eller tycker att den är viktig. Så länge det inte finns en omvänd ordbok mellan finska och isländska är det viktigare att få förslag på ord där det går. Här borde det dock också nämnas att avsaknaden av en finsk-isländsk ordbok tas upp av tre informanter i en fritt formulerad slutkommentar.

Informanterna svarar också att 87 % av dem tittar på den finska översättningen (”alltid” 34 %, ”ofta” 38 %, ”ibland” 15 %). Av de 4 informanterna som svarar att de aldrig tittar på finskan har bara en person finska som modersmål. Dessa informanter verkar först och främst använda ISLEX som finsk-isländsk ordbok eller som isländsk-svensk.

Om vi tittar närmare på svenska som målspråk så är det något oväntat hur liten skillnad det finns mellan svaren angående svenska och finska: 80 % av informanterna svarar att de ”alltid” (27 %), ”ofta” (41 %) eller ”ibland” (12 %) tittar på den svenska versionen av ISLEX (jämfört med 87 % finska).

Efter de tre stora språken isländska, finska och svenska kommer sedan alternativet att titta på alla målspråken samtidigt: 45 % av informanterna svarar att de ”alltid” (6 %), ”ofta” (15 %) eller ”ibland” (24 %) använder denna funktion. Med tanke på ISLEX bakgrund som ett nordiskt samarbetsprojekt som ska främja den nordiska språkförståelsen är detta ett positivt resultat.

Till sist vill jag påpeka att informanterna i den finska enkätundersökningen anger att de även använder de övriga språken som målspråk. Om man slår ihop de som svarar ”alltid”, ”ofta” eller ”ibland” svarar 14 % att de ser på den färöiska förklaringen, 26 % på bokmål, 23 % på nynorska och 29 % på danska. Tittar man när-

mare på de som använder dessa målspråk ser man att det är ingen homogen grupp: bland dem ser man språkstuderanden och andra användare i olika åldrar.

Om jag försöker dra en slutsats utgående från den bild som framträder i tabell 6 kan jag konstatera att finska ISLEX-användare utnyttjar ordboken på fler sätt än enbart som ett isländskt-finskt lexikon. Svaren indikerar att finskspråkiga finnar och användare med andra modersmål använder ordboken som en multifunktionell ordbok mellan språk i Norden. Siffrorna visar också att det finns ett visst mervärde i att integrera den finska delen i ISLEX istället för att skapa en egen isländsk-finsk nätordbok som också hade varit ett rimligt alternativ, speciellt med tanke på att finskan är ett icke-indoeuropeiskt språk.

4.3. Hur nöjda är användarna med ordboken och vilka är deras önskemål om vidareutveckling?

Följande fyra frågor handlar om informanternas erfarenhet av ISLEX och hur de tycker att man skulle kunna förbättra ordboken. Den sista frågan är öppen och ger informanten en möjlighet att fritt formulera sina egna tankar.

Först tillfrågades informanterna hur ofta de hittar de ord de söker. Tabell 7 visar hur svaren fördelas.

	N	%
alltid	11	10,6 %
ofta	86	82,7 %
ibland	6	5,8 %
sällan	1	1 %
aldrig	0	0 %
totalt	104	100 %

Tabell 7: Sökning som ger träffar.

Som framgår av tabell 7 hittar informanterna rätt ofta det de söker i ISLEX. Av 104 användare säger 97 att deras sökning ”ofta” eller ”alltid” ger utslag. Endast 6 individer svarar att de bara hittar orden ibland och en person säger att det sker sällan. Med tanke på att användarna så ofta söker på målspråken kan man säga att resultatet är oväntat bra.

Den positiva responsen återspeglas även i följande fråga: Vad tycker du om ISLEX? Användarna ombads uppskatta sitt intryck på en femdelad skala, från ”utmärkt” till ”dålig”. Svaren återges i tabell 8.

	N	%
utmärkt	42	40,8 %
bra	51	49,5 %
ganska bra	7	6,8 %
inte så bra	2	1,9 %
dålig	1	1,0 %
totalt	103	100 %

Tabell 8: Utvärdering av ISLEX.

Den största delen av de som svarar på enkäten ger ISLEX betyget ”utmärkt” eller ”bra”, dvs. 93 informanter av 103. Sju informanter anser att ordboken är ”ganska bra” medan tre stycken ger ordboken betyget ”inte så bra” eller ”dålig”.

Om man tittar på de som är missnöjda med ISLEX är det en brokig grupp: två unga språkstuderanden som bor i Finland och en finne i åldern 41–67 bosatt på Island. Den som anser att ordboken är ”dålig” svarar även att hen ”sällan” hittar de ord hen söker (en ung student som går en nybörjarkurs), medan de två andra som är missnöjda anger att de ”ofta” respektive ”ibland” hittar det de söker. Dessa informanter ger inga närmare förklaringar på sitt missnöje i det öppna kommentarfältet så det är svårt att veta vad det är som inte fungerar. Den äldre personen

som är bosatt på Island verkar dock använda ISLEX först och främst som en finsk-isländsk ordbok vilket kan förklara problemet.

Om man tittar på det som andra informanter har skrivit kan man också få en viss bild av vad redaktörerna skulle kunna göra för att förbättra ISLEX. En del av kommentarerna handlar om böjning av isländska ord. Ibland verkar användarna inte riktigt förstå hur ordboken fungerar och vilken information den ger. Som en respons på detta kunde redaktionen fundera på om det går att göra ordboken tydligare för finskspråkiga användare, i synnerhet nybörjare. Eftersom finskan är ett icke-indoeuropeiskt språk är det inte orimligt att finskspråkiga nybörjare behöver mer hjälp för att komma igång med isländska än skandinaver. En annan möjlig lösning är att sökfunktionen för ISLEX utvidgas så att ordboken föreslår ord medan användaren matar in i sökfältet eller att de finska översättarna skriver korta och tydliga användarinstruktioner som riktas särskilt till finskspråkiga nybörjare.

5. Avslutning

I denna artikel har jag dryftat frågor kring den finska versionen av den isländsk-skandinaviska ordboken ISLEX. Syftet var att diskutera anpassningen av den finska delen till de övriga språken i ISLEX, dels med tanke på finskans struktur, dels med tanke på användarnas syn på ordboken.

I avsnitt 3 diskuterade jag själva redigeringen. När arbetet med den finska delen inleddes hade de svenska, norska och danska översättarna redan slutfört sitt arbete och ordbokens isländska del hade redan tagit form. Det är bland annat på grund av detta som de finska översättarna inte har haft samma möjlighet som de andra att be den isländska redaktionen om att justera strukturen på ordboksartiklarna, dvs. indelning i underkategorier och val av

exempel. I artikeln har jag diskuterat tre strukturella problem: 1) genus och personliga pronomen, 2) negationer angivna inom parentes och 3) fakultativa led. Slutsatsen som man kan dra av denna diskussion är att den största utmaningen som översättarna har råkat ut för än så länge gäller idiom, kollokationer och språkprov där den isländska redaktionen anger möjliga variationer med vinkelparentes (=led kan bytas ut) eller med parentes (=led kan strykas). Oftast har arbetet dock gått bra och det har varit en stor fördel för översättarna att ha tillgång till de övriga ISLEX-språken.

I avsnitt 4 har jag redogjort för resultatet av en finsk enkätundersökning av ISLEX. Resultaten visar bl.a. att användarna i stort sett är nöjda med ordboken. De tycker att de ofta hittar de ord de söker. Undersökningen visar även att informanterna inte bara använder isländska som sökspråk utan också finska, svenska och i någon mån de andra språken också. Samtidigt visar undersökningen att användarna också tittar på fler än ett målspråk och ungefär en fjärdedel väljer att titta på alla språken samtidigt. Följaktligen kan man säga att flera av de finska brukarna närmar sig ordboken som ett internordiskt och bidirektionellt lexikon.

Slutsatsen som jag har kommit fram till efter att ha diskuterat finska ISLEX både rent språkligt och ur användarnas synvinkel är att det går att göra en ordbok som inkluderar samtliga statsbärande språk i Norden och att det finns ett tydligt mervärde med att inkludera den finska delen i ett nordiskt sammanhang. En egen ordbok mellan isländska och finska hade visserligen tillåtit mer justeringar i källspråksmaterialet men samtidigt hade användarna då gått miste om den nordiska dimensionen.

Litteratur

Hannesdóttir, Anna Helga (2012): Ekvivalensteorier i tvåspråkig lexikografi. I: *LexicoNordica* 19, 39–58.

Hannesdóttir, Anna Helga (2015): What is a Target Language in an Electronic Dictionary? I: Iztok Kosem, Miloš Jakubíček, Jelena Kallas & Simon Krek (red.): *Electronic lexicography in the 21st century: linking lexical data in the digital age*. Proceedings of the eLex 2015 conference, 11-13 August 2015, Herstmonceux Castle, United Kingdom. Ljubljana/Brighton: Trojina, Institute for Applied Slovene Studies/Lexical Computing Ltd, 236–249.

Hilmisdóttir, Helga & Nina Martola (2016): Flerspråkig lexikografi och obesläktade språk – erfarenheter från finska Islex. I: Asgerd Gudiksen & Henrik Hovmark (red.): *Nordiske Studier i Leksikografi* 13. København: Nordisk forening for leksikografi, 143–154.

Hilmisdóttir, Helga & Marjakaisa Matthíasson (2018): Natur och digitala, tvåspråkiga ordböcker: en studie av finska som sökspråk i ISLEX. I: Ásta Svavarsdóttir, Halldóra Jónsdóttir, Helga Hilmisdóttir & Þórdís Úlfarsdóttir (red.): *Nordiska studier i leksikografi* 14. Reykjavík: Nordisk forening for leksikografi, 120–128.

ISLEX = *ISLEX orðabókin*. Þórdís Úlfarsdóttir (huvudred.). Stofnun Árna Magnússonar í íslenskum fræðum. <www.islex.is> (februari 2019).

Jónsdóttir, Halldóra & Þórdís Úlfarsdóttir (2011): ISLEX – en flersproget nordisk ordbog. I: Birgit Eaker, Lennart Larsson & Anki Mattisson (red.): *Nordiska studier i lexikografi* 11. Lund: Nordisk förening för lexikografi, 353–366.

Jónsdóttir, Halldóra & Þórdís Úlfarsdóttir (2016): Hvem bruger ISLEX og hvordan? I: *LexicoNordica* 23, 89–104.

Järvelä, Tuomas. (2008): *Islanti-suomi-islanti sanakirja*. Helsinki: Yliopistopaino.

Karlsson, Fred (2009): *Finsk grammatik*. Helsinki: Suomalaisen Kirjallisuuden Seura.

Ljunggren Andreas (2011): *Från isländsk-svensk till isländsk-skandinavisk ordbok – några skärvor som kan fogas till en bild av*

vägen fram till ISLEX tillkomst. Rapport om ISLEX-projektet skriven av Sveriges ambassadör i Reykjavik den 14 november 2011. <http://islex.fi/greinar/Ljunggren-nov2011.pdf> (mars 2019).

Rauset, Margunn, Anna Helga Hannesdóttir & Aldís Sigurðardóttir (2011): Ein-, to- eller fleirspråkleg ordbok? I: Birgit Eaker, Lennart Larsson & Anki Mattisson (red.): *Nordiska studier i lexikografi* 11. Lund: Nordisk förening för lexikografi, 512–523.

Úlfarsdóttir, Þórdís (2013): ISLEX – norræn margmála orðabók. I: *Orð og tunga* 15, 41–71.

Helga Hilmisdóttir
forskningsdocent, ph.d.
Árni Magnússon-institutet för isländska studier
Laugavegur 13
IS-101 Reykjavík
helga.hilmisdottir@arnastofnun.is

Google Translate: trussel eller redning for oversættelsesordbøger?

Patrick Leroyer & Henrik Køhler Simonsen

This article discusses the use of Google Translate together with translation dictionaries. We carried out an experiment involving ten test subjects, who were asked to record their post-editing of a Google Translate generated translation. The data show that the test subjects did not get the help they needed from translation dictionaries. In this light, we argue that translation dictionaries should also be designed for post-editing, and we present a function-based model.

1. Indledning

Da det allerførste nummer af *LexicoNordica* udkom i 1994, var temaet "Nordiske ordbøker i samfunnsmessig perspektiv". De enkelte samfund har brug for at overvinde sprogbarrierer for at kommunikere med omverden, og blandt de ti artikler var der én ved Gellerstam (1994), der netop satte fokus på leksikografisk støtte til den interlingvale kommunikation. Gellerstam (1994:51) skrev

> [...] och vilken roll kommer att spelas av lexikonet för nya syften som maskinöversättning, kontrollsystem för grammatik, stil, stavning, avstavning? 'Ordboken' håller på att bli ett alltmer heterogent begrepp.

Allerede dengang blev der udtrykt tvivl om ordbøgerne alene kunne slå til. Sprogteknologien var på fremmarch og Gellerstam (1994:51) udtrykte begrundet tvivl, om hvorvidt ordbogen kunne bevare sin homogene form og funktioner og stå distancen i forhold til teknologien, herunder også maskinoversættelse. Fejl-

behæftede oversættelser var fra starten af et anerkendt fænomen, hvilket Flanagan (1994) påpegede.

De nordiske lande har altid været afhængige af samhandel og har altid haft et udpræget internationalt udsyn og dermed også et stort behov for at kommunikere på andre sprog end de nordiske. Følgelig har den nordiske leksikografi helt fra starten haft en stærk tradition for at udvikle bilingvale ordbøger til og fra de nordiske sprog, som primært bruges til at oversætte med, specielt i faglige sammenhænge.

Det er vigtigt at påpege, at oversættelse i stigende grad erstattes af maskinoversættelse. Vi har derfor interesseret os for, hvilken rolle danske oversættelsesordbøger spiller i den nye virkelighed med oversættelser genereret af Google Translate (herefter refereret til som GT), samt for, hvorvidt der er behov for at opdatere oversættelsesordbøger til det nye GT-paradigme.

Vi har endvidere interesseret os for, hvordan brugere af oversættelsesordbøger mellem dansk og engelsk anvender disse til at efterredigere GT-genererede oversættelser med, og i hvilken grad de får den fornødne leksikografiske hjælp. Tarp (2004) opstiller ganske vist korrektur som en leksikografisk funktion, men vi ønsker at undersøge, om der i realiteten ikke er brug for en efterredigeringsfunktion med en særlig selektion og præsentation af datakategorier, der i højere grad kan hjælpe brugeren med at efterredigere GT-genererede oversættelser.

De spørgsmål, der har styret vores empiriske og teoretiske arbejde, er følgende:

1. Hvordan efterredigerer forsøgspersonerne GT-genererede oversættelser mellem L1 (dansk) og L2 (engelsk)?
2. Hvordan inddrager forsøgspersonerne ordbøger og eksterne tekstressourcer, og med hvilke resultater?
3. I hvilket omfang kan man pege på muligheder for effektivisering af efterredigering?

4. Hvilke fagleksikografiske datakategorier kræves særligt til efterredigering af GT-genererede fagoversættelser?

2. Metode og data

I løbet af efteråret 2018 blev 13 forsøgspersoner rekrutteret på kandidatuddannelsen i International Business Communication på Copenhagen Business School. Forsøget involverede studerende, som i deres studium arbejder med fagsproglig kommunikation. De studerende blev inddelt i to hold. Der var tale om 9 kvinder og 4 mænd mellem 25 og 30 år. Tre forsøg blev sorteret fra pga. mangelfulde skærmoptagelser. Forsøgspersonerne kom fra to forskellige klasser.

Forsøgspersonerne på hold A fik adgang til en kreativ tekst, og forsøgspersonerne på hold B fik adgang til en faglig nyhedstekst. Begge tekster havde et omfang på ca. 1 A4-side og var på dansk. Den kreative tekst stammer fra en blog af den danske influencer Lotte Reimar, jf. Reimar (2018), og den faglige nyhedstekst stammer fra Københavns Lufthavn, jf. Cph.dk (2012). Sværhedsgraden matcher sværhedsgraden af de tekster, som de studerende arbejder med på kandidatstudiet.

Formålet med at inddele forsøgspersonerne i to hold med hver deres type tekst var at undersøge hvilke udfordringer de to tekst-typer indebar. Da både kreativt sprogbrug og faglige tekster stadig volder GT vanskeligheder, jf. bl.a. Bundgaard (2017), har vi også tekstmæssigt kunnet udfordre GT på to forskellige måder. Begge hold blev bedt om først at oversætte den danske tekst til engelsk ved hjælp af GT og derefter at efterredigere oversættelsen med danske oversættelsesordbøger og andre online-ressourcer efter eget valg. Forsøgspersonerne fik oplyst, at de frit kunne bruge og havde adgang til alle de oversættelsesordbøger mellem dansk og engelsk, de plejer at bruge på studiet, samt til eksterne ressourcer

og til Google Translate. Forsøgspersonerne valgte dog kun at anvende ordbogen Vinterberg & Bodelsen (V&B), hvilket formentlig skyldes en vis grad af vanetænkning. V&B ses her som eksempel på konventionelle oversættelsesordbøger mellem dansk og engelsk. Begge hold fik instruktioner om at tænde for deres skærmoptager og optage hele processen, og om at fremsende både skærmoptagelse og den oversatte, efterredigerede tekst.

Forsøget er et kvalitativt studium, som er blevet udviklet til at analysere et specifikt emne med en specifik population i en specifik kontekst, og kvaliteten af dette studium er blevet sikret ved at fokusere på hhv. validitet og reliabilitet, som i øvrigt defineres af Saunders, Lewis & Thornhill (2009:156-157):

> Validity deals with questions like whether or not data can be trusted, and whether or not the subsequent findings actually are related to the research questions and finally whether or not the author has been able to deal with the data in an unbiased and critical way.

samt

> Reliability deals with questions like whether or not the research method used, is used in a consistent and structured way to ensure that what is measured is measured consistently.

Den anvendte metodologi sikrer, at vores data både er valide og relevante, og at de er behandlet konsistent og ensartet.

Resultaterne af et kvalitativt studium kan som sådan ikke generaliseres eller replikeres, men vi vil fremhæve, at de udvalgte studerende er repræsentative for kategorien af danske studerende, der studerer engelsk virksomhedskommunikation på et universitet i Danmark, idet der i deres studier indgår oversættelse med og uden

Google Translate. Forsøget er dermed også relevant og realistisk i forhold til deres studieaktiviteter og undervisning, og vores metodologiske tilgang sikrer både ekstern validitet og reliabilitet.

Forsøget gav i alt ti skærmoptagelser af ca. en times varighed samt ti oversatte, efterredigerede tekster på ca. en side, hvilket udgør vores empiriske grundlag.

3. Oversættelse under opbrud

For at sætte forsøget i relief er det for det første nødvendigt at se på oversættelsesdisciplinen. Begge forfattere har undervist i og arbejdet professionelt med oversættelse af fagsproglige tekster i en række organisationer og virksomheder de sidste 20 år, og vi har set oversættelsesdisciplinen hurtigt ændre karakter. Det er et velkendt faktum, at oversættelsesdisciplinen har været udsat for radikal innovation af nye teknologier, der nedbryder branchens traditionelle forretningsmodeller, hvilket f.eks. Bundgaard (2017) også diskuterer. Selve behovet for oversættelse generelt er faldet, mens de oversættelser, der laves, i stadigt stigende grad gennemføres under anvendelse af GT, ligesom virksomheder og enkeltpersoner bruger GT til at oversætte ikke kun dele af tekster men hele tekster med. Mange oversættelser skal ikke udgives, men bruges internt i virksomheden som informationskilde, og dertil er GT særlig velegnet. Dette har skabt en ny situation for de oversættere, som sidder med GT-genererede oversættelser.

Brugere efterredigerer GT-genererede oversættelser, fordi disse halter på en række punkter, først og fremmest ved oversættelse af termer og deres kollokationer, da GT mangler faglig viden. Hvis GT skal yde sit bedste skal den være koblet op på en termbase i et oversættelseshukommelsessystem som f.eks. Trados, jf. (TRADOS) eller Memsource, jf. (MEMSOURCE). Trados og Memsource er oversættelseshukommelsessystemer med indbyggede term-

baser og både oversættelsessegmenter samt termer er valideret af professionelle oversættere, som kvalitetssikrer segmenterne og termregistreringer vha. deres faglig viden.

Hermed står oversætteren over for nye udfordringer, da der kræves endnu højere sprogekspertise for at gennemskue fejl og mangler, der umiddelbart ikke "springer i øjnene". Spørgsmålet er, om traditionelle oversættelsesordbøger uændret kan bruges ved efterredigering, for opgaven består nu primært i at efterredigere maskinoutput og ikke i at oversætte fra bunden.

Vi argumenterer for, at oversættelsesordbøger med fordel kan udstyres med en efterredigeringsfunktion, som er baseret på nøjagtig de datakategorier, der er nødvendige for at kunne bistå brugerne hurtigst muligt under posteditering. Da GT sparer kostbar tid, specielt ved stor volumen, bør efterredigering kunne ske hurtigt for at optimere tidsgevinsten.

4. Efterredigering ifølge TAUS-retningslinjer

For at sætte forsøget yderligere i relief, er det ligeledes nødvendigt at fastslå, hvad efterredigering er. Efterredigering defineres i ISO 18587 (2017) som

> [...] the process of the analysis and correction of text resulting from an automatic or semiautomatic translation to ensure its compliance with grammar, punctuation, spelling and meaning.

Der findes en del gode teoretiske overvejelser om efterredigering, herunder f.eks. Carl et al. (2011), som diskuterer den proces, som man som oversætter gennemgår i forbindelse med efterredigering. Forfatterne dokumenterer processen ved hjælp af key logging og eye tracking. Et andet bidrag er Koglin & Cunha (2019), som dis-

kuterer resultaterne af et forsøg med maskinoversatte tekster, der er rige på metaforer samt hvilke kognitive og kommunikative udfordringer oversætteren har i denne proces. Endelig sammenligner Carl, Jia & Wang (2019) kvaliteten af maskinoversatte tekster med tekster oversat af mennesker helt fra bunden.

Dernæst er det nødvendigt at diskutere, hvordan GT-genererede tekster systematisk kan efterredigeres i henhold til retningslinjer, der tager højde for den opnåede tidsgevinst, f.eks. TAUS-retningslinjer efter Translation Automation User Society, jf. TAUS (2019). Vi foreslår i det følgende en efterredigeringsfunktion, som tager sit udspring i disse retningslinjer, og som i stor udstrækning sikrer det mest velegnede kvalitetsniveau, så maskinoutput udnyttes mest og bedst muligt. Vi tager udgangspunkt i de to sæt retningslinjer, som benævnes som henholdsvis TAUS 1 og TAUS 2 (2019):

TAUS 1 – Guidelines for achieving "good enough" quality:
Aim for semantically correct translation

- Ensure that no information has been accidentally added or omitted.
- Edit any offensive, inappropriate or culturally unacceptable content.
- Use as much of the raw MT output as possible.
- Basic rules regarding spelling apply.
- No need to implement corrections that are of a stylistic nature only.
- No need to restructure sentences solely to improve the natural flow of the text.

TAUS 2 – Guidelines for achieving quality similar or equal to human translation:
Aim for grammatically, syntactically and semantically correct translation

- Ensure that key terminology is correctly translated and that untranslated terms belong to the client's list of Do Not Translate Terms.
- Ensure that no information has been accidentally added or omitted.
- Edit any offensive, inappropriate or culturally unacceptable content.
- Use as much of the raw MT output as possible.
- Basic rules regarding spelling, punctuation and hyphenation apply.
- Ensure that formatting is correct.

Vi har nu etableret den teoretiske forståelse af de nye vilkår for oversættelse samt efterredigering og kvalitetssikring vha. TAUS, og vi kan gå videre til en diskussion af artiklens forskningsspørgsmål på basis af de empiriske data.

5. Resultater og analyse

Efter afsluttet forsøg analyserede vi skærmoptagelserne samt de tilsvarende oversatte og efterredigerede tekster. Ved at registrere og analysere opslag og indtastninger under efterredigeringen kunne vi identificere empirisk funderede svar på vores forskningsspørgsmål. I det følgende vil vi beskrive og diskutere forsøgspersonernes adfærd under efterredigeringsprocessen og uddrage de vigtigste indsigter, som har ført til udvikling af teorien for den nye funktion.

For hver af forsøgspersonerne oprettede vi en logbog, hvor de enkelte handlinger blev tidsregistreret, beskrevet og analyseret. Det gav os indsigt i deres adfærd under processen. Som det vil fremgå, har forsøgspersonerne skullet tackle lignende udfordringer, hvor der ved efterredigeringen af den kreative tekst var særlige

udfordringer ved enkelte ord og ved talesprogs- og slangudtryk. Ved den faglige nyhedstekst var det de faglige termer, der voldte vanskeligheder.

På det første forskningsspørgsmål om, hvordan forsøgspersonerne efterredigerede de oversatte tekster, kan vi på basis af vores analyser slå fast, at ni ud af ti forsøgspersoner ikke synes at efterredigere i fornødent omfang. Det gør de hverken efter TAUS 1 "good enough" quality, eller TAUS 2 "similar or equal to human translation". Ved analyse af de færdige tekster kunne det konstateres, at kvaliteten svigter. Ingen stavekontrollerede f.eks. den efterredigerede tekst.

På det andet forskningsspørgsmål om, hvordan forsøgspersonerne anvender ordbøger og eksterne tekstressourcer, kan vi på basis af analysen af især skærmoptagelserne slå fast, at ni ud af ti forsøgspersoner hverken søger hjælp i de leksikografiske ressourcer, ej heller implementerer de fundne leksikografiske data i fornødent omfang. Ingen af forsøgspersonerne slår f.eks. op i en monolingval L2-ordbog for at kontrollere betydningen af en tvivlsom ækvivalent.

5.1. Oversættelse af kreativ tekst

I den kreative tekst var der udfordringer med at oversætte og efterredigere ordene "magelighed", "kraftedeme" og "listepik". Dette kan f.eks. ses i følgende eksempel, hvor forsøgsperson 1 ved tidsstempel 01:20 vil efterredigere den af GT foreslåede ukorrekte oversættelse "Hit by Magic" til "Ramt af magelighed!":

L1: Ramt af magelighed!
L2: Hit by Magic!

Eksempeludsnit 1: Forsøgsperson 1.

Forsøgsperson 1 åbner V&B, slår "magelighed" op, kigger forskellige steder i artiklen, men synes ikke at finde det vedkommende har brug for og ændrer ikke den ukorrekte oversættelse af "magelighed". Det er først senere ved tidstempel 06:22, at forsøgsperson 1 ændrer "Magic" til "Comfort".

Et andet eksempel, som illustrerer de udfordringer som forsøgspersonerne har haft med at oversætte og efterredigere almensproglige termer er, når de giver op og bruger nettet i stedet for. Både forsøgsperson 1, 2 og 5 slår først op i V&B og derefter konstaterer de, at det, som de leder efter øjensynlig ikke findes, hvorefter de slår op i andre ressourcer på nettet. Forsøgsperson 2 slår f.eks. "kraftedeme" op i V&B ved tidstempel 05:55 og får intet resultat. Derefter slår forsøgsperson 2 op i GT som anvendes som ordbog og får resultatet "fucking". Heller ikke dette resultat anvendes og forsøgsperson 2 udelader helt ordet i efterredigeringen:

L1: Den kommer jo kraftedeme snigende ...
L2: Because it sneaks up at you ...

Eksempeludsnit 2: Forsøgsperson 2.

Endelig viser et tredje eksempel, at oversættelse og efterredigering af udtrykket "listepik" i særlig grad har været vanskeligt for forsøgspersonerne. I dette tilfælde har det formentlig været manglende omverdensviden, der har gjort at hverken GT eller forsøgspersonerne har kunnet håndtere oversættelsen. Forsøgsperson 5 f.eks. slår ved tidstempel 07:08 først termen "listepik" op i V&B, men får intet resultat. Derefter søger vedkommende i Google men får heller ikke noget resultat. Derefter slås der op i URBAN, hvor forsøgsperson 5 finder termen "sleazeball", som vedkommende vælger at anvende i efterredigeringen:

L1: som Claus Listepik Elming til et Afterparty!
L2: like a sleazeball at an afterparty!

Eksempeludsnit 3: Forsøgsperson 5.

Der er endvidere adskillige eksempler fra vores data på, at flere forsøgspersoner ændrer noget korrekt til noget ukorrekt, hvilket formentlig skyldes en form for "perception bias", altså tendensen til at være udpræget subjektiv i forhold til GT. De fleste studerende er jo igennem deres studietid blevet fortalt, at man ikke kan stole på resultater fra GT. Den indgroede skepsis kan være en årsag til, at visse forsøgspersoner ændrer korrekte forslag til ukorrekte formuleringer i teksten. Dette kan f.eks. ses i følgende eksempel, hvor forsøgsperson 1 ved tidstempel 02:48 ændrer placeringen af adverbialen "in many ways":

L1: Selvom jeg på mange måder føler, at ...
L2: Although I in many ways feel that ...

Eksempeludsnit 4: Forsøgsperson 1.

GT har placeret adverbialfrasen "på mange måder" korrekt, men forsøgsperson 1 vælger alligevel at ændre den til ukorrekt placering mellem subjekt og verbum.

Som allerede nævnt, kan vi også se, at ni ud af ti forsøgspersoner kun i meget ringe omfang anvender de forhåndenværende leksikografiske ressourcer, og at de oftest slet ikke finder det, de søger efter. I stedet søger de på internettet, hvor de synes at finde det, de leder efter. De vælger ressourcer som f.eks. Google, hvor de søger efter konkrete websites som f.eks. www.andeby.dk eller URBAN. Det er også tydeligt, at forsøgspersonerne har svært ved at forstå betydningen af den danske tekst også selvom sværhedsgraden er passende, jf. Reimar (2018), men de bruger ikke L1-ordbog.

5.2. Oversættelse af faglig tekst

I den faglige tekst var der udfordringer med de faglige termer "flåde", "prinsgemal", "serviceklasser" og "kinesiske kagemur" og selvom sværhedsgraden var passende, jf. Cph.dk (2012), så havde GT og forsøgspersonerne vanskeligheder. En generel iagttagelse var, at selvom GT åbenlyst foreslår en forkert oversættelse, så lader forsøgspersonerne ofte fejlen stå eller også fejlretter de den ved at parafrasere. Dette kan ses i følgende eksempel fra forsøgsperson 6 som ved tidstempel 04:15 slår "flåden" op i V&B, leder lidt efter den rigtige ækvivalent, men vælger at parafrasere den engelske tekst i stedet for:

L1: flåden af langdistancefly
L2: amount of long distance planes

Eksempeludsnit 5: Forsøgsperson 6.

Et andet eksempel er forsøgsperson 7, som ved tidsstempel 06:55 vil efterredigere den af GT foreslåede forkerte ord-for-ord oversættelse af den danske term "prinsgemal" til "prince mill". Forsøgsperson 7 slår op i V&B, finder "Prince Consort" og indarbejder det umiddelbart efter i efterredigeringen:

L1: lød spådommen fra prinsgemalen
L2: the prediction from the prince mill sounded

Eksempeludsnit 6: Forsøgsperson 7.

GT fejler på det terminologiske niveau, fordi den mangler faglig viden og skaber problemer for de brugere, der ikke er fortrolige med fagområdet. I dette eksempel går det dog, idet forsøgsperson 7 er opmærksom på, at "prince mill" lyder bizart.

Et sidste eksempel er forsøgsperson 8, der møder følgende udfordring med oversættelsen af termen "service classes":

L1: med 245 sæder fordelt på tre serviceklasser: Business, Economy Extra og Economy Class
L2: with 245 seats distributed on three service classes: Business, Economy Extra and Economy Class

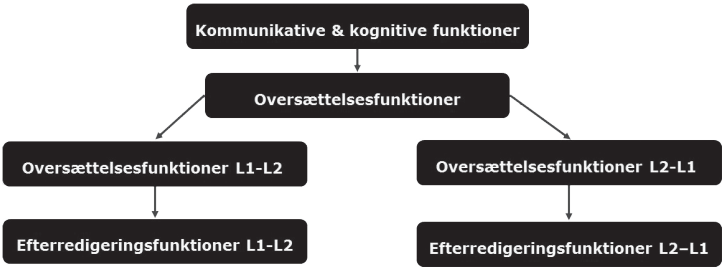
Eksempeludsnit 7: Forsøgsperson 8.

Forsøgspersonen ændrer "distributed on" til "divided into" og vælger simpelt hen at udelade oversættelsen af service classes.

Der synes at være empirisk belæg for at slå fast, at der er behov for udvikling af både almensproglige og fagsproglige ordbøger, der i højere grad, end tilfældet er nu, tilvejebringer løsninger på de efterredigeringsproblemer, brugerne møder. Ordbøgerne skal benytte sig af særlige datakategorier, der præcist er målrettet efterredigeringssituationer og dertil knyttede behov for informationer i forbindelse med GT-genererede oversættelser. Derfor har vi udviklet et koncept til efterredigering, som præsenteres i næste afsnit.

6. Koncept til efterredigeringsordbøger

Vores data viser, at GT fejler i mange tilfælde og at brugerne ofte ikke finder den information, de har brug for i den konventionelle oversættelsesordbog. Der er brug for funktioner og datakategorier i oversættelsesordbøger, som kan understøtte efterredigering. Afhængigt af situation, optræder efterredigeringsfunktioner både som kommunikative og kognitive situationer, som vist i figur 1.



Figur 1: Kommunikative/kognitive efterredigeringsfunktioner.

I kommunikative situationer er der tale om efterredigering af oversættelser i erhvervssammenhæng, både til intern og ekstern brug. Brugeren har behov for semantiske data, der hurtigt kan fjerne tvivl om GT har fået fat i de rigtige ækvivalenter. Derved kan tidsgevinsten maksimeres.

I kognitive situationer er der tale om efterredigering i læringsøjemed, hvor tid spiller en mindre rolle. Her har brugeren behov for semantiske data, der er mere omfangsrige og kan bidrage til at understøtte læringsprocessen ved træning i efterredigeringsmetoder. Denne skelnen er i øvrigt også anvendt i Leroyer & Simonsen (2018), der taler om tekstproduktion af en særlig art. Med efterredigering kan man i analogi hermed derfor tale om fagoversættelse af en særlig art.

Efterredigering er en sammensat proces, der består af flere indbyrdes afhængige delfaser, og til hver enkelt af dem er en særlig ordbogskombination den mest velegnede. Dette er illustreret i tabel 1.

Særlig reception på L1	monolingval L1-ordbog
Særlig reception på L2	monolingval L2-ordbog
Evaluering af ækvivalens L1>L2	bilingval L1-L2 ordbog
Evaluering af L2 acceptabilitet	monolingval L2-ordbog

Tabel 1: Efterredigeringsfaser og ordbogskombinationer.

På basis af de enkelte faser og tilsvarende ordbogskombinationer kan vi kombinere disse tanker med TAUS-retningslinjer, som vi diskuterede ovenfor. Vi kan hermed afstikke rammerne for en ændret selektion og præsentation af datakategorier, som oversættelsesordbøger med fordel kan indeholde for at kunne assistere brugerne.

Særlig reception L1 og særlig reception L2 peger på definitioner som centrale datakategorier, se tabel 2, hvor korte definitioner matcher de kommunikative situationer (den professionelle oversætter skal blot hurtigt tjekke semantikken), mens de lange matcher de kognitive situationer (brugerne følger undervisning i posteditering). Det skal bemærkes, at datakategorier passer både til TAUS 1 og TAUS 2. I begge tilfælde er det semantikken, der skal evalueres og tjekkes.

TAUS 1 og 2
Enkelt- og flerordslemmata på L1 Definitioner på L1 (kort/lang) Enkelt- og flerordslemmata på L2 Definitioner på L2 (kort/lang)

Tabel 2: Evaluering af særlig reception L1 og L2.

Overførsel fra L1 til L2 peger på ækvivalenter som central datakategori (se tabel 3), hvor TAUS 1 kræver et tjek af semantikken, mens TAUS 2 også kræver et tjek af syntaks på kollokationsniveau. Denne datapræsentation ligner den gængse præsentation i konventionelle fagoversættelsesordbøger. Forskellen er blot, at der skelnes mellem en reduceret præsentation med ækvivalenter og en udvidet præsentation med tilføjelse af kollokationer og deres oversættelser. I overensstemmelse med *Den Danske Ordbog* (DDO) definerer vi her kollokationer som "forbindelse af to eller flere ord der sædvanligvis optræder hyppigt sammen" og som ifølge Bergenholtz & Tarp (1994:119ff.) udgør en semantisk specifik kombination i det pågældende fagsprog.

TAUS 1
Enkelt- og flerordslemmata på L1 L2-ækvivalenter
TAUS 2
Kollokationer til L1-lemmata Oversættelse af L1-kollokationer til L2

Tabel 3: Evaluering af ækvivalens L1>L2.

Evaluering af acceptabilitet (se tabel 4) er på sin vis en afspejling af den særlige L2-reception, som vi skitserede ovenover. Forskellen er blot, at der inddrages synonymi til L2-lemmata på TAUS 1-niveau og til kollokationer på TAUS 2-niveau. Dette peger på følgende datakategorier:

TAUS 1
Enkelt- og flerordslemmata på L2 Synonymer til L2-lemmata Definitioner af L2-lemmata (kort/lang)
TAUS 2
Kollokationer til L2-lemmata Synonymer til L2-kollokationer

Tabel 4: Evaluering af acceptabilitet på L2.

Samler vi alle datakategorierne, når vi frem til følgende dataselektion:

TAUS 1
Enkelt- og flerordslemmata på L1 Enkelt- og flerordslemmata på L2 Definitioner af L1-lemmata (kort/lang) Definitioner af L2-lemmata (kort/lang) L2-ækvivalenter Synonymer til L2-lemmata

TAUS 2
Kollokationer til L1-lemmata Oversættelse af L1-kollokationer til L2 Kollokationer til L2-lemmata Synonymer til L2-kollokationer
Datapræsentationen kan filtreres ud fra:
Brugersituationen – kommunikativ eller kognitiv Ønsket TAUS niveau – 1 eller 2 Redigeringsfasen – evaluering og tjek af L1, L1>L2, eller L2

Tabel 5: Dataselektion TAUS 1 og 2, samlet.

Dermed vil datamængden falde betragteligt og risikoen for informationsoverload tilsvarende.

Gellerstam (1994) sluttede sin artikel med at påpege en stigende heterogenitet i ordbøger. Et postredigeringsmodul som skitseret her bekræfter denne udvikling. GT baner vej for heterogene leksikografiske moduler, der kan implementeres på leksikografiske platforme, hvorved dataselektion og datapræsentation kan tilpasses brugersituationen. Det vil så være muligt at udnytte tidsgevinsten ved maskinoversættelse og samtidig minimere mængden af leksikografiske data, der kræves til posteditering. Efterredigeringsmodulet kan med fordel implementeres som udvidelse af de termbaser, der sammen med GT, er integreret i hukommelsesdrevne oversættelsesplatforme.

7. Konklusion

Norden er en region med tætte internationale forbindelser og med stort behov for oversættelse til og fra de nordiske sprog. Gode oversættelsesordbøger og samlinger af bilingvale data (korpora og oversættelseshukommelser) er stadig yderst relevante værktøjer, akkurat som Gellerstam påpegede for 25 år siden. Vores undersø-

gelse peger på, at oversættelsesordbøger i en GT-æra bør ændres, så de i højere grad kan understøtte efterredigering med relevante, situationstilpassede datakategorier. En anden vigtig faktor er de studerendes manglende tiltro til GT. Dette kræver, at de studerende i fremtiden lærer at anvende GT, herunder hvornår de skal gribe ind og hvornår de kan stole på det output, GT leverer.

Vores data viste for det første, at forsøgspersonerne i ringe grad arbejdede med at efterredigere oversættelser fra L1 til L2. Det kom frem, at deres søge- og implementeringsstrategier lod en del tilbage at ønske, og at de kun i ringe grad var i stand til at anvende den konventionelle oversættelsesordbog eller til at implementere de fundne leksikografiske data optimalt. Der er følgelig brug for decideret undervisning i efterredigering i uddannelsen af translatører. En anden meget vigtig faktor er de studerendes ordbogskundskaber. Havde forsøgspersonerne lært at anvende forhåndenværende ordbøger systematisk, så ville resultatet formentlig have været bedre. Dette fjerner dog ikke behovet for efterredigeringsfunktioner.

For det andet viste det sig, at forsøgspersonerne gjorde udstrakt brug af eksterne online ressourcer, og at de både fandt og implementerede relevante data. Undervisning i leksikografisk informationssøgning bør derfor styrkes.

Som svar på det tredje spørgsmål anbefaler vi en leksikografisk efterredigering i overensstemmelse med TAUS. Dette vil kunne sikre, at efterredigeringen bliver mindre tidskrævende.

Som svar på det fjerde forskningsspørgsmål foreslår vi et koncept med målrettede efterredigeringsfunktioner. Vi anbefaler to definitionsformater – kort og langt – afhængigt af den overordnede brugersituation.

Udbredelsen af GT kan meget vel vise sig at være en stærk drivkraft i udviklingen af særlige leksikografiske data af høj kvalitet til brug i oversættelsessystemer. I den forstand kan vi fastslå, at GT snarere bør anskues som en redning end en trussel, for der er et stigende behov for udvikling af nye oversættelsesordbøger og mul-

tilingvale termbaser, hvor efterredigering i langt højere grad sættes i fokus.

Litteratur

Ordbøger og oversættelseshukommelse

DDO = *Den Danske Ordbog*. Online: <ordnet.dk/ddo> (juli 2019).

MEMSOURCE = *Memsource*. <www.memsource.com> (marts 2019).

TRADOS = *SDL Trados*. <www.trados.com> (marts 2019).

Urban = *Urban Dictionary*. <https://www.urbandictionary.com> (marts 2019).

V&B = Herman Vinterberg & C. A. Bodelsen: *Dansk-engelsk. Gyldendals Røde Ordbøger*. København: Gyldendal. <http://ordbog.gyldendal.dk.esc-web.lib.cbs.dk> (marts 2019).

Anden litteratur

Bergenholtz, Henning & Sven Tarp (1994): *Manual i fagleksikografi. Udarbejdelse af fagordbøger – Problemer og løsningsforslag*. Aarhus: Systime.

Bundgaard, Kristine (2017): (Post-)Editing – A Workplace Study of Translator-Computer Interaction at TextMinded Danmark A/S. PhD dissertation. Aarhus BSS, Aarhus University: Department of Management.

Carl, Michael, Barbara Dragsted, Jakob Elming, Daniel Hardt & Arnt Lykke Jakobsen (2011): The Process of Post-Editing: A pilot study. I: Bernadette Sharp, Michael Zock, Michael Carl & Arnt Lykke Jakobsen (eds): *Proceedings of the 8th International NLPCS Workshop. Special Theme: Human-Machine Interaction in Translation*. Copenhagen: Samfundslitteratur, 131-142.

Carl, Michael, Yanfang Jia & Xiangling Wang (2019): How does the post-editing of Neural Machine Translation compare with from-scratch translation? A product and process study. I: *The Journal of Specialised Translation* 31, 60-86.

Cph.dk (2012): Prins Henrik åbnede Shanghai-rute. <https://www.cph.dk/om-cph/presse/nyheder/2012/3/prins-henrik-abnede-shanghai-rute> (marts 2019).

Flanagan, Marian (1994): Error classification for MT evaluation. I: *Proceedings of the Association for Machine Translation in the Americas* (*AMTA 1994*), 65-72. <http://mt-archive.info/AMTA-1994-Flanagan.pdf> (marts 2019).

Gellerstam, Martin (1994): Produktionsordböcker – vad är det? I: *LexicoNordica* 1, 43-52.

ISO 18587 (2017): Translation services – Post-editing of machine translation output – Requirements. Geneva: International Organization for Standardization. <https://www.iso.org/standard/62970.html> (marts 2019).

Koglin, Arlene & Rossana Cunha (2019): Investigating the post-editing effort associated with machine-translated metaphors: a process-driven analysis. I: *The Journal of Specialised Translation* 31, 38-59.

Leroyer, Patrick & Henrik Køhler Simonsen (2018): Fagleksikografi i spændingsfeltet mellem viden og kommunikation. I: *LexicoNordica* 25, 115-133.

Reimar, Lotte (2018): Ramt af magelighed! <http://lottereimar.dk/2018/02/11/ramt-af-magelighed/> (marts 2019).

Saunders, Mark, Philip Lewis & Adrian Thornhill (2009): *Research methods for business students*. Harlow: Pearson Education.

Tarp, Sven (2004): Korrektur og retning som leksikografiske funktioner. I: *Hermes, Journal of Linguistics* 33, 117-147.

TAUS (2019): MT Post-editing Guidelines. An agreed set of post-editing best practices. <https://www.taus.net/academy/best-practices/postedit-best-practices/machine-translation-post-editing-guidelines> (marts 2019).

Patrick Leroyer
lektor, PhD
Aarhus Universitet
Jens Chr. Skous Vej 4
DK-8000 Aarhus C
pl@cc.au.dk

Henrik Køhler Simonsen
ekstern lektor, PhD, MBA
Copenhagen Business School
Dalgas Have 15
DK-2000 Frederiksberg
hks.msc@cbs.dk

“Hvordan siger man ...?” Om direkte brugerkontakt i ordbogsarbejdet

Jacqueline Levin & Frederikke Ulv

This article describes how the authors, in their role as editors, interact via live chat and email with users of the ordbogen.com online dictionaries, and how this interaction benefits both users and dictionaries. Not only do we help users with their questions regarding language, translation or the use of the dictionaries, but we also gain insights into potential improvements to our dictionaries and subsequently revise entries, within the technical constraints of each dictionary. The article concludes with perspectives on “the dictionary of our dreams”.

1. Baggrund, afgrænsning og begrebsafklaring

Websitet ordbogen.com blev startet i 2003 med en omfattende engelsk-dansk ordbog over almen- og fagsprog, som var udarbejdet og løbende blev udvidet af translatør Jørgen Rohde på baggrund af brugernes søgninger og forespørgsler. I årene 2006-2008 udarbejdedes i samarbejde med bl.a. studerende og undervisere fra Syddansk Universitet i Odense fem egne ordbøger. Først udviklede man en engelsk-dansk ordbog på basis af kun de mest søgte ord i den eksisterende ordbog (den præcise andel har aldrig været offentliggjort). Logningen viste nemlig at størstedelen af ordene i praksis aldrig blev brugt, så dem var der ikke grund til at bruge tid på – en pointe også fremført af Bergenholtz & Johnsen (2005:122). I 2007-2009 udviklede Ordbogen A/S sine egne tysk-, fransk-, spansk- og svensk-danske ordbøger på basis af de mest frekvente ord på sprogene. Siden udgivelsen er ordbøgerne løbende blevet udvidet og revideret ud fra logger over brugernes søgninger samt

direkte forespørgsler. Ifølge Abel & Meyers klassifikation (2013) vil det sige at ordbøgerne revideres på baggrund af indirekte brugerbidrag af både implicit karakter (logning) og eksplicit (brugerforespørgsler).

Denne artikel omhandler brugerhenvendelser der vedrører Ordbogens egenproducerede engelsk-, fransk- og tysk-danske ordbøger, da det er disse forespørgsler forfatterne bag artiklen mest har beskæftiget sig med. Besvarelsen af brugerhenvendelser er dels led i det daglige arbejde i sig selv, dels indgår emnerne for forespørgslerne i de redaktionelle overvejelser. I den forbindelse er det på sin plads med en afklaring af begreberne *sprogmedarbejder* og *redaktør* som bruges i artiklen:

Sprogmedarbejdere er alle fuldtids- og deltidsansatte, inklusive studentermedhjælpere, i Ordbogen A/S' sprogafdeling. Disses arbejde involverer udarbejdelse, revision eller korrektur af ordbogsartikler eller grammatikøvelser til websitet grammatip.com på et eller flere sprog. Desuden involverer det bemanding af ordbogen.-coms sprogchat vedrørende dansk, engelsk, tysk, fransk eller spansk. De vagthavende sprogmedarbejdere fordeler besvarelsen af chatten imellem sig efter sproglig formåen, men de fleste kan besvare de mest grundlæggende spørgsmål til ordbøgerne uanset sprog.

Redaktørerne af en given ordbog vil derimod kun være medarbejdere med særlige forudsætninger og/eller uddannelse inden for det pågældende sprog. Forfatterne til denne artikel håndterer i egenskab af sprogmedarbejdere chathenvendelser på alle fem nævnte sprog, som eksemplificeret i kapitel 3, mens vi i egenskab af redaktører udfører vores primære arbejde på enten engelsk- eller franskordbogen.

2. Direkte brugerkontakt

Som det fremgår af afsnit 1, har brugernes søgninger på ordbogen.com været logget siden sitet blev lanceret i 2003, og de data

redaktørerne får herfra, udgør fortsat et vigtigt råmateriale til de løbende revisioner af de egenproducerede ordbøger. Her ses sort på hvidt hvilke ord flest brugere savner, og dermed også hvilke artikler redaktørerne skal fokusere på for at tilgodese flest mulige af en ordbogs brugere, som pointeret af Bergenholtz & Johnsen (2005:122). Ord tilføjes uafhængigt af emne, fagområde eller stilleje. Ordbøgerne er deklarerede som almensproglige, men savner brugerne fx tekniske gloser eller slangudtryk, tilføjes disse også i det omfang redaktørerne kan verificere deres brug og oversættelser.

Herudover kan brugerne henvende sig til ordbogen.com via e-mail (herunder gennem kontaktformularer) og siden 2011 også via chat. Henvendelser vedrørende sprog er forbeholdt brugere som er logget ind, dvs. tilknyttet et abonnement, mens alle kan henvende sig vedrørende ikkesproglige forespørgsler, fx priser, abonnementer eller teknisk support. I praksis er det dog kun sprogchatten der er lukket for øvrige besøgende på sitet. Benyttes de generelle feedbackkanaler til sproglige spørgsmål, kontrolleres det ikke om afsenderen er betalende abonnent eller ej.

I det følgende gennemgås kort nogle af kendetegnene for de henvendelser som Ordbogens sprogmedarbejdere modtager via chat og mail. Disse henvendelser danner grundlag for henholdsvis kapitel 3 og 4 herunder.

2.1. Henvendelser på chatten

De fleste henvendelser på chatten bærer præg af at brugerne ønsker et svar her og nu fordi de skriver mens de er i færd med at løse en opgave ved hjælp af ordbogen.com. Cirka tre femtedele af forespørgslerne på hverdage kommer fra brugere der benytter det statslige uddannelseslogin UNI-Login, primært grundskoleelever. Dette optælles ikke automatisk, så tallet er baseret på en manuel gennemgang af chatloggen for november 2017 (november fordi denne måned er fri for både skoleferier og eksamensperioder, og 2017 fordi registreringen for november 2018 er ufuldstændig pga.

tekniske problemer). Af de 2.328 chats optalt for den måned, kom 1.360 gennem UNI-Login.

De øvrige forespørgsler fordeler sig på brugere som logger på via privat-, erhvervs- eller øvrige uddannelsesabonnementer eller ad kanaler som ikke kræver identifikation (login).

For en ordens skyld bør det nævnes at langt fra alle chats munder ud i reel eller seriøs brugerkontakt. Omtrent halvdelen af alle chats (1.175) afsluttes ved at brugeren logger ud igen før vedkommende har stillet (eller fået besvaret) et spørgsmål. Af den halvdel af chatbrugerne der starter en dialog med en sprogmedarbejder, er der derudover godt en fjerdedel (302) som klassificeres som "useriøse", fx fordi der er tale om skoleelever der kun prøver at lokke et svar ud af en sprogmedarbejder frem for selv at slå ordene op i ordbogen. Desværre modsiger dette forhåbningen udtrykt af Törnqvist (2010) om at en begrænsning til registrerede eller betalende brugere kan nedbringe antallet af useriøse henvendelser. I stedet har ordbogen.com måttet lave en blokeringsfunktion (som Törnqvist også foreslår) til netop gruppen af brugere registreret via UNI-Login.

De reelle forespørgsler på chatten er dels forespørgsler på ord eller fraser som man ikke har kunnet finde i ordbogen, dels spørgsmål om hjælp til grammatik, formuleringer og lignende. I kapitel 3 beskrives arbejdet med chatten med et fokus på hvordan og hvorfor ovennævnte skoleelever behandles forskelligt fra ikke-skolebrugere.

2.2. Mailhenvendelser

Mailhenvendelser spænder fra simple forespørgsler om at få et ord føjet til ordbogen (dette især når brugerne benytter kontaktformularerne) til mere udførlige formuleringer vedrørende mere komplekse problematikker som brugerne ønsker hjælp til eller dialog om. Uanset om brugerne sender en mail eller benytter en

formular, er der tale om hvad Abel & Meyer (2013:185) definerer som *free form feedback* idet formularerne er baseret på et tekstfelt til fri formulering.

Specifikke gloser som forespørges i de simple henvendelser, tilføjes, på samme måde som hyppigt loggede søgninger, i det omfang redaktørerne kan verificere deres brug og oversættelser. Af pladshensyn beskrives behandlingen af disse typer mail ikke særskilt i denne artikel.

Der kan også, som på chatten, være tale om at brugerne starter en dialog parallelt med at de prøver at løse en sproglig opgave. Disse håndteres i dialog med brugerne parallelt med chatdialogerne beskrevet i kapitel 3, og beskrives ligeledes heller ikke videre her.

Endelig bruges mail også til ønsker af mere generel karakter, fx informationstyper som ikke – eller ikke konsekvent – er med i en ordbog (forklaringer, eksempler, grammatiske oversigter, lydskrift m.m.). Disse mails er oftest ikke skrevet i selve brugssituationen, men af brugere som har taget sig tiden til at skrive til Ordbogen ved en senere lejlighed. Det er også i denne ende af spektret man finder de brugere der ikke skriver på egne vegne, nemlig lærere der skriver med henblik på deres elevers brug af ordbøgerne. Arbejdet med lærerhenvendelser beskrives i kapitel 4.

3. Sproghjælp på chatten

Som tidligere nævnt hjælper sprogmedarbejderne på chatten brugere som har tvivlsspørgsmål her og nu, men som det uddybes i afsnit 3.3, afhænger besvarelsen af hvem der henvender sig. Dette eksemplificeres ved repræsentanter for to fremherskende brugertyper på chatten, en erhvervskunde og en skoleelev. Begge henvender sig med varianter af spørgsmålet "Hvordan siger man ...?" som udtryk for at de søger hjælp til at formulere sig korrekt på et fremmedsprog.

3.1. Erhvervskunden

Her følger et anonymiseret uddrag med tidsstempler fra en chat mellem en sprogmedarbejder fra Ordbogen og en bruger fra en erhvervsvirksomhed som bad om hjælp til frasen *tilbud på*.

Benævnelserne "KUNDENAVN" og "Kundenavn" er brugerens egne anonymiseringer.

> **10:21 Bruger**: Hej. Jeg er ved at oversætte et tilbud til en kunde - fra DK til EN. Tilbud oversætter jeg med Quote, men jeg er i tvivl om, hvordan jeg korrekt siger tilbud på x. Sætningen er: "Som aftalt sender vi her et tilbud på en foranalyse af KUNDENAVNs infrastruktur". Indtil videre har jeg oversat det til: As agreed, we send this quote on a pre-analysis of the infrastructure of KUNDENAVN.
> [...]
> **10:22 Ordbogen**: Jeg mener man siger "for". Jeg dobbelttjekker lige.
> **10:24 Ordbogen**: Jo, det skal være "for" :)
> **10:27 Bruger**: As agreed, we send this quote for a pre-analysis of the infrastructure of Kundenavn.
> [...]
> **10:28 Ordbogen**: "Her" eller "hermed" oversættes ofte som en -ing-form: "As agreed, I am sending you a quote ..."
> **10:29 Ordbogen**: Resten har jeg ingen bemærkninger til.
> **10:32 Bruger**: Tak for hjælpen.

Eksempel 1: Brugerhenvendelse fra en erhvervskunde.

Brugeren henvender sig specifikt med henblik på at få kontrolleret sin oversættelse af frasen *tilbud på* med *quote on* (10:21). Sprogmedarbejderen giver sit bud på den korrekte formulering og slår efter om den fremgår af ordbogen (10:22). Gør den ikke det, noteres den til eventuel tilføjelse. (Frasen var dog allerede med). Sprog-

medarbejderen bekræfter sin formulering (10:24), og brugeren genformulerer med den korrekte præposition (10:27). Sprogmedarbejderen supplerer med et grammatisk forbedringsforslag til en anden del af teksten (10:28) og indikerer at teksten som helhed derefter er korrekt (10:29). At *we* i samme replik fejlagtigt byttes ud med *I*, lader heldigvis ikke til at forvirre brugeren.

3.2. Skoleeleven

I følgende eksempel gengives en chat med en skoleelev der, ligesom erhvervskunden i afsnit 3.1, beder om hjælp til en formulering. Her antager hjælpen dog en anden form:

> **08:41 Bruger**: Hej, hvordan siger man jeg går i 9 klasse?
> [...]
> **08:41 Ordbogen**: Hvis du starter med at slå op i ordbogen og prøve selv, skal jeg nok kigge på det.
> **08:42 Bruger**: ich gehe zu 9. klasse?
> **08:43 Ordbogen**: Prøv at slå "klasse" op og kigge på eksemplerne.
> **08:44 Bruger**: ich in die 9. Klasse gehen?
> [...]
> **08:47 Ordbogen**: Du har bare oversat "at gå i 9. klasse", men det skal jo være nutid. Prøv at kigge på eksemplet med "Meine Tochter". Der bliver lavet om på rækkefølgen af ordene i nutid.
> **08:48 Bruger**: Ich gehe in die neunte Klasse?
> **08:48 Ordbogen**: Perfekt! :)
> **08:48 Bruger**: tak:)

Eksempel 2: Brugerhenvendelse fra en skoleelev.

Skolebrugeren beder, ligesom erhvervsbrugeren, om hjælp til en konkret formulering (*jeg går i 9. klasse*), men denne gang udløser

det standardformuleringen "Hvis du starter med at slå op i ordbogen ..." (08:41) som er gemt som hurtigsvar i sprogmedarbejderens chatprogram. Efter tilsyneladende at have gættet (08:42) hjælpes eleven på rette vej idet sprogmedarbejderen fortæller hvilken ordbogsartikel vendingen kan findes i (08:43). Parallelt med chatten med erhvervskunden slår sprogmedarbejderen altid efter for at sikre sig at opgaven kan løses ved hjælp af ordbogen, og hvis ikke, noteres det så den kan blive revideret efterfølgende. Eleven begår dernæst den klassiske fejl at benytte oversættelsen i sin opslagsform (08:44) hvorefter sprogmedarbejderen hjælper ved at påpege hvilke elementer i sætningen der mangler opmærksomhed (08:47). Til sidst forstår eleven sammenhængen og er i stand til at formulere sætningen grammatisk korrekt (08:48).

3.3. Chattens rolle afhænger af brugeren

Chattens rolle afhænger af brugerens kommunikationssituation. I praksis antages det at erhvervskunder henvender sig under udførelsen af deres erhverv, mens skoleelever skriver i forbindelse med skolens sprogundervisning.

I sproghjælpen til erhvervskunden prioriteres brugerens behov for at få formuleret et korrekt engelsksproget tilbud til en kunde højere end instruktion i brug af ordbogen. Derfor vælger sprogmedarbejderen i dette tilfælde fx ikke at meddele at vendingen *tilbud på* allerede var med i ordbogen, men bekræfter blot den korrekte formulering, og derfor supplerer sprogmedarbejderen også med et grammatisk forbedringsforslag til en del af teksten som brugeren ikke havde spurgt til.

Hvor ordbogen for erhvervskunden er et arbejdsredskab, er den for skoleeleven et læremiddel som indgår i undervisningen i fremmedsprog eller dansk. For at ordbøgerne kan udvikle sig til meningsfulde arbejdsredskaber for eleverne, er det derfor også en disciplin i sig selv i sprogundervisningen at lære at forstå og få

det fulde udbytte af informationerne i ordbøgerne. Det kræver en indsats fra elevernes side, men hvis de er villige til at prøve, som eleven i dette eksempel, hjælper vi gerne. Kontakt fra elever der håber på at en sprogmedarbejder bare vil forære dem svarene, sker dog hyppigt nok til at hurtigsvar som det nævnte ("Hvis du starter med at slå op ...") er praktiske for sprogmedarbejderne.

For at opsummere kan man sige at chatten for erhvervsbrugeren supplerer ordbogsopslagene med korrektur, sparring og som en art facitliste. Over for skoleelever er fokus derimod på brug af ordbogen, men aldrig ud over elevens evner og aldrig som facitliste.

4. Metahjælp på mail: lærerhenvendelser

I sammenligning med brugere som henvender sig til ordbogen.com med aktuelle tvivlsspørgsmål, adskiller sproglærerne sig ved at skrive om mere generelle problemstillinger og på vegne af deres elever. Dermed er deres forespørgsler, selv når de omhandler konkrete ord og vendinger, snarere møntet på ordbøgerne som redskaber eller hjælpemidler i undervisningen. Det betyder også at selv om lærerne i forhold til de øvrige brugergrupper udgør en meget lille andel, er de særligt interessante for os som redaktører fordi deres mails giver os vigtige indsigter. De er den gruppe der har størst direkte indflydelse på udformningen af franskordbogen på overordnet niveau for at gøre den mere velegnet som hjælpemiddel.

Ser man nærmere på de konkrete mails, tegner der sig det mønster at Ordbogen stort set aldrig modtager spørgsmål eller kommentarer fra lærere angående engelskordbogen, sjældent angående tyskordbogen, men jævnligt angående franskordbogen. Der kan være mange grunde til den store forskel i antal henvendelser, men en væsentlig faktor er at elevernes sproglige niveau er meget forskelligt i de forskellige fremmedsprog. Engelsk- og tyskunder-

visningen foregår på et højere niveau end franskundervisningen og er i højere grad baseret på formulering af længere tekster. Vi kan da også se af engelsk- og tyskelevernes spørgsmål på chatten at de derfor oftere opsøger ordbøgerne med henblik på vendinger og grammatik, mens franskeleverne mere basalt efterspørger enkeltstående gloser. Det fremgår af lærernes mails at mange elever som er begyndere og let øvede i fransk, har svært ved at vælge den rigtige glose.

Denne opgave egner mange artikler i franskordbogen i sin nuværende udformning sig ikke til. På den ene side blev den i lighed med engelsk- og tyskordbogen udviklet til brugere med mere avancerede franskkundskaber, og på den anden side blev den ikke udvidet med sætningseksempler efter lanceringen som de to andre ordbøger. Informationstyper som brugseksempler eller verbalbøjninger nævnes af Atkins & Rundell (2008:33) som muligt indhold der skal tages stilling til i forbindelse med brugerundersøgelser før ordbogsarbejdets påbegyndelse, men her foregår tingene i omvendt rækkefølge idet franskordbogen først efter sin udgivelse fik kundekredsen udvidet til skoler.

De fleste henvendelser fra fransklærere vedrører en af tre informationstyper som vil blive beskrevet nærmere i de følgende afsnit: tilføjelse af ord og vendinger til ordbogen (afsnit 4.1), tilføjelse af uddybende oplysninger (4.2) samt tilføjelse af bøjningsskemaer (4.3).

4.1. Tilføjelse af ord og vendinger

Da skoleeleverne udgør en væsentlig andel af brugerne på ordbogen.com, er det vigtigt for virksomheden at tilpasse ordbøgerne til deres behov som de observeres af deres lærere. Når lærerne henvender sig med ønsker om rettelser eller tilføjelser til ordbøgerne, gør de det nemlig på baggrund af observationer som dem der kommer til udtryk i forespørgsler som nedenstående:

> [Jeg kunne] godt tænke mig at følgende ord/udtryk blev optaget i ordbogen, da eleverne ofte har svært ved dem / laver fejl i dem: ... "tage" ... i betydningen at tage et sted hen ... [Og] ofte mener eleverne "hvile sig" når de bruger udtrykket "slappe af" ...

Andre gange ønsker lærerne specifikke ord eller oversættelser føjet til ordbøgerne fordi de skal bruges i arbejdet med en konkret tekst. I de tilfælde vægter vi det højt at tilføje ordene så eleverne kan bruge ordbøgerne til at løse opgaver i undervisningssammenhæng.

4.2. Tilføjelse af uddybende oplysninger

For en skoleelev kan artikler uden eksempler være stort set ubrugelige som det fremgår af en lærerhenvendelse som denne:

> Det ville være dejligt, om der var flere eksempler i jeres ordbogsopslag, for eleverne har ingen chance for at gennemskue nuancer på de forskellige ordforslag og får tit og ofte så valgt noget vrøvl. Forståelige, enkle eksempler ...

Artiklerne i franskordbogen er forsynet med forklaringer eller emneangivelser, men disse er ofte for abstrakte at forholde sig til for den let øvede bruger hvis de ikke er ledsaget af konkrete sætningseksempler. Samtidig er eleverne også vant til at arbejde med engelsk- og tyskordbøgerne som indeholder eksempler til stort set alle oversættelser. Vi har derfor valgt at prioritere at forsyne flest mulige højfrekvente ord med enklere forklaringer og med meningsfulde eksempler.

4.3. Tilføjelse af bøjningsskemaer

Den sidste del af forespørgslerne fra lærerne omhandler et længe næret ønske om bøjningsskemaer i franskordbogen. Redaktørerne hos Ordbogen har stor forståelse for ønsket, men efterspørgslen har indtil for nylig ikke været stor nok til at retfærdiggøre investeringen. Igennem den sidste tid er der dog registreret en støt stigende efterspørgsel, dels fordi antallet af gymnasier blandt kunderne stiger, og dels fordi Ordbogens kundekonsulenter sørger for at gøre mere opmærksom på den mulighed for medbestemmelse lærerne har igennem at sende feedback.

I stedet for bøjningsskemaer er de franske verber forsynet med et antal systematiske bøjningsformer hvoraf verbets øvrige former principielt kan udledes. Det kræver dog kendskab til bøjningssystemet eller franskkundskaber på et højere niveau end elevernes. Franskordbogens søgelogger indeholder da også mange bøjningsformer, hvilket kan antyde at eleverne ikke kan ræsonnere sig frem til verbernes infinitiv, og at de fra de andre ordbøger på ordbogen.com er vant til at kunne søge på bøjede former.

Da det fremgår af søgeloggen at nogle relativt få (især uregelmæssige) bøjningsformer er overrepræsenterede i søgningerne, er redaktørerne begyndt at indeksere de hyppigst søgte former så brugerne vises den relevante ordbogsartikel. Søgeloggen for januar 2018 (måneden før dette arbejde påbegyndtes) viser således at 14 ud af de 15 hyppigste forgæves søgninger i franskordbogen var bøjningsformer af franske verber. I januar 2019, efter et års arbejde, var der til sammenligning kun 3 franske verbalbøjninger blandt de øverste 15 i søgeloggen. Denne løsning giver dog ikke brugerne svar på hvilken bøjningsform der er tale om, og løser derfor kun problemet halvt. Og hvis verbets grundform ser meget anderledes ud end den søgte bøjningsform, kan eleverne endda opfatte det som om de ved en fejl får vist et helt andet verbum. Vi arbejder derfor fortsat på at få bøjningsskemaer i ordbogen.

4.4. Indirekte indflydelse på ordbogens indhold

Ud over den direkte feedback via mail og chat og indirekte via søgeloggerne øver lærere og elever endnu en form for indirekte indflydelse på franskordbogen. Sideløbende med ordbøgerne udgiver Ordbogen A/S nemlig grammatikopgaver gennem undervisningsportalen grammatip.com til bl.a. grundskolen som det er tanken at eleverne skal kunne løse med hjælp fra ordbøgerne.

For at det skal være muligt at slå alle øveordene op i ordbogen, gennemgås de relevante ordbogsartikler for at sikre, for eksempel, at ordene er forsynet med letforståelige forklaringer eller eksempler (jf. afsnit 4.2). Tilsvarende kan opgaveforfatterne i visse øvelsestyper vælge at bruge den betydning der står øverst i ordbogsartiklerne, og specificere i opgaveformuleringen at eleverne skal bruge den øverste oversættelse.

5. Fra brugerkontakt til ordbogsindhold

I de foregående kapitler har vi gennemgået mange af de typer brugerhenvendelser der kan lede til tilføjelser af ord, vendinger, oversættelser osv. i Ordbogens egenproducerede ordbøger, men i praksis er det kun et fåtal ud af det samlede antal mails og chats der resulterer i ordbogsrevisioner. Der er dagligt brugere der får ord optaget i ordbogen fordi de savnede et ord eller en vending som uden videre kunne tilføjes til glæde for andre brugere. Men en forespørgsel kan være for generel til at den kan honoreres fra dag til dag (bøjningsskemaer i tyskordbogen blev efterspurgt i lang tid før der blev sat midler af til projektet), eller den kan være så specifik at der ikke kan generaliseres til en ordbogsartikel, som nævnt i afsnit 5.1 herunder.

De to eksempler på chatdialoger i kapitel 3 udgør en tredje type forespørgsel som er svær at yde hjælp til igennem ordbogsartik-

ler. De kunne have ledt til revisioner af eksisterende artikler, hvis ordforbindelserne *tilbud på* og *gå i n'te klasse* ikke allerede havde været med, men det var de. Og netop det kendetegner mange henvendelser: Brugerne søger hjælp til opgaver de ikke kan løse ved hjælp af ordbogen, eller som de ikke stoler på at de har løst rigtigt, og hvor tvivlen ikke som sådan skyldes mangler i ordbogen. En løsningsmulighed der løbende arbejdes på i Ordbogens udviklingsafdeling, er en mere intelligent søgefunktion som er bedre til at gennemskue hvad brugerne mener, hvis de søger på længere tekststrenge.

Tilbage står dog stadig en række tvivlssituationer for brugerne som ikke kan afhjælpes for indeværende. Vi har ikke selv svar på hvordan alle situationerne kan håndteres gennem forbedringer af ordbogen.com, men forsøger at komme med nogle bud i afsnit 5.2.

5.1. Spørgsmål der ikke kan løses med ordbogsartikler

Blandt de hyppigst tilbagevendende typer spørgsmål som ikke kan besvares igennem ordbøgernes artikler, kan nævnes følgende:

Forklarende oversættelser: I kapitel 2 nævntes det at efterspurgte fagord tilføjes i det omfang redaktørerne kan verificere dem. Kan vi, som redaktører, ikke finde en autoriseret oversættelse, kan vi i stedet hjælpe med at formulere en forklarende oversættelse. Det hjælper brugeren i den aktuelle kommunikationssituation, men oversættelsen vil som regel ikke blive føjet til ordbogen.

Omskrivninger: Nogle grammatiske strukturer ændres eller bortfalder i oversættelser, hvilket kan være svært både at forklare og illustrere så den uindviede ikke kan blive i tvivl. Desuden kan det være svært for brugeren at gennemskue hvilket ord man skal slå op under for at søge svaret.

Grammatik: Brugerne har også spørgsmål af mere overordnet grammatisk karakter, fx vedrørende kongruens i tal, person, køn eller kasus, som ikke kan indplaceres under opslagsordene. (Under

hvilket ord burde en bruger fx kunne slå op om det hedder *she is* eller *she are*?).

Kreativ skrivning/professionel sparring: Nogle gange kan brugeren ikke bruge en ordret oversættelse fordi der også skal rammes en bestemt "smart" tone, et bogstavrim, et sloganpræg el.lign., fx i salgsmateriale. Eller en professionel oversætter kan simpelthen føle at en formelt set korrekt oversættelse rammer skævt. Her agerer sprogmedarbejderne kreative sparringspartnere efter bedste evne, men resultatet vil stort set altid være for kontekstspecifikt til at kunne føjes til ordbogen.

Fejlstavning: Ordbøgernes søgefunktion kan til en vis grad håndtere fejlstavning af ord og foreslå op til 10 opslagsord med stavemåder i nærheden af det søgte. Det gælder dog både på fremmedsprogene og på dansk at hvis brugerens gæt på en stavemåde er for langt fra den korrekte, vil den ikke blive genkendt, også selv om der arbejdes på at forbedre forslagsmekanismen ved at inddrage information om hvilke ord brugeren efterfølgende endte med at slå op.

6. Perspektiver

Vi har i denne artikel forsøgt at beskrive nogle af de gavnlige effekter af kontakten med brugerne af ordbogen.com. På den ene side får brugerne hjælp til brug af ordbogen, sparring om oversættelser og svar på tvivlsspørgsmål. På den anden side får vi som redaktører input og ideer til revision af ordbogsartiklerne og til udvidelse af ordbogens funktionalitet. På den måde er brugerkontakten med til at forbedre både den enkelte ordbog og ordbogssitet som helhed. Og selv når det ikke er muligt at give brugerne det de beder om i den givne situation, går de sjældent tomhændede fra dialogen. På den måde fungerer brugerkontakten også som en udvidelse af ordbogen, en indbygget spørgefunktion der kan hjælpe med det ordbøgerne alene ikke kan give svar på.

Når det så er sagt, ser vi utallige muligheder for forbedringer som ikke kan gennemføres aktuelt fordi hver ny funktionalitet som skal udvikles, har en økonomisk omkostning som skal kunne forsvares i et overordnet budget. Men afslutningsvis vil vi her give os selv lov til at drømme frit.

Der tales ofte om at ordbøger vil forsvinde ud af syne og blive indbygget i de redskaber brugerne benytter, fx skriveprogrammer eller e-bogslæsere der ud fra konteksten selv kan servere den rigtige oversættelse i den rigtige form. Mange arbejder hen imod den type løsninger, også hos Ordbogen A/S, hvor man har udviklet Write Assistant, et redskab som understøtter skrivning på engelsk ved at foreslå ord og tilbyde oversættelser fra den skrivendes modersmål ved hjælp af en sprogmodel baseret på korpus- og ordbogsmateriale som beskrevet af Tarp, Fisker & Sepstrup (2017). I skrivende stund kan en testversion for dansk-, tysk- og spansktalende brugere downloades fra writeassistant.com.

Den type sproghjælp vil være til stor gavn for brugere som kommunikerer på fremmedsprog som led i deres arbejde (som fx erhvervskunden i afsnit 3.1), men ikke for brugere som arbejder på at lære et fremmedsprog, som fx skoleeleverne omtalt flere steder i denne artikel. Når man lærer et sprog, har man tværtimod behov for selv at tilegne sig den viden som ordbøger og grammatikker indeholder, såsom bøjninger, syntaks, kongruens m.v. Den viden kan kun indøves ved hjælp af bevidstgørelse om og synliggørelse af mønstrene og undtagelserne i sproget. Bøjningsskemaer for de franske verber er her et oplagt eksempel. Det gælder stadig at der vil være meget vundet ved at integrere ordbøgerne i de redskaber brugerne benytter, men i en læringssammenhæng skal redskaberne ikke skjule sprogenes logik, de skal synliggøre den.

Det kan for eksempel være når de føromtalte skriveprogrammer og e-bogslæsere giver adgang til fulde ordbogsopslag, inklusive fx betydningsangivelser og brugseksempler, så brugeren selv skal vælge den relevante oversættelse i konteksten, men det kan

også være en yderligere integration mellem ordbogsopslag, frasesøgning, relevante grammatiske forklaringer og øvelser. I vores drømmeordbog ville en bruger for eksempel kunne skrive *she is eller she are* i søgefeltet i bedste googlestil og så både få det korrekte svar på det implicitte spørgsmål og mulighed for at læse begrundelsen, øve sig i reglerne og til sidst teste sin indlæring.

Men brugerne skulle stadig kunne skrive til Ordbogen hvis de følte behov for at få svar fra en sprogkyndig. Det kunne fx være hvis opgaven indeholdt et kreativt aspekt som nævnt ovenfor, men det skulle også være fordi vi tror på at jo mere udbredt det bliver i fremtiden at få hjælp til selvhjælp af kunstig intelligens, jo mere vil man værdsætte den (sjældne?) luksus at få hjælp fra et menneske hvor maskinerne ikke slår til.

Litteratur

Abel, Andrea & Christian M. Meyer (2013): The dynamics outside the paper: user contributions to online dictionaries. I: *Proceedings of the 3rd Biennial Conference on Electronic Lexicography (eLex 2013)*. Ljubljana: Trojina, Institute for Applied Slovene Studies/Tallinn: Eesti Keele Instituut, 179-194.

Atkins, B. T. Sue & Michael Rundell (2008): *The Oxford Guide to Practical Lexicography*. Oxford: Oxford University Press.

Bergenholtz, Henning & Mia Johnsen (2005): Log Files as a Tool for Improving Internet Dictionaries. I: *Hermes* 34, 117-141.

Grammatip.com: <http://www.grammatip.com> (maj 2019).

Ordbogen.com: <http://www.ordbogen.com> (maj 2019).

Tarp, Sven, Kasper Fisker & Peter Sepstrup (2017): L2 write assistants and context-aware dictionaries: New challenges to lexicography. I: *Lexikos* 27, 494-521.

Törnqvist, Lars (2010): Brukarmedverkan i webbordböcker. I: Kristinn Jóhannesson, 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: Göteborgs universitet, 383-390. <https://www.cercurius.se/filer/Brukarmedverkan.pdf> (maj 2019).

Write Assistant: <http://writeassistant.com> (juni 2019).

Jacqueline Levin
specialkonsulent, cand.mag.
Ordbogen A/S
Billedskærervej 8
DK-5230 Odense M
jrl@ordbogen.com

Frederikke Ulv
sprogmedarbejder, cand.ling.merc.
Ordbogen A/S
Billedskærervej 8
DK-5230 Odense M
fbu@ordbogen.com

Somaliska tvåspråkiga ordböcker i Norden: utbud och kvalitet

Morgan Nilsson

There is a great need for good Somali dictionaries in the Nordic countries due to a large Somali immigrant population. After a brief general presentation of Somali and its lexicographic tradition, this article first gives a short survey of the existing Somali-Nordic language dictionaries. This is followed by a review of the largest dictionaries with respect to two important parameters: regional variation in the Somali vocabulary and the representation in the dictionaries of vocabulary typical of core school subjects. In summing up, some major areas for improvement are pointed out.

1. Inledning

Somaliskan utgör i dag ett av de största minoritetsspråken i Norden och behovet av högkvalitativa ordboksresurser mellan somaliskan och språken i Norden är därför stort. I akademiska sammanhang får somaliskan däremot väldigt lite uppmärksamhet, inte bara i Norden, utan i världen i stort. En viktig fråga är därmed om de somaliska ordböcker som vi har i Norden är tillräckliga, inte bara rent kvantitativt, utan också kvalitativt. Denna artikel söker utvärdera den rådande situationen.

Som bakgrund presenteras somaliskan kort i avsnitt 2, och i avsnitt 3 sägs lite om den somaliska lexikografitraditionen. I avsnitt 4 följer så en inventering av somaliska ordböcker i Norden och i avsnitt 5 presenteras de ordböcker som valts ut för granskning. I avsnitt 6 undersöks den regionala variationen i ordförrådet och i avsnitt 7 skolämnenas ordförråd. Avsnitt 8 berör e-ordböckernas layout, varpå undersökningen sammanfattas i avsnitt 9. Slutligen

diskuteras i avsnitt 10 vilka utvecklingsbehov som finns för somaliska ordboksresurser i Norden.

2. Somaliskans ställning

Somaliskan är ett kusjitiskt språk inom den afroasiatiska familjen. Det talas av omkring 25 miljoner människor, vilket gör somaliskan till det 9:e största inhemska språket i Afrika och det ungefär 70:e största språket i världen. Språkområdet på Afrikas horn omfattar inte bara Somalia och Somaliland, utan även de sydöstra delarna av Djibouti samt de östra delarna av Etiopien och Kenya. Ytan motsvarar ungefär Sverige, Norge och Finland tillsammans. Somaliskan är officiellt språk i Somalia och Somaliland, medan det är regionalt officiellt språk i den etiopiska Somaliregionen. I Djibouti och Kenya är somaliskans ställning svag och språket används nästan bara i muntlig kommunikation. Se vidare Nilsson (2018a).

Ett stort antal somalisktalande lever även på andra platser i världen. Enligt statistik från SCB och dess nordiska motsvarigheter kan man utgå från att 1 % av befolkningen i Sverige talar somaliska, medan motsvarande siffra för Norge är 0,8 % och för Finland och Danmark 0,4 %. Enligt Skolverkets statistik talar 2 % av barnen i den svenska grundskolan somaliska, vilket gör språket till det tredje största efter svenskan och arabiskan (Nilsson 2018b:7–8).

3. Lexikografisk tradition

Redan innan somaliskan etablerats som skriftspråk skapades något dussin ordböcker mellan somaliskan och europeiska språk. Den äldsta är de Larajasses (1897) *Somali-English and English-Somali Dictionary* på 301 sidor. Här anges uppslagsordens ordklass, substantivens genus, bestämda artikel och pluralform, liksom verbens

transitivitet. Många exempel ges också på ordens användning. Den första somalisk-tyska ordboken är Reinisch (1902) *Die Somali-Sprache. Vol. II. Wörterbuch* med 540 sidor. Här är informationen ännu något rikare än i föregångaren. I tillägg anges ordens betoning liksom kognater i arabiskan, hebreiskan, amhariskan och flera kusjitiska språk.

Den somaliska lexikografin har alltså en ganska lång tradition, men alla de ordböcker som skrevs fram till mitten av 1960-talet är författade av européer. Då de hade skolning i hantverket är några av verken rätt sofistikerade. Numera är de dock föga kända då de inte är skrivna med dagens ortografi.

Under 1960-talets andra hälft växer en standardiserad ortografi fram och somaliskan blir 1973 Somalias officiella språk. Samtidigt inleds också en ny era för den somaliska lexikografin. Sedan 1969 har det producerats fem enspråkiga ordböcker och över hundra tvåspråkiga. Däremot finns ännu ingen ordbok som bygger på en korpus, då den första av dagens tre somaliska korpusar[1] lanserades först 2015.

De flesta moderna ordböcker är skapade av somalier och av resultatet att döma har författarna sällan någon större skolning i lexikografi. Utöver ekvivalenterna anger många publikationer på sin höjd ordklasstillhörighet och noggrannheten brister ofta i en beklaglig omfattning. Stavfel är vanliga, men också en hel del sakfel. Ett fåtal undantag finns dock i form av bra ordböcker.

Den allra första moderna ordboken är en rysk-somalisk-rysk ordbok (Stepanjenko & Osman 1969), enkel till sin utformning, men innehållet håller ändå en hög kvalitet. Den första enspråkiga ordboken (Keenadiid 1976) gavs ut relativt snabbt av den nyinstiftade Somaliska konst- och vetenskapsakademin. Den har en enkel struktur och innehåller inga flerordsuttryck. Den enda grammatiska informationen är substantivens genus.

Arbetet på en mera omfattande enspråkig ordbok fortsat-

1 För mer information om korpusarna se Nilsson (2018b:150–151).

te, men avbröts av inbördeskriget. Det återupptogs senare vid universitetet i Rom och resulterade i *Qaamuuska Af Soomaaliga* (Mansuur & Puglielli 2012) på 945 sidor. Inte heller denna ordbok innehåller några flerordsuttryck, med undantag för partikelverb. Den för somaliskan så vanliga konstruktionstypen med två substantiv i juxtaposition, t.ex. *kubbadda cagta* 'fotboll' (eg. 'bollen foten'), lyser helt med sin frånvaro, precis som alla andra typer av fraser. Ordboken är ändå den hittills mest vetenskapligt utformade och den utgör ett kollektivt arbete som involverat ett stort antal redaktörer. Information ges om ordklasstillhörighet, substantivens genus, verbens transitivitet, böjningsinformation för relevanta ordklasser samt synonymer. Däremot saknas exempel på ordens användning.

Mot bakgrund av somaliskans relativt svaga lexikografiska tradition i modern tid inställer sig förstås frågan hur det är ställt med kvaliteten i de somaliska ordböckerna i Norden.

4. Somaliska ordböcker i Norden

Innan den somaliska invandringen till Norden tagit sin början kring 1990 fanns det knappast något behov av somaliska ordböcker till eller från de nordiska språken. Med invandringen blev behovet snabbt påtagligt, i synnerhet från de nordiska språken till somaliska. Någon kompetens för att hantera denna situation fanns inte från förlagens och samhällets sida. I stället kom olika eldsjälar i den somaliska gruppen att själva utveckla ordböcker för sina landsmäns behov.

4.1. Ordböcker från somaliska till nordiska språk

Att författa en ordbok från sitt modersmål till ett främmande språk innebär en större utmaning än att författa en ordbok till sitt

modersmål och behovet av en sådan ordbok är ofta mindre än från det främmande språket till modersmålet. Då både användarna och ordboksförfattarna främst är somalisktalande har ordböcker från somaliska inte prioriterats. Idag finns endast fyra sådana ordböcker: *Qaamuus Soomaali-Iswiidish* (Soyan 1992), 8 000 ord; *Kort somalisk-svensk ordlista* (Nilsson 2019), 8 800 ord; *Somalisk-dansk ordbog* (Nuur 1996), 7 600 ord; *Somali-suomi sanakirja* (Arale 2017a) 146 sidor.

Ännu finns ingen somalisk-norsk ordbok, men i *Norsk-somali ordbok* (Guuleed 2008) finns ett somaliskt index. De tre elektroniska ordböckerna från finska (Kotus 2017), norska (Lexin 2005) och svenska (Lexin 2006b) möjliggör även sökning på somaliska ord, men ordböckerna är från början skapade med det nordiska språket som källspråk.

4.2. Ordböcker från nordiska språk till somaliska

Från svenska till somaliska finns tre mindre[2] ordböcker skrivna av två olika författare: *Svensk-somalisk ordlista* (Hassan 1991), *Svensk-somalisk ordbok* (Hassan 1997) och *Svensk-somaliskt lexikon* (Badil 1998). De båda författarna har i samarbete även författat den somaliska texten i *Svensk-somaliskt lexikon* (Lexin 2006a) som utgör en av språkversionerna inom det svenska Lexin-projektet. Denna ordbok omfattar 28 500 ord. Förutom i tryckt form finns den även i en nätbaserad version (Lexin 2006b) där den är sökbar i båda riktningarna.

Från danska finns två mindre, lite äldre ordböcker och en nyare mellanstor[3]: *Dansk-somalisk ordbog* (Nuur 1997), *Dansk-somalisk ordbog* (Xayir 1997) och *Dansk-somali ordbog* (Hussein, Macaani & Elmi 2015) med 16 000 ord. Dessvärre finns ingen elektronisk ordbok.

2 Jfr Svenséns (2004:33) grupp 1: 5 000–15 000 lemman.
3 Jfr Svenséns (2004:33f.) grupp 2: 15 000–35 000 lemman.

Från norska till somaliska finns två mindre ordböcker som är utarbetade av samma person: *Norsk-somali ordbok med somali-norsk register* (Guuleed 2008) med 14 700 ord och *Norsk-somalisk illustrert ordbok* (Lexin 2017) med 11 000 ord. Den senare utgör en del i det norska Lexin-projektet och finns även tillgänglig som e-ordbok på webben (Lexin 2005) där den är sökbar i båda riktningarna.

Från finska finns två traditionella, mindre ordböcker: *Suomi-somali sanakirja* (Zinger, Calderon & Axmed 1999) och *Suomi-somali sanakirja* (Arale 2017b). Vidare finns den mellanstora, webbbaserade *Suomi-somali-sanakirja* (Kotus 2017) vid Institutet för de inhemska språken. Här är även de somaliska ekvivalenterna sökbara. I början av 2019 omfattade denna ordbok 24 630 uppslagsord till och med bokstaven T. Någon tryckt version planeras inte.

5. De undersökta ordböckerna

För en kvalitativ undersökning valdes endast ordböcker från de nordiska språken till somaliska eftersom det i den motsatta riktningen finns ytterst få publikationer och alla är mycket begränsade till sitt omfång. För varje källspråk valdes den elektroniska ordboken respektive den största ordboken. För svenskan och finskan innebär det samma ordbok, medan det för norskan finns en tryckt ordbok som är större än den elektroniska och det för danskan saknas en elektronisk ordbok. Därmed kom tre elektroniska (sv, no, fi) och två tryckta ordböcker (da, no) att ingå i undersökningen:

SLexin = *Svensk-somaliskt lexikon* (Lexin 2006b), en elektronisk ordbok med 28 500 uppslagsord,

DaSo = *Dansk-somali ordbog* (Hussein, Macaani & Elmi 2015), en tryckt ordbok med 16 000 uppslagsord,

NLexin = *Lexin: bokmål-somali-bokmål* (Lexin 2005), en elektronisk ordbok med 11 000 uppslagsord,

NoSo = *Norsk-somali ordbok* (Guuleed 2008), en tryckt ordbok med 14 700 uppslagsord,

FiSo = *Suomi-somali-sanakirja* (Kotus 2017), en elektronisk ordbok med 24 630 uppslagsord till och med bokstaven T.

För den kvalitativa undersökningen valdes några aspekter som ofta lyfts fram av somaliska modersmålslärare såsom viktiga för de elever som använder ordböckerna: Hur behandlas den regionala variationen i det somaliska ordförrådet i ordböckerna? I vilken grad har centrala begrepp för skolans ämnen inkluderats? Hur noggrant översätts eller förklaras dessa begrepp? Dessutom bedömdes också användarvänlighet och akribi i största allmänhet.

6. Regional variation i ordförrådet

Mellan de språkliga varieteterna i de tre sinsemellan oberoende geografiska områdena med en somaliskspråkig administration (Somalia, Somaliland och den etiopiska Somaliregionen) råder relativt små skillnader i ljudsystem, formsystem och satsbyggnad. Förutom en regelbunden variation mellan /ɖ/ och /r/ efter vokal kommer de största skillnaderna till uttryck i ordförrådet. Främst rör det sig om vardagsnära ord för t.ex. färger och mat, medan skillnaderna är små inom fackordförrådet (Nilsson 2018a).

Vid granskningen undersöktes i vilken utsträckning ordböckerna ger två alternativa ekvivalenter, hur de i så fall presenteras grafiskt i förhållande till varandra och huruvida man favoriserar nordliga eller sydliga synonymer eller former.

Som utgångspunkt för undersökningen av den regionala variationen i det somaliska ordförrådet togs en ännu opublicerad pilotstudie genomförd vid Göteborgs universitet 2016 med 62 somalisktalande deltagare. Ur denna studie valdes 41 begrepp där somaliskan har två (eller ibland flera) i stort sett likvärdiga synonymer med varierande regional distribution. De undersökta svenska

uppslagsorden listas här alfabetiskt med motsvarande norska, danska och finska uppslagsord inom parentes i de fall då de har en annan form: *apa* (*ape*, *abe*, *apina*), *banan* (*banaani*), *bil* (*auto*), *bröd* (*brød*, *leipä*), *dörr* (*dør*, *ovi*), *fisk* (*kala*), *fladdermus* (*flaggermus*, *flagermus*, *lepakko*), *flicka* (*jente*, *pige*, *tyttö*), *fråga* (*spørsmål*, *spørgsmål*, *kysymys*), *fönster* (*vindu*, *vindue*, *ikkuna*), *grön* (*grønn*, *grøn*, *vihreä*), *gul* (*keltainen*), *handduk* (*håndkle*, *håndklæde*, *pyyheliina*), *hungrig* (*sulten*, *nälkäinen*), *hus* (*talo*), *hyena* (*hyene*, *hyæne*, *hyeena*), *halva/hälft* (*halvpart*, *halvdel*, *puolikas*), *här* (*her*, *tässä/täällä*), *igelkott* (*pinnsvin*, *pindsvin*, *siili*), *katt* (*kat*, *kissa*), *kanna* (*kjele*, *kedel*, *kahvipannu*), *kniv* (*veitsi*), *kyckling* (*kylling*, *kana*), *kök* (*kjøkken*, *køkken*, *keittiö*), *melon* (*vannmelon*, *meloni*), *mjöl* (*mel*, *jauho*), *morot* (*gulrot*, *gulerod*, *porkkana*), *näsduk* (*lommetørkle*, *tørklæde*, *nenäliina*), *pojke* (*gutt*, *dreng*, *poika*), *räv* (*rev*, *ræv*, *kettu*), *röd* (*rød*, *punainen*), *salt* (*suola*), *skjorta* (*skjorte*, *paita*), *skugga* (*skygge*, *varjo*), *smuts* (*skitt*, *skidt*, *lika*), *svar* (*vastaus*), *tisdag* (*tirsdag*, *tiistai*), *titta* (*kikke*, *kigge*, *katsoa*), *törstig* (*tørst*, *tørstig*, *janoinen*), *väg* (*vei*, *vej*, *tie*), *ägg* (*egg*, *æg*, *muna*).

I SLexin och NoSo återfinns samtliga undersökta uppslagsord. I FiSo och DaSo saknas ordet *hyeena/hyæne*, vilket inte förvånar då ordböckerna avspeglar nordiska förhållanden. I FiSo saknas ytterligare fyra uppslagsord då ordboken i nuläget bara sträcker sig till bokstaven T, men hoppet bör vara gott att de ska komma med. I NLexin, som är den minsta av de undersökta ordböckerna, saknas sex av orden: *halvpart*, *hyene*, *skygge*, *spørsmål*, *skitt*, *tørst*.

Uppslagsordet *morot* (*gulrot*, *gulerod*, *porkkana*) kan tjäna som ett representativt exempel. Det är det enda av de undersökta begreppen där samtliga ordböcker ger två synonyma ekvivalenter: SLexin: *dabacase*, *karooto*; DaSo: *karooto* (*dabacasse*); NLexin: *dabacase* (”*kaarooto*”); NoSo: *dabacase; kaarooto*; FiSo: *dabacase*, *karooto*. Två av ordböckerna har dock satt den ena synonymen inom parentes och en av dem dessutom inom citationstecken.

Att man sätter den ena synonymen inom parentes förekommer vid sammanlagt 8 uppslagsord i SLexin, vid 4 uppslagsord i DaSo och vid 2 uppslagsord i NLexin. Det är olyckligt att göra så eftersom det antyder att den ena synonymen skulle ha företräde framför den andra. Parentes används dock inte vid något av de undersökta uppslagsorden i NoSo och FiSo.

En enda ekvivalent ges i NLexin vid 30 av de undersökta uppslagsorden, i DaSo vid 18 uppslagsord och i SLexin vid 14 uppslagsord. Talen är oroande höga och antyder att man inte insett hur stor roll den regionala variationen spelar för att en somalisk ordbok ska vinna gillande och acceptans bland användare från alla delar av språkområdet. De återstående två ordböckerna har nöjt sig med en ekvivalent bara i ett litet antal fall, FiSo vid 7 uppslagsord och NoSo vid 8 stycken. Då de två norska ordböckerna är författade av en och samma person är det glädjande att se en så stor förbättring mellan versionerna. Man frestas att misstänka att instruktionerna för NLexin kan ha stipulerat att endast en ekvivalent skulle anges i normalfallet.

I exemplet ovan innehåller DaSo också en felstavning (*s* dubbeltecknas aldrig i standardsomaliskan) medan NLexin och NoSo innehåller en ovanlig stavning med *-aa-* i stället för det vanligare *karooto*. Det sammanlagda antalet stavfel i de undersökta orden är i DaSo 5, i NoSo 4, i SLexin 3, i NLexin 3 och i FiSo 2. Båda fallen i FiSo rör namnet på en veckodag som ges med liten begynnelsebokstav trots att det finns tydliga regler om motsatsen, något som även bekräftas av korpusdata.

Granskningen, som sammanfattas i tabell 1, ger också vid handen att nordsomaliska ekvivalenter överväger tydligt i SLexin, medan innehållet i DaSo, NLexin och NoSo har en klart sydsomalisk prägel. Endast i FiSo gör innehållet ett neutralt intryck och synonyma ekvivalenter verkar systematiskt ges i alfabetisk ordning.

	SLexin	DaSo	NLexin	NoSo	FiSo
Antal uppslagsord	41	40	35	41	36
2+ ekvivalenter	46 %	45 %	9 %	80 %	81 %
1(1) ekvivalent	20 %	10 %	6 %	–	–
1 ekvivalent	34 %	45 %	86 %	20 %	19 %
stavfel & småfel	5 %	15 %	6 %	7 %	3 %
dominerande	nord	syd	syd	syd	–

Tabell 1: Regional variation i de granskade ordböckerna.[4]

7. Skolordförråd

När det gäller de undersökta ordböckernas behandling av sådana begrepp som är typiska för skolans ämnen är följande citat från Lexin (2005) intressant:

> LEXIN-ordbøkene er laget spesielt for minoritetsspråklige elever i grunnopplæringen. Oppslagsordene er valgt ut spesielt for denne målgruppen.

Frågan är i vilken grad detta löfte verkligen infrias.

För somaliskans del förekommer en del fördomar med avseende på termer och fackordförråd. Man kan få höra att somaliskan är ett så ungt språk att det ännu inte finns någon etablerad terminologi, och även om det finns en del termer så kan man inte använda dem för nästan ingen känner till dem. Naturligtvis ligger det en del i dessa påståenden, men på det stora hela är de överdrivna.

Under åren 1971–1976 producerades läroböcker på somaliska för undervisning i somaliska, matematik, samhällskunskap, historia, geografi, biologi, kemi och fysik till och med gymnasienivå.

4 Procenttalen är beräknade på det antal ord som återfinns i respektive ordbok.

2001 producerade UNICEF/UNESCO nya skolböcker i somaliska, matematik, naturkunskap, samhällskunskap för grundskolans årskurs 1–8. Under de senaste fem-sex åren har Somalia, Somaliland och Etiopien producerat var sin egen serie av läromedel för grundskolan. All ingående terminologi är nästan helt enhetlig över tid och rum. För somaliskan finns alltså en etablerad terminologi som täcker det stoff som behandlas i skolan. Denna terminologi har varit i bruk i över 40 år. Att termerna inte är allmänt kända beror naturligtvis främst på okunskap som i sin tur beror på att bara en mindre del av den somaliskspråkiga befolkningen haft möjlighet att gå i skolan, och ännu färre har gjort det på somaliska. Att mot denna bakgrund utelämna etablerade termer ur ordböcker med motiveringen att nästan ingen känner till dem är definitivt inte rätt väg att gå. Hur många svensktalande känner till exakt vad en *romb* är?

7.1. De 73 undersökta begreppen

Den kvalitativa undersökningen av fackordförrådet i de fem ordböckerna grundar sig på 73 begrepp som valts ur somaliska skolböcker för årskurs 3. Följande ord valdes inom ämnet **naturkunskap**: *björn*, *skalbagge*, *reptil/kräldjur*, *däggdjur*, *groddjur/amfibie*, *gälar*, *fena*, *ryggrad*, *ryggradslös*, *sinne*, *hörsel*, *syn*, *känsel*, *lukt*, *smak*, *känselspröt/antenn*, *förstoringsglas*, *mikroskop*, *kompass*, *Europa*, *Medelhavet*, *ekvator*, *meridian/längdgrad*, *breddgrad*, *klimat*, *fast form*, *flytande form/vätska*, *gas*, *magnet*, *fotosyntes*, *klorofyll*, *grundämne*, *koldioxid*. Inom **matematik** valdes: *addition*, *term*, *addend*, *summa*, *subtraktion*, *minuend*, *subtrahend*, *differens*, *multiplikation*, *multiplikand*, *multiplikator*, *faktor*, *produkt*, *division*, *dividend*, *divisor*, *kvot*, *bråk*, *täljare*, *nämnare*, *ekvation*, *area*, *cylinder*, *kon*, *kub*, *kvadrat*, *ellips/oval*, *passare*, *rektangel*, *romb*, *rätvinklig*, *triangel/trekant*, *fyrkant*. Inom ämnesområdet **språk** valdes orden: *adjektiv*, *plural/flertal*, *pronomen*, *singular/ental*, *substantiv*, *verb*, *preposition*.

Den första frågan är förstås om orden återfinns i ordböckerna. För de ord som förekommer kartlades huruvida de var försedda med en markör som angav det aktuella ämnesområdet och om det på somaliska ges en ekvivalent, en förklaring eller bådadera. Avseende faktainnehållet bedömdes slutligen varje ekvivalent och varje förklaring som antingen fullgod, oprecis eller felaktig. Vidare noterades om det förekom stavfel eller andra formella fel.

7.2. Undersökningens utfall

FiSo utmärker sig genom att innehålla nästan alla de undersökta orden. Endast tre av orden saknas i den del av ordboken som föreligger, nämligen *sammakkoeläin* 'groddjur', *selkärangaton* 'ryggradslös' och *neljäkäs* 'romb'. Nio finska ord hör hemma i den del av ordboken som ännu saknas, men hoppet verkar gott för att de ska komma med. I SLexin saknas 13 ord, i DaSo 22, i NLexin 37, men i NoSo bara 21. Vilka ord det rör sig om redovisas i Nilsson (2019b).

I SLexin och NoSo upptäcktes inga ologiska luckor i naturliga grupper av ord. I NLexin saknas dock *lengdegrad/meridian*, medan *breddegrad* finns med, i FiSo saknas *singulaari*, men *pluraali* finns med, och i DaSo saknas *breddegrad/breddekreds*, *addition*, *subtraktion* och *nævner*, medan *længdegrad*, *division*, *multiplikation* och *tæller* finns med. I några fall saknas det sökta uppslagsordet, men i stället finns en homonym, t.ex. NoSo:s "**produkt et** *sb.* waxyaalaha la soo saaro", vilket avser något som tillverkas och inte den matematiska termen *produkt*.

Genomgående är artiklarna i FiSo särskilt väl utformade. En typisk artikel återges i figur 1. Den innehåller ett tydligt framhävt uppslagsord, uppgift om ordklass, en länk (KS) till morfologisk information om uppslagsordet med ordförklaring på finska, ämnesmarkör (kasv., kem. [botanik, kemi]), tydlig somalisk ekvivalent i fet stil, samt förklaring på somaliska med diskret grå text.

klorofylli [5] *substantiivi* KS

kasv., kem. **cagaariye**
cagaariyuhu waa molikiyuul orgaanig ah oo suuraggeliya footoosintasiska dhirta

Figur 1: Termen *klorofylli* i FiSo.

Överlag är det dock så att ämnesmarkör ofta saknas i ordböckerna. Ibland ges ingen ekvivalent, utan bara en förklaring på somaliska, trots att en fullgod ekvivalent finns. Många gånger ges i stället den somaliska ekvivalenten utan någon förklaring, vilket i sig inte är något märkligt för en ordbok, men i somaliskans fall vore det sannolikt nyttigt med förklaringar vid de flesta termer och fackuttryck, då många är föga kända bland allmänheten.

Andra brister som förekommer är stavfel, t.ex. SLexins *isleeg* för *ekvation* i stället för det korrekta *isle'eg*, eller ordbildningsfel, t.ex. SLexins *saddex jibbaar* ('tre exponenter') som ekvivalent till den geometriska termen *kub* i stället för det korrekta *saddexjibbaarane* (ungefär 'treexponentare'). Slutligen förekommer det, i synnerhet i DaSo, att ekvivalenten placeras inom parentes efter förklaringen, t.ex. "**bjørn** *sub.* -en, -e, -ene - Xayawaan dhogor leh oo ku nool dhalalka qabow (madax kuti)" ['Djur med päls som lever i kalla länder (björn)']. Detta ger intryck av att *madax kuti* inte skulle vara en fullgod ekvivalent, vilket är missvisande.

I flera fall är ekvivalenten inte direkt felaktig, men heller inte idealisk. I NLexin anges *oorso* som ekvivalent till *bjørn*. Detta somaliska ord är inte den vedertagna termen inom naturvetenskapen, utan ett regionalt ord med begränsad spridning. I SLexin anges både *madaxkuti*, som är det standardspråkliga ordet, och *buti*, som är ett annat regionalt ord med begränsad spridning.

Ett tydligt exempel på en helt felaktig ekvivalent finns i SLexin under uppslagsordet *passare*. Som somalisk ekvivalent anges *jiheeye*, vilket betyder *kompass*. Felet kan antas bero på att *passare* på engelska heter *drawing compass*.

När det gäller oprecisa förklaringar utgör termen *gälar* ett gott exempel. I SLexin ges en förklaring på somaliska som i översättning lyder 'plats som fiskar andas på', i NLexin anges 'cell för andning hos fisk', och i DaSo 'hålet som fiskar andas med'. I stort sett fyller nog de här förklaringarna sitt syfte, men de hade kunnat formuleras bättre.

Som exempel på felaktiga förklaringar fungerar termen *kvadrat*. Den förklaring som ges i NLexin innebär 'en fyrkant som är ett parallellogram', och i SLexin 'en fyrkant med raka vinklar'. Dessvärre är dock *parallellogram* en hyperonym till *kvadrat* och en rak vinkel (180°) är något helt annat än en rät vinkel (90°). Sannolikt beror felaktigheterna helt enkelt på bristande fackkunskap hos författarna.

De olika problem som kan iakttas i de olika ordböckerna summeras i tabell 2.

	SLexin	DaSo	NLexin	NoSo	FiSo
Antal ord som finns	60	51	36	52	61
Bara homonym finns	5 %	6 %	6 %	6 %	2 %
Ämnesmarkör saknas	100 %	84 %	97 %	77 %	52 %
Ingen ekvivalent	12 %	8 %	13 %	2 %	2 %
Oprecis ekvivalent	15 %	33 %	8 %	19 %	8 %
Felaktig ekvivalent	5 %	14 %	3 %	2 %	2 %
Ingen förklaring	33 %	69 %	78 %	71 %	59 %
Oprecis förklaring	8 %	14 %	14 %	4 %	0 %
Felaktig förklaring	3 %	2 %	0 %	0 %	0 %
Andra formella brister	12 %	25 %	3 %	2 %	0 %

Tabell 2: Skolordförråd i de granskade ordböckerna.[5]

5 Procenttalen är beräknade på det antal ord som återfinns i respektive ordbok.

Det är särskilt slående att ämnesmarkör saknas i mycket stor utsträckning i samtliga ordböcker samt att NLexin bara innehåller hälften av de undersökta orden. Vidare är SLexin i betydligt högre grad än de andra hjälpsam med ordförklaringar på somaliska, medan FiSo håller den genomgående högsta kvaliteten på innehållet, följd av de norska ordböckerna. SLexin och i all synnerhet DaSo innehåller en oroande mängd felaktig eller oprecis information.

8. E-ordböckernas layout

När det gäller de elektroniska ordböckerna har FiSo en klart mer läsarvänlig layout än SLexin och NLexin. Som framgår av figur 1 har FiSo ett väl avvägt bruk av teckenstorlek, fetstil, kursiv och färg, liksom vita ytor mellan olika element. SLexin utnyttjar inte färger, kursiv och teckenstorlek. NLexin färgmarkerar söksträngen i träffarna och ger grammatisk information i en mindre teckenstorlek, vilket knappast gör artiklarna tydligare. I SLexin används ytterst lite vita ytor mellan olika element, vilket gör artiklarna kompakta och svåröverskådliga, medan NLexin i stället innehåller onödigt mycket tomrum mellan olika element i artikeln. Dessutom visar NLexin alltid ekvivalent och förklaring på engelska före motsvarande information på somaliska, vilket leder till att enkla ordboksartiklar ofta inte ryms på skärmen.

I FiSo visas inte någon finsk förklaring eller morfologisk information om det finska uppslagsordet i den primära vyn. Denna information nås i stället via en knapp som öppnar ett separat fönster. Systemet bidrar starkt till att fokusera på somaliskan då informationen blir mer avskalad och överskådlig.

9. Sammanfattning

FiSo är utan tvekan den av de undersökta ordböckerna som håller högst kvalitet. Den har det bredaste ordförrådet, det mest tillförlitliga innehållet och den tydligaste layouten. NLexin är minst, den innehåller få fackuttryck och ger sällan regionala synonymer. Den privatutgivna NoSo utgör en tydlig utbyggnad av NLexin och är påtagligt mer innehållsrik. Det är beklagligt att författaren inte har haft möjlighet att göra detta arbete inom ramen för en vidareutveckling av Lexin, då den tryckta NoSo nu är betydligt bättre än NLexin, men mera svårtillgänglig. SLexin innehåller störst andel slarvfel, medan DaSo innehåller störst andel faktafel.

Kort sagt visar FiSo i den här studien det bästa resultatet och DaSo det sämsta, medan SLexin, NLexin och NoSo befinner sig däremellan. En uppenbar förklaring till FiSo:s höga kvalitet är att ordboken tillkommit inom ett mångårigt projekt vid Institutet för de inhemska språken där man haft möjlighet att anställa och internt utbilda två somaliska ordboksredaktörer.

10. Utvecklingsbehov

De utvecklingsbehov som direkt följer av genomgången ovan är att SLexin behöver fackgranskas och utökas en del. NLexin behöver granskas och utökas kraftigt, framför allt med fler uppslagsord, men även med mer information om de befintliga uppslagsorden. I båda fallen tycks det befogat med ett särskilt fokus på skolordförrådet, men den regionala variationen behöver också uppmärksammas ordentligt för att undvika favorisering av en viss varietet. För DaSo behövs en omfattande granskning, och en omvandling till elektroniskt format vore önskvärd.

I motsatt riktning, dvs. från somaliska till de nordiska språken är situationen mycket bristfällig. Här vore en ny samordnad

ordbok liknande ISLEX (2011) något att eftersträva. Med definitioner på somaliska och grammatisk information om uppslagsorden skulle den vara användbar även som en enspråkig somaliska elektronisk ordbok, något som i nuläget saknas. En sådan ordbok bör naturligtvis bygga på en tillräckligt stor, representativ korpus, och att utveckla en sådan är därför också synnerligen angeläget. Bland annat behöver korpusen innehålla så mycket som möjligt av befintliga skolböcker och annan facklitteratur för att kunna kartlägga vilka fackuttryck som faktiskt används i modern somaliska.

I en modern ordbok vore det också önskvärt att kunna säga något om regionala somaliska ords geografiska distribution. För att få en objektiv grund för sådan markering skulle crowdsourcing genom enkäter i sociala medier vara en framkomlig väg. Även när det gäller att belägga termer skulle crowdsourcing kunna tillämpas, men då begränsad till fackfolk, exempelvis universitetsstudenter i olika ämnen. Här skulle man kunna tillämpa SIL:s (2019) Rapid Word Collection Methodology där ett antal personer i grupp diskuterar sig fram till ideala ekvivalenter för listor av tematiskt ordnade ord på ett välbekant källspråk, i somaliskans fall lämpligen engelska.

Litteratur

Ordböcker

Arale, Farah (2017a): *Somali-suomi sanakirja*. Espoo: Tarka. 146 s.

Arale, Farah (2017b): *Suomi-somali sanakirja*. Espoo: Tarka. 244 s.

Badil, Sahardid Mahamoud (1998): *Svensk-somaliskt lexikon. Qaamuuska Iswidhishka iyo Soomaaliga*. Stockholm: Författares bokmaskin. 397 s.

de Larajasse, Evangelist (1897): *Somali-English and English-Somali Dictionary*. London: Kegan Paul, Trench, Trübner & Co. 301 s. <hdl.handle.net/2307/5177> (mars 2019).

Guuleed, Cismaan Cabdi (2008): *Norsk-somali ordbok med somali-norsk register*. 4:e uppl. Ås: elexi.no. 491 s.

Hassan, Mohamed Musse (1991): *Svensk-somalisk ordlista. Qaamuuska Afka Iswiidhishka-Soomaaliga*. Angered. 243 s.

Hassan, Mohamed Musse (1997): *Buug'ereyeed Iswiidhish Soomaali. Svensk-somalisk ordbok*. Angered: Abille. 451 s.

Hussein, Abdulahi Mukhtar, Mohamed Abdulahi A. Macaani & Mohamed Hassan Elmi (2015): *Dansk-somali ordbog. Qaamuus deenish-soomaali.* 5. oplag. Herning: Special-pædagogisk forlag. 460 s.

ISLEX (2011) = *ISLEX*. Reykjavík: Stofnun Árna Magnússonar í íslenskum fræðum. <islex.is> (juni 2019).

Keenadiid, Yaasiin C. (red.) (1976): *Qaamuuska Af-Soomaaliga.* Mogadishu & Firenze. 498 s. <hdl.handle.net/2307/2021> (mars 2019).

Kotus (2017) = Liban Ali Hersi, Abdi Awil Musse (somalisk text), Ulla Horstia, Liisa Nuutinen & Riitta Eronen (finsk text). *Suomi-somali-sanakirja*. Helsinki: Kotimaisten kielten keskus. <kaino.kotus.fi/somali> (juni 2019).

Lexin (2005) = *Lexin: bokmål-somali-bokmål / nynorsk-somali-nynorsk*. Oversettelse: Osman Gulled. Oslo: Utdanningsdirektoratet. <lexin.udir.no> (mars 2019).

Lexin (2006a) = *Svensk-somaliskt lexikon. Qaamuuska Iswidhishka iyo Soomaaliga.* Översättning: Sahardid Mahmoud B. & Mohammed sh. Hassan. Stockholm: Institutet för språk och folkminnen. 1 140 s.

Lexin (2006b) = *Lexin: svenska-somaliska.* Översättare: Sahardid Mahmoud B. & Mohammed sh. Hassan. Stockholm: Institutet för språk och folkminnen & Kungliga tekniska högskolan. <lexin.nada.kth.se> (mars 2019).

Lexin (2017) = *Lexin: norsk-somalisk illustrert ordbok / qaamuuska sawirrada leh ee af-Norwiijiga iyo af-Soomaaliga.* Oversettelse: Osman Gulled. 2. utgave. Oslo: Utdanningsdirektoratet. 537 s.

Mansuur, Abdalla & Annarita Puglielli (red.) (2012): *Qaamuuska Af-Soomaaliga.* Rom: Roma Tre Press. 947 s. <hdl.handle.net/2307/720> (mars 2019).

Nilsson, Morgan (2019): *Kort somalisk-svensk ordlista.* Göteborg: Göteborgs universitet. 182 s. <morgannilsson.se/Kort somalisk-svensk ordbok.pdf> (juni 2019).

Nuur, Ali Ahmed (1996): *Somalisk-dansk ordbog. Qaamuus Soomaali-Deenish.* Nuur Konsulent Service. 266 s.

Nuur, Ali Ahmed (1997): *Dansk-somalisk ordbog. Qaamuus Deenish-Soomaali.* 2. oplag. Nuur Konsulent Service. 344 s.

Reinisch, Leo (1902): *Die Somali-Sprache. Vol. II. Wörterbuch.* Wien: Alfred Hölder. 540 s. <hdl.handle.net/2307/1705> (mars 2019).

Soyan, Abdigani M. (1992): *Qaamuus Soomaali-Iswiidish. Somalisk-Svenska.* Sunne. 202 s.

Stepanjenko, D. I. & M. H. Osman (1969): *Abwan urursan Af Soomaali iyo Rusha, Rush iyo Af Soomaaliya. Kratkij somali-russkij i russko-somali slovar'.* Moskva: Sovetskaja entsiklopedija. 319 s.

Xayir, Xersi Faarax (1997): *Dansk-somalisk ordbog. Qaamuus deenish-soomaali.* 2. udgave. Somalisk bøger. 290 s.

Zinger, Mika, Anna-Maija Aquilera Calderon, Warsame Mohamed Axmed (1999): *Suomi-somali sanakirja.* Helsinki: Somalian community in Finland. 143 s.

Annan litteratur

Nilsson, Morgan (2018a): Somali as a pluricentric language: Corpus-based evidence from schoolbooks. I: Rudolf Muhr & Benjamin Meisnitzer (red.): *Pluricentric Languages and Non-Dominant Varieties Worldwide: New Pluricentric Languages – Old Problems*. Frankfurt a.M.: Peter Lang, 75–91.

Nilsson, Morgan (2018b): *Lite om somaliskan.* Göteborg: Göteborgs universitet. <morgannilsson.se/Lite om somaliskan.pdf> (juni 2019).

Nilsson, Morgan (2019b): Dataunderlag till artikeln Somaliska tvåspråkiga ordböcker i Norden. <morgannilsson.se/data/NFL2019.xlsx> (juni 2019).

SIL (2019) = SIL International. *Rapid word collection.* <rapidwords.com> (juni 2019).

Svensén, Bo (2004): *Handbok i lexikografi. Ordböcker och ordboksarbete i teori och praktik.* 2:a uppl. Stockholm: Norstedts Akademiska Förlag.

Morgan Nilsson
universitetslektor, fil.dr
Göteborgs universitet
Institutionen för språk och litteraturer
Box 200
SE-405 30 Göteborg
morgan.nilsson@gu.se

Bokmålsordboka og *Nynorskordboka* – einegga, toegga eller siamesiske tvillingar?

Margunn Rauset

Bokmålsordboka and *Nynorskordboka* represent the two official Norwegian written varieties, Bokmål and Nynorsk. This article discusses whether their current digitized versions should be considered as two separate dictionaries or as one dictionary including two varieties. User feedback indicates that their common digital platform at <ordbok.uib.no> might raise some expectations of features associated with a bilingual dictionary. How does that influence the lexicographer's choices when revising the two dictionaries in parallel in an ongoing five-year project?

1. Innleiing

Dei første trykte utgåvene av dei to norske standardordbøkene *Bokmålsordboka* og *Nynorskordboka* kom i 1986. Er dei, godt og vel 30 år seinare, framleis å rekne som to ordbøker, eller blir dei digitale utgåvene på <ordbok.uib.no> og i appen Ordbøkene etter kvart oppfatta som éin språkleg ressurs på to målformer?

Artikkelen blir innleidd med eit språkhistorisk og språkpolitisk tilbakeblikk, og vi får forklaringa på kvifor desse to ordbøkene har blitt omtala som *tvillingordbøker*. Andre del tek føre seg korleis brukarane oppfattar og nyttar ordbøkene i dag – med særleg vekt på kva vi kan lese ut av kontakt med ordbokbrukarane via e-post og brukarloggar. Siste del handlar om den pågåande revisjonen av *Bokmålsordboka* og *Nynorskordboka*, korleis vi i Revisjonsprosjektet tolkar mandatet vårt, og korleis vi vurderer dei to ordbøkene med tanke på om dei er sjølvstendige ressursar eller éin noko atypisk tospråkleg ressurs.

2. Formålet med dei to ordbøkene

2.1. Tvillingnemninga

Definisjons- og rettskrivingsordbøkene *Bokmålsordboka* og *Nynorskordboka* blei i si tid utvikla av to redaksjonar under same tak ved Universitetet i Oslo, begge i samarbeid med Norsk språkråd. Så tett var det redaksjonelle samarbeidet i starten at bokmålsredaktør Lars Andreas Kulbrandstad omtala bokmåls- og nynorskutgåva som eitt verk: «Ei ny ettbindsordbok blir nå utarbeidd [...]. Boka blir redigert på bokmål og nynorsk samtidig, og den vil bli gitt ut under tittelen «Norsk handordbok»» (Kulbrandstad 1976:7). Manuskripta til bokstavane *a*–*k* og *v* blei først utarbeidde av nynorskredaksjonen, *l*–*u* og *w*–å av bokmålsredaksjonen, og så bytte dei manuskript ved halvgått løp. Slik kunne dei byggje på arbeida til kvarandre (BOB 1986:V).

Sjølv om ordbøkene blei gjevne ut på kvar sine forlag, *Bokmålsordboka* på Universitetsforlaget (seinare på Kunnskapsforlaget) og *Nynorskordboka* på Det Norske Samlaget, fekk førsteutgåvene lik grafisk utforming. Seinare har dei òg følgt kvarandre tett i utvikling og redigering, og dette har gjort at *Bokmålsordboka* og *Nynorskordboka* har blitt kalla *tvillingordbøker*. Nynorskredaktør Dagfinn Worren (1998:59, 63) beskriv dei som to likeverdige ordbøker utforma etter den same malen, og han omtalar arbeidet med dei som eit «toegga tvillingprosjekt».

Worren utdjupar ikkje kva han legg i denne nemninga, men i ein norsk kontekst førestiller eg meg siamesiske ordboktvillingar som heilt parallelle i struktur og innhald, og at skilnadane som finst mellom målformene, berre hadde blitt uttrykte i ortografi og morfologi. Om eit leksem ikkje kan dokumenterast brukt i både bokmål og nynorsk, hadde det heller ikkje blitt eit oppslagsord i den siamesiske tvillingordboka. Eit einegga ordbokpar ville ikkje hatt same parallellitetskrav til lemmalistene, og dei to ressursane

kunne i noko større grad spegla ulikskapane som er i bokmål og nynorsk. Innhaldet ville likevel vere meir prega av alt som er felles mellom målformene, enn det som skil dei. Med ei toegga tvillingordbok meiner eg at dei to ressursane er parallelle i struktur og langt på veg i innhald, men at dei likevel har som føremål å beskrive målformene med sine særeigne trekk, på sjølvstendig grunnlag. Det inneber òg at lemmalistene blir noko ulike.

Nøyaktig kvar lista skal liggje for parallellitet mellom *Bokmålsordboka* og *Nynorskordboka*, er ikkje klart definert, men i eit notat til Kulturdepartementet slår Språkrådet fast at det er viktig at ordbøkene held fram med å bli utvikla parallelt:

> I nettversjonen er det nå sånn at når du søker på et ord, får du treff i begge ordbøkene (hvis ordet finnes i begge), med mindre du eksplisitt ber om søk i bare en av dem. Denne parallelliteten har vært viktig for Språkrådet og viktig for norsk språkpolitikk. Språkrådet ønsker å beholde ordbøkene som «tvillinger». (Språkrådet 2014:10)

I Noreg står *Bokmålsordboka* og *Nynorskordboka* i ein særskild posisjon, då dei er den viktigaste kanalen eigaren Språkrådet har for å formidle informasjon om gjeldande offisiell rettskriving til publikum (Språkrådet 2014:9).[1] Norsk rettskriving var frå starten av 1900-talet prega av hyppige og omfattande reformer. Slike blir vedtekne på departementsnivå. Frå 1950-åra har det vore Norsk språknemnd og etterfølgjarane Norsk språkråd og Språkrådet som har forvalta normene og stått for dei løpande endringane. For desse organa har oppdaterte ordlister vore svært viktige, særleg for å nå fram til skulane på ein effektiv måte (Hovdenak & Ims 2016:157). Det er med andre ord ein lang tradisjon i Noreg

1 I 2016 overtok Universitetet i Bergen (UiB) ansvaret for m.a. *Bokmålsordboka* og *Nynorskordboka* frå Universitetet i Oslo. Ordbøkene blir no drifta av UiB og Språkrådet i fellesskap.

for å bruke ordlister for å spreie informasjon om normeringa av målformene. Historia til *Bokmålsordboka* og *Nynorskordboka* viser at dei òg blei til i ein tydeleg språkpolitisk kontekst, og at dei var tenkte å tene spesifikke språkpolitiske formål.

2.2. Bakgrunnen for at ordbøkene blei til

Gjennom store delar av førre hundreår var den norske språkdebatten oppheta, og det kulminerte i meir og mindre organisert motstand mot den offisielle språkpolitikken slik han blei forvalta av Norsk språknemnd (1952–1972). I hovudtrekk gjekk denne politikken ut på at dei to målformene bokmål og nynorsk skulle sameinast i eitt språk, samnorsk, noko mange sentrale røyster kjempa innbite mot. I alle leirar blei motiva og hensiktene til motstandarane mistenkjeleggjorde (Rambø 1999:40), og mange hadde ikkje tillit til at Norsk språknemnd var organet som kunne bringe utviklinga av norsk språk vidare:

> I begynnelsen av sekstiårene så det nærmest ut til at språkstriden hadde låst seg fast i en ufruktbar stillingskrig mellom myndighetenes samnorsklinje og den private sektors ytterliggående riksmålslinje. (Haugen 1966:228)

Som eit ledd i å roe gemytta blei det sett ned ein komité med rektor ved Universitetet i Oslo, Hans Vogt, som leiar. Vogtkomiteen, eller «Språkfredskomiteen» som han òg har blitt kalla, skulle vurdere den politiske og rettslege sida ved den norske språkpolitikken og gjere framlegg om tiltak som kunne tene til å ta vare på og utvikle den norske språkarven (Innstilling 1966:3). Frå politisk hald var det viktigaste no å ta vare på rikdommen i begge målformer og i dialektene i bygd og by – og slik demme opp for den felles fare som truga norsk utanfrå (Haugen 1966:230), altså engelsk.

Eitt av framlegga til Vogtkomiteen var å opprette Norsk lek-

sikografisk institutt ved Universitetet i Oslo, då komiteen peika på eit stort behov for gode, praktiske eittbindsordbøker for bokmål og nynorsk. Dei viste til at det fanst mange ordlister, så alle nordmenn for ein billig penge kunne slå opp korleis eit ord skulle stavast og bøyast, men ikkje kva ordet betyr, eller korleis det skal brukast (Innstilling 1966:13). Marit Hovdenak (2014:231) har peika på at det òg var ein bakanforliggande motivasjon for utgjevinga at dansk og svensk hadde nyare morsmålsordbøker av liknande type, og i 1977 kom det dessutan ei tilsvarande ordbok for den uoffisielle normalen riksmål, *Riksmålsordboken* (Guttu 1977).

Språkpolitisk var det viktig at det blei slått fast at det skulle utarbeidast to likeverdige allmennordbøker for dei to offisielle formene av norsk skrift. Både argumentasjonen for behovet for handordbøker og den seinare understrekinga av parallellitet i utarbeidinga av *Bokmålsordboka* og *Nynorskordboka* var ein manifestasjon av ein politisk vilje til å vise at målformene, med sine særpreg, er og skal vere jamstilte.

2.3. Kvifor er tvillingane ulike?

Trass i den parallelle redigeringa av førsteutgåvene er det ein god del forskjellar mellom *Bokmålsordboka* og *Nynorskordboka*. Ein grunnleggjande forskjell er at det er ordbøker for kvar si målform: «Norsk har to skriftspråksnormalar som har svært mykje sams, men som òg har kvar sine skriftkulturar» (Hovdenak 2014:233). Det har på ulike måtar prega redigeringa av ordbøkene.

I dei opphavlege planane for prosjektet *Norsk handordbok* var det meininga av dei to ordbøkene skulle bli like i omfang: mellom 600 og 700 sider og rundt 60 000 oppslagsord (Kulbrandstad 1976:8). Dei skulle òg ha med dei same informasjonskategoriane: alle tillatne variantar av ortografien og morfologien til lemmaa, ordklasse, uttale for ein del framandord, etymologi, definisjonar og bruksdøme. For nynorsken sin del var det nødvendig å ta med

ein god del samansetjingar fordi målforma mangla eit verk som svara til *Tanums store rettskrivningsordbok*, som inneheld mange slike på bokmål. Normeringsbehovet i nynorsk var altså større enn i bokmål, og for nynorskutgåva av *Norsk handordbok* var det viktig å få med flest mogleg ord, då dette var den første allmenne eittspråklege ordboka på denne målforma (Worren 1998:61). Dermed var tanken at nynorskutgåva skulle ha med noko fleire samansetjingar og følgjeleg eit høgare tal oppslagsord, medan dei gjennomsnittlege artiklane i bokmålsutgåva skulle vere noko lengre (Kulbrandstad 1976:8).

Det blei tidleg klart at dei to redaksjonane hadde ulike mål for arbeidet. I brukarrettleiinga til *Bokmålsordboka* (1986:VII) står det at boka skal ha med «det sentrale ordforrådet i moderne bokmål». Det blir utdjupa som ord vi møter gjennom radio og fjernsyn, i aviser og allmenne tidsskrift, i skjønnlitteratur og i lærebøker for grunnskulen og den vidaregåande skulen. I utgangspunktet ser det ut til å ha vore tanken for begge utgåvene av *Norsk handordbok*: «Boka skulle langt på vei dekke ordforrådet i moderne norsk slik en møter det i litteratur og presse» (Kulbrandstad 1976:8). I nynorskutgåva forlét ein delvis denne skriftspråksdominerte linja då redaksjonen ønskte å løfte fram den tette relasjonen mellom nynorsk og dialektene, eller *målføra*, som nynorsken byggjer på. Eit viktig poeng her er at det ikkje er noko skarpt skilje mellom dialektene og det nynorske standardspråket, og som ein praktisk regel for å velje ut dialekttilfang tok nynorskredaksjonen med dialektord som var registrerte i tre fylke eller meir (eller to fylke i Nord-Noreg) i ordsamlingane til *Norsk Ordbok* – sjølv om dei er lite brukte i litteraturen (Hovdenak 2014:234).

Worren set dette i samanheng med korleis nynorsk som minoritetsspråk heile tida må legitimere seg i høve til brukarane sine. Den ideologiske grunngjevinga for nynorsken som den skriftlege samnemnaren for målføra krev at ei nynorsk ordbok har med eit breitt tilfang frå dialektene: «[N]ynorsken som alternativt språk

må marknadsføre seg som skriftmålet for målføremerkt ordtilfang» (Worren 1998:63). Dessutan fløymde ei dialektbølgje over Noreg på 1970-talet. Ideologien, ikkje minst i nynorskorganisasjonen Noregs Mållag, var at alle talemålsformer er fullverdige og like gode, uavhengig av tidlegare litterær bruk, og dette var noko særleg nynorskredaksjonen måtte ta omsyn til (ibid.).

Det blei ei utfordring å halde seg til den oppsette planen med 60 000 oppslagsord samstundes som ein både fekk med det særskilde målføretilfanget og det moderne ordtilfanget. Ifølgje Worren såg ein dette likevel ikkje som eit problem i nokon av redaksjonane, for ein tenkte at særbokmålske innlån frå tysk og dansk skulle vege opp for dialektorda i nynorsk. Slik gjekk det ikkje, og *Nynorskordboka* (1986) enda opp med rundt 90 000 oppslagsord, medan *Bokmålsordboka* (1986) hadde 65 000 oppslagsord. Då ein etter datamaskinell teljing i 1984 blei klar over den store differansen, gjekk ein òg vekk frå fellesnemninga *Norsk handordbok* og kalla ordbøkene *Bokmålsordboka* og *Nynorskordboka* (Worren 1998:64).

Vegane skildest ytterlegare åt i seinare trykte utgåver, då dei blei gjevne ut til ulike tider og i ulikt antal utgåver. *Bokmålsordboka* kom i tredje og siste trykte utgåve i 2005, etter den siste rettskrivingsreforma i bokmål same år, medan *Nynorskordboka* kom i fjerde og siste trykte utgåve i 2006.

2.4. Digitale tvillingar?

I dag finst *Bokmålsordboka* og *Nynorskordboka* både i nettversjon og som app. Standardordbøkene kom på nett så tidleg som i 1994, og begge hentar ordklasse og bøying frå Norsk ordbank (Grønvik & Kyrkjebø 2018:83f.). Frå starten låg ordbøkene side om side og med felles søkjegrensesnitt, og slik har det halde fram gjennom ulike oppgraderingar. Det er sjølvsagt eit medvite val frå gamle og nye eigarar av ordbøkene at dei blir presenterte på ei felles digital

plattform: «Enkelt søk i begge ordbøkene er altså standard, av pedagogiske og språkpolitiske grunner. Brukerne får se hva som er likt og ulikt i målformene» (Hoel 2010).

Tabell 1 viser antal søk på nettsida i 2017 og 2018, fordelt på fellessøk i begge ordbøkene, søk berre i *Bokmålsordboka* og søk berre i *Nynorskordboka*. Dei fleste brukarane brukar standardoppsettet og søkjer i begge ordbøkene samstundes:

2017	
Fellessøk	23 132 540
Berre BOB	4 750 263
Berre NOB	6 549 480
Totalt	**34 432 283**

2018	
Fellessøk	22 954 837
Berre BOB	4 356 708
Berre NOB	6 680 883
Totalt	**33 992 428**

Tabell 1: Bruksstatistikk for nettutgåva ordbok.uib.no.

Naturleg nok blir innhaldet i ordbøkene i større grad samanlikna når ein får dei opp ved sida av kvarandre på ein skjerm, enn då brukarane måtte slå opp i to trykte bøker for å sjekke kva som er likt og ulikt. Den digitale parallelliteten har fungert så godt at vi skal sjå i del 3 at brukarane oppfattar ordbøkene som éin ressurs i større grad enn redigeringa av dei i si tid tok høgde for.

Den viktigaste tekniske nyvinninga dei siste åra er utviklinga av appen Ordbøkene for Android og iOS. Sidan lanseringa i september 2017 har over 80 000 brukarar lasta ned appen, og i 2018 hadde han 3400–3800 daglege brukarar på tvers av plattform. Som i nettutgåvene kan ein velje kva målform ein ønskjer å få treff på, jf. øvste linja i figur 1. Merk likevel at fellessøket ikkje er standardoppsett her, men éin av tre likestilte søkjemodus.

I 2018 blei det utført 11,4 millionar søk i appen. Eit gledeleg trekk ved appbruken er at han legg seg på toppen av dei om lag 34 millionar søka på nettsida ordbok.uib.no i 2018, eit tal som heldt seg nokså stabilt samanlikna med året før, jf. tabell 1. Iallfall så langt ser det ut til at appen ikkje utgjer ein konkurrent til nett-

Figur 1: Tilskoren skjermdump frå appen Ordbøkene i Android.

sida, men at ordbøkene når nye og andre brukarar med appen. Ser vi på nettsida og appen samla, blir det i gjennomsnitt utført nærare 125 000 søk i dei digitale ordbøkene i døgnet, og ordbøkene gjekk frå å ha 34,5 millionar søk i 2017 til 45 millionar søk i 2018.

Dessverre har vi ikkje oversikt over om appbrukarane søkjer annleis enn brukarane av nettsida. Det kan tenkjast at det mindre formatet på skjermen gjer at ikkje like mange vel å få søkjeresultat på begge målformer som tilfellet er på nett. At fellessøket ikkje er standardoppsett i appen, kan òg spele inn.

Ei mogleg utvikling av ordbøkene er å leggje inn relasjonar mellom lemmaa som gjer at eit fellessøk på bokmålsforma *deltakelse* òg gjev treff på nynorskekvivalenten *deltaking*, og at søk på nynorskforma *ause* òg gjev bokmålsekvivalenten *øse*. Det er noko som blir vurdert, og som ingen så langt har ytra prinsipiell motstand mot. Ein slik funksjon ville gjort ordbøkene meir som ei tospråkleg ordbok og fått fram parallellitet òg når det ikkje er ortografisk likskap.

3. Brukarane: søkjeloggar og brukarkontakt

Søkjestatistikkane og kontakten vi har med brukarane, viser at dei digitale utgåvene av *Bokmålsordboka* og *Nynorskordboka* i høg

grad blir brukte til å slå opp i begge samstundes og å samanlikne innhaldet i dei. Brukarkontakten gjev eit godt innblikk inn i korleis forskjellar og lakuner i ei av ordbøkene blir oppfatta av mange brukarar.

3.1. Brukarspørsmål til nettsida

Nedst på sida til nettordbøkene står det ei e-postadresse brukarane kan vende seg til, og det er Språkrådet som har ansvar for å svare på spørsmåla og tilbakemeldingane som kjem dit. Eg fokuserer på innsende brukarspørsmål som belyser tvillingstatusen til ordbøkene, altså i kva grad *Bokmålsordboka* og *Nynorskordboka* blir oppfatta som sjølvstendige eller som éin ressurs.

Forkortelsen d.m. - ikke gangbar på bokmål?

to 21.03.2019 00:14

Til: 'ordbokene@uib.no';

Hei!

Sært spørsmål: Stemmer det at d.m. er ok som forkortelse for "denne månaden" på nynorsk, men at den ikke kan brukes på bokmål?

Figur 2: Brukarspørsmål om oppslagsord i *Nynorskordboka* som ikkje står i *Bokmålsordboka*.

Utgangspunktet for brukaren i figur 2 er at hen har skrive inn forkortinga *d.m.* som fellessøk i dei digitale ordbøkene og fått treff i *Nynorskordboka*, men ikkje i *Bokmålsordboka*.[2] Brukaren blir usik-

2 Rett nok har artikkelen **d.m.** 'denne måned' blitt oppretta i ettertid i *Bokmålsordboka*, som eit resultat av denne tilbakemeldinga.

ker på om det betyr at forkortinga dermed ikkje er brukande i bokmål, og skriv inn for å få oppklart korleis den manglande oppføringa skal tolkast.

Dette er ein type reaksjon på ordboklakuner som går igjen i mange brukartilbakemeldingar. Dersom éi ordbok har med eit lemma, med ikkje den andre, blir mange usikre på om det betyr at det er restriksjonar i målforma mot å bruke ordet. I ein situasjon som den norske, der det er lang tradisjon for å bruke ordlister og ordbøker for å fortelje språkbrukarane kva som til ei kvar tid er «innanfor» og «utanfor» i to rettskrivingsnormalar som har vore gjennom fleire omfattande endringar, er det til ei viss grad forståeleg.

Andre flittige brukarar av ordbøkene har fått med seg at manglande oppføringar ikkje nødvendigvis betyr det same som at ord og uttrykk ikkje er brukande i den målforma som ikkje har oppføring, men sender melding om ymse dei meiner opplagt manglar i den eine eller andre ordboka. Følgjande brukarspørsmål handlar om ein frase i *Bokmålsordboka* som ikkje står i *Nynorskordboka*: «I den elektroniske Nynorskordboka finn eg ikkje dette uttrykket: Ta mål av seg til å + infinitiv. Kvifor ikkje? Kjem det med i Nynorskordboka, slik det står i Bokmålsordboka?» Spørsmålet illustrerer korleis ordbokbrukarane gløttar til den andre ordboka om dei ikkje finn det dei er på utkikk etter, i den ordboka dei primært leita etter noko i. At brukaren tek seg bryet med å skrive ein e-post, kan moglegvis vere eit teikn på at hen ønskjer at frasen skal stå i begge ordbøkene for at vedkomande kan kjenne seg heilt trygg på at han er gangbar mynt i begge målformene.

For leksikografane er det alltid ei vurderingssak kva kollokasjonar og faste frasar ein tek med, og det må dessutan seiast at det empiriske grunnlaget for utvalet i førstegongsredigeringa på 1970- og 1980-talet var veldig avgrensa samanlikna med dei store korpusa som er tilgjengelege i dag. Med den jamførande praksisen brukarane legg for dagen, er det god grunn til å velje dei same kollo-

kasjonane og faste frasane i *Bokmålsordboka* og *Nynorskordboka* – sjølvsagt så lenge det er empirisk grunnlag for det i målformene.

Ein type brukarspørsmål det kjem fleire av enn ein kanskje skulle tru, er av meir allmenn art. Ein brukar var på jakt etter nynorske ekvivalentar til to bokmålske kollokasjonar: «Eg studerer pedagogikk, og lurer på kva den mest fornuftige omsetjinga av *klassisk betinging* og *operant betinging* er. Bøkene mine er på bokmål, og eg skriv nynorsk.» *Klassisk betinging* og *operant betinging* står heller ikkje i *Bokmålsordboka*, men studenten vender seg til e-postadressa til ordbøkene i håp om å få eit kvalifisert svar på kva variantar som kan brukast på nynorsk for bokmålsterminologi hen har møtt i læreboka si på bokmål. Brukaren ber ikkje spesifikt om at kollokasjonane kjem inn i ordbøkene, men det kan tenkjast at vedkomande treng hjelp i ein eksamenssituasjon eller til oppgåveskriving.

Spørsmåla over illustrerer at mange brukarar har ei rimeleg vid forståing av kva ei ordbok er og kan hjelpe dei med. For ordbokredaksjonen i Bergen er det til uvurderleg hjelp at det er Språkrådet som sit i førstelinjetenesta og svarer på alle slike spørsmål frå ålmenta, medan redaksjonen kan konsentrere seg om å fornye og forbetre innhaldet i ordbøkene.

3.2. Brukarloggar som frå tospråklege ordbøker?

Brukarloggane frå nettordbøkene har ein litt annan profil enn kva ein kanskje kunne vente av to eittspråklege morsmålsordbøker. Frå andre språk veit vi at dei informasjonskategoriane som oftast blir søkte etter i L1-ordbøker, er tydingar, staving, uttale og synonym – og når det gjeld tydingane, er det dei «vanskelege» orda brukarane slår opp, ikkje dei vanlege (referert hos Svensén 2004:549ff.). Difor er det noko overraskande at det blant dei 50 mest frekvente søkjeorda i *Bokmålsordboka* og *Nynorskordboka* er nær 80 % heilt vanlege norske verb med både bokmålsk og nynorsk ortografi: *skrive*, *vere*, *være*, *bli*, *gi*, *komme*, *bruke*, *finne, sjå*, *se* og *gå*. Blant substan-

tiva finn vi *ordbøker*, *menneske*, *nynorsk*, *tekst*, *bok* og *bilde*. Det kan tenkjast at *ordbøker* og *nynorsk* hamnar blant topp 20-søkjeorda fordi søk på desse orda i Google og andre søkjemotorar fører språkbrukarane til ordbøkene. Dei to bokmålspronomena *jeg* og *de* er dei einaste frå den ordklassa på topp 50-lista. På topp 100-lista over dei mest frekvente søkjeorda har delen av verb minka til 65 %, vel 20 % er substantiv. Det dukkar dessutan opp nokre ord frå andre ordklasser: subjunksjonen *hvis*, adjektiva *liten* og *mye*, adverba *bare*, *hvor* og *hvordan* og determinativa *mange* og *dette*. Ikkje eit einaste ord blant dei 100 mest frekvente søkjeorda kan reknast som framandord i norsk.

At så mange søkjer på mykje brukte (sterke) verb og det ein må rekne som kjernevokabularet i eit språk, kan peike i retning av at fleire brukar dei norske standardordbøkene til skriftleg produksjon enn til resepsjon (jf. Svensén 2004:15f.). Truleg er det for å sjekke bøyingane at mange slår opp høgfrekvente verb, og det er vanlegare blant brukarar av L2-ordbøker og L2→L1-ordbøker å søkje grammatikkopplysingar enn det er blant brukarar av L1-ordbøker (Svensén 2004:549).

Ei forklaring kan vere at standardordbøkene i dag har fleire brukarar som ikkje har norsk som morsmål, enn det som var tilfellet for nokre tiår tilbake. Frå andre undersøkingar veit vi at bruken av norske ordbøker heng saman med morsmålet til elevane: Elevar med eit anna førstespråk enn norsk nyttar digitale ordbøker hyppigare enn elevar med norsk som førstespråk (Hovdenak & Ims 2016:160). Det kan òg vere ei forklaring på kvifor fleirtalet av dei som kommenterer og gjev stjerner til appen Ordbøkene, har namn som ein tradisjonelt ikkje har oppfatta som norske, slik ein mellom anna kan sjå det på kommentarane til appen i Google Play.

Ein annan samanheng kan vere at fleire slår opp på sidemålet enn på hovudmålet sitt, som er den målforma ein vel som opplæringsspråk og primært bruksspråk. Maks 15 % av den norske befolkninga har nynorsk som hovudmål, men likevel viste tabell 1 at

det er *Nynorskordboka* som blir hyppigast slegen opp i blant dei som vel å berre søkje i ei av ordbøkene.

Tabell 2 viser kor ofte brukarane klikkar seg vidare og ser på bøyingsopplysingane i dei to ordbøkene.

Søk i bøyingsparadigme	2018
Bøying *Bokmålsordboka*	2 508 582
Bøying *Nynorskordboka*	5 849 577

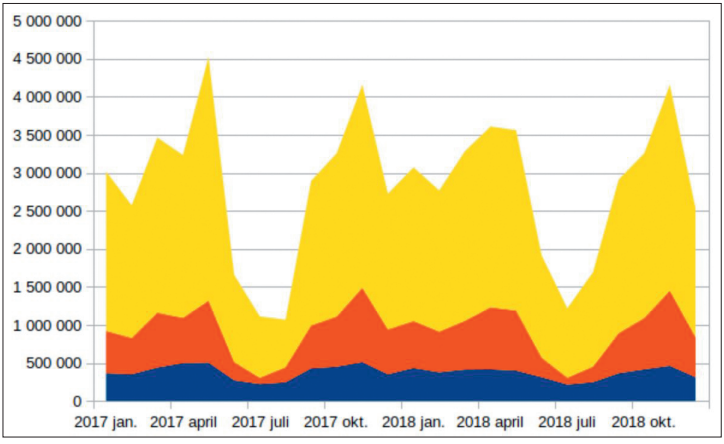
Tabell 2: Søk i bøyingsparadigma til BOB og NOB.

Meir enn dobbelt så ofte sjekkar brukarane dei morfologiske opplysingane i nynorsk som i bokmål. Det kan òg slå ut på statistikken at nynorskbrukarane, som dagleg blir språkbada i bokmål, slår opp oftare enn bokmålsskrivande:

> Det kan òg sjå ut til at lærarar ved skular som har nynorsk som opplæringsmål eller hovudmålform, i større grad nyttar digitale ordbøker i undervisninga enn lærarar på skular der bokmål er mest nytta. (Hovdenak & Ims 2016:161)

Figur 3 viser at bruken varierer sterkt gjennom året. Tala i venstremargen er antal søk i nettordbøkene. Det gule feltet viser søk i begge ordbøkene. Søk i *Nynorskordboka* er markerte med oransje, og søk i *Bokmålsordboka* er markerte med blått. Legg merke til at mengda søk i *Bokmålsordboka* held seg atskilleg meir stabil enn mengda søk i *Nynorskordboka* og i begge. Særleg bruken av *Nynorskordboka* toppar seg markant i eksamensmånaden mai og i heildagsprøvemånaden november.

Figur 3 kan vanskeleg tolkast på andre måtar enn at svært mange elevar brukar ordbøkene når dei skriv eksamen og heildagsprøve på sidemålet sitt. Sjølv om elevar med nynorsk som hovudmål slår opp oftare enn bokmålselevane, er dei i utgangspunktet så mykje færre enn elevane med bokmål som hovudmål at det i seg



Figur 3: Bruken av nettsida gjennom året i 2017 og 2018. Kjelde: Ordbokstatistikk (2019).

sjølv ikkje kan forklare svingingane i bruken av *Nynorskordboka* i prøvemånadane.

4. Revisjonsprosjektet

Revisjonen av *Bokmålsordboka* og *Nynorskordboka* (Revisjonsprosjektet) starta opp for fullt ved Universitetet i Bergen 1. september 2018. I 5 ½ år skal dei to ordbøkene reviderast frå A til Å, og for første gong skal dei same redaktørane arbeide både med bokmål og med nynorsk.

Prosjektet skal sikre at ordbøkene beheld posisjonen sin som fasiten for ortografi og morfologi i bokmål og nynorsk, som er dei sidene ved det norske språket som Språkrådet har normeringsansvar for (Språkrådet 2015:13). Dei viktigaste oppgåvene til ordbokredaksjonen blir å revidere definisjonane så dei er i tråd med dagens språkbruk, gjere dei to ordbøkene likare i omfang og innhald, og ta inn nye ord og tydingar.

4.1. Mandatet til prosjektet

Bokmålsordboka og *Nynorskordboka* blei utvikla i eit tolv år langt toegga tvillingprosjekt, der ordbøkene i kraft av seg sjølve skulle vise at målformene er likeverdige. Med vektlegginga av dialektorda i *Nynorskordboka* ser ein samstundes at det har vore viktig for nynorskredaksjonen å få vist fram korleis nynorsken skil seg frå bokmålet, òg for å gje legitimitet til målforma. Med den store differansen i talet på oppslagsord og ulike språklege profilar er det markante skilnadar mellom dei to standardordbøkene slik dei ligg føre i dag.

Den felles digitale plattforma gjer at innhaldet i ordbøkene blir samanlikna på ein heilt annan måte enn ein kunne føreseie på 1970- og 80-talet, og vi har sett døme på korleis brukarane tolkar meir eller mindre tilfeldige lakuner i ei av ordbøkene som reelle målformforskjellar. Spørsmålet no er om *Bokmålsordboka* og *Nynorskordboka* skal halde fram med å vere to sjølvstendige ordbøker, eller om redaksjonen skal tenkje meir på dei som éi ordbok med to målformer, slik det i utgangspunktet var tenkt med to utgåver av det felles verket *Norsk handordbok*.

Det finst gode argument for at *Bokmålsordboka* og *Nynorskordboka* bør behaldast og vidareutviklast som to ulike verk. Språkpolitisk er det truleg mindre viktig å eksponere parallellitet no enn før. Den offisielle tilnærmingspolitikken, som stegvis skulle fjerne forskjellane mellom målformene, er for lengst forlaten. I praksis skjedde det med innstillinga til Vogtkomiteen, men det blei først formelt vedteke i 2002. I dag er den offisielle politikken at målformene skal normerast uavhengig av kvarandre, utan mål om at dei skal bli likare. Det har vore nokså djuptgripande normendringar i begge målformene etter førstegongsredigeringa av ordbøkene, og særleg i bokmål har endringane delvis gått i motsett retning av tilnærming.

Likevel er det bestemt at det er ei viktig revisjonsoppgåve å gjere ordbøkene likare kvarandre. I samarbeidsavtalen mellom Universitetet i Bergen og Språkrådet frå 2016 blir det slått fast at ordbøkene for bokmål og nynorsk så langt det er mogleg, skal følgje kvarandre i innhald og utviklast parallelt. Dei umotiverte «hola», lemma som finst i den eine ordboka, men ikkje i den andre, skal tettast, og påfallande diskrepansar mellom korleis eitt og same ord er definert i dei to ordbøkene, skal jamnast ut (Språkrådet 2017:4). Det er ingen grunn til at lemmaet *blåbær* skal ha ulike tydingar i dei to ordbøkene, og det gjeld det store fleirtalet av oppslagsord. I nye artiklar utarbeider ikkje redaktørane éin definisjon basert på bokmålsmaterialet og ein annan basert på nynorskmaterialet, men ein felles definisjon blir utforma på begge målformer, vel å merke så lenge det er grunnlag for å opprette artikkelen i begge ordbøkene.

I staden for å snakke om parallellitet som eit språkpolitisk mål skildrar Språkrådet i prosjektsøknaden til Kulturdepartementet parallell redigering av ordbøkene som eit fornuftig økonomisk grep:

> BOB og NOB er og skal være tvillingordbøker, dvs. at de skal være parallelle i innhold bortsett fra når forhold som er særegne for den ene eller den andre målformen, tilsier noe annet. For å sikre framdriften i revisjonsprosjektet er det avgjørende at arbeidet skjer parallelt i de to ordbøkene. Vi legger til grunn at prosjektet rekrutterer fagpersoner med bakgrunn i nordisk språkvitenskap og tilstrekkelig kompetanse i både bokmål og nynorsk. (Språkrådet 2017:7)

Når det er sagt, skal redaksjonen sjølvsagt ikkje gjere målformene likare kvarandre, sjølv om ordbøkene blir det. Både bokmål og nynorsk er slitesterke bruksspråk, og nynorsken treng ikkje drag-

hjelp i form av ei ordbok for å ha legitimitet. Det er ein fin tanke at folk frå heile landet kan finne igjen målføreord frå sitt område i nynorsken, men samstundes har redaksjonen avgrensa kjeldegrunnlag for å vurdere kva som faktisk er i munnleg bruk rundt om i dag. Det kan diskuterast om det gagnar nynorsken, som den minst brukte og mest pressa målforma, om *Nynorskordboka* blir ein oppsamlingsstad for eit målføretilfang dei færraste kjenner seg igjen i.[3] Trass alt er mandatet til Revisjonsprosjektet å dokumentere dei to *skrift*språka: «[Ordbøkene skal vere] kjelde til kvalitetssikra og oppdatert informasjon om normert rettskriving og om vanleg språkbruk for skriftspråka nynorsk og bokmål» (Språkrådet & Universitetet i Bergen 2016:1, pkt. 3.1). I den samanhengen er det òg eit viktig poeng at *Norsk Ordbok* i mellomtida er fullført som trykt verk, og det er dit brukarane primært bør vende seg om dei er interesserte i målføretilfanget. *Norsk Ordbok* utgjer saman med *Det Norske Akademis ordbok* dei dokumenterande ordbøkene for bokmål og nynorsk, og tvillingordbøkene kan i dag i høgare grad enn i 1986 konsentrere seg om kjernen i dei moderne norske skriftspråka.

Det står ikkje skrive nokon stad at særprega til dei to ordbøkene skal vidareførast, men av respekt for dei ulike tradisjonane og historiene målformene er berarar av, er det ein fin balansegang å ivareta målformskilnadane slik dei har blitt presenterte i ordbøkene, og prinsippet Språkrådet har knesett: «[I] prinsippet bør de fleste oppslagsord kunne gjenfinnes både i NOB og BOB» (Språkrådet 2017:3). Det ligg i korta både at gammalmodig litterært tilfang (til dømes *amtmanninne*) kan forsvinne ut av *Bokmålsordboka*, og at noko dialekttilfang med svakt skriftspråksbelegg (til dømes *andkræmeleg*/*andkræneleg*) forsvinn frå *Nynorskordboka*. Det har redaksjonen støtte for i søknadsteksten, sjølv om talemålsbelegg òg kan vektleggjast:

3 Problemet er ikkje at artiklane tek opp plass i ei digital ordbok, men tida det tek for redaktørane å vurdere svakt dokumentert, gammalt talemålstilfang og finne belegg for å oppdatere artiklane, kan gå på kostnad av å opprette nye artiklar som gjerne er meir aktuelle for fleire.

> Oppslagsord som mangler tilstrekkelig tilfang i nyere korpus av norsk tekst, kan vurderes tatt ut av ordbøkene. [...] I tillegg til korpusfrekvens må attesterte talemålsbelegg tas med i betraktning. Merk at «stryking» i denne sammenhengen ikke betyr at oppslagsordet blir slettet fra basene, bare at artikkelen ikke blir publisert. (Språkrådet 2017:6)

Ein skal vere forsiktig med å tukle med arvematerialet, og *Bokmålsordboka* og *Nynorskordboka* både har vore og kjem til å vere toegga tvillingar. Mandatet til Revisjonsprosjektet er å gjere ordbøkene likare i omfang og innhald, men likevel skal vi ikkje framstille målformene som likare enn korpusbelegg av nyare skrift kan dokumentere at dei er. Sjølv om redaksjonen i samråd med eigarane skulle kome til at vi ønskjer å kople ordbøkene tettare saman med ekvivalensrelasjonar, vil begge ordbøkene likevel stå støtt på eigne bein.

Litteratur

Ordbøker

BOB (1986–) = *Bokmålsordboka*. Oslo: Kunnskapsforlaget. <ordbok.uib.no> (mars 2019).

Det Norske Akademis ordbok (2017). Oslo: Kunnskapsforlaget. <naob.no> (mars 2019).

Guttu, Tor (hovedred.) (1977): *Riksmålsordboken*. Oslo: Kunnskapsforlaget.

NOB (1986–) = *Nynorskordboka*. Oslo: Det Norske Samlaget. <ordbok.uib.no> (mars 2019).

Norsk Ordbok. Ordbok over det norske folkemålet og det nynorske skriftmålet (2016). Oslo: Det Norske Samlaget. <no2014.uib.no> (mars 2019).

Tanums store rettskrivningsordbok (1964–). Oslo: Kunnskapsforlaget.

Annan litteratur

Grønvik, Oddrun & Rune Kyrkjebø (2018): Språksamlingane og leksikografisk arbeid ved Universitetet i Bergen. I: Ásta Svavarsdóttir, Halldóra Jónsdóttir, Helga Hilmisdóttir & Þórdís Úlfarsdóttir (red.): *Nordiske Studier i Leksikografi* 14. Reykjavík: Nordisk Forening for Leksikografi, 80–86.

Haugen, Einar (1966): *Rikksspråk og folkemål. Norsk språkpolitikk i det 20. århundre.* Omsett av Dag Gundersen. Oslo: Universitetsforlaget.

Hoel, Jan (2010): Edle årganger i ny tapning. I: *Språknytt* 3. <https://www.sprakradet.no/Vi-og-vart/Publikasjoner/Spraaknytt/Arkivet/Spraknytt-Spraknytt-2010/32010/Edle-arganger-i-ny-tapning/> (mars 2019).

Hovdenak, Marit (2014): Bokmålsordboka og Nynorskordboka gjennom ein generasjon. I: Ruth Vatvedt Fjeld & Marit Hovdenak (red.): *Nordiske studier i leksikografi* 12. Oslo: Nordisk forening for leksikografi, 229–246.

Hovdenak, Marit & Ingunn Indrebø Ims (2016): Bruk av digitale ordbøker i norsk skule. I: Asgerd Gudiksen & Henrik Hovmark (red.): *Nordiske Studier i Leksikografi* 13. København: Nordisk Forening for Leksikografi, 155–164.

Innstilling (1966): *Innstilling om språksaken fra komitéen til å vurdere språksituasjonen m.v. oppnevnt ved kongelig resolusjon 31. januar 1964*. Oslo: Kirke- og undervisningsdepartementet. <https://www.nb.no/items/URN:NBN:no-nb_digibok_2007041801131> (mars 2019).

Kulbrandstad, Lars Anders (1976): Norsk handordbok. I: *Språknytt* 2, 7–8.

Ordbokstatistikk (2019). Universitetet i Bergen. <ordbok.uib.no/perl/ordbokstatistikk.cgi> (mars 2019).

Rambø, Gro-Renee (1999): *Bokmålsreformen i 1981 – med særlig vekt på Særutvalgets arbeid*. Kristiansand: Høgskolen i Agder.

Revisjon av *Bokmålsordboka* og *Nynorskordboka* (2019): Revisjonsprosjektet. <www.uib.no/revisjonsprosjektet> (mars 2019).

Språkrådet (2014): *Norsk ordbokpolitikk. Notat fra Språkrådet til Kulturdepartementet 3. mars 2014*. <https://www.sprakradet.no/upload/Aktuelt/Notat%20om%20norsk%20ordbokpolitikk.pdf> (mars 2019).

Språkrådet (2015): *Retningslinjer for normering av bokmål og nynorsk*. <www.sprakradet.no/globalassets/spraka-vare/norsk/retningslinjer-for-normering-nynorskversjon.pdf> (mars 2019).

Språkrådet (2017): *Prosjektsøknad – revisjon av Bokmålsordboka og Nynorskordboka*. Upublisert.

Språkrådet & Universitetet i Bergen (2016): Samarbeidsavtale. Upublisert.

Svensén, Bo (2004): *Handbok i lexikografi. Ordböcker och ordboksarbete i teori och praktik*. Andra upplagan. Stockholm: Norstedts Akademiska Förlag.

Worren, Dagfinn (1998): Om å avgrense eit ordtilfang – soga om målføreorda i Nynorskordboka. I: Ruth Vatvedt Fjeld & Boye Wangensteen (red.): *Normer og regler. Festskrift til Dag Gundersen 15. januar 1998*. Oslo: Kunnskapsforlaget, 59–70.

Margunn Rauset
hovudredaktør for revisjonen av *Bokmålsordboka* og *Nynorskordboka*
Universitetet i Bergen
Postboks 7805
NO-5020 Bergen
margunn.rauset@uib.no

Kva bruker vi minoritetsspråksordbøker til? Ein studie av brukarloggane for tolv tospråklege ordbøker

Trond Trosterud

The article analyses the user logs of twelve different bilingual minority language dictionaries. The logs show that the dictionaries are in frequent use, with lookup in the direction towards the minority language (production use) dominating by approximately 60% of the total queries. Query recall is high, around 90%, despite the dictionaries being medium-size to small, and most failed queries are due to orthographical errors. The dictionaries translating into the minority language are used for text production to a varying degree. This is clearly shown in the profile of the most commonly used query words.

1. Introduksjon

Artikkelen analyserer brukarloggen til tolv ulike e-ordbøker under ordbokportalen Neahttadigisánit for til saman fem ulike minoritetsspråk: nordsamisk, sørsamisk, enaresamisk, skoltesamisk og kvensk.

Kapittel 2 presenterer kort dei fem minoritetsspråka og andre tilgjengelege e-ordbøker for desse språka. Kapittel 3 presenterer ordbokportalen Neahttadigisánit og dei ordbøkene som blir handsama i denne artikkelen. Kapittel 4 viser korleis bruk av ordbøkene blir loggført, og gjev ein analyse av loggdata for dei tolv e-ordbøkene frå 2018, med vekt dels på dei mest frekvente oppslagsorda og dels på karakteristiske drag ved dei mislukka søka. Til slutt kjem ein konklusjon.

2. Bakgrunn

2.1. Dei fem minoritetsspråka som blir handsama

Temaet for artikkelen er ordbøker for fem ulike minoritetsspråk i to ulike land, Noreg og Finland. Tabell 1 viser kor mange talarar dei ulike språka har ifølgje Ethnologue, og kor livskraftige dei er målt etter EGIDS-skalaen. EGIDS-verdiane i tabellen er som følgjer: 2 = Provincial, 6b = Threatened, 7 = Shifting. For nærare presentasjon av skalaen, sjå EGIDS; for ei drøfting, sjå Lewis & Simons (2010).

	Noreg			Finland		
Språk	nord-samisk	sør-samisk	kvensk	nord-samisk	enare-samisk	skolte-samisk
Talarar	20 000	500	5000	1700	300	300
EGIDS	2	6b	7	6b	2	7

Tabell 1: Talarar og poeng på EGIDS-skalaen (Kjelde: Ethnologue).

Sør- og nordsamisk er vestsamiske språk, mens enare- og skolte-samisk er austsamiske. Som grannespråk er likevel nord- og enaresamisk meir eller mindre innbyrdes forståelege. Sør- og skolte-samisk ligg lenger unna. Alle dei fem språka har ulike ortografiske prinsipp. Sørsamisk har digrafar etter mønster av norsk, mens dei andre samiske språka bruker diakritiske teikn. Kvensk har same ortografiske prinsipp som finsk, men skil seg frå finsk både morfofonologisk og leksikalsk. Dei ortografiske prinsippa for enaresamisk er i hovuddrag dei same som for finsk, men også her er det skilnader. Det er med andre ord ikkje mogleg å skrive dei samiske språka med ortografiske reglar frå dei ulike majoritetsspråka.

2.2. Andre e-ordbøker

I tillegg til e-ordbøkene som blir analysert her, har brukarane for eitt av språkpara tilgang til fleire e-ordbøker.

For nordsamisk er den største e-ordboka *Nordsamisk-norsk-nordsamisk digital ordbok* på nett (NNNDO). Den er gjeven ut av forlaget Davvi girji og har 51 000 lemma i kvar retning. Denne e-ordboka bygger på papirordbøkene SNO (1995) og SNSO (2000), men inneheld berre lemma og omsetjing, ikkje forklaringane eller eksempelsetningane frå papirordbøkene.

Det finst Wiktionary-ordbøker for fleire samiske språkpar, men berre nordsamisk har fleire enn 1000 lemma (til fransk 220 129, til engelsk 3867, til finsk 3060[1]). To andre e-ordbokportalar tilbyr også ordbøker mellom nordsamisk og norsk bokmål, Glosbe og DinOrdbok. Desse ordbøkene inneheld primært leksikografisk materiale frå Wiktionary eller frå Neahttadigisánit. DinOrdbok viser i tillegg eksempelsetningar frå eit parallellkorpus.[2]

For språkparet nordsamisk–norsk finst det med andre ord alternativ til e-ordbøkene til Universitetet i Tromsø – Noregs arktiske universitet (UiT). For dei andre språkpara som blir handsama her, finst det ikkje andre e-ordbøker enn Neahttadigisánit.

3. UiT sine samiske og kvenske ordbøker

3.1. E-ordboka Neahttadigisánit

Forskingsgruppa for samisk språkteknologi ved UiT, Giellatekno, har det seinaste tiåret laga ein infrastruktur for e-ordbøker for morfologirike minoritetsspråk. Denne infrastrukturen kombinerer ei maskinlesbar ordbok med ein morfologisk analysator og tilgang til eit tekstkorpus. Med utgangspunkt i arbeid gjort av Sjur Moshagen integrerte vi først ei nordsamisk-norsk ordbok i ordbokapplikasjonen for Macintosh (jf. Antonsen et al. 2009), og

1 Data er henta frå dei respektive sidene på Wiktionary 3.1.2019. Det høge talet for fransk kjem av at kvar bøyingsform har sin eigen lemmaartikkel. Nordsamisk-fransk er likevel den klårt største Wiktionary-ordboka for nordsamisk.

2 Parallelltekstane er primært henta frå korpuset OPUS.

seinare for andre operativsystem via ordbokformatet StarDict. Desse e-ordbøkene var basert på ein kompileringsrutine der det for kvar bøyingsform for kvart lemma vart generert ein eigen ordbokartikkel, som inneheldt bøyingsforma sjølv, grunnforma og ein peikar til hovudartikkelen for lemmaet. Denne ordboka, Vuosttaš digisánit («Dei første digi-orda»), kunne lastast ned på alle datamaskinar og mobiltelefonar og brukast utan nettilgang. Deretter laga vi ein ny ordbokinfrastruktur, Neahttadigisánit, eit oppsett der ordboka er tilgjengeleg på ein nettserver. Når brukaren slår opp ei ordform, enten frå ordbokgrensesnittet eller i ein separat klikk-i-tekst-funksjon, blir ordforma sendt til ein morfologisk analysator, som finn grunnforma. Denne blir så sendt til ordboka, som returnerer lemmaartikkelen, med lenkjer til setningseksempel i eit korpus og til ein morfologisk generator som genererer bøyingsparadigmet for kvart lemma.

Dei ulike ordbøkene er tilgjengelege online (sjå under Neahttadigisánit i litteraturlista). I tillegg til den opprinnelege ordboka frå nordsamisk er det no 80 andre språkpar fordelt på 11 ulike ordbokportalar. Neahttadigisánit er presentert i Johnson, Antonsen & Trosterud (2013).

3.2. Ordbøkene

Ordbøkene som blir analysert i denne artikkelen, er nordsamisk↔norsk, nordsamisk↔finsk, sørsamisk↔norsk, enaresamisk↔finsk, skoltesamisk↔finsk og kvensk↔norsk. Referanse til dei ulike ordbøkene står under Neahttadigisánit i litteraturlista.

Nordsamisk↔norsk bygger opphavleg på Nils Jernslettens *Álgosátnegirji* (Jernsletten 1991), men er seinare nesten tidobla i omfang. Ordboka frå norsk til nordsamisk er utarbeidd av Giellatekno. *Álgosátnegirji* var, som namnet ('Nybyrjarordbok') viser, ei ordbok for dei med nordsamisk som andrespråk. Ho inneheldt

det samiske grunnordforrådet og forklarte korleis det vart brukt, for studentar på nybyrjarnivå. Av den grunn var også avvikande bøyingsformer sett opp i lemmalista, med referanse til grunnforma (til dømes står *boađán* 'eg kjem' i lemmalista, med referanse til infinitiven *boahtit*). I dei elektroniske ordbøkene er det mogleg å søke på alle ordformer, så tilvisingsformene frå papirordboka er fjerna, men ordboka har framleis eit fokus på L2-brukarar.

Ordboka nordsamisk↔finsk bygger på eit grunnordforråd frå den etymologiske databasen Álgu ved Forskningscentralen för de inhemska språken i Finland, men er seinare utvida. Spesielt gjeld dette frekvente finske ord i den finsk-nordsamiske delen. Sørsamisk↔norsk bygger på leksikografisk arbeid gjort av Albert Jåma og Tove Brustad, publisert i Jåma (2001). Enaresamisk↔finsk bygger på ein tidlegare versjon av Olthuis & Valtonen (2019). Både dei sørsamiske og dei enaresamiske ordbøkene er seinare utvida ein god del. Skoltesamisk↔finsk bygger på Pekka Sammallahti og Jouni Mosnikoff si ordbok *Suomi-koltansaame sanakirja* (1991). Kvensk↔norsk bygde opprinneleg på ordbøker laga av Terje Aronsen (jf. Aronsen 2010) og Eira Söderholm (jf. Söderholm 2012), der ordforrådet er tilpassa ortografien vedtatt av Kvensk språkting. E-ordboka er meir enn dobbelt så stor som dei opprinnelege kjeldene, og det er primært Kvensk institutt som har arbeidd med å utvide ordboka. Jf. Haavisto et al. (2014) for ein presentasjon.

Storleiken på dei ulike ordbøkene går fram av tabell 2 nedanfor. Lmaj viser til dei respektive majoritetsspråka norsk og finsk, og Lmin til eitt av dei samiske språka eller kvensk. Ordbøker i dei ulike retningane mellom to språk er sett opp på kvar sin rad i same kolonne slik at den første raden viser tal for ordbøkene med minoritetsspråk som lemma og den andre raden viser tal for ordbøkene med majoritetsspråk som lemma. Tabellen viser med andre ord at nordsamisk-norsk ordbok har 39 862 lemma, norsk-nordsamisk har 26 021 lemma, og så bortetter.

	Lmaj = norsk			Lmaj = finsk		
Lmin	nord-samisk	sør-samisk	kvensk	nord-samisk	enare-samisk	skolte-samisk
Lemma = Lmin	39 862	11 174	9104	13 175	21 740	33 594
Lemma = Lmaj	26 021	14 754	8599	11 301	17 752	30 217

Tabell 2: Lemma i tolv av Giellateknos nettordbøker (Neahttadigisánit), januar 2019.

3.3. Dei morfologiske analysatorane attom ordbøkene

Som referert til ovanfor er det for alle ordbøkene mogleg å slå opp på kva slags ordform som helst. Oppslaget blir sendt til ein morfologisk analysator som gjev både lemma og morfologisk analyse. Lemmaet blir brukt til å slå opp i ordbokartikkelen, og den morfologiske analysen blir presentert for brukaren på sida av sjølve ordbokartikkelen. Alle lemma i alle ordbøkene er inkludert i dei ulike analyseprogramma. Med unntak av språkparet nordsamisk–finsk, der det er mogleg å tenkje seg både samisk- og finskspråklege brukarar, vil brukarane ikkje vere interesserte i morfologisk informasjon om majoritetsspråket (norsk eller finsk). I og med at ordbøkene har ein klikk-i-tekst-funksjon, er ordbøkene likevel utstyrt med ein morfologisk analysator også for finsk og norsk. Brukarar i Noreg er nok i stand til å identifisere delane i eit ord som til dømes *partimedlemmene*, og tilsvarande for brukarar i Finland med det tilsvarande ordet *puoluejäsenet*, men med hjelp av den morfologiske analysatoren kan dei berre klikke på orda, eller lime dei inn i ordboka, og for delane *parti* og *medlem* (finsk: *puolue* og *jäsen*) få ut dei respektive nordsamiske omsetjingane *bellodat* og *miellahttu*. Dermed inneheld loggdata informasjon om både søkeord og lemma, som vist i neste avsnitt.

4. Analyse av loggdata

4.1. Logganalyse som forskingsområde

Herbert E. Wiegand set opp forsking på ordbokbruk som ein av dei fire deldisiplinane innafor metaleksikografien (Wiegand 1998), og det er innafor leksikografien ein rik forskingstradisjon på ordbokbruk. *LexicoNordica* 15 (2008) hadde til dømes ordbokbruk i Norden som tema, og fleire av artiklane der analyserte loggdata. Ordbokbruk kan bli undersøkt via observasjon og via spørjeundersøkingar. Med innføringa av e-ordbøker har ein ny innfallsport til kunnskap om ordbokbruk opna seg: loggoppføring av kvart oppslag gjort i ordboka.

Den mest opplagte måten å bruke loggdata på er til å forbetre lemmatilfanget i sjølve ordboka. Ifølgje de Schryver & Joffe (2004:187) er det få rapportar om slik bruk av loggdata. For utviklinga av Neahttadigisánit har loggdata tvert imot vorte brukt på nett denne måten. Haavisto et al. (2014) er eit godt døme på analyse av loggdata for å forbetre lemmatilfanget i ei ordbok. Då artikkelen vart skriven, inneheldt den kvensk-norske ordboka 4702 lemma, og loggføringa hadde vore operativ i kort tid. Loggen inneheldt berre 691 oppføringar for begge retningane. Likevel kunne loggen vise at sentrale ord som *olla* 'å vere' og *vasen* 'venstre' mangla i ordboka, og både dei og andre manglande ord kunne bli lagt til. Tilsvarande arbeid har blitt gjort for alle dei andre Neahttadigisánit-ordbøkene.

Med informasjon om den einskilde brukaren er det mogleg å undersøke søkemønster for individuelle brukarar. Almind (2008) inneheld ei drøfting av ulike aspekt ved ordbokbruken som det er mogleg å få ut av ei loggfil. Sentralt for han er i kor stor grad det er mogleg å seie ikkje berre kva det blir søkt etter, men også i kva grad ordboka oppfyller ein spesifikk funksjon (Almind 2008:36). I loggane som blir brukt i denne artikkelen, har vi ikkje tilgang til

den einskilde brukaren, så målsetjinga her er å analysere kva brukarane som kollektiv har søkt etter.

4.2. Loggføring

Bruken av Neahttadigisánit blir loggført. Som illustrasjon viser tabell 3 nedanfor fire tilfeldige loggoppføringar, dvs. loggen for fire autentiske ordbokoppslag. Første kolonne gjev forma brukaren har søkt etter. Andre kolonne fortel om søket var vellukka (T, true) eller ikkje (F, false). Tredje kolonne fortel kva lemma (eitt eller fleire) søket får tilslag på. For dei første to oppslaga er det meir enn eitt lemma: Finsk *minä* er tvitydig mellom nominativ av *minä* 'eg' og essiv av *mikä* 'kva', og norsk *for* er fleirtydig mellom verba *fare* og *fore* og mellom substantivet, preposisjonen, adverbet og konjunksjonen *for*.

Fjerde kolonne gjev omsetjing. Der søket ikkje er vellukka (F), viser loggen naturleg nok verken lemma eller omsetjing. Andre opplysningar som kjem fram i lemmaartikkelen (eksempelsetningar, bruksrestriksjonar og så bortetter), er ikkje med i loggen. Dei neste to kolonnane identifiserer språkparet brukaren har søkt i. Siste kolonne viser dato og klokkeslett.

Søk	TF	Lemma	Omsetjing	L1	L2	Dato, tid
minä	T	mikä, minä	mii; mun	fin	sme	2018-12-29 T15:55:22
for	T	fare, fore, for	luoittihit, livkit; illativ...	nob	sme	2018-12-29 T15:55:23
fuotni	T	fuotni	dårlig	sme	nob	2018-12-29 T15:55:37
forma-sjon	F			nob	sme	2018-12-29 T15:55:49

Tabell 3: Eksempel på loggoppføringar. Språka er identifisert med ISO 639-3-kodar (SIL International 2007): fin = finsk, nob = norsk bokmål, sme = nordsamisk.

Tabell 4 viser talet på alle loggførte oppslag i dei tolv ordbøkene i 2018 (i perioden 1.1.–29.12.2018). Data for ordbøkene der lemma er på minoritetsspråket, står i raden «Ordbokoppslag på Lmin», og data for ordbøkene der lemma er på majoritetsspråket, står i raden «Ordbokoppslag på Lmaj». Siste rad viser fordelinga språkretningane imellom: 40,7 % av oppslaga i ordbøkene mellom nordsamisk og norsk er til dømes gjort for språkparet norsk–nordsamisk (Lmaj–Lmin). Oppslaga frå nordsamisk til norsk utgjer med andre ord 59,3 % av oppslaga for dette språkparet (i og med at 59,3 % + 40,7 % = 100 %).

	Lmaj = norsk			Lmaj = finsk		
Lmin = ...	nord-samisk	sør-samisk	kvensk	nord-samisk	enare-samisk	skolte-samisk
Ordbok-oppslag på Lmin	953 807	317 452	70 449	247 590	176 662	340 494
Ordbok-oppslag på Lmaj	655 928	206 378	42 529	273 459	127 654	222 793
Lmaj / alle oppslag	40,7 %	39,4 %	37,6 %	52,5 %	41,9 %	39,6 %

Tabell 4: Loggførte oppslag i 2018 for dei tolv ordbøkene.

Som det går fram av tabellen, er ordbøkene mykje i bruk. Målt i oppslag per morsmålstalar er bruksfrekvensen langt høgare enn for e-ordbøker på majoritetsspråka. Med unntak av paret finsk–nordsamisk, der oppslag frå finsk til nordsamisk er i knapt fleirtal, går rundt rekna fire av ti ordbokoppslag frå majoritetsspråket til minoritetsspråket.

Loggdøma i tabell 3 viser at det blir loggført kor mange av ordbokoppslaga som ordbøkene kan svare på. Tabell 5 viser kor stor prosentdel av oppslaga som ikkje blir funne i dei ulike ordbøkene (dvs. av tala i tabell 4). Ordbøkene til og frå majoritets- og minoritetsspråka står på kvar si linje, på same måte som i tabell 4.

	Lmaj = norsk			Lmaj = finsk		
Oppslag utan svar	nord-samisk	sør-samisk	kvensk	nord-samisk	enare-samisk	skolte-samisk
Ordbokoppslag på Lmin	10,7 %	77 %	10,2 %	12,8 %	11,3 %	2,7 %
Ordbokoppslag på Lmaj	11,8 %	6,8 %	10,8 %	22,4 %	9,1 %	4,5 %

Tabell 5: Prosent av oppslaga som ikkje resulterte i treff.

Éi ordbok peiker seg ut med langt høgare prosent av oppslag utan svar enn dei andre, nemleg finsk-nordsamisk, med 22,4 % oppslag utan treff (tabell 5). Dette korresponderer med storleiken på ordboka: Den finsk-nordsamiske ordboka er berre halvparten så stor som den norsk-nordsamiske (jf. tabell 2) og har dermed også dårlegare treffprosent for omsetjing til same målspråk.

4.3. Analyse av søk som ikkje gjev treff

No kan det vere fleire grunnar til at eit oppslag ikkje finst i ordboka. Éin er at ordet manglar, ein annan er at det er skrive feil. Eg har sett på dei orda som ikkje gav søketreff i ordbøkene, og kontrollert dei opp mot ein stavekontroll for kvart av dei relevante språka. Resultatet går fram av tabell 6, der «Ordbokoppslag med skrivefeil» viser til oppslag som ikkje blir kjent att av stavekontrollar for dei respektive språka.

	Lmaj = norsk			Lmaj = finsk		
Ordbokoppslag med skrivefeil	nord-samisk	sør-samisk	kvensk	nord-samisk	enare-samisk	skolte-samisk
Ordbokoppslag på Lmin	73,7 %	63,1 %	89,0 %	56,3 %	83,9 %	74,4 %
Ordbokoppslag på Lmaj	39,5 %	32,1 %	34,8 %	11,6 %	34,0 %	30,8 %

Tabell 6: Feilskrivingsprosent i søkeorda som ikkje fekk treff.

Tabell 6 viser svært stor variasjon. Det mest slåande er at ordbokbrukarane er langt støare (mellom to og fire gonger så stø) i ortografien til majoritetsspråket enn i ortografien til minoritetsspråket. Delar av skilnaden i feilskriving må også bli sett i lys av treffprosent. Når ein stor del av oppslaga ikkje finst i ordboka, blir feilskrivingsprosenten lågare, som for finsk-nordsamisk, mens der ein låg prosent av oppslaga manglar, er prosenten av skrivefeil høgare, som for finsk-skoltesamisk.

Skilnadene i feilskriving frå språk til språk er likevel store. Den største feilskrivingsprosenten har kvensk, som også har det yngste skriftspråket, og der avgrensinga mot finsk er vanskeleg å dra. Vi skal òg merke oss at resultata for andre ordbøker også viser at feilskriving er den viktigaste kjelda til mislukka søk: For ei undersøking av 500 vilkårleg valde feilsøk i *Bokmålsordboka* viser Nygaard & Fjeld (2008) at så mykje som 46 % kjem av skrivefeil.

Tala for minoritetsspråk er likevel høgare enn for majoritetsspråk. At over 56 %, og i dei fleste tilfella over 70 %, av dei mislukka oppslaga på minoritetsspråket inneheld skrivefeil, bør få konsekvensar for korleis vi presenterer lemma i minoritetsspråklege ordbøker. Eitt mogleg svar på dette problemet finn vi i den norsk-nordsamiske ordboka SNSO. Som vist i Trosterud (2001) har denne ordboka lemmaartiklar strukturert som i ei nordsamisk-norsk ordbok. Det vil seie at ordbokartiklane ikkje er inndelt i tydingar og tydingsgrupper med norske bruksdøme og nordsamiske omsetjingar av dei. I staden listar ordboka opp alle dei samiske omsetjingane alfabetisk. Denne alfabetiseringa kan i lys av brukarloggen presentert her bli sett på som ein strategi for å gje ordbokbrukarane tilgang til ordartiklar for lemma dei ikkje veit korleis dei skal skrive.

I Trosterud (2001) illustrerte eg problemet med redigeringsprinsippet for SNSO ved å sitere artikkelen for verbet *gå*. *Gå* hadde 28 omsetjingar, og desse var altså lista opp alfabetisk. Viss vi heller ser SNSO som ei nordsamisk rettskrivingsordbok, kan sa-

miskspråklege brukarar i Noreg som lurer på korleis *vázzit* blir skrive, slå opp på ordet *gå* og lese seg nedover til V, der dei finn den sentrale omsetjinga, *vázzit*. At det er *gå* dei skal slå opp på for å finne *vázzit*, veit brukarane fordi dei er tospråklege. Fordelen med å slå opp i den norskspråklege ordboka framfor direkte i ei samiskspråkleg ordbok er at ein kjem unna rettskrivingsordbok-paradokset: For å finne ordet du ikkje veit korleis blir stava, må du kunne stave det.

Problemfri er denne metoden likevel ikkje. For enkle ordpar går det fint, men i mange tilfelle er det ikkje opplagt kva norske oppslagsord det samiske ordet ein er interessert i, står under. For å hjelpe e-ordbokbrukarane til å bruke ordbøkene med nord-, enare- og skoltesamisk som L1 har vi bygd inn ein viss toleranse i systemet: I tillegg til den vanlege e-ordboka som blir evaluert her, har Neahttadigisánit også ein tolerant inputmodus som aksepterer bruk av grunnbokstavar utan diakritiske teikn, slik at ordboka for kvar av bokstavane *acdnstz* i input undersøker om det å bruke *áčđŋštž* i staden vil gje treff på eit oppslag (til dømes **odas* for korrekt *ođas* 'ny'). Oppslag der brukaren vel denne modusen, er elles haldne utanfor analysen i denne artikkelen, men dei utgjer omtrent 9 % av den totale ordbokbruken. 12,6 % av dei vellukka oppslaga som er loggført for nordsamisk ved bruk av den tolerante modusen, ligg utanfor den normative grammatikken, og utan bruk av tolerant modus ville dei ikkje resultert i treff i ordboka. Bruken av ein slik tolerant modus vil med andre ord auke treffprosenten tilsvarande.

Ein spesiell form for toleranse er inkludert i den kvensk-norske ordboka. Kvensk er standardisert i tre ulike dialektar: Porsanger-, Varanger- og Nordreisa-dialekt. Skilnadene er primært morfologiske og morfofonologiske, men dei er svært djuptgripande, og store delar av ordforrådet har systematiske skilnader i bøyingsmorfologien. I ordboka er dette løyst ved at den morfologiske analysatoren inneheld morfofonologien for alle tre dialektane.

Ordboka er i utgangspunktet skriven på Porsanger-dialekt, og den morfologiske generatoren, som genererer ordformene til paradigma som blir presentert, genererer også berre Porsanger-former. Det er dermed mogleg å slå opp på Varanger-forma *pimeä* og få vite at ordet betyr *mørk*, og at lemmaet (Porsanger-forma) er *pimmee*. På denne måten blir ei ordbok som opprinneleg er skriven for éin dialekt, i stand til å gje omsetjingar av ord frå tre ulike dialektar.

4.4. Dekningsgrad

Det er stor variasjon i dekningsgraden til dei ulike ordbøkene. Tabell 7 viser dekningsgraden for dei 1000 og 10 000 mest vanlege lemmaa for kvart språk (skoltesamisk manglar fordi eit relevant korpus ikkje var tilgjengeleg). Frekvensdata for dei samiske språka og kvensk er henta frå SIKOR. Sjangrane tekstane er henta frå, er primært administrative tekstar. For nordsamisk er det også avisspråk og for enaresamisk også tidsskriftsartiklar av blanda innhald. For norsk brukte eg NOWAC og for finsk frekvenslister frå Kielipankki (sjå litteraturlista for referansar).

	Lmaj = norsk			Lmaj = finsk	
Dekningsgrad for topp-1000	nord-samisk	sør-samisk	kvensk	nord-samisk	enare-samisk
Ordbokoppslag på Lmin	84,0 %	85,5 %	79,9 %	78,5 %	87,6 %
Ordbokoppslag på Lmaj	87,8 %	74,3 %	66,0 %	88,3 %	86,8 %
Dekningsgrad for topp-10 000	nord-samisk	sør-samisk	kvensk	nord-samisk	enare-samisk
Ordbokoppslag på Lmin	68,0 %	40,5 %	28,8 %	40,8 %	61,6 %
Ordbokoppslag på Lmaj	57,5 %	32,9 %	32,2 %	46,7 %	48,9 %

Tabell 7: Dekningsgrad for dei 1000 og 10 000 vanlegaste orda i 10 ordbøker.

Den beste dekningsgraden har finsk-nordsamisk, noko som kjem av at den same frekvenslista for finsk som ligg attom denne tabellen, vart brukt også i utarbeidinga av sjølve ordboka. Elles illustrerer det brå fallet i dekningsgrad frå topp-1000 til topp-10 000 at alle ordbøkene (bortsett frå nordsamisk-norsk) er relativt små (jf. tabell 3).

4.5. Grunnform versus bøyingsform som søkeord

I og med at alle ordbøkene er knytt til ein morfologisk analysator, er det mogleg å slå opp også på alle ordformene i bøyingsparadigmet til kva som helst lemma i ordboka. I loggen er søk på oblike former kjenneteikna ved at søkeord (kolonne 1 i tabell 3) og lemma (kolonne 3 i tabell 3) er ulike. Tabell 8 viser at dei fleste ordbokoppslaga i ordbøkene frå minoritetsspråk til majoritetsspråk blir gjort på lemma. Språket med flest oppslag på oblike former er nordsamisk, og dette er også språket med mest tekst tilgjengeleg elektronisk.

	Lmaj = norsk			Lmaj = finsk		
Ordbok Lemma = Lmin	nord-samisk	sør-samisk	kvensk	nord-samisk	enare-samisk	skolte-samisk
Oppslag ≠ lemma	42,0 %	35,8 %	26,5 %	37,5 %	31,4 %	22,5 %

Tabell 8: Ordbokoppslag på oblike former i ordbøker med lemma på minoritetsspråket, i prosent av alle oppslaga.

Som vist i Antonsen et al. (2009:278) er delen av ord i grunnform i løpande nordsamisk tekst 7,9 %. Viss den dominerande bruken hadde vore lesing av e-tekst med klikk-i-tekst-funksjonen, hadde prosenten av oblike former med andre ord vore langt høgare enn dei rundt 40 % vi har. Ein rimeleg konklusjon er at ordbøkene ikkje primært blir brukt til resepsjon av minoritetsspråkstekst. I og med at dei fleste oppslaga går frå minoritets- til majoritetsspråket

(jf. tabell 4), er det heller ikkje mogleg å seie at dei primært blir brukt til produksjon. Loggen viser ikkje i kor mange av oppslaga brukarane går vidare til å generere bøyingsparadigme for ordet eller å slå opp i korpus, men ei mogleg forklaring på dataa i desse to tabellane kan vere at det er L1-brukarar som slår opp i ordboka for å få opplysningar om rettskriving og bøying.

4.6. Innhaldet i loggane

Ein annan innfallsvinkel til loggdata er å sjå nærare på orda brukarane slår opp. Tabell 9 viser dei 15 vanlegaste bokmålsorda som brukarane slår opp i ordbøkene til kvensk, sørsamisk og nordsamisk.

Rang	Norsk-nordsamisk	Norsk-sørsamisk	Norsk-kvensk
1	skulle	måtte	vel
2	ha	få	jeg
3	kunne	forstå	hus
4	gå	ha	være
5	se	samisk språk	hei
6	bli	gjøre	gå
7	komme	se	kaste
8	få	snø	bli
9	dra	møte	kunne
10	være	komme	du
11	fare, fore, for	språk	se
12	det	nå	eie
13	snakke	løpe	følge etter
14	som	høre	ha
15	på	være ordstyrer	skulle

Tabell 9: Dei 15 vanlegaste norske oppslagsorda.

For den nordsamiske ordboka er dei 10 vanlegaste oppslagsorda hjelpeverb eller andre viktige verb, og dei andre orda er sentrale funksjonsord. Legg særleg merke til det nest vanlegaste ordet, verbet *ha*, som ikkje har noko direkte motsvar på nordsamisk. Brukaren treng med andre ord hjelp til å produsere habitive setningar, og slår opp på *ha*. Samla sett består lista av verb som opptrer i ein stor del av setningane i språket, ord med ulike tydingar i ulike samanhengar. Inntrykket loggen gjev, er at han er dominert av brukarar som har norsk og ikkje nordsamisk som morsmål, og at dei bruker ordboka for å produsere nordsamisk tekst.

For sørsamisk ser vi same tendens, sjølv om den ikkje er like klår. Også her dominerer hjelpeverb og andre sentrale verb. Verbet *ha* er ikkje like høgt oppe, men sørsamisk skil seg frå nordsamisk ved å ha eit verb som svarer til det norske *ha*. Eit par substantiv har også snike seg med. Oppslagsorda *språk* og *samisk språk* bør nok bli sett i samanheng med verbet *forstå* (som i *forstå samisk*), og den høge plasseringa av *snø* kan tyde på at ein stor del av brukarane er meir interesserte i eksotiske drag ved språket enn i å verkeleg produsere sørsamisk tekst. Det 15. oppslaget er støy. I det heile er inntrykket også her at ordboka er ei produksjonsordbok, men delen av brukarar som produserer normal prosa, ser likevel ut til å vere lågare enn for nordsamisk.

For kvensk er det vanskelegare å dra nokon konklusjonar. Hovudinntrykket er mangelen på sentrale verb. Verba *ha*, *skulle* og *vere* er med, men ikkje like høgt oppe som for nordsamisk. Derimot er *eie* med, noko som sannsynlegvis er uttrykk for at brukarar er på utkik etter ein alternativ strategi for å uttrykke eigarkonstruksjonen. Kvensk har eit verb *omistaat* 'å eige', men den vanlege måten å uttrykke at nokon har noko, er, som på nordsamisk, med på-kasus og verbet for å vere («Eg har ein båt» blir uttrykt som *Minula oon venet* 'På meg er båt'). I det heile er det vanskelegare å dra konklusjonar basert på materialet for norsk-kvensk. At det empiriske materialet berre er ein brøkdel av det det er for nord-

samisk (42 000 mot 655 000, jf. tabell 4), gjer også at rangeringa av oppslagsord på frekvenslista blir meir tilfeldig. Det er likevel klårt at den kvenske ordboka i mindre grad enn dei to samiske blir brukt til produksjon av L2-tekst.

Mønsteret for finsk-nordsamisk (tabell 10) minner om mønsteret for dei to ordbøkene frå norsk til høvesvis nord- og sørsamisk. Bortsett frå *voida* manglar dei modale hjelpeverba, men ut over det er listene dominert av dei mest sentrale finske verba. For finsk-enaresamisk er det fleire verb enn for finsk-nordsamisk, noko som kan spegle det aktive revitaliseringsarbeidet for enaresamisk. På begge listene er det også substantiv, men det er sentrale substantiv til bruk i språkproduksjonen. Også her ber lemmautvalet preg av at brukarane har dei samiske språka som L2.

Rang	Finsk-nordsamisk	Finsk-enaresamisk
1	olla 'vere'	olla
2	saada 'få'	saada
3	fiinnis 'fin'	tulla
4	se 'den, det'	voida
5	tulla 'kome'	pitää
6	voida 'kunne'	lapsi 'barn'
7	mennä 'gå, dra'	tehdä
8	tehdä 'gjere'	ruoka 'mat'
9	aika 'tid'	mennä
10	työ 'arbeid'	haluta
11	katsoa 'sjå'	tarvita 'trenge'
12	haluta 'ville'	käydä 'vitje'
13	pitää 'halde, like'	kirjoittaa 'skrive'
14	tämä 'denne, dette'	syödä 'ete'
15	ihminen 'menneske'	se

Tabell 10: Dei 15 vanlegaste finske oppslagsorda for to ordbøker.

5. Konklusjon

Ein analyse av loggane for tolv ulike ordbøker mellom fem ulike minoritetsspråk og dei respektive majoritetsspråka deira viser at ordbøkene er mykje brukt. Dei fleste oppslaga går frå minoritetsspråket til majoritetsspråket, men skilnaden er ikkje overveldande (40 % vs. 60 %). Dei ulike språksamfunna bruker til ein viss grad ordbøker på ulike måtar. Ordbokbrukarane for dei samiske ordbøkene slår opp sentrale verb, hjelpe- og modalverb. For kvensk er mønsteret meir tilfeldig og dei sentrale verba ikkje like dominerande.

Oppslaga på minoritetsspråkslemma er dominert av søk på grunnform; over 60 % av søka er av denne typen. Dette står i kontrast til løpande tekst, der grunnformene utgjer mindre enn 10 % av ordformene. Det ser med andre ord ikkje ut til at bruk av ordbøkene som resepsjonsordbøker er den dominerande bruken.

Graden av feilskriving av oppslagsord er ulik for minoritets- og majoritetsspråksoppslag. Blant mislukka søk på minoritetsspråka er prosenten av feilskrivne ord så høg som mellom 60 og 90. For majoritetsspråka er prosenten også høg, mellom 30 og 40, men altså langt lågare. Dette resultatet viser ikkje berre at oppslagsord svært ofte blir skrivne feil i e-ordbøker, men også at dette er eit stort problem for minoritetsspråk. Måtar å hjelpe ordbokbrukarane på slik at dei finn oppslagsorda dei er på jakt etter, bør med andre ord få høg prioritet ved vidare arbeid med ordbokgrensesnitt.

Litteratur

Ordbøker

Álgu = Etymologische Datenbank für die saamischen Sprachen. Forskningscentralen för de inhemska språken i Finland. <http://kaino.kotus.fi/algu/> (mars 2019).

Aronsen, Terje (2010): *Kvensk-norsk-kvensk elektronisk ordbok*. Redigert av Verena Schall & Trond Trosterud. Universitetet i Tromsø.

Bokmålsordboka. Oslo: Kunnskapsforlaget. <ordbok.uib.no> (mars 2019).

DinOrdbok. <http://dinordbok.no> (mars 2019).

Glosbe = Glosbe – the multilingual online dictionary. <http://glosbe.com> (mars 2019).

Jernsletten, Nils (1991): *Álgosátnegirji sámi-dáru sátnegirji = samisk-norsk ordbok*. Kárášjohka: Davvi girji.

Jåma, Albert (2001): *Norsk-sydsamisk ordliste. Gærjiste vaalteme*. Hemnes sameforening. <http://www.ruovatsijte.no/gaerjiste-vaalteme_2001.pdf> (mars 2019).

Neahttadigisánit. E-ordbøker frå Giellatekno ved Universitetet i Tromsø – Noregs arktiske universitet: nordsamisk: <http://sanit.oahpa.no> (mars 2019), sørsamisk: <http://baakoah.oahpa.no> (mars 2019), enaresamisk: <http://saanih.oahpa.no> (mars 2019), skoltesamisk: <http://saan.oahpa.no> (mars 2019) og kvensk: <http://sanat.oahpa.no> (mars 2019). Felles ordbokportal: <http://dicts.uit.no/> (mars 2019). Kjeldekode for grensesnittet: <https://gtsvn.uit.no/langtech/trunk/apps/dicts/nds/> (mars 2019).

NNNDO = *Nordsamisk-norsk-nordsamisk digital ordbok på nett*. Kárášjohka: Davvi girji. <https://533.davvi.no/> (mars 2019).

Olthuis, Marja-Liisa & Taarna Valtonen (2019): *Suomi-inarinsaame sanakirja*. Saamelaiskäräjät.

Sammallahti, Pekka & Jouni Mosnikoff (1991): *Suomi-koltansaame sanakirja*. Ohcejohka: Girjegiisa.

SNO = John Henrik Eira, Johan Jernsletten, Brita Kåven, Ingrid Nordal & Aage Solbakk (1995): *Sámi-dáru sátnegirji = Samisk-norsk ordbok*. Kárášjohka: Davvi girji.

SNSO = Brita E. Kåven (red.) (2000): *Stor norsk-samisk ordbok / Dáru-sámi sátnegirji*. Kárášjohka: Davvi girji.

StarDict. <http://www.huzheng.org/stardict/> (mars 2019).

Söderholm, Eira (2012): *Sannoi ja haamui. Porsangin ja Raisin variantti. Kainun sana- ja sanahaamulista Aikamatkaa varten. Kvensk ord- og ordformliste til Aikamatka*. Manuskript. Universitetet i Tromsø – Noregs arktiske universitet. <https://tinyurl.com/y6xmfh45> (mars 2019).

Wiktionary. <http://wiktionary.org> (mars 2019).

Annan litteratur

Almind, Richard (2008): Søgemønstre i logfiler. I: *LexicoNordica* 15, 33–55.

Antonsen, Lene, Ciprian-Virgil Gerstenberger, Sjur N. Moshagen & Trond Trosterud (2009): Ei intelligent elektronisk ordbok for samisk. I: *LexicoNordica* 16, 271–283.

de Schryver, Gilles-Maurice & David Joffe (2004): On How Electronic Dictionaries are Really Used. I: G. Williams & S. Vessier (eds.): *Proceedings of the Eleventh EURALEX International Congress*. EURALEX 2004, Lorient, France, July 6–10, 2004. Lorient: Faculté des Lettres et des Sciences Humaines, Université de Bretagne Sud, 187–196.

EGIDS = Language Status. I: David M. Eberhard, Gary F. Simons & Charles D. Fennig (eds.) (2019): *Ethnologue: Languages of the World. Twenty-second edition*. Dallas, Texas: SIL International. <https://www.ethnologue.com/about/language-status> (mars 2019).

Ethnologue = David M. Eberhard, Gary F. Simons & Charles D Fennig (eds.) (2019): *Ethnologue: Languages of the World. Twenty-second edition*. Dallas, Texas: SIL International. <http://www.ethnologue.com> (mars 2019).

Haavisto, Mervi, Kaisa Maliniemi, Leena Niiranen, Pirja Paavalniemi, Tove Reibo & Trond Trosterud (2014): Kvensk ordbok på nett – hvem har nytte av den? I: Ruth Vatvedt Fjeld & Marit

Hovdenak (red.): *Nordiske studier i leksikografi* 12. Oslo: Nordisk forening for leksikografi, 176–192. <https://tidsskrift.dk/nsil/article/view/20997> (mars 2019).

Johnson, Ryan, Lene Antonsen & Trond Trosterud (2013): Using finite state transducers for making efficient reading comprehension dictionaries. I: *Proceedings of the 19th Nordic Conference of Computational Linguistics (NODALIDA 2013), May 22–24, 2013, Oslo University, Norway.* NEALT Proceedings Series 16, 59–71. <https://munin.uit.no/bitstream/handle/10037/12884/paper_5.pdf> (mars 2019).

Kielipankki = Kielipankin konkordanssihakuohjelma. <http://korp.csc.fi> (mars 2019).

Lewis, M. Paul & Gary F. Simons (2010): Assessing endangerment: Expanding Fishman's GIDS. I: *Revue Roumaine de Linguistique* 55(2), 103–120. <https://www.lingv.ro/RRL%202%202010%20art01Lewis.pdf> (mars 2019).

NOWAC = *Norwegian web as a corpus*. <https://www.hf.uio.no/iln/tjenester/kunnskap/sprak/korpus/skriftsprakskorpus/nowac/index.html> (mars 2019).

Nygaard, Lars & Ruth Vatvedt Fjeld (2008): Analyse av søkelogger for bedre søkemuligheter i elektroniske ordbøker. I: *LexicoNordica* 15, 57–72.

OPUS. <http://opus.nlpl.eu> (mars 2019).

SIKOR. Universitetet i Tromsø – Noregs arktiske universitets og det norske Sametingets samiske tekstsamling, versjon 06.11.2018. <http://gtweb.uit.no/korp/> (samiske språk) og <http://gtweb.uit.no/f_korp> (austersjøfinske språk) (begge mars 2019).

SIL International (2007): ISO 639-3, *Codes for the representation of names of languages – Part 3: Alpha-3 code for comprehensive coverage of languages*. <https://iso639-3.sil.org/> (september 2019).

Trosterud, Trond (2001): *Stor norsk-samisk ordbok / Dáru-sámi sátnegirji* [bokmelding]. I: *LexicoNordica* 8, 283–306.

Wiegand, H.E. (1998): *Wörterbuchforschung: Untersuchungen zur Wörterbuchbenutzung, zur Theorie, Geschichte, Kritik und Automatisierung der Lexikographie.* 1. Teilband. Berlin: Walter de Gruyter.

Trond Trosterud
professor, dr.art.
Institutt for språk og kultur
UiT – Noregs arktiske universitet
Postboks 6050 Langnes
NO-9037 Tromsø
trond.trosterud@uit.no

IKKE-TEMATISK BIDRAG

Revision av konstruktionsuppgifter i *Svensk ordbok utgiven av Svenska Akademien*

Kristian Blensenius

The article presents an overview of considerations made in the process of revising verbal constructional patterns in the monolingual defining dictionary *Svensk ordbok utgiven av Svenska Akademien* ('The Contemporary Dictionary of the Swedish Academy'). The digital format of the dictionary makes it possible to provide extended, and at the same time more user-friendly, information on different combinatorial patterns, and also on their relation to syntactic examples.

1. Inledning

En rad informationskategorier reviderades inför utgivningen av den första utgåvan av definitionsordboken *Svensk ordbok utgiven av Svenska Akademien* (2009, hädanefter SO). För en genomgång av viktigare förändringar mellan SO och föregångaren *Nationalencyklopedins ordbok* (NEO, 1995–1996), se Malmgren (2009a). Inför den andra utgåvan av SO, SO2, som är tänkt att publiceras inom något år, genomgår flera informationskategorier ytterligare en revision. Se Sköldberg (2017) och Sköldberg & Hannesdóttir (2017) för en genomgång av några av revisionspunkterna. Inte minst kommer de informationskategorier som innefattar flerordsuttryck att behandlas inför utgivandet av SO2, och i föreliggande artikel ska framför allt arbetet med konstruktionsuppgifterna beskrivas och problematiseras.

I SO ges, grovt sett, information om tre sorters flerordsuttryck inom ordboksartiklarna. Nedan ordnas sorterna efter grad av kompositionalitet.

- Idiom, t.ex. *som fisken i vattnet* (under **fisk**), som har ganska låg grad av kompositionalitet – idiomets betydelse kan i princip inte härledas från de ingående ordens betydelser – och som återfinns som särskilda fraser, en sorts underlemman, i slutet av ordboksartikeln.
- Mer eller mindre fasta fraser inklusive kollokationer (se inledningen i SO, s. XVII), som ofta återfinns ”frilagda” först i exempelramsorna, men som också kan återfinnas inuti de längre, satsformade språkexemplen. Ett typfall av det senare är kollokationen *fatta beslut*, som är inbäddad i ett satsformat exempel i **fatta 4**: *Finansinspektionen har fattat beslut om att ge företaget en varning*.
- Konstruktionsuppgifter (valens)[1], dvs. upplysningar som kan betraktas som formaliserade mönster eller abstraktioner av fria ordkombinationer som associeras med uppslagsordet (Howarth 1998:28). Här är uppslagsordet fixerat, medan flera andra led i uttrycket ofta är variabla. Konstruktionsuppgifterna konkretiseras ofta, men inte alltid, med språkexempel. En konstruktionsuppgift som den till **fatta 1**, *fatta ngn/ngt (med ngt)*, exemplifieras av frasen *fatta spaken med höger hand*. Som synes innehåller konstruktionsuppgifterna en kombination av syntaktisk information och enkel semantisk information, bl.a. om huruvida den nominala referenten är en person, *ngn*, eller en icke-person, *ngt*.

1 Termen *valens*(*uppgift*) används ibland synonymt med den generellare *konstruktionsuppgift*, men den förra termen för också med sig ett teoretiskt perspektiv, nämligen att det är det enskilda huvudordet, typiskt ett verb, som bestämmer vilka komplement det realiseras med (se t.ex. Herbst 2014). I denna artikel kommer den relativt teorineutrala termen *konstruktionsuppgift* att användas, bl.a. eftersom det är den term som används i SO (SO, s. XVIII).

I fokus här är som sagt den sistnämnda typen av information om flerordsuttryck, nämligen konstruktionsuppgifter för verben. Ansatsen är mer praktisk än teoretisk. För historisk överblick och mer teoretiska överväganden i fråga om framför allt verbala konstruktionsuppgifter i ordböcker som SO hänvisas till Toporowska Gronostaj (1996) och Malmgren & Toporowska Gronostaj (2009). En valensteoretisk historik ges av bl.a. Faulhaber (2011:3–6), och en genomgång av valens i förhållande till alternativa perspektiv ges av exempelvis Herbst (2014).

2. Konstruktionsuppgifter

I det följande ges först en något mer ingående beskrivning av bl.a. SO:s konstruktionsuppgifter och sedan ges några övergripande revideringsåtgärder i 2.2. I 2.3 och framåt redovisas ett antal mer specifika revisionspunkter som är planerade att genomföras i SO2.

2.1. Konstruktionsuppgifter i SO och i några andra ordböcker

Det går att, baserat på vad det står i SO:s inledning, urskilja två huvudsakliga funktioner hos konstruktionsuppgifter. För det första ska ordboken ”ge så fullständiga upplysningar som möjligt om alla viktiga ord i det svenska ordförrådet”, och detta generella syfte omfattar även konstruktionsuppgifterna. För det andra har konstruktionsuppgifterna också en mer uttalat praktisk funktion, nämligen att svara på frågor som ”Heter det *i stor utsträckning* eller *i hög utsträckning*?” (SO, s. IX). Nedan ges tre exempel på konstruktionsuppgifter till olika typer av verb i första upplagan av SO:

växa igen *växa igen*
förorsaka *förorsaka* (*ngn*) *ngt*
upplysa 2 *upplysa ngn* (*om ngn*/*ngt*/SATS)

Ovan illustrerar **växa igen** konstruktionsuppgiften för ett intransitivt (partikel)verb, medan **förorsaka** illustrerar konstruktionsuppgiften för ett bitransitivt verb med optionellt indirekt objekt (*ngn*), som i *förseningen förorsakade honom åtskilligt bryderi*. Slutligen exemplifierar **upplysa** (i betydelse **2**) konstruktionsuppgifterna för ett monotransitivt verb med optionellt bundet adverbial i form av prepositionsfras med *om* som huvudord. Som stöd åt bl.a. konstruktionsuppgifterna innehåller ordboken redigerade, ”semi-autentiska”, och redaktionella exempel (Malmgren 2009a:16), vilket är i linje med vad som ofta anses lämpligt för just illustrativa exempel i ordböcker (Atkins & Rundell 2008:457).

Bland de nordiska enspråkiga ordböckerna har *Den Danske Ordbog* (DDO) utmärkt sig för att ha grundliga konstruktionsuppgifter (se t.ex. Malmgren 2009a:16). Konstruktionsuppgifter är för svenskans del sedan tidigare inte bara företrädda i SO (och i föregångaren NEO), utan står också att finna i konstruktionsordboken *Svenskt språkbruk. Ordbok över konstruktioner och fraser* (2003, hädanefter SSB) och i den svenska delen av den elektroniska inlärarordboken *Lexins svenska lexikon*. Den förstnämnda, SSB, har emellertid fått kritik för att uppvisa luckor i inventariet av uppslagsord och innehålla alltför oredigerade, språkligt tveksamma och nyckfullt utvalda konstruktionsuppgifter och språkprov (Larsson 2004; Malmgren 2004). SSB finns heller inte i elektronisk form, och det finns, såvitt kan förstås, inga planer på att uppdatera eller digitalisera ordboken.

Föregångaren till *Lexins svenska lexikon*, den enspråkiga svenska inlärarordboken *Svenska ord*, innehöll konstruktionsuppgifter redan 1984, vilket den var tidigast med i Norden (Malmgren & Toporowska Gronostaj 2009:183). Lexins svenska ordböcker är emellertid speciellt anpassade för undervisningen i svenska som andraspråk, och antalet uppslagsord är därför i princip begränsat till de vanligaste orden: totalt innehåller *Lexins svenska lexikon* drygt 30 000 uppslagsord (se Loenheim 2012:300), vilket kan jäm-

föras med SO:s 65 000. Konstruktionsuppgifter ges för verb, men bara för de viktigaste adjektiven och substantiven (Malmgren & Toporowska Gronostaj 2009:187). Någon kontinuerlig uppdatering av konstruktionsuppgifterna sker inte.

SO får alltså sägas fylla en viktig roll som informationskälla även för konstruktionsuppgifter i modern svenska. Den är en uppdaterad och omfattande allmänordbok med en vid målgrupp – den vänder sig till ”dem som vill lära sig mer om det svenska ordförrådet”, (SO s. IX) – en målgrupp som också inkluderar avancerade inlärare (Sköldberg 2017:123). Rollen blir inte mindre viktig av att SO dessutom verkar sakna omedelbar konkurrens även från nätresurser i fråga om att bidra med omsorgsfullt utvalda konstruktionsuppgifter tillsammans med redigerade exempel.

2.2. Övergripande åtgärder i SO:s konstruktionsuppgifter

Arbetet med SO2 har präglats av fortsatt strävan mot ett tydligt beskrivningsspråk och användarvänlighet. Denna strävan har bl.a. kommit till uttryck i ett antal utrymmeskrävande åtgärder. Eftersom ordboken är planerad för att enbart utkomma i elektroniskt format, kommer dessa ganska sällan att orsaka problem. Många förkortningar har exempelvis kunnat lösas upp, och för konstruktionsuppgifternas del har pronominella variabler som *ngn* och *ngt* expanderats och framförs i klartext med kapitäler som *NÅGON* respektive *NÅGOT*. Även förkortningar som *V* har expanderats till *VERB*. (Upplösningen av förkortningar har både Malmgren (2012) och Loenheim (2012) tidigare framlagt som önskemål för *Lexins svenska lexikon*.)

I huvudsak har befintliga metaspråkliga uttryck för variablerna i konstruktionsuppgifterna behållits. Även SO2 är tänkt att differentiera bl.a. *BELOPP* (för bl.a. objekt och rektioner som betecknar pris), *SÄTT* (för bl.a. sättsadverbial) och *TAL* (för bl.a. räkneord).

Några smärre tillägg har också införts: exempelvis har argumentrollen DJUR introducerats för att signalera att referenten nödvändigtvis är ett djur. Djurreferenter etiketteras i första upplagan av SO som *ngt*, och genom att byta ut beteckningen kan det i många fall adderas en nyans i konstruktionsuppgifterna som tidigare inte var synlig i denna informationskategori. Förutom detta kan betydelsen illustreras på ett semantiskt adekvatare sätt. Exempelvis har **dressera 1** betydelsebeskrivningen 'metodiskt öva in (viss färdighet) hos djur' i huvudbetydelsen. Konstruktionsuppgiften, *dressera ngn/ngt*, delas med underbetydelsen som avser personer, 'öva in själlös disciplin hos'. Denna konstruktionsuppgift delas nu i stället upp på huvud- och underbetydelse som *dressera* DJUR respektive *dressera* NÅGON (mer om uppdelningen i 2.4 nedan).

2.3. Subjekt och finit verb

En informationslucka i konstruktionsuppgifterna i SO är avsaknaden av angivelse av subjektet i aktiv påståendesats. Både *Lexins svenska lexikon* och SSB inkluderar denna uppgift (SSB dock bara ibland). Att SO inte inkluderar subjektet är ofta inget större problem, särskilt inte för modersmålstalaren av svenska, som ganska enkelt torde kunna förutsäga subjektet. Men luckan märks desto tydligare i artiklar där betydelseindelningen åtminstone delvis kan hänföras till subjektets form och/eller funktion. Ett typexempel är verbet **regna**, för vilket konstruktionsuppgiften i SO ser ut på följande sätt:

KONSTRUKTION: *regna*

Uppgiften tillför förvisso informationen att verbet används intransitivt, men i övrigt bidrar upplysningen inte med så mycket för den som vill veta hur *regna* används syntaktiskt. Med väderverb är det inte sällan dessutom så att de konstrueras på två sätt,

vilket signaleras av artikelstrukturen: dels nödvändigtvis med expletivt subjekt, *det* (som i huvudbetydelsen), dels med bildlig betydelse (som i underbetydelserna) och då typiskt med referentiellt subjekt, t.ex. *bladen regnar*.

Andra typer av subjekt som med fördel kan anges i konstruktionsuppgifterna är mer eller mindre obligatoriska formella subjekt, såsom i artikeln för **duga**. I SO antyder konstruktionsuppgiften *duga* (*att*+V) att infinitivfrasen är ett objekt till *duga*, men i ett av exemplen, *det duger inte att sova bort hela förmiddagen*, framgår det att det är frågan om en postponerad infinitivfras med det expletiva *det* som subjekt.

Av bl.a. ovannämnda skäl är subjektet tänkt att anges i konstruktionsuppgifterna i SO2. Inspirationskällan och i viss mån utgångspunkten har varit *Lexins svenska lexikon*, som i huvudsak använder sig av tre sorters subjektsangivelser: *ngn* (om personer), *ngt* (om icke-personer) och expletivt *det*. Notera att subjektsangivelsen även i SO2 är begränsad till dessa tre huvudtyper (inklusive *NÅGRA* för pluralt subjekt). Subjektet realiseras i form av en nominalfras (jfr Toporowska Gronostaj 1996:157), och subjekt i form av infinitivfraser eller bisatser specificeras inte utan inordnas i *NÅGOT*.

Notera att SO2, i likhet med *Lexins svenska lexikon*, inte gör någon skillnad mellan olika referenter. Ett verb som *arrestera* representeras exempelvis i tidigare Lexinupplagor som *Svenska ord* på detta sätt: ⟨*A arresterar B*⟩, där *A* och *B* betecknar olika referenter. I SO2 blir representationen alltså fortsatt av typen *NÅGON arresterar NÅGON*. Även om *A* och *B* tydliggör referensskillnaden, torde ändå få användare tilldela de båda *NÅGON* samma referens. I de fall objektet nödvändigtvis betecknar samma referent som subjektets referent används också i regel reflexivpronomen i konstruktionsuppgiften. För ett partikelverb som **ta på** *gör man alltså i SO2 en skillnad mellan NÅGON tar på (NÅGON) NÅGOT* (som i *han tog på henne/sig overallen*) och den absoluta användningen *NÅGON tar på sig* (t.ex. *han tog på sig och gick*).

Subjektsuppgifterna i *Lexins svenska lexikon*:s konstruktionsuppgifter har automatiskt överförts till SO:s databas för de verb som delas med SO. Detta ger redaktören en ledning om vad subjektet kan vara, men skillnaderna i artikelstruktur mellan de båda ordböckerna medför att en bedömning ändå får göras i det enskilda fallet. Konstruktionsuppgifterna för ett verb som *inleda* kan alltså inte utan vidare föras över från *Lexins svenska lexikon*, eftersom verbet där bara har en enda betydelse, 'påbörja, starta', parad med konstruktionsuppgiften *ngn inleder ngt* och exempel som *ordföranden inledde diskussionen*. Motsvarande artikel i SO lånar uppgiften om personsubjekt, *ngn*, enbart till den första huvudbetydelsen: konstruktionsuppgiften är i sin helhet *NÅGON inleder (NÅGOT) (med NÅGOT/att+V)*. Denna huvudbetydelse innefattar emellertid i SO också en underbetydelse som gäller 'även om företeelse', nämligen 'utgöra början till', som i *en kort bakgrundsteckning inleder artikeln*. Denna matchar inte konstruktionsuppgiften med personlig subjektsreferent, så underbetydelsen får tilldelas konstruktionsuppgift manuellt (mer om detta i 2.4 nedan).

Den subjektsangivelse som införs i SO2 har också lett till att redaktören uppmärksammas på och i en del fall förändrar befintlig artikelstruktur. Ett återkommande problem är vilken sorts subjekt av *NÅGON* eller *NÅGOT* som ska anges, vilket kan illustreras av partikelverbet **räkna ut**. Det första subjekt som infinner sig till detta verb är väl för de flesta ett mänskligt, *NÅGON*, vilket i detta fall också får stöd av SO:s betydelsebeskrivning 'komma fram till [...] *genom tankearbete*' (min kursivering). Men i likhet med många andra mänskliga aktiviteter sköter maskiner och datorer en del av uträknandet (bl.a. vid uträknandet av väderprognoser), dock inte, som i detta fall, med tankekraft. I fall som dessa kan redaktören använda Språkbankens verktyg Korp (Borin, Forsberg & Roxendal 2012) med målet att försöka identifiera de frekventaste typerna av subjekt. I fallet **räkna ut** visar sig *NÅGON* vara ett rimligt val för huvudbetydelsen, då denna typ av subjekt är klart frekventare än

andra typer. Men även icke-personer förekommer, vilket måste leda till en översyn av såväl huvudbetydelsens beskrivning som artikelstrukturen. En lösning är att införa det maskinella uträknandet som en underbetydelse med egen konstruktionsupplysning.

Redaktören kommer i många fall en god bit genom att söka efter ”<vilket ord som helst> + <det aktuella verbets lemgram, dvs. böjningsformer>”. En sådan rudimentär sökning har dock sina begränsningar, bl.a. den att man bara letar efter just ordningen subjekt + verb i resultaten. Det kan också finnas behov av ett sökverktyg som tar hänsyn till associationsmått (se Herbst 2018), bl.a. för att man inte ska riskera att fånga in ord som råkar vara frekventa överlag i korpusen. Här kan man använda Korps funktion Ordbild, som ordnar bl.a. ett verbs subjekts- och objektsargument utifrån måttet Lexicographer’s Mutual Information (se *Swe-Clarins handbok*). Redaktörens språkliga intuition krävs dock, för det är inte säkert att det som redovisas som subjekt till ett verb är subjektet i den aktiva satsen. Exempelvis är de högst rankade subjekten till verbet *tvätta* i Korps korpusar GP och Webbnyheter *citron* och *kläder*. I flera fall är de subjekt i passiva satser (t.ex. *här tvättas kläder i 40 grader*). I fallet *regna* är det också viktigt att vara uppmärksam; här ger den utökade sökningen ”<vilket ord som helst> + <*regna* som lemgram>” det förväntade resultatet att subjektet *det* är frekvent tillsammans med *regna* i olika former i korpusen. I ordbilden ges dock istället *granat* och *raket* som de högst rankade, och *det* listas inte.

2.4. Placering i mikrostrukturen

Ett önskemål som framförs av bl.a. Nikula (2010:368) är att konstruktionsuppgifterna i SO placeras högre upp i ordboksartikeln, ovanför språkexemplen. Detta är tänkt att uppfyllas i SO2, eftersom konstruktionsuppgifterna är just abstraktioner av generella mönster som är tänkta att konkretiseras av exemplen och dessutom tjäna som riktmärke för betydelseindelningen.

En idé som lånats av DDO (se t.ex. Lorentzen & Trap-Jensen 2005:254) är att ge konstruktionsuppgifter även för underbetydelserna i tillämpliga fall. I första upplagan av SO ges endast en samlad konstruktionsinformation i slutet av huvudbetydelsen, och det blir upp till ordboksanvändaren att koppla en viss konstruktionsuppgift till rätt underbetydelse. Det är lätt att inse att detta system kan vålla problem för användaren, varför SO2 kommer att följa principen om att (minst en av) huvudbetydelsens konstruktionsuppgifter ska gälla även för underbetydelserna. Om så inte är fallet ska en ny konstruktionsuppgift ges för underbetydelsen. Ett exempel är [1]**blåsa 1**, vars tänkta SO2-artikel ges nedan i förkortad form:

[1]**blåsa** verb
1 (vanligen i opersonlig konstruktion) naturligt strömma fram luft (parallellt med jordytan) på grund av lufttrycksskillnader
KONSTRUKTION: *det blåser* (MÅTT)
EXEMPEL: *det regnade och blåste*; *det blåste full storm*

- ibland med tonvikt på förändring, särskilt ökning (vanligen med partikel, särskilt **på**, **upp**)
  KONSTRUKTION: *det blåser* (*upp*) (*till* NÅGOT), *det blåser* (*på*)
  EXEMPEL: *det blåste upp till storm*; *på kvällen vred vinden till väst och det blåste på ordentligt*
- ofta även (vanligen med partikel, t.ex. **bort**, **igen**, **ner**, **omkull**) följa med eller drivas av strömmande luft
  KONSTRUKTION: NÅGOT *blåser* (*bort*/*igen*/*ner*/*omkull*)
  EXEMPEL: *papperen blåste bort*; *fönstret blåste igen*; *trädet har blåst ner*; *flaggstången blåste omkull*
- även bildligt råda uppståndelse
  EXEMPEL: *det har alltid blåst kring den berömde advokaten*

- även (med partikeln **över**) sluta råda uppståndelse
  KONSTRUKTION: NÅGOT *blåser över*
  EXEMPEL: *skandalen valsar runt i tidningarna just nu, men det blåser snart över*

Som synes gäller olika konstruktionsuppgifter för olika underbetydelser, och om detta inte är fallet (som i den bildliga underbetydelsen 'råda uppståndelse' ovan) hänvisas läsaren till huvudbetydelsens konstruktionsuppgifter.

2.5. Språkprov som stöd åt konstruktionsuppgifterna

Konstruktionsuppgifternas tillgänglighet har diskuterats tidigare (bl.a. av Atkins & Rundell 2008:400–402 och Malmgren & Toporowska Gronostaj 2009). Hur gör man dem begripliga utan att de förlorar beskrivningskraft? I arbetet med SO2 har inte förenkling av konstruktionsuppgifterna i sig prioriterats, utan i stället är en strävan att varje konstruktionsuppgift, i första hand dem för nyinlagda uppslagsord, ska tillgängliggöras genom att förses med konkretiserande språkexempel. Detta gäller även för led som anges vara optionella, dvs. står inom parentes. I andra hand måste de konstruktionsuppgifter som inte kompletteras av exempel utfyllas med sådana.

Inledningsvis behålls informationsuppdelningen med ett informationsfält för samtliga konstruktionsuppgifter och ett annat för samtliga språkexempel, men i ett senare skede kan man tänka sig ett närmande mot det system som används i specialordböcker som SSB eller i allmänordböcker som *Collins COBUILD English Dictionary Online*. I båda dessa ges konstruktionsuppgifter i omedelbar anslutning till språkexemplen. Detta system har framförts som önskemål en tid (Cantell & Martola 1996:217), men den tryckta pappersordbokens utrymmesbegränsningar har förmodligen varit med och bevarat det rådande systemet.

2.6. Expandering av konstruktionsuppgifter

Flera studier har uppmärksammat att vissa konstruktionsuppgifter kan vara svårtolkade; se t.ex. Bergenholtz & Vrang (2005:176) för situationen i DDO, Malmgren (2009b:99–100) för NEO och Malmgren & Toporowska Gronostaj (2009) för SO.

Malmgren & Toporowska Gronostaj (2009:186) uppskattar förvisso att flertalet konstruktionsuppgifter i SO är ganska enkla: exempelvis avser drygt 2 000 konstruktionsuppgifter monotransitiva verb med sakobjekt, t.ex. **besiktiga** med konstruktionsuppgiften *besiktiga ngt*. Samtidigt kan det konstateras att några tusen verbala konstruktionsuppgifter ändå kan tydliggöras. Exempelvis innehåller drygt 1 200 verbala konstruktionsuppgifter minst tre snedstreck för alternativa led, som i *signalera (till/åt ngn/ngt) (att+V/SATS)*.

I fall som det närmast ovanstående är det gissningsvis såväl kompakthet som grammatisk notation som gör konstruktionsuppgiften potentiellt svårtillgänglig. Ett steg på vägen mot att tillgängliggöra förståelsen av den grammatiska notationen skulle kunna vara att göra varje led i konstruktionsuppgifterna klickbara och försedda med en, gärna ganska verbspecifik, beskrivning av vad de betyder. Om detta blir en realitet är inte klart, men något som har implementerats är att skriva ut de kombinationer som konstruktionsuppgifterna rymmer. Eftersom SO tillämpar ett kompakt notationssystem, med parenteser för optionella led och snedstreck för alternativa led, kan det även för den avancerade användaren vara svårt att föreställa sig vilka mönster som kan produceras. Konstruktionsuppgifterna för verbet **hemställa** ('formellt begära') ser exempelvis ut så här i första upplagan av SO:

hemställa (hos/till ngn/ngt) (om) ngt/att+V/SATS

Bakom denna kortfattade information döljer sig ett stort antal olika konstruktionsuppgifter som kan vecklas ut i redaktionsgräns-

snittet (om det blir implementerat på nedanstående sätt i SO2 är just nu inte klart):

NÅGON *hemställer* (hos NÅGON) (*om*) NÅGOT
NÅGON *hemställer* (hos NÅGON) (*om*) *att*+VERB
NÅGON *hemställer* (hos NÅGON) (*om*) SATS
NÅGON *hemställer* (hos NÅGOT) (*om*) NÅGOT
NÅGON *hemställer* (hos NÅGOT) (*om*) *att*+VERB
NÅGON *hemställer* (hos NÅGOT) (*om*) SATS
NÅGON *hemställer* (till NÅGON) (*om*) NÅGOT
NÅGON *hemställer* (till NÅGON) (*om*) *att*+VERB
NÅGON *hemställer* (till NÅGON) (*om*) SATS
NÅGON *hemställer* (till NÅGOT) (*om*) NÅGOT
NÅGON *hemställer* (till NÅGOT) (*om*) *att*+VERB
NÅGON *hemställer* (till NÅGOT) (*om*) SATS

Förutom att generera bakomliggande konstruktionsvarianter – dock utan att ge alla möjliga kombinationer av optionella led – klargör expanderingen vad som hör till vad, bl.a. att *att+V/SATS* inte avser antingen infinitivfras med infinitivmärket *att* eller bisats inledd av *att*, utan antingen infinitivfras med *att* eller bisats oavsett inledare.

2.7. Sökbara konstruktionsuppgifter

Ett valensperspektiv på konstruktionsuppgifter har sina fördelar, exempelvis att det förklarar enskilda verbs ofta oförutsägbara komplementrestriktioner (Herbst 2018), men det har också sina nackdelar (se t.ex. Lyngfelt 2008). För att komma ett steg närmare att fånga generaliseringar om verbs betydelser och användningar (jfr Levin 1993), bör man därför kunna supplera den lexikala informationen med sökningar efter särskilda mönster. En fråga kan vara: Hur kan överföringsbetydelse konstrueras, eller översatt till en av de möjliga syntaktisk-semantiska strukturerna, vilka verb

kombineras enligt mönstret *NÅGON + VERB + NÅGOT (till NÅGON)*? I detta fall är svaret att det aktuella syntaktiska mönstret aktualiseras med verb som *förmedla*, *hyra ut* och *räcka*. Denna typ av sökningar har hittills inte låtit sig göras inom ramen för den i huvudsak tryckta första upplagan av SO.

3. Några återstående frågor

I avsnitt 2 ovan beskrivs ett antal revisionspunkter som är tänkta att i någon utsträckning genomföras inför publiceringen av SO2. Flera frågor återstår dock att överväga, och ett par av dem redovisas i detta avsnitt.

3.1. Hur många mönster ska ges?

En fråga som ofta gör sig påmind är vilken sorts syntaktiska operationer som ska illustreras med konstruktionsuppgifterna. Konstruktionsväxlingar ges emellanåt i SO:s konstruktionsuppgifter, bl.a. för reciproka verb som **samsas**: *samsas med ngn/ngt* (*om ngt*), men även *ngra samsas* (*om ngt*). Mer regelbundna grammatiska konstruktionsväxlingar som passivering behandlas däremot inte, och det är kanske få som invänder emot detta. Men det kunde finnas skäl att ange undantagen, exempelvis för verb som i vissa betydelser typiskt är passiva. Ett exempel är **fördjupa**, som i huvudbetydelsen 'göra djupare' försetts med den reviderade konstruktionsuppgiften *NÅGON fördjupar NÅGOT*. Redan i första underbetydelsen, 'göra mer inträngande', gör sig emellertid en nästan obligatorisk passivkonstruktion gällande i exemplet *deras vänskap fördjupades med åren*. I den andra underbetydelsen, med kommentaren "i pass. och perf. particip ofta i väderrapporter och dylikt", med exemplet *lågtrycket fördjupas*, blir passiven inte mindre ofrånkomlig. Ett omvänt exempel illustreras av jämförelseverbs-

konstruktioner som *NÅGON/NÅGOT liknar NÅGON/NÅGOT*, där passiv motsvarighet inte vore acceptabel (**NÅGON/NÅGOT liknas av NÅGON/NÅGOT*).

Man kan också fråga sig hur syntaktiskt specifika mönstren ska vara, framför allt i fråga om satser. Att bisatser av olika typer, bl.a. *att*-satser och interrogativa bisatser, förs samman under etiketten *SATS* i SO kan innebära en viss informationsförlust, som såväl Malmgren & Toporowska Gronostaj (2009:189) som Nikula (2010:368) konstaterar. I flera fall har konstruktionsuppgiften preciserats i jämförelse med föregångaren NEO, bl.a. vid verb som *undra* i *jag undrar om det inte är dags för kaffe nu*. Här anges numera *om SATS* (och inte bara *SATS*, som i NEO) som ett möjligt komplement. I många andra fall kan man dock fortsätta preciseringen, t.ex. för verbet *verka*. Konstruktionsuppgiften *verka som SATS*, till betydelsen 'ge intryck av att vara', följs exempelvis även av *som* + konditional bisats (t.ex. *det verkade som om hon var arg*). En slutlig detalj är att *SATS* i ett antal fall visserligen innefattar satsledsfungerande men huvudsatsformade satser (som i *Svenska Akademiens grammatik*, SAG 4:844–846, benämns "citatmeningar"). Exempel inkluderar anföringsverb som **anmärka 2** '(i korthet) yttra', med konstruktionsuppgiften *anmärka SATS* och exempel som *"De skjuter dåligt", anmärkte generalen mitt under attentatet*. Här skulle man kanske kunna införa rollen *YTTRANDE* (som i den tänkta konstruktionsuppgiften *anmärka YTTRANDE*) för att signalera att det inte är vilket satsformat satsled som helst det är frågan om.

En tänkbar utveckling av de expanderade konstruktionsuppgifterna (avsnitt 2.6 ovan) vore slutligen att på ett eller annat sätt tydliggöra brukliga ledföljder inom ramen för en subjekt-verb-ledföljd. Exempelvis kan man formulera sig i enlighet med konstruktionsuppgiften under **proppa**, *NÅGON proppar NÅGON/NÅGOT (ADJ) (med NÅGOT)*, men inte gärna varianter som *NÅGON proppar (ADJ) (med NÅGOT) NÅGON/NÅGOT* (*de proppade väskan full med böcker*, men inte *de proppade full med böcker väskan*). Av värde

kunde också vara att tydliggöra notationen i fråga om hur optionella led kan användas. Detta gäller bl.a. för konstruktionsuppgifter som SO:s *köpa ngt (av ngn) (till/åt ngn) (för* BELOPP*)*. Här representerar konstruktionsuppgiften en fullt grammatisk sats som dock inte är alldeles lätt att föreställa sig om samtliga (alltså även optionella) led används i den angivna ordningen. Om det är ordbokens uppgift att ange konstruktionsuppgifter av denna detaljeringsgrad kan förvisso diskuteras.

3.2. Normativa eller deskriptiva konstruktionsuppgifter?

Med den avancerade inläraren som presumtiv användare av SO, är det inte oviktigt att fundera över ordbokens övergripande deskriptiva karaktär. Att SO överlag har en deskriptiv inriktning är ställt utom tvivel, men frågan är om deskriptiviteten gäller fler kategorier än stavning, böjning och ordbildning, där SAOL 14 kan komplettera med mer normativa upplysningar. Eller, kan man tänka sig att även konstruktionsuppgifterna innefattas? Jämför t.ex. med den normativa inriktningen hos konstruktionsordboken *Svensk handordbok. Konstruktioner och fraseologi* (1966), som uttryckligen samordnades med och kompletterade den då aktuella SAOL 9 på fraseologins område. Inledningen till nämnda konstruktionsordbok (s. VII) meddelar att ett syfte med boken är att ”varna för felaktigt eller mindre lämpligt språkbruk”.

Som Larsson (2004:211) konstaterar i en recension av SSB, kan bl.a. ordboksanvändare med annat modersmål än svenska förvänta sig att ordboken ska vara normativ, att den ska ge vägledning i ”det goda språkbruket” (några studier av vad användarna förväntar sig av konstruktionsuppgifter i en ordbok känner jag förvisso inte till). Möjligen är SO också i praktiken mer normativ i hanteringen av konstruktionsuppgifter än i hanteringen av exempelvis stavningsvarianter. Kanske delvis eftersom konstruktionsuppgifterna, till skillnad från bl.a. stavningsvarianter, saknar bruklig-

hetsmarkeringar. Detta skulle kunna bidra till att konstruktionsuppgiftskandidater får invänta acceptans i den neutrala sakprosan innan de kan lanseras. Utelämnandet av infinitivmärket efter *komma* har exempelvis uppmärksammats ganska länge: i grammatikböcker sedan åtminstone 1980-talet (se t.ex. Jörgensen & Svensson 1987:77). Det har emellertid dröjt till 2009 innan konstruktionsuppgiften för hjälpverbet *komma* (närmare bestämt **²komma 6**) uppdaterats, från NEO:s *komma att+INF*, dvs. med obligatoriskt infinitivmärke, till SO:s *komma* (*att+*)*V*, med optionellt infinitivmärke. Som jämförelse kan nämnas att *Språkriktighetsboken* (2005:360), som tidsmässigt publicerades mellan NEO och SO, konstaterar att konstruktioner med *komma* utan *att* kommit att bli ”fullt acceptabla”.

Frågan är om SO fortsättningsvis ska införliva frekventa men i alla texttyper kanske inte fullt etablerade konstruktionsmönster. Är det dags att förteckna ofta stigmatiserade verbkombinationer som *visa på* NÅGOT och öppna upp NÅGOT, och bör pekpinnarna användas mot fraser som *hata på* NÅGON? Bruklighetsmarkeringar även i konstruktionsuppgifterna kunde vara ett sätt att hantera frågan.

4. Sammanfattning

Som framgått ovan kommer en del märkbara förändringar att genomföras i SO:s konstruktionsuppgifter, medan andra ännu är önskemål från artikelförfattaren. Många av ändringarna tillkommer tack vare det elektroniska formatets utrymmesfördelar och möjligheter till länkning mellan informationskategorier. Detta innebär naturligtvis inte att all tänkbar konstruktionsinformation därför kan eller ska ges för alla verb, vare sig specifikare information om exempelvis alla typer av tänkbara subjekt (se 2.3 ovan) eller mer generell grammatisk information såsom passivering (se 3.1 ovan). Vidare kan det diskuteras hur öppna redaktörerna ska

vara för att inkludera konstruktionsuppgifter med mer inskränkt bruklighet (se 3.2 ovan).

I alla händelser bör det påpekas att förändringarna i SO2 är små om man jämför med det gedigna grundarbete som redan har utförts i föregångaren. Det är fråga om finjusteringar snarare än om omvälvningar.

Litteratur

Ordböcker

Collins COBUILD English Dictionary Online. <https://www.collinsdictionary.com/dictionary/english> (januari 2019).

DDO = *Den danske ordbog*. <https://ordnet.dk/ddo> (januari 2019).

Lexins svenska lexikon = *Svenska ord*, 4 uppl. <http://lexin.nada.kth.se/lexin> (januari 2019).

NEO (1995–1996) = *Nationalencyklopedins ordbok*. Höganäs: Språkdata Göteborg & Bokförlaget Bra Böcker AB.

SAOL 9 (1950) = *Svenska Akademiens ordlista*. 9 uppl. Stockholm: Svenska Bokförlaget/Norstedts.

SAOL 14 (2015) = *Svenska Akademiens ordlista*. 14 uppl. <https://svenska.se/saol> (januari 2019).

SO (2009) = *Svensk ordbok utgiven av Svenska Akademien*. <https://svenska.se/so> (januari 2019).

SSB (2003) = *Svenskt språkbruk. Ordbok över konstruktioner och fraser*. Utarbetad av Svenska språknämnden. Stockholm: Norstedts ordbok.

Svensk handordbok. Konstruktioner och fraseologi (1966). Ture Johannisson & Karl Gustav Ljunggren (red.). Stockholm: Esselte Studium.

Svenska ord (1984) = *Svenska ord – med uttal och förklaringar*. Stockholm: Skolöverstyrelsen/Esselte studium.

Annan litteratur

Atkins, B. T. Sue & Michael Rundell (2008): *The Oxford Guide to Practical Lexicography*. Oxford: Oxford University Press.

Bergenholtz, Henning & Vibeke Vrang (2005): Den Danske Ordbog bind 2 (E–H) og 3 (I–L) – en ordbog for folket eller for akademikere? I: *LexicoNordica* 12, 169–191.

Borin, Lars, Markus Forsberg & Johan Roxendal (2012): Korp – the corpus infrastructure of Språkbanken. I: *Proceedings of LREC 2012*. Istanbul: ELRA, 474–478.

Cantell, Ilse & Nina Martola (1996): Nationalencyklopedins ordbok. I: *LexicoNordica* 3, 209–222.

Faulhaber, Susen (2011): *Verb Valency Patterns. A Challenge for Semantics-Based Accounts*. Berlin/New York: De Gruyter.

Herbst, Thomas (2014): The valency approach to argument structure constructions. I: Thomas Herbst, Hans-Jörg Schmid & Susen Faulhaber (red.): *Constructions – Collocations – Patterns*. Berlin/Boston: De Gruyter, 167–216.

Herbst, Thomas (2018): Is language a collostructicon? A Proposal for Looking at Collocations, Valency, Argument Structure and Other Constructions. I: Pascual Cantos-Gómez & Moisés Almela-Sánchez (red.): *Lexical Collocation Analysis. Quantitative Methods in the Humanities and Social Sciences*. Cham: Springer, 1–22.

Howarth, Peter (1998): Phraseology and second language proficiency. I: *Applied Linguistics* 19/1, 24–44.

Jörgensen, Nils & Jan Svensson (1987): *Nusvensk grammatik*. Malmö: Liber.

Larsson, Lennart (2004): Svenskt språkbruk. Ordbok över konstruktioner och fraser. Utarbetad av Svenska språknämnden. I: *Språk och stil* NF 14, 209–219.

Levin, Beth (1993): *English Verb Classes and Alternations: A Preliminary Investigation*. Chicago: University of Chicago Press.

Loenheim, Lisa (2012): Lexins svenska lexikon. I: *LexicoNordica* 19, 295–316.

Lorentzen, Henrik & Lars Trap-Jensen (2005): Grammatiske oplysninger i Den Danske Ordbog. I: Ruth Vatvedt Fjeld & Dagfinn Worren (red.): *Nordiske studiar i leksikografi* 7. Oslo: Nordisk foreining for leksikografi, 252–266.

Lyngfelt, Benjamin (2008): Huvuddriven valens – konstruktionsdriven argumentstruktur? I: Kristinn Jóhannesson, Hans Landqvist, Aina Lundqvist, Lena Rogström, Emma Sköldberg & Barbro Wallgren Hemlin (red.): *Nog ordat? Festskrift till Sven-Göran Malmgren den 25 april 2008*. Göteborg, 272–279.

Malmgren, Sven-Göran (2004): En ny svensk konstruktionsordbok. I: *LexicoNordica* 11, 271–294.

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, 14–20.

Malmgren, Sven-Göran (2009b): On production-oriented information in Swedish monolingual defining dictionaries. I: Sandro Nielsen & Sven Tarp (eds.): *Lexicography in the 21st Century: In honour of Henning Bergenholtz*. Amsterdam/Philadelphia: John Benjamins, 93–102.

Malmgren, Sven-Göran (2012): Från Svenska ord (Lexin) 3 till Svenska ord 4. I: Birgit Eaker, Lennart Larsson & Anki Mattisson (red.): *Nordiska studier i lexikografi 11*. Lund: Nordiska föreningen för lexikografi, 454–465.

Malmgren, Sven-Göran & Maria Toporowska Gronostaj (2009): Valensbeskrivning i svenska ordböcker – och några andra. I: *LexicoNordica* 16, 181–196.

Nikula, Kristina (2010): Svensk ordbok – en guldgruva för språkintresserade. I: *LexicoNordica* 17, 351–375.

SAG = Ulf Teleman, Staffan Hellberg & Erik Andersson (1999): *Svenska Akademiens grammatik*. Del 1–4. Stockholm. <https://svenska.se/grammatik/> (augusti 2019).

Sköldberg, Emma (2017): Innehållet i *Svensk ordbok utgiven av Svenska Akademien* – eller kampen mellan norm och bruk. I: Saga Bendegard, Ulla Melander Marttala & Maria Westman (red.): *Språk och norm. Rapport från ASLA:s symposium, Uppsala universitet 21–22 april 2016*. Uppsala: ASLA, 123–129.

Sköldberg, Emma & Anna Helga Hannesdóttir (2017): Svenska ord – men vilka? Om uppslagsorden i *Svensk ordbok utgiven av Svenska Akademien*. I: Emma Sköldberg, Maia Andréasson, Henrietta Adamsson Eryd, Filippa Lindahl, Sven Lindström, Julia Prentice & Malin Sandberg (red.): *Svenskans beskrivning* 35. Göteborg: Göteborgs universitet, 329–340.

Språkriktighetsboken (2005). Stockholm: Norstedts.

Swe-Clarins handbok. Avsnitt 2: ”Ordbild – Ordens bästa vänner”. <https://sweclarin.se/swe/handbok#__RefHeading___Toc469052362> (januari 2019).

Toporowska Gronostaj, Maria (1996): *Integrerad valensbeskrivning. Mot ett formaliserat verbvalenslexikon*. Göteborg: Göteborgs universitet.

Kristian Blensenius
postdoktor
Institutionen för svenska språket
Göteborgs universitet
Box 200
SE-405 30 Göteborg
kristian.blensenius@gu.se

ANMELDELSER

To store leksikografer – om Svend Eegholm-Pedersens *Mothstudier*

Ebba Hjorth

Svend Eegholm-Pedersen: *Mothstudier. Kildegrundlaget for den første store danske ordbog*. København: Universitets-Jubilæets danske Samfund 2018. 221 sider. Pris (indbunden): 200 DKK.

1. *Mothstudier*s tilblivelseshistorie

I 2018 udgav Universitets-Jubilæets danske Samfund leksikografen[1] Svend Eegholm-Pedersens *Mothstudier. Kildegrundlaget for den første store danske ordbog*[2]. Værket er omfattende med fire hovedafsnit: I Indledning, II Almindelige rigssprogskilder (ordbøger og ordlister), III Specielle kilder og IV Afslutning. Dertil en fyldig Bibliografi. I denne anmeldelse forkorter jeg i den løbende tekst Svend Eegholm-Pedersen SEP, og hans værk forkortes MSt.

I forordet skriver SEP at han i en artikel til Poul Lindegård Hjorths 70-års festskrift *Ord, Sprog oc artige Dict* (Eegholm-Pedersen 1997) fik lejlighed til at belyse forskellige sider af arbejdsgangen ved Moths pionerværk, herunder spørgsmålet om mulige forbilleder og inspirationskilder.

SEP genoptog kildestudierne i 2010'erne, og han har koncentreret sig om (s. 5) den ene del af Moths værk, verbalordbogen (dvs. den sproglige ordbog), mens en undersøgelse af kildegrund-

1 I 1981-1988 udgav Svend Eegholm-Pedersen (sammen med Aage Hansen) *Holberg-Ordbog. Ordbog over Ludvig Holbergs Sprog*, bd. 1-5. Han har desuden været en meget aktiv redaktør ved *Ordbog over det danske Sprog* og *Ordbog over det Danske Sprog. Supplement*.

2 Om Moth og hans ordbog kan der læses på <https://mothsordbog.dk/om-ordbogen>.

laget for realleksikonnet (dvs. det encyklopædiske leksikon) overlades til "andre(s!) nye kildestudier".

SEP kommer i indledningen også kort ind på Mothmanuskriptets udnyttelseshistorie frem til den digitale publicering muliggjort af en Carlsbergfondbevilling til Det Danske Sprog- og Litteraturselskab (DSL) i anledning af Selskabets 100-års jubilæum i 2011.

2. Min private historie med Moth

Her følger nu min private udgave af Moths udnyttelseshistorie.

Moth har siden min studiestart i 1960'erne været et ganske særligt begreb, i begyndelsen naturligvis fuldstændig ukendt for mig, men efterhånden kom der flere og flere elementer til i mit kendskab til Moth. Det begyndte vel med at jeg fra tid til anden stødte på kildebetegnelserne Moth., A-Ø, Moth.[1]A-Ø og Conv.A-Ø i *Ordbog over det Danske Sprog* uden at jeg dog forstod den vidtrækkende betydning af disse kildeangivelser der angav henvisninger til dansk leksikografis største og mest epokegørende enmandsværk, Matthias Moths verbalordbog og realleksikon.

Da jeg blev studentermedhjælper på Københavns Universitets Institut for Navneforskning (nu Afdeling for Navneforskning, Institut for Nordiske Studier og Sprogvidenskab (NorS)), blev jeg ved et symposium mellem danske og svenske navneforskere præsenteret for udvalgte bind af Matthias Moths 62 foliobind store håndskrevne ordbogsværk og fik samtidig en beskrivelse af værket ved Aage Rohmann, en Moth-forsker som SEP henviser til adskillige gange i MSt.

Jeg var dybt fascineret og samtidig glad for at mit arbejde som afskriver af navne i markbøger fra 1681-1683 for navneforskerne gjorde mig i stand til at læse i Moths egenhændige værk!

I et undervisningsforløb i leksikografi ved John Kousgård Sø-

rensen var Moth også på programmet. I kurset stiftede vi bekendtskab med den afskrift af Moths ordbogsværk der blev fremstillet under 2. verdenskrig. Afskriften var tænkt dels som en sikring af værket, dels som en beskæftigelsesforanstaltning, og det var tanken at afskriften skulle være forarbejde til en trykt udgave. Afskriften fylder 11.522 sider, og den er indbundet i 40 bind, verbalordbogen fylder de 27 af de 40 bind og fagleksikonnet de 13 øvrige. Afskriften befinder sig på DSL's bibliotek.

Der blev imidlertid ikke råd til at trykke Moths ordbog, og originalen og afskriften var stadig de eneste muligheder der var for at udnytte det righoldige materiale.

Endnu et initiativ blev taget til at mangfoldiggøre ordbogsværket som jeg, som en kuriositet, også lige vil nævne. Man planlagde at nedfotografere håndskriftsiderne til microfiche-format (gennemsigtige 10,5x14,8 cm kort), der kunne læses i et specielt apparat. Det blev heldigvis ikke til noget!

Derimod gav – som nævnt ovenfor – en generøs jubilæumsgave fra Carlsbergfondet til DSL i anledning af Selskabets 100-års jubilæum i 2011 mulighed for en digital udgivelse af Moths store værk. Jeg glæder mig over sammen med Lars Trap-Jensen at have plantet idéen til denne gave hos Fondet.

Æren for realiseringen, for digitaliseringen og opbygningen af netstedet mothsordbog.dk tilfalder imidlertid alene Marita Akhøj Nielsen. Siden er et fremragende og velstruktureret værktøj til studier i verbalordbogen i både translittereret og faksimileret form, og dertil kommer – ud over beskrivelse af Moths værk og søgevejledninger – adgang til trykt og til digital Mothlitteratur og til dokumenter, der er vigtige i Mothsammenhænge. Også det her anmeldte værk er tilgængeligt fra hjemmesiden i digital form, og de digitale søgemuligheder øger i meget høj grad nytteværdien af SEP's MSt. Disse søgemuligheder vil blive brugt i en del af denne omtale af MSt. Til Moth-ordbogens udgivelseshistorie hører også at den er udgivet som app og kan downloades frit.

3. Strukturen i *Mothstudier*

3.1. Indledning

Første hovedafsnit Indledning har tre underafsnit: I A: Ordbogen og ordbogssamlingerne, I B: Kilderne og I C: Plan for fremstillingen.

I afsnit I A giver SEP en kort beskrivelse af Moths liv og virke, og han gennemgår derefter minutiøst det samlede ordbogsværks indhold fordelt på fem ordbogskategorier: en dansk-latinsk verbalordbog med betydningsbeskrivelser og latinske ækvivalenter, en realordbog, en almindelig latinsk-dansk ordbog, en latinsk nøgle til realordbogen og endelig en mindre polyglotordbog.

I afsnit I B beskrives beskedent begrænsningerne i SEP's kildestudier: Udeladt er realordbogens kilder og kilderne til det omfattende klassiske latinske ækvivalent- og sublemma/citatstof med undtagelse af Erasmus Roterodamus og Peder Laale. Den gennemførte og meget detaljerede undersøgelse af Moths hjemlige, nationalsproglige og nordiske kilder og hjælpemidler er ikke foretaget tidligere, og SEP beskriver ærligt undersøgelsens vanskeligheder: De forskellige, noget afvigende redaktioner har måttet inddrages, og tidligere tilløb til kildebeskrivelser af Moths værk har ikke på samme måde inddraget kilder og hjælpemidler.

Moths kilder er både trykte og forgængeres utrykte ordbogsværker, men epokegørende er det materiale som Moth indsamlede ved en rundsendt skrivelse i 1697 til Danmarks og Norges biskopper, der opfordrede dem til at sætte stifternes præster i gang med at indsamle ord efter nærmere beskrevne retningslinjer: både rigssprogsord og dialektalt ordstof var et ønske. Desuden blev fagord og dyre- og plantenavne efterspurgt. Blandt det indsamlede stof er især fyldige jyske og et par norske ordlister bevaret. Af utrykt kildemateriale spiller også stof fra præsterne og filologerne Laurids Kok og Peder Syv en vigtig rolle.

I afsnit I C præsenteres dispositionen i fremstillingen i SEP's kildestudier.

3.2. Almindelige og specielle kilder

I kapitel II foretages først en analyse af de almindelige rigssprogskilder i form af trykte og utrykte ordbøger og ordlister med hovedvægt på Laurids Kok (II A). Dernæst behandles de bevarede indsamlinger fra 1697 og frem (II B) og endelig (II C og D) Poul Jensen Coldings (*Dictionarium Herlovianum*) og Peder Syvs (*Prøve Paa en Dansk og Latinsk Ord-Bog*) værker.

I kapitel III analyseres en række specielle kilder og hjælpemidler: sprogstof med excerpter af ældre dansk (III A), ordsprog (III B), dialektord fra indsamlingen (III C), norske kilder med bidrag fra indsamlingen (III D), islandske kilder herunder ordbøger (III E) og endelig faglige kilder til verbalordbogen (III F).

Kapitel II analyserer altså de almindelige kilder (Laurids Kok, Indsamlet materiale, Poul Jensen Colding og Peder Syv) og kapitel III beskæftiger sig med specielle kilder (excerpter af ældre dansk litteratur, ordsprog, dialekter, norske (indsamlede) kilder, islandske og faglige kilder). I alt beskrives 10 forskellige typer af kilder. Det kræver dog nogen fordybelse i værket at holde de 10 gruppers materiale ude fra hinanden.

4. Hvordan arbejdede Moth?

Jeg vil i det følgende se på Moths udnyttelse af særlige kilder ud fra mine egne særinteresser og aktive fagområder og med moderne leksikografiske briller på ordbogsforfattelse.

4.1. Anders Sørensen Vedel

Anders Sørensen Vedel udnyttes af Moth til dokumentation af ældre, arkaisk ordstof.

I forarbejder og materialesamlinger til ordbogen findes excerpter af Vedels Saxooversættelse fra 1575 med titlen "Nogle *rare* [sjældne] Danske ord *colligerede* af *Saxonis* Grammatici Danske *Version Andreæ Velleji*". Denne ordsamling er ikke udført af Moth. Skriftkarakteren og ortografien i de ledsagende ordforklaringer peger på en anden ophavsmand (MSt, s. 90), og ikke alle ordene opført i listen er medtaget i ordbogen. Interessen i Vedel-listen er koncentreret om det leksikografiske, om specielle ord og udtryk, men også afvigende former og grammatiske forhold er beskrevet i listen. SEP diskuterer om stof fra denne liste er medtaget direkte af Moth eller om vejen fra Vedel til Moth går gennem andre kilder. Men det er sikkert at ord som *krage*, med betydningen 'kaldes et grenefuldt træ, hvôraf grenene halfdelen er afhuggen, for at klyve op på' tilføjet i margenen, *kriggerrig* 'Er den, som har stôr lyst og attrâ til krîg' og *krogsindet* 'Er den som har et ligt og forvendt sînd' har fundet vej fra Vedel til Moths ordbog.

Vedels Saxooversættelse er således en del af korpusset til Moths ordbog selv om excerperingen ikke er foretaget af Moth selv, men med en anden som mellemled.

4.2. Peder Laale

Peder Laale udnyttes som eksempelstof af Moth i dokumentationen af ordsprog og talemåder.

Der findes ikke excerpter til Peder Laales berømte ordsprogssamling fra middelalderen, som foreligger i tidlige tryk fra 1506 og 1515, men sådanne excerpter har formodentlig eksisteret. SEP's undersøgelser beskriver sammenhængen mellem de forskellige ordsprogssamlinger som er citeret i Moth, og i en fodnote (s. 120) viser

han hvilke metoder han også tager i brug i sine kildestudier. Han skriver her: "Med de foreliggende elektroniske søgefaciliteter i den i 2013 netpublicerede Mothordbog kan der nu gives konkrete tal for citeringen af de med faste forkortelser anførte latinske forfattere." På en liste over de 25 hyppigst citerede ligger Peder Laale på en tolvteplads med 1071 forekomster (til sammenligning er Erasmus Roterodamus citeret 5970 gange (se også s. 197)).

4.3. Anders Arrebo

Anders Arrebo udnyttes af Moth som tekstkorpus.

Fra Anders Arrebos *Hexaëmeron* fra 1661 er verbet *brøste* excerperet, og et eksempel på dette ord optræder i førsteredaktionen i form af et margencitat "og klokken hannem brøster [han mangler klokken, véd ikke hvad klokken er] arrebo 212". Dette citat er dog ikke medtaget i den endelige version af *brøste*. Samme skæbne overgår substantivet *medvand*, der i førsteredaktionen betydningsbeskrives "kaldes det når strømmen er med den seilende" og kildeangives "Med vand Ar. 101". Nogle specielle arreboske vindtypetermer, *Hafgul* 'kaldes en vind, som skal reise sig af sŷdvest, og blæse til Høstmanedt, i stadshavet i Norge' og *Hombakke vind* 'kaldes de svale vestlige vinde, som blæser i hundedagene', optræder i både første- og slutredaktion, men uden kildeangivelse til Arrebo. Det skyldes formodentlig at de tydelige referencer til Arrebo har været støttet af norske ordlistebidrag fra Jonas Ramus i Norderhov og af Jakob Bork indsendt til Mothsamlingen. Anders Arrebo var en periode biskop over Trondheim stift.

Moths udnyttelse af Hexaëmeron er altså leksikografisk set både direkte brug af værket som korpus og indirekte brug via indsendte ordlister med leksikografisk udvalgt interessant ordforråd.

4.4. Peder Syv

Peder Syv udnyttedes af Moth som leksikografisk leverandør til flere forskellige oplysningstyper.

Forholdet mellem Peder Syv og Moth er kompliceret, og det er beskrevet af SEP i flere kapitler i MSt. Moth har udnyttet Syvs trykte ordbogsfragment fra 1692 *Prøve paa en Dansk og Latinsk Ord-Bog* indeholdende bogstaverne A og Å. Et eksempel blandt flere på inspiration fra Peder Syv er opslagsordet *Amen*, der dog hos Moth er fyldigere behandlet med danske definitioner end hos Peder Syv, men de to opslagsord har dog den kendte talemåde *det er så vist, som Amen i kirken* tilfælles (MSt, s. 71). Afhængighedsforholdet mellem Syv og Moth kompliceres yderligere af andre ordbogsaktører i tiden (Laurids Kok, Ole Worm og Poul Jensen Colding), men SEP gør omhyggeligt rede for hvem der kan have skrevet af efter hvem. Da Moth i slutningen af 90'erne for alvor gik i gang med sit store ordbogsprojekt og indsamlede materiale, bad han Syv om at få adgang til resten af hans ordbogsmanuskript, og kildehenvisningen *P. S.*, der optræder flere gange i Moths arbejdsmanuskript, viser at han fik adgang også til andre bogstaver i Syvs manuskript.

Lister over gamle og ellers lidet brugelige ord i Syvs *Nogle betenkninger om det Cimbriske Sprog* 1663 og excerpter af hans ordsprogsamling fra 1682 og 1688 var til rådighed for Moth ved beskrivelsen af det ældre ordstof.

Syv var altså inspirationskilde til flere forskellige oplysningstyper for Moth både via ordbog(sfragmenter), ordlister og gennem ordsprogssamlinger og visestof.

5. Konklusioner

Efter punktstudierne, som SEP stærkt eufemistisk og selvudslet-

tende benævner dem på s. 196, i kapitel II og III følger 12 siders *Afslutning* i kapitel 4.

I afslutningen præciserer SEP at det er kildegrundlaget for verbalordbogens lemmavalg der har været i fokus, mens fagsproget og målsprogets latin med de citerede latinske forfattere ikke har været ærindet.

5.1. Leksikografiske tanker

Når man som leksikograf læser i MSt, er det slående hvor mange og vigtige leksikografiske problemstillinger og overvejelser der er fælles for Moth og en nutidig leksikograf. Set med en moderne leksikografs øjne er det interessant at SEP kan konstatere at Moth har udnyttet sine leksikografiske hjælpemidler – både de trykte og de utrykte – "energisk og ganske tæt". Denne tætte udnyttelse eller genbrug kan konstateres både med hensyn til lemmaudvalg og lemmaudformning, eksempler og definitioner. Udnyttelsen har været tæt for både det utrykte materiale der var stillet til rådighed til fri afbenyttelse, og for de trykte værker.

Det har tilsyneladende i Moths samtid været helt uproblematisk at anvende denne meget frie citeringspraksis. Men SEP's sammenligning mellem førsteredaktionen og slutredaktionen viser dog at teksten i slutredaktionen tit er ændret og rettet til så kildeafhængigheden ikke er helt så tydelig eller næsten forsvundet.

Moth fremstillede ved redaktionen af de enkelte bogstavers begyndelse lister over sine meddelere, men normalt nævnte han dem ikke. En interessant undtagelse udgør forkortelsen "P. S." (dvs. Peder Syv), der ses mange steder ved glosetilføjelserne i førsteredaktionens arbejdseksemplar.

Af andre danske kilder henvises der blandt flere andre til ordsprogssamlingen med "PL" (dvs. Peder Laale). Til Arrebos *Hexaëmeron* henvises der med "arrebo" eller "Ar.".

Kritisk vurdering af de sproglige kilder man som leksikograf

bygger sit arbejde på, er nødvendig. SEP skriver at Moth ikke altid har øvet kritik af sine kilders oplysninger og heller ikke altid har haft mulighed og forudsætning for det. Det gælder især det norrøne og gammeldanske ordforråd, og det har medført adskillige fejl og mangler i værket.

5.2. Afslutning og anbefaling

SEP's værk afsluttes som tidligere nævnt med 12 siders Bibliografi delt op i I Håndskriftkilder, II Trykte kilder. Andre kildeværker og III Sekundærlitteratur.

SEP har med MSt begået et fantastisk grundigt og beundringsværdigt værk, der afdækker Matthias Moths arbejdsmetoder. Som den fremragende leksikograf – med ordbogsværker som *Holbergordbogen* og *Ordbog over det danske sprog* bag sig – og som den omhyggelige og vedholdende forsker han er, har han haft forudsætningerne for at gennemføre disse meget minutiøse studier og at give os moderne leksikografer indblik i at vi har store forgængere inden for leksikografien og meget at leve op til, både i forhold til Matthias Moth og i forhold til Svend Eegholm-Pedersen. Som leksikograf bør man fordybe sig i *Mothstudier* og glæde sig over at være en del af så spændende og udfordrende et fag.

Litteratur

Eegholm-Pedersen, Svend (1997): Fra Moths værksted. Lidt om den første store danske ordbog. I: Flemming Lundgreen-Nielsen, Marita Akhøj Nielsen & John Kousgård Sørensen (red.): *Ord, Sprog oc artige Dict. Et overblik og 28 indblik 1500-1700. Festskrift til Poul Lindegård Hjorth*. København: C.A. Reitzels Forlag, 485-505.

Holberg-Ordbog. Ordbog over Ludvig Holbergs Sprog (1981-1988). Aage Hansen & Svend Eegholm-Pedersen (red.). Bind 1-5. Netpubliceret 2011 af Det Danske Sprog- og Litteraturselskab: <holbergordbog.dk> (juli 2019). Også tilgængelig som app.

Moths Ordbog. <https://mothsordbog.dk/> (august 2019).

Ordbog over det danske Sprog (1919-1956). Bind 1-28. Netpubliceret af Det Danske Sprog- og Litteraturselskab sammen med *Ordbog over det danske Sprog. Supplement* (bind 1-5): <ordnet.dk/ods/> (august 2019). Også tilgængelig som app.

Ebba Hjorth
ledende redaktør, mag.art.
Det Danske Sprog- og Litteraturselskab
Christians Brygge 1
DK-1219 København K
eh@dsl.dk

Norsk banneordbok – ei omfattande samling med eit noko vidt banneomgrep

Tor Erik Jenstad

Ruth Vatvedt Fjeld: *Norsk banneordbok*. Oslo: Humanist forlag 2018. 510 s. Pris 429 NOK.

1. Banning i vinden

Banning har stått meir og meir fram som eit seriøst og stuereint granskingsobjekt den siste mannsalderen. I nordisk samanheng kan ein nok seia at dette starta for alvor med dei populærvitskaplege bøkene til Magnus Ljung (Ljung 1984 og 1987). Det finst sjølvsagt enkelte arbeid om banning som er mykje eldre enn dette, til dømes Hansen (1911) og Jespersen (1911) som noko av det eldste. Men det er først i mykje seinare tid at banning har blitt eit verkeleg ekspansivt forskingsfelt. Dei siste 20 åra har vi såleis fått eit skandinavisk forskarnettverk (SwiSca – Swearing in Scandinavia), eit banneleksikon for barn (Uri, Nilsen & Østby 2011), undervisningsmateriell i banning (Hestnes 2007) og ei bok om «nestenbanning» (Hasund 2005), for å nemne noko. Elevar i vidaregåande skule skriv fordjupingsoppgåver om banning, og ikkje minst er banning stadig eit pirrande emne for media, særleg kvar sommar i agurktida. Samstundes har banneord og andre tabuord fått gradvis meir skikkeleg og seriøs behandling i ordbøker.

Ruth Vatvedt Fjeld har vore sentral i banneforskinga i nordisk samanheng, med arbeid som Fjeld (2001 og 2002). Ho er professor i språkvitskap ved Universitetet i Oslo med norsk ordforråd som spesialområde og treng elles neppe nærare presentasjon, som sentral person både i Nordisk foreining for leksikografi og Euralex gjennom ei årrekke.

Meldaren er leksikograf med. Han har arbeidd på *Norsk Ordbok* (NO) og *Det Norske Akademis ordbok* (NAOB), og er no tilknytt Språksamlingane i Bergen i samband med oppstart for revisjon og digitalisering av NO band 1–5. Han har elles mellom anna gjeve ut *Trønderordboka* saman med Arnold Dalen og fire populærvitskaplege bøker om skjellsord.

2. Generelt

Norsk banneordbok er eit arbeid con amore, som resultat av det redaktøren sjølv har kalla sin «språkbaby», nemleg å samle på banneuttrykk (*ILNytt* 2011). I forordet (s. 7) kallar ho det eit hobbyprosjekt.

Etter det eg kan forstå, er boka einestående i nordisk samanheng. Tanken går til den svenske *Stora fula ordboken* (Dagrin 2013), men denne er likevel utarbeidd etter eit ganske anna konsept.

Tilfanget i banneordboka er henta frå mange ulike kjelder, slik det er gjort greie for i brukarrettleiinga. Av ordbøker er først og fremst *Norsk Riksmålsordbok* (NRO) og NO utnytta, og berre i liten grad NAOB, som kom først i innspurten av arbeidet med banneordboka. Ikring ti elektroniske korpus er nytta; desse er lista opp på s. 58–59. Elles har Fjeld brukt sitt eige banneordforråd og notert banneord etter andre gjennom mange år. Totalt har dette gitt eit imponerande stort materiale med over 1000 lemma.

3. Oppbygginga av boka

Banneordboka startar med eit kort forord om bakgrunnen for arbeidet. Ein fyldig introduksjonsdel (s. 9–48) inneheld ei utgreiing om banninga si historie, ulike typar banning, kven som bannar, kvifor vi gjer det, og kva måtar vi bannar på. Denne bolken har

ei nyttig og viktig opprydding i mange omgrep og kan lesast uavhengig av sjølve ordbokdelen. Her finst instruktive forklaringar av omgrep som *alterbanning*, *egobanning* og *vanebanning*.

Deretter kjem ei brukarrettleiing (s. 49–62), der det er gjort greie for utvalet av oppslagsord og oppbygginga av artiklane. Vi får som nemnt vita kva for korpus som er brukt, og rettleiinga inneheld ei forkortingsliste og ei ordliste over faguttrykk. Hovuddelen (s. 63–506) er ein alfabetisk gjennomgang av norske banneord frå a til å. Boka blir avslutta med ei litteraturliste og forslag til vidare lesing.

Her tillet eg meg å skyte inn ein liten kommentar til omtalen av Arthur Arntzens klassiske døme på gledesbanning (s. 27), der det heiter at guten som etter mange års fråvær kom heim til si kjære mor, i pur glede ikkje fann andre ord enn «du mor, du mor, du helvetes mor». Det er litt unøyaktig attgitt. Slik Arntzen (1986:31) sjølv fortel det, hadde guten vore til sjøs og var heimkommen til jul og skulle overrekke mora eit silkesjal i julegåve. I det høvet fann han ikkje andre ord enn dei nemnte.

4. Artikkelstruktur

Kvar artikkel har eit oppslag med feit skrift. Også fleirordsuttrykk er ført som oppslag, på alfabetisk plass, i glattalfabetisk makrostruktur, til dømes:

du milde
du milde Johannes
du skal faen ikke
du slette tid
du store alpakka

Oppslaga er altså konsekvent sortert etter første bokstav, og ikkje

etter det semantisk tyngste ordet. Eit uttrykk som *gi jamt faen* er såleis å finne under bokstaven g, og ikkje under f. Uttrykket er heller ikkje nemnt i oppslaget *faen*.

Rekkjefølgja i fleirordsuttrykk er elles styrt av det første ordet i uttrykket, slik at til dømes *fy tusan* kjem før både *fysj* og *fytte grisen*.

I oppslaget er det også påført grad av tabu, med frå eitt til fire lyn. Oppslagsord som redaktøren oppfattar som heilt avtabuiserte, får tabugrad 0, altså inga markering. Døme på oppslag med tabugrad 0 er *gid*, *gå fem på*, *gå i baret* og *hele sulamitten*. Tabugrad 1 har uttrykk som *fillern*, *gi dunderen* og *jøssenam*. Vurdert til tabugrad 2 er til dømes *bitterdød*, *bry seg sytten/søtten om* og *kjøss meg bak*. Hakket sterkare, med tabugrad 3, er *dæven dundre*, *for svarte*, *faen* og *satan*. Høgste tabugrad, 4, får dei grovaste kjønnsorda, som *fitte* med samansetningar og *kukksuger*. Tabugraderinga er kommentert nærare nedanfor (under punkt 11).

Sjølve artikkelen blir innleidd med ein definisjon med kommentarar, ofte med etymologiske opplysningar og med tilvisingar til andre oppslag. Deretter kjem ein bolk med døme og kjelder (korpus, ordbok, database, eigen intuisjon eller eiga røynsle). Nedst i artikkelen er det ei linje med opplysning om ordklasse, der «frase» og «interjeksjon» naturleg nok er dei mest frekvente, deretter «tema», med sortering etter emnefelt (som «djevel», «Gud», «sex», «sjukdom») og til slutt vurdering av bruksfrekvens, med kategoriar som «vanlig», «sjelden», «lite brukt» og «lokal bruk». Det siste skal fange opp dialektuttrykk.

Som døme på ein typisk artikkel kan vi ta for oss *gubbevaremegvel* (med tabugrad 1, markert som del av oppslaget). Her får vi opplysning om opphavet, som er ei leksikalisering av bøna eller ønsket *måtte Gud bevare meg vel!* Definisjonen seier at «ordet brukes som uttrykk for lett forferdelse eller sterk begeistring». Som variantform blir ført opp *bevare meg (vel)!* Deretter følgjer fire døme, i dette tilfellet henta frå korpus og bloggar. Til slutt er

det opplysning om ordklasse (her interjeksjon), tema («Gud») og bruksfrekvens («vanlig»).

I det følgjande vil kvar enkelt del av artiklane i ordboka bli kommentert noko nærare.

5. Utvalet av oppslagsord

Eit gjennomgåande problem med boka er at banning aldri blir heilt klart avgrensa og definert. Det blir uansett operert med eit temmeleg vidt banneomgrep. Fjeld er klar over det og grunngir utvalet blant anna etymologisk og funksjonelt (til dømes s. 51). Ho seier også at nokre vil meine at ein del av oppslagsorda ikkje er banning i det heile, men held fram at det er relativt kva den enkelte oppfattar som banning, og at det kan vera uttrykk som historisk sett er banning, men som ikkje lenger er tabu. Det er nok likevel mange som vil undrast på at interjeksjonar som *hei*, *heisann*, *au*, *fysj* og *isjameg* har fått plass, men ikkje verken *huff* (berre i sambandet *huff a meg* / *huffameg*) eller *æsj*. Også ord som *innmari* og *svinaktig* blir nok sjeldan oppfatta som banning. Kanskje dette kunne vore kommentert og klårgjort noko meir inngåande i innføringsdelen. Eitt kriterium Fjeld nemner for å ta med reine forsterkarar som ikkje i seg sjølve er tabulada, er at dei blir brukt nesten berre i uttrykk der ein bannar. Men i fleire artiklar med forsterkande uttrykk er det døme som på ingen måte viser banning, til dømes *hu hei*.

Reine skjellsord utan vesentleg tabulading, som *burugle*, er det også diskutabelt å ta med i ei banneordbok etter mitt syn. Slike ord kan for all del vera stygge og vondsinna, men det er ikkje eit tabuelement i dei. Skulle ein ta med alt slikt, ville boka svulme urimeleg opp.

Sjølv om tabuladinga varierer både i tid og mellom folk, trur eg det hadde vore ein fordel å knyte utvalet meir tydeleg nettopp til

tabu. Tabu er ein fundamental komponent i banning – utan tabu, inga banning.

Eg vil knyte nokre kommentarar og spørsmål til enkelte oppslag:

Fittihælvete står som oppslag med rett etymologi, men det burde nok også vore ein artikkel på det vel så vanlege *fyttihælvete*.

Kvifor (i huleste!) er det oppslag både på *i svarteste* og *svarteste*, men berre på *huleste* og ikkje *i huleste*? Alle døma i artikkelen *huleste* viser sambandet *i huleste*.

Skal først *trykkfeilsdjevelen* vera med, burde det også vore rom for det synonyme *prentesvarten*, som har artikkel både i NO og i NAOB. Det usamansette *svarten* har fått oppslag, med tabugrad 3.

Eitt og anna banneord kan ein sakne, slik som *hålede/håladø*. Det trønderske utbrotet *dærsomitt* manglar (sjå NO 1:380, artikkel *dersom* tyding 2 b). Likeeins saknar eg den jærske eufemismen *fadosten*.

NO er brukt, både den trykte boka og setelarkivet, men her hadde det nok vore meir å hente. Her er eit einskilt døme: *fyt(t)ehuren* (NO 3:1167) frå Brunlanes, rett ovanfor *fytterakkaren*, som eg går ut frå har vore konsultert i NO. Ein systematisk gjennomgang kunne ha gitt meir dialektmateriale og betre fordeling utover landet. Uansett er det ikkje råd å få fullstendig dekning; det ville ikkje ein heil redaksjon klart heller.

6. Oppslagsform

Det er valt former som ligg nær opp til kjeldene, og utan streng normering. Prinsippet er nok sunt i seg sjølv, men det fører samtidig til at ein misser samanheng mellom former som eigentleg er same ordet, som *djevlig*, *jevlig* og *jævlig*. Alle har fått kvart sitt oppslag. Samanhengen kan likevel i stor grad bergast gjennom tilvisingar (sjå nedanfor).

Når oppslagsformene er så strengt bundne av førelegga som

dei i mange tilfelle er, kan det få ein del ikkje heilt vellykka utslag. For eksempel blir *jøjje* ført som eige oppslag (variantform av *jøye*) på bakgrunn av nokre (tilfeldige?) skriftlege belegg. *Jeija* har også fått ei litt uheldig oppslagsform. Det naturlege hadde vore *jeia*, og dømet med «fy jegjei» høyrer meir naturleg inn under eit oppslag *jeie*. NO (5:1466) har fyldig omtale av desse formene i artiklane *II jei*, *II jeia* og *jeie*.

Stundom stussar ein på oppslagsformene. Kvifor er *din djevel* valt som oppslagsform, når seiemåten også finst som døme i artikkelen *djevel*? Og kvifor er det apostrof i *filler'n meg* (s. 190), når ingen av døma har apostrof og det heller ikkje er apostrof i artikkelen *fillern* (s. 192). Etter mi meining burde også *filler'n meg* vore flytta og sett inn mellom *fillern* og *fillern også*.

7. Tilvisingar

Boka inneheld mange tilvisingar, noko som er til stor hjelp for å orientere seg i materialet, men det kunne godt vore fleire. Eg skal nemne nokre døme:

Det hadde vore naturleg med tilvisingar mellom oppslaga *djevelen hakke meg* og *faen hakke meg*, og likeeins mellom *gi beng* og *beng*. På same måte kunne og burde *blanke faen* og *gi blanke faen* vore samordna; alle døma i det første oppslaget inneheld verbet *gi*. Ein burde også hatt ei kopling mellom *fytterakkernsinn* og *fyttirakkern skinn*.

Tomme eller villeiande tilvisingar er det dessverre ein del døme på. I artikkelen *fy tusan* er det vist til *tusan*, men det er ingen artikkel på det. Oppslaget *bætterdø* har tilvising til *betterdø*, som ikkje finst som oppslag, og til *bitterdø* (oppslagsforma her er *bitterdød*). Den ganske vanlege forma *bætterde* er ikkje nemnt. Her kan ein nok også ha innvendingar mot frekvenskarakteristikken «sjelden» og kommentaren «lite brukt i moderne norsk» (s. 95).

8. Etymologiske merknader

Etymologiane er jamt over utførlege og solide, men somme av dei er kanskje i dristigaste laget. Fjeld har hatt som generell rettesnor å vera varsam, men som ho sjølv seier: «Det er [...] ikke til å unngå at ordboka inneholder noen private og udokumenterte påstander, til tross for bevisste forsøk på å utvise faglig forsiktighet» (s. 58). Det er også heilt korrekt, slik ho seier på same stad, at det kan vera vanskeleg å bruke tradisjonelle etymologiske metodar på eit materiale som dette.

Litt uklart er det som står om *fanden*, der teorien om samband med gammalfrisisk presens partisipp *fandian* 'freistande' kanskje kan trekkast i tvil (men jf. likevel etymologien i NAOB, artikkel *fanden*). Norrønt *fjándi*, bunden form *fjándinn* 'fiende', er det høgst sannsynlege opphavet, og det er ikkje sikkert dette skal knytast til frisisk.

Under oppslaget *fy vøle* er det vist til norrønt *vǫrðr* 'vaktar, vaktmann'. Dette er vanskeleg å forlike med vanleg lydutvikling. Rettare er det gjort i artikkelen *tvi vøre deg!* (kvifor med «deg», sidan døma syner berre «tvi vøre»?), der det er kopla til norrønt *verða*. Sjå NO 12:974, underoppslaget *tvi vorte* i artikkelen *verta*, tyding C 3. *Vøle* og *vøre* er her variantar av same ordform, den første i målføre med overgang frå *rð* til retrofleks *l*. I somme tilfelle kan nok *vøre* også vera konjunktivform av *vera*, jf. NO 12:899, artikkel *I vera*, tyding C 5.

Fillan (frå Hallingdal) er nok inga omforming av stadnamnet Fillan på Hitra. Det forslaget burde vore stroke, slik at «dialektform av *fillern*» hadde fått stå aleine som etymologi her.

Eg har lita tru på at *bella-don* skal vise til planten belladonna, og at *belle* skal vera ei kortform av dette (s. 81). Det siste kjem visst berre føre i sambandet *belle de* og *belle don* (NO 1:522, artikkel *III belle*). Elementet *don* finst også i andre kraftuttrykk, som *dyre don*. Kanskje går *belle don* og *belle de* i siste instans tilbake på *bitre død*.

Andskota burde vel hatt ei tilvising til *helskota* (det er tilvising andre vegen). Her kunne ein også supplert med målføreopplysning om *andskote* frå Leka i Namdalen hos Hans Ross (1895:6).

Når det gjeld *mutt putt* i *mutt putt aleine*, er eg redd det blir ei overtolking å sette det til *putt* 'pytt' og kople det til *helvetes pøl*. Etter mi vurdering kan det like gjerne vera ein variant av *mutters aleine*, med innlagt rim. Dermed blir jo også banneelementet høgst diskutabelt, men så har det riktig nok fått tabugrad 0.

Om interjeksjonen *pytt* blir det slått fast som sikkert at han går tilbake på ei eldre tyding 'svovelpøl'. NAOB seier opphavet er uvisst og nemner også som ein mogleg teori at det rett og slett er eit lydord som skal gi att spytting.

Oppslaget *faren min er glassmester* er sikkert rett oppfatta som «fordekt banning» med omforming av *faen*. Men her kunne det kanskje også vore knytt til ein kommentar om uttrykket *far din var ikkje glasmeister/glasmakar* – sagt til nokon som skyggjer for ein (sjå NO 4:340).

9. Døme

Døma er autentiske, henta frå korpus eller ymse nettkjelder, som bloggar og nettaviser, stundom også notert ned direkte frå «folkets munn». Dette er ein særs verdifull del av det framlagte materialet. I somme høve kan nok døma vera feilplassert. Såleis er eit døme med *bort i natten* ført under oppslaget *bort i alle staur og vegger*. I det aller siste oppslaget, årre steikje, viser to av døma *årre heite*. Dette burde vel hatt eige oppslag.

10. Bruksfrekvens

Kategoriane er påført med støtte i korpusmateriale, men det er sjølvsagt eit vanskeleg arbeid å vurdere kor frekvente dei ulike ut-

trykka er i faktisk bruk. Dette kjem jo også mykje an på ståstad og bakgrunn, ikkje minst geografisk. I innleiinga (s. 57) blir det hevda at *farken* blir brukt nærmast heilt marginalt i dag. Dette vil nok ein trønder protestere på, og det går da heller ikkje i hop med opplysninga «ofte brukt» i artikkelen *farken*. Ein gjennomsnitts trønder vil nok heller ikkje vera samd i at *bætterdø* blir brukt «sjelden».

11. Litt meir om tabugradering

Tabugraderinga er ei nyskaping i norsk leksikografi og representerer ei anna tilnærming enn dei litt tilfeldige bruksmarkeringane «vulg», «nedsetj» o.l., som har vore nytta hittil. Ei fin drøfting av dette står på s. 54–55, der det mellom anna er vist til granskingsforsøk der ein bed informantar vurdere «styrken» av ulike kraftuttrykk. Sjølvsagt vil dette vera langt på veg subjektivt og individuelt, noko Fjeld også er fullt klar over, og ho nemner sjølv problematiske sider ved slike undersøkingar. Graderinga i boka er vel da i stor monn gjort ut frå eit subjektivt inntrykk. Likevel kan slike opplysningar tene til råd og rettleiing om i kor stor monn eit uttrykk er passande eller ikkje. Eg noterer meg at *fitte* med grad 4 er rekna som kraftigare enn *faen* og *helvete*, som begge må nøye seg med grad 3. Kjønnsrelatert banning blir altså vurdert som sterkare enn banning på religiøst grunnlag, noko eg slett ikkje er sikker på vil gjelde allment i norsk. *Jævlig* er nede på grad 2, medan *fy fabian* så vidt kjem inn på skalaen med grad 1. Og mange oppslag får som nemnt ikkje gradering i det heile, som *gå fem på*.

12. Vurdering

Trass i dei kritiske merknadene ovanfor er *Norsk banneordbok* eit stort og imponerande arbeid der eit svært omfattande materia-

le blir lagt fram på systematisk vis. Fjelds leksikografiske klo avspeglar seg i det grundige arbeidet og det systematiske oppsettet. Banning er eit mytebelagt område der det er stor trong for etterretteleg informasjon. Boka fyller eit tomrom i norsk leksikografi og fell naturleg inn mellom slangordbøker og skjellsordbøker (som Jenstad 1992 og 1999). Den utførlege og instruktive innleiinga, med grei opprydding i mange omgrep, gjer ho også til mykje meir enn ei rein ordbok. Dei inkonsistensane som finst, er kanskje langt på veg eit resultat av at arbeidet med boka har gått føre seg over lang tid. Boka bør absolutt også komma i ein digital versjon.

Litteratur

Ordbøker

Dagrin, Bengt G. (2013): *Stora fula ordboken. Försummade och förtalade ord i full frihet*. 5:e upplagan. Stockholm: Bengt G. Dagrin och Carlssons Bokförlag.

Jenstad, Tor Erik (1992): *Den store norske skjellsordboka*. Trondheim: Marcus forlag.

Jenstad, Tor Erik (1999): *Nye skjellsordboka*. Trondheim: Tapir / Norske Gram Publishing.

NAOB (2017) = *Det Norske Akademis ordbok*. Det Norske Akademi for Språk og Litteratur / Kunnskapsforlaget. <https://www.naob.no/> (april 2019).

NO (1966–2016) = *Norsk Ordbok. Ordbok over det norske folkemålet og det nynorske skriftspråket.* Oslo: Det Norske Samlaget. <http://no2014.uib.no/> (april 2019).

NRO (1937–1995) = *Norsk Riksmålsordbok*. Oslo: Det Norske Akademi for Språk og Litteratur / Kunnskapsforlaget.

Ross, Hans (1895): *Norsk Ordbog*. Kristiania: Albert Cammermeyers Forlag.

Annan litteratur

Arntzen, Arthur (1986): *Æ lyg ikkje*. Oslo: Det Norske Samlaget.

Fjeld, Ruth Vatvedt (2001): Banning og andre kraftuttrykk. I: Dag Gundersen (red.): *Språkvett. Skriveregler, grammatikk og språklige råd fra a til å*. Oslo: Kunnskapsforlaget, 64–66.

Fjeld, Ruth Vatvedt (2002): Om banning og sverting. I: *Maal og minne* 2, 152–166.

Hansen, Hans Peter (1911): Folk bander. I: *Sprog og Kultur* 11, 1–32.

Hasund, Ingrid Kristine (2005): *Fy farao! Om nestenbanning og andre kraftuttrykk*. Oslo: Cappelen Damm.

Hestnes, Hæge (2007): Teaching Swearing. I: *Språkfag 1. Idéhefter for lærere*. Trondheim: Tapir akademisk forlag, 53–60.

ILNytt (2011). Internavis for Institutt for lingvistiske og nordiske studium, Universitetet i Oslo. Utgåve publisert 13.5.2011.

Jespersen. Otto (1911): Om banden og sværgen. Iagttagelser om edernes sproglære. I: *Festskrift til H.F. Feilberg fra nordiske sprog- og folkemindeforskere på 80 års dagen den 6. august 1911*. Kristiania: Bymaalslagets forlag, 33–40.

Ljung, Magnus (1984): *Om svordomar: i svenskan, engelskan og arton andra språk*. Stockholm: Förlaget Akademilitteratur.

Ljung, Magnus (1987): *Banning i norsk, svensk og 18 andre språk*. Oslo: Universitetsforlaget.

Uri, Helene, Jostein Sand Nilsen & Andreas Østby (2011): *Fakta faen*. Oslo: Cappelen Damm.

Tor Erik Jenstad
forskar, dr.art.
Institutt for språk og litteratur, NTNU
NO-7491 Trondheim
tor.jenstad@ntnu.no

En nyliberal ordlista

Lennart Larsson

Hanna Alvage, Elin Lilleman Eriksson & Saga Gärde (red.). *Nyliberal ordlista*. Bokförlaget Atlas 2018. 191 sidor. Pris från 170 SEK.

Att den kraftiga högersväng som präglat inte bara den ekonomiska politiken utan hela samhällsutvecklingen under de senaste 40 åren – det som ofta sammanfattas under benämningen *nyliberalism* – gjort avtryck i språket är inget att undra över. I spåren av avregleringarna, privatiseringarna och nedskärningarna i den offentliga sektorn har också språkbruket ändrats: nya ord och uttryck har kommit i svang, andra har hamnat i skymundan, medan ytterligare andra har fått nya användningar eller konnotationer. Det här finns det förstås all anledning att uppmärksamma, så initiativet att – som det är formulerat i baksidestexten till *Nyliberal ordlista* (fortsättningsvis NO) – försöka ”synliggöra hur den nyliberala ideologin förändrar språket och styr våra ordval” är utan tvekan lovvärt.

Men vad är NO för slags bok egentligen? Ska man gå efter titeln så är det en ordlista, och den innehåller också mycket som kännetecknar en sådan: här finns ett åttiotal alfabetiskt ordnade artiklar med fetstilta lemman följda av ordklass och böjningsangivelser, definitioner (ibland med flera numrerade betydelsemoment) samt autentiska språkprov. Å andra sidan kan man på pärmens baksida läsa att det ”är en politisk och poetisk stridsskrift”, och huvuddelen av bokens textmassa går också vida utöver de gränser som brukar kännetecknа lexikografiska produkter. Tyngdpunkten ligger på de vidhängande kommentarerna till artiklarnas lemman – texter av varierande omfång som till allra största delen härrör från redaktörerna men i sammanlagt nio fall är skrivna av andra författ-

tare (till de mera namnkunniga hör poeterna Athena Farrokhzad och Jila Mossaed liksom dramatikern Suzanne Osten).

Att det i första hand handlar om en stridsskrift, ett politiskt debattinlägg med vänsterförtecken, får sin bekräftelse när man läser det knappa tiotalet recensioner av NO från våren 2018 som finns att läsa i databasen Mediearkivet (innehållande dagstidningstext från ca 1990 och framåt). Även om omdömena om boken går starkt isär så har recensenterna en sak gemensamt: bortsett från lemmaurvalet (se vidare nedan) är de lexikografiska aspekterna av boken helt frånvarande i deras texter.

NO kan alltså beskrivas som en politisk stridsskrift som har formen av en ordlista. Och det är framför allt i den senare egenskapen, som lexikografisk produkt, som den kommer att avhandlas här.

Att det lexikografiska innehållet i NO är lånat från annat håll är ingenting som redaktörerna (verksamma som journalister och kulturarbetare) sticker under stol med. Men när de i eftertexten skriver att ”[d]efinitionerna av orden är hämtade ur *Svenska Akademiens ordlista*” (s. 191) så är det en sanning med avsevärd modifikation. I själva verket handlar det om ett hopplock från webbplatsen svenska.se där Akademiens samtliga tre ordböcker – förutom *Svenska Akademiens ordlista* (SAOL) även *Svensk ordbok utgiven av Svenska Akademien* (SO) och *Svenska Akademiens ordbok* (SAOB) – i varierande utsträckning har fått bidra. En illustrativ inblick i hur redaktörerna gått till väga ger redan den inledande artikeln *andrum* (bild 1; artikeln är här återgiven i sin helhet för att ge en uppfattning om ordlistans uppläggning).

Bortsett från lemmats inledande versal är den första raden helt och hållet överensstämmande med SAOL, och detsamma gäller det första betydelsemomentets synonymdefinition ”frist”. Det andra momentet – inlett av det malplacerade ”tillgång till luft bereds” som ju beskriver en verbhandling – återfinns däremot inte i SAOL. Här är källan i stället SAOB, vars första moment (med en

And|rum substantiv ~*met*
1 frist 2 tillgång till luft bereds; lufttillgång;
gm fritt utrymme berett tillfälle att andas
I anv.: "Sverige behöver ett andrum", Stefan Löfven, 24/11 2015.

Under en presskonferens sammankallad av Sveriges statsminister Stefan Löfven den 24 november 2015 motiverades beslutet att begränsa Sveriges asyllagstiftning till EU:s miniminivå, med att det skulle "skapa en bättre situation för de människor som redan är här" och ge Sverige ett *andrum*. Åtgärderna omfattade bland annat inskränkningar av anhöriginvandringen, vilket resulterade i att familjer splittrades samt att den enda lagliga och säkra vägen för personer som söker asyl togs bort. Den nya lagen som trädde i kraft den 20 juli 2016 syftar till att, såväl praktiskt som signalpolitiskt, kraftigt minska antalet asylsökande samt uppmana människor på flykt att söka asyl i andra länder.

Bild 1: Artikeln *andrum*.

rumslig betydelse som redan när artikeln trycktes 1898 betecknades som "numera mindre br[uklig]") lyder: "öppning l. fritt rum, hvarigenom tillfälle till andning l. tillgång till luft beredes; lufttillgång; gm fritt utrymme beredt tillfälle att andas."

Det är inte helt lätt att förstå hur det blivit så här. Har redaktörerna inte riktigt förstått ordboken (SAOB är ju långtifrån lättläst)? Eller har de bara inte gett sig tid att tänka över vad orden står för? Hur som helst blir fortsättningen genom alfabetet inte mycket bättre, utan titt som tätt möts man av bristande akribi, av förbiseenden och förvanskade citat. Ett ytterligare exempel är artikeln *flexibilitet* vars lexikografiska del återges i bild 2.

Flex·ibil·itet substantiv *~en*
• anpassningsbar, föränderlig; böjlig; som på ett sätt kan anpassas efter omständigheterna
I anv.: "Ge exempel på när du var flexibel", cv-tips från vuxenutbildningsföretaget Eductus 2018.

Bild 2: Artikeln *flexibilitet*.

Också här är det SAOL som fått släppa till definitionens första del med de tre synonymerna; helt i sin ordning – frånsett det faktum att det som definieras här inte är substantivet *flexibilitet* utan adjektivet *flexibel*. Den följande relativsatsen är däremot hämtad från SO, men återigen i en märkligt stympad form som gör texten mer eller mindre obegriplig; i originalet lyder satsen "som (på ett smidigt sätt) kan anpassas efter omständigheterna".

Ett tredje exempel på NO:s oövertänkta användning av svenska.se ger artikeln *hushållsnära tjänst* (bild 3).

I det här fallet följer definitionen av *tjänst* nästan till punkt och pricka det första betydelsemomentet i SO; de enda avvikelserna är dels att förlagan har "avsedd" och inte "avsedda" (tydligen har redaktörerna tolkat adjektivet som syftande på "arbetsuppgifter och kvalifikationskrav" och inte på "ställning"), dels att det som här står inom parentes i SO står i petitstil. Betydligt värre är att den betydelse av *tjänst* som föreligger i uttrycket *hushållsnära tjänst* ju är en annan, den som finns i SO:s andra betydelsemoment och förklaras som "utövande av arbete inom ramen för viss befattning".

Ett återkommande lexikografiskt missgrepp i NO är också att lemmat inte överensstämmer med det autentiska språkprovet: så har *bidragsnarkoman* ett språkprov som talar om "bidragsmissbrukare", och på samma sätt illustreras *kostnad* med "utryck-

Hus·hålls|nära tjänst oböjligt adjektiv, substantiv *~en ~er*
Av **hushåll** • husligt arbete som är en förutsättning för det normala livet i ett hem o. **nära** • kort avstånd o. **tjänst** ställning i arbetslivet med väldefinierade arbetsuppgifter och kvalifikationskrav, avsedda att innehas av en person (särsk. om statliga och kommunala arbeten)
I anv.: "Ge bort städning och andra hushållsnära tjänster i julklapp för halva priset via HemQ.se. Även jultomten är avdragsgill, men bara om han också städar, passar barnen eller lagar maten", pressmeddelande företaget HemQ 30/11 2012. Se *RUT.*

Bild 3: Artikeln *hushållsnära tjänst*.

ningskostnaden", *produktivitet* med "produktiva" och *välfärdskonsument* med "välfärdskunderna".

Hur ska man förklara denna närmast nonchalanta lättvindighet som redaktörerna uppvisar i förhållande till NO:s lexikografiska sida? Antagligen är det av samma orsak som att ingen av recensenterna ägnat den någon uppmärksamhet: ordlistekonceptet har setts som ett strukturerande ramverk utan egentlig betydelse, en ingång till något annat och viktigare, som lite av en dekorativ utsmyckning. Men för att fylla sin funktion bör också en ingång eller en utsmyckning utformas med omsorg; om inte väcker det bara tvivel och misstro hos den uppmärksamma läsaren.

Man blir misstänksam. Och tyvärr tvingas man konstatera att också de uppföljande kommenterande texterna i många styck-

en lider av brist på eftertanke, av oklara formuleringar och rena faktafel. Även här kan *flexibilitet* tjäna som exempel. I kommentarens inledning påstås att ”[o]rdet förekom i etablerade medier första gången 2006”. Oavsett om det gäller substantivet eller adjektivet så är påståendet falskt. I Kungliga bibliotekets databas Svenska dagstidningar härstammar det äldsta belägget på *flexibel* från *Dagligt allehanda* 1780, och även om betydelsen i de tidigare beläggen handlar om böjlighet i fysisk bemärkelse (i språkprovet 1780 saluförandet av ”en flexible Silfwer Mans Catheeter”) så dyker de första exemplen på betydelsen ’anpassningsbarhet’ upp långt tidigare än 2006. Exempelvis kan man bland de 256 träffarna på ”flexib*” i Svenska dagstidningar 1963 hitta en anställningsannons i *Dagens Nyheter* (ett etablerat medium om något!) där det bland de önskvärda kvalifikationerna bl.a. nämns ”Samarbetsförmåga (det skattar vi högt). Flexibilitet och så förstås noggrannhet.” (Dagens Nyheter 7/10 1963, s. 32; citerat från DN.arkivet). Så varför just 2006? Bara för att det var det år alliansregeringen kom till makten? Ett indicium på att så är fallet är att också artikeln *energitjuv* (med den utvidgade betydelsen av någon som lägger beslag på ens psykiska energi) har den nästan identiska formuleringen ”[f]örekommer i etablerade medier för första gången år 2006”. Ändå kan man t.ex. i *Aftonbladet* 1999 läsa: ”Så står du emot energitjuvarna: Tänk positivt! Skriv ner ett positivt budskap till dig själv.” (Aftonbladet 6/5 1999, s. 13; citerat från Mediearkivet.)

Den enda lexikografiska aspekt som tas upp av NO:s recensenter gäller urvalet av lemman. Och de tål förvisso att undra över. Som flera av recensenterna påpekar – även sådana som i övrigt uttrycker sig uppskattande om boken – gör redaktörerna en mycket vid tolkning av begreppet *nyliberalism*. Vad har t.ex. de redan nämnda *andrum* och *energitjuv* med nyliberalismen att göra? Eller för den delen *svenska värderingar*, *tandröntgen* och *åldersbedömning*? Här finns på sin höjd ett indirekt samband, genom att de här uttryckssätten skriver in sig i eller bara råkar

anknyta till den högernationalistiska diskurs som ju i mångt och mycket kan ses som en reaktion på nyliberalismens verkningar.

Ett diskutabelt fall av annat slag är *arbetsgivare* vars (inom parentes sagt något luddigt formulerade) kommentartext ser ut så här:

> Kring användningen av ordet arbetsgivare har det skett en förskjutning inom styrkeförhållandet mellan den som köper respektive säljer arbete. Begreppet arbetsköpare implicerade en part som förväntas betala ett pris motsvarande arbetssäljarens prestation medan *arbetssäljaren* var den som sålde sin arbetskraft till arbetsköparen [...] Arbetet presenteras som en gåva. Mottagaren har i sin tur att vara tacksam och förhålla sig till givarens villkor. (S. 13 f.; originalets kursiv.)

Även om man ur ett vänsterperspektiv gärna håller med om att ordet *arbetsköpare* bättre än *arbetsgivare* fångar relationen mellan arbetsmarknadens två parter så innebär det inte att det senare ordet hör hemma i en *nyliberal* ordlista. Ända sedan de här orden etablerades i svenskan under senare delen av 1800-talet så har *arbetsgivare* (liksom *arbetstagare*) hela tiden varit den ojämförligt mest frekventa beteckningen, medan *arbetsköpare* (och i än högre grad *arbetssäljare*) fört en relativt undanskymd tillvaro (något som enkelt kan konstateras genom en sökning i Svenska dagstidningar). Så även om ordet *arbetsgivare* kan kritiseras för att fördunkla relationen mellan arbete och kapital är det långt ifrån något uttryck för en specifikt nyliberal tankemodell.

Till detta kan läggas att ett antal ord och uttryck som har en tydlig koppling till ekonomisk politik av nyliberalt snitt och därmed mycket väl hade platsat i en dylik ordlista inte har beretts utrymme i NO. Några exempel är *kund*, *affärsområde* och *leverera* använt inom områden som vård, omsorg och skola, *hyvling* och *precisionsbemanning* på arbetsplatserna, den utvidgade an-

vändningen av *varumärke* på inte bara individuell utan också t.ex. kommunal nivå, de sedan början av 1990-talet högfrekventa *sunda statsfinanser* och *stålbad*, liksom ordet *kommersialisering* använt med positiva konnotationer.

Som sagt: en ordlista som förtecknar de språkförändringar som följt i spåren av den nyliberala strömkantringen är utan tvekan ett lovvärt initiativ, och det är bara att beklaga att resultatet blivit som det blivit. För även om det finns inslag i NO som är klart läsvärda – så exempelvis Mara Lees analys av ordet *utanförskap* liksom Jila Mossaeds dikt Flyktingström – kan man inte komma ifrån att boken som helhet framstår som ett fuskverk, inte bara som lexikografisk produkt utan också som stridsskrift.

Litteratur

DN.arkivet. <https://arkivet.dn.se/> (mars 2019).

Mediearkivet (Retriever Research). <http://web.retriever-info.com/services/archive?> (mars 2019).

SAOB (1898–) = *Svenska Akademiens ordbok*. I: svenska.se.

SAOL (2015) = *Svenska Akademiens ordlista*. 14:e upplagan. I: svenska.se.

SO (2009) = *Svensk ordbok utgiven av Svenska Akademien*. I: svenska.se.

Svenska dagstidningar. Kungliga biblioteket. <https://tidningar.kb.se/> (mars 2019).

svenska.se. Svenska Akademiens ordböcker. <https://svenska.se/> (mars 2019).

Lennart Larsson
docent vid Institutionen för nordiska språk i Uppsala, redaktör på Svenska Akademiens ordbok
Institutionen för nordiska språk
Uppsala universitet
Box 527
SE-751 20 Uppsala
lennart.larsson@nordiska.uu.se

NE:s stora italienska ordbok – ny og utvidet utgave av den største ordboken mellom svensk og italiensk

Geir Lima

NE:s stora italienska ordbok. Italiensk-svensk/svensk-italiensk. Hovedredaktører: Riccardo Guglielmi og Verner Egerland. NE Nationalencyklopedin, Stockholm, 2017. 1365 sider (italiensk-svensk del: 717 sider; svensk-italiensk del: 607 sider; appendiks: 27 sider). Pris for papirutgaven: fra 468 SEK til 576 SEK. Elektronisk utgave på <www.ne.ord.se> (abonnementstjeneste som koster 29 SEK per måned).

1. Innledning

NE:s stora italienska ordbok er en videreutvikling av Norstedts forlags italienske ordbok, som kom ut første gang i 1994 i to bind (*Norstedts italiensk-svenska ordbok* og *Norstedts svensk-italienska ordbok*). De to bindene ble slått sammen til ett bind i 1998 da ordboken kom ut i nytt opplag med tittelen *Norstedts italienska ordbok*. I forordet informerer forlaget oss om at den nåværende utgaven er utvidet med nærmere 10 000 ord og fraser for å dekke de siste tiårenes teknologiske og samfunnsmessige utvikling. Ordboken skal dermed til sammen ha 121 000 ord og fraser og er den største ordboken til og fra italiensk på det svenske markedet. Arbeidet med oppdateringen av den nye utgaven har først og fremst blitt utført av fil.dr. Riccardo Guglielmi og professor Verner Egerland ved Lunds universitet.

Den nye utgaven av *NE:s stora italienska ordbok* er blant de største ordbøkene mellom italiensk og skandinaviske språk som er tilgjengelige på markedet, sammen med Gyldendals *Italiensk-Dansk*

Ordbog og *Dansk-Italiensk Ordbog* i Danmark og Kunnskapsforlagets *Italiensk blå ordbok* i Norge.

I Italia finnes det ingen forlag som har noen større ordbøker mellom italiensk og skandinaviske språk på markedet. Den siste store ordboken i denne kategorien var en italiensk utgave av *Norstedts italiensk-svenska ordbok* og *Norstedts svensk-italienska ordbok* fra 1994, som var utgitt av forlaget Zanichelli. Den var identisk med de svenske ordbøkene, men med italiensk omslag.

Når forlaget NE (Nationalencyklopedin) nå har kommet med en ny, utvidet og oppdatert utgave av sin største italienske ordbok, er det noe som mange vil kunne ha stor nytte av, ikke bare i Sverige, men også i andre nordiske land, og naturligvis også i Italia.

Denne anmeldelsen tar utgangspunkt i papirutgaven, men den elektroniske utgaven vil også bli nevnt, først og fremst der den skiller seg ut fra papirutgaven. Selv om innholdet i de to versjonene er det samme, er det store forskjeller mellom dem når det gjelder lesbarhet, og dette vil bli kommentert i anmeldelsen. Når ikke annet er presisert, er det papirutgaven som kommenteres.

2. Målgruppe

2017-utgaven har ingen informasjon om målgrupper eller ordutvalg, men i italiensk-svensk bind fra 1994 fantes det et forord med opplysninger om hvilke tanker som lå bak utvalget av oppslagsord:

> Ordboken tar i första hand upp det aktuella italienska allmänspråket, men den innehåller även en mängd vardagliga ord och uttryck, liksom ett brett urval termer från olika fackområden. Äldre språk har även beaktats i viss mån, med tanke på läsning av äldre texter. (*Norstedts italiensk-svenska ordbok* 1994:III)

I 2017-utgaven er det fremdeles det aktuelle allmennspråket som utgjør grunnstammen i ordboken, og selv om eldre språk ikke har noen dominerende plass, finnes det flere eksempler på slike ord og uttrykk. Studenter og andre som leser eldre italiensk litteratur på originalspråket, vil ha utbytte av ordboken, men de utgjør ikke noen hovedmålgruppe.

I hovedsak fremstår den nye utgaven av ordboken som en moderne, allmennspråklig ordbok som er tilpasset en bred brukergruppe. Det er naturlig å anta at forlaget har sett for seg at brukerne av ordboken vil være skoleelever, studenter, forskere, oversettere, turister eller andre som har noe med Italia å gjøre. Uansett vil det være snakk om en sammensatt gruppe der noen i utgangspunktet vil ha gode språkkunnskaper og lang erfaring med å bruke ordbøker, mens andre vil befinne seg på nybegynnernivå.

En målgruppe som vi ikke må glemme når vi skal vurdere behovet for en stor italiensk-svensk/svensk-italiensk ordbok, er den store gruppen av språkbrukere som har italiensk som morsmål, og som har behov for en ordbok som kan hjelpe dem til å finne seg til rette i det svenske språket og i det svenske samfunnet. Det finnes flere universiteter i Italia hvor det blir gitt undervisning i svensk språk og litteratur, og det finnes et ikke ubetydelig antall italienere som av en eller annen grunn har behov for å kunne kommunisere på svensk. Det finnes mange italienere som bor og arbeider i Sverige, og en ordbok mellom svensk og italiensk bør også ta hensyn til denne brukergruppen.

Slik ordboken fremstår, er den først og fremst beregnet på svenske brukere. Innledningen er bare på svensk, og alle forklaringene er tilpasset svenske brukere. Bakerst i boken finnes det informasjon om det italienske språket, en oversikt over de italienske regionene med kart over Italia og en oversikt over bøyning av italienske verb. Det finnes derimot ingen informasjon som er tilpasset italienske brukere om det svenske språket eller svensk geografi eller lignende.

I hovedsak kan vi med andre ord definere italiensk-svensk del som en resepsjonsordbok for svenske brukere, og på samme måte er svensk-italiensk del først og fremst en produksjonsordbok. Selv om ordboken ikke er symmetrisk, betyr det naturligvis ikke at brukere med italiensk som morsmål ikke vil kunne ha nytte av den, men med noen enkle grep (som for eksempel å oppgi kjønn for de svenske substantivene og å gi informasjon om bøyning av svenske verb) kunne nytten ha vært enda større.

At en ordbok av denne typen bare er tilpasset brukerne i det landet der ordboken kommer ut, er ikke uvanlig. Det er dette som har vært det vanligste i Skandinavia, og det er absolutt ikke noe som bare gjelder denne ene ordboken mellom svensk og italiensk. Ikke desto mindre kunne det være en idé å tenke mer på den store gruppen av potensielle brukere med italiensk som morsmål i fremtiden. Det bør også forlagene kunne ha økonomisk interesse av.

3. Struktur

3.1. Generelt

2017-utgaven viderefører strukturen i utgavene fra 1994 og 1998. I papirutgaven er teksten satt opp i to spalter, der oppslagsordene står i halvfete typer, mens opplysninger om ordklasse og kjønn står i kursiv. Homografer står oppført som separate lemmaer, og når de har forskjellig opphav, er de adskilt med arabiske tall:

1 pesca [-ɛ-] *f* persika [...]
2 pesca [-e-] *f* fiske [...]

Dersom det dreier seg om lemmaer som har samme opphav, men som tilhører forskjellige ordklasser, brukes romertall. På samme måte skilles det også mellom transitive og intransitive verb:

pesto [-e-] **I** *adj* sönderstött [...] **II** *m* [...] pesto [...]
saltare I *vb tr* **I** hoppa över [...] **II** *vb itr* **1** hoppa [...]

Eksempler og uttrykk er markert med halvfet skrift og kursiv.

I italiensk-svensk del er substantivene markert med *m* for hankjønn og *f* for hunkjønn, mens i svensk-italiensk del er substantivene markert med *s* for substantiv. Det hadde som nevnt vært til stor hjelp for italienske brukere om det hadde vært oppgitt kjønn for de svenske substantivene også.

Til stor hjelp for både svenske og italienske brukere er det utvilsomt at det også for de svenske verbene er oppgitt om de brukes transitivt eller intransitivt. Dette er med på å gjøre ordboken både oversiktlig og brukervennlig.

Polysemi markeres med arabiske tall på denne måten:

fraschetta 1 liten gren, kvist **2** lättsinnig kvinna

Noen steder finnes det markører som hjelper brukeren til å finne frem til den riktige ekvivalenten når det er snakk om polysemi, og dette gjelder særlig i svensk-italiensk del:

groda 1 djur ranocchio, rana **2** misstag granchio, cantonata

Markørene skiller seg ut, som vi ser, ved at skriftstørrelsen er mindre enn for den øvrige teksten.

Ved delvis ekvivalens mellom betydninger, som vi kan se et eksempel på i oppslagsordet *skog* i svensk-italiensk del, brukes det på samme måte markører med en annen skriftstørrelse for å forklare betydningsforskjellene. Her brukes det ikke tall for å skille betydningene fra hverandre, men semikolon:

skog foresta, selva; mindre bosco

3.2. Uttalemarkering

I ordbokens innledning finnes det oversiktlig informasjon om uttalereglene i italiensk. Det opplyses at det vanligste er at trykket ligger på nest siste stavelse, og at det ikke oppgis noen form for trykkmarkering i ordboken dersom et ord har trykk på denne stavelsen. I andre tilfeller blir den trykksterke vokalen markert med en prikk under bokstaven. I tilfeller der den trykksterke vokalen er en *e* eller en *o*, og brukeren må få vite om vokalen er åpen eller lukket, står den nødvendige informasjonen i hakeparentes bak oppslagsordet:

fuorigioco [-ɔ-]
fuorilegge [-e-]

Dette skal forstås slik at i det første tilfellet skal den trykksterke vokalen uttales som en åpen *o*, mens i det andre tilfellet skal den uttales som en lukket *e*. Det er mulig at den informasjonen brukeren her får, er tilstrekkelig for å forstå hvordan disse ordene skal uttales, men det kan også tenkes at noen brukere vil ha behov for mer fullstendig uttalemarkering. Man kunne kanskje ha oppgitt mer fullstendige uttaleopplysninger, for eksempel på denne måten:

fuorigioco /fworiˈdʒɔko/
fuorilegge /fworiˈleddʒe/

Nå må det umiddelbart legges til at den elektroniske utgaven på en utmerket måte gjør det mulig å høre uttalen av ordene, og at uttalemarkeringen på den måten ikke blir like viktig, mens det i papirutgaven naturligvis dreier seg om plassbesparing. Dersom alle oppslagsordene skulle ha fullstendige uttaleopplysninger, ville boken ha blitt mye større, så det er mulig å forstå forlagets prioritering her.

Ord som er fremmedord i italiensk, og ord med uregelmessig uttale, får fullstendige uttaleopplysninger:

équipe [eˈkip]
vedette [veˈdɛt]

Måten uttalen angis på i papirutgaven, er med andre ord tradisjonell. Den viderefører en konvensjon som vi kjenner fra ordbøker mellom italiensk og andre europeiske språk, og den viderefører nøyaktig samme system som fantes i 1994-utgaven. Det er altså ingen innovasjon på dette området i den nye papirutgaven, men konvensjonen fungerer tross alt ganske godt. Det er relativt sjelden at lemmaene i italienske ordbøker er forsynt med fonetisk skrift, av den enkle grunn at behovet for det ikke er like stort i italiensk som for eksempel i fransk og engelsk.

I papirutgaven står det ingenting om hvordan de svenske ordene skal uttales, men til gjengjeld får de italienske brukerne svært mye hjelp i den elektroniske utgaven når det gjelder dette.

3.3. Eksempler og uttrykk

Det skilles ikke mellom idiomatiske uttrykk og eksempelsetninger i det grafiske oppsettet. Mange av eksemplene er idiomatiske uttrykk eller setningskonstruksjoner som er oppført med verb i infinitiv. Vi finner eksempelsetninger med finitte verb og uttrykk med infinitte verb side om side med samme grafiske utforming, slik som for eksempel under oppslagsordet *deporre*:

la gallina depone le uova hönan lägger ägg
deporre qu. da un grado degradera ngn

Dette er ikke nødvendigvis noe problem i seg selv, og særlig ved oppslagsord hvor det bare er noen få eksempler, fungerer dette

fint. Litt verre kan det bli ved lemmaer med svært mange eksempler. I slike tilfeller kunne det kanskje ha vært en fordel med en form for systematisering av eksemplene og uttrykkene.

Det er relativt få eksempelsetninger med finitte verb i ordboken sammenlignet med uttrykk med infinitte verb. For eksempel har oppslagsordet *besök* flere uttrykk som viser konstruksjoner som er svært nyttige for brukeren (som *få besök av ngn*, *avlägga besök hos ngn*), men i tillegg til disse konstruksjonene kunne man kanskje ha ønsket seg eksempelsetninger slik at man hadde sett hvordan konstruksjonene faktisk brukes (av typen *mina föräldrar har varit på besök hos oss över helgen*). Flere slike eksempelsetninger ville ha vært til stor hjelp, særlig for nybegynnere.

Partikkelverbene i svensk-italiensk del står oppført for seg selv til slutt i artiklene, noe som bidrar til god orden og oversikt, særlig i den elektroniske utgaven. Partiklene er oppført i alfabetisk rekkefølge, slik at det blir lett å finne frem, også ved verb i svensk-italiensk del der artiklene er svært lange, som *gå*, *göra*, *stå*, *vara* osv. Denne systematiseringen er forbilledlig, og stort mer oversiktlig kan det ikke bli.

I italiensk-svensk del er det imidlertid ikke alltid like lett å finne frem blant eksemplene. Under oppslagsordet *essere* (verb) finnes det 34 eksempler og uttrykk, og det er vanskelig å se om det er noe system i dem. De er ikke oppført i alfabetisk rekkefølge, og det ser heller ikke ut som om det er noen form for tematisk inndeling. Til tross for at det grafiske oppsettet både i papirutgaven og i den elektroniske utgaven er ryddig og oversiktlig, blir det vanskelig å finne frem blant eksemplene når det er svært mange av dem.

Artikkelen *mer* i svensk-italiensk del har 15 eksempler og uttrykk, og som vi ser er det i utgangspunktet ikke umulig å orientere seg blant dem, selv om det tilsynelatende ikke er noe åpenbart system i organiseringen av dem:

mer *pron* o. *adv* più, ancora; bättre meglio; lite till un poco di più; ***en*** ~ uno in più, ancora uno; ***finns det ~?*** ce n'è ancora?; ~ ***eller mindre*** più o meno; ~ ***och*** ~ sempre più; ***gör inte så ~!*** non fare più così!; ***inte*** ~ längre non più; ***ingen*** ~ ***än han*** solo lui; ***ingenting*** ~ nient'altro; ***lite*** ~ ancora un poco; ***mycket*** ~ molto di più; ***så mycket*** ~ ***som*** tanto più che; ***ni vet*** ~ ***än jag*** sapete più di me; ***det är*** ~ ***än jag vet*** non ne ho idea; ***han är*** ~ ***än 20 år*** ha più di vent'anni; ~ ***än någon annan*** più di qualsiasi altro

Figur 1: Oppslagsordet *mer* i svensk-italiensk del av papirutgaven.

mer

pronomen, adverb

più, ancora; bättre meglio; lite till un poco di più

en mer uno in più, ancora uno

finns det mer? ce n'è ancora?

mer eller mindre più o meno

mer och mer sempre più

gör inte så mer! non fare più così!

inte mer längre non più

ingen mer än han solo lui

ingenting mer nient'altro

lite mer ancora un poco

mycket mer molto di più

så mycket mer som tanto più che

ni vet mer än jag sapete più di me

det är mer än jag vet non ne ho idea

han är mer än 20 år ha più di vent'anni

mer än någon annan più di qualsiasi altro

Figur 2: Oppslagsordet *mer* i svensk-italiensk del av den elektroniske utgaven.

Innholdet er, som man ser, nøyaktig det samme i de to versjonene, men lesbarheten er betydelig bedre i den elektroniske utgaven. Her er forkortelsene utvidet, og oppslagsordet, som er erstattet med tilde i papirutgaven, er tatt med i alle eksemplene i den elektroniske utgaven, hvor søkeordet også er understreket.

4. Ordforråd

4.1. Generelt

Som nevnt dekker ordboken i hovedsak allmennspråket, men det er også tatt med fagtermer fra ulike fagområder både i italiensk-svensk del og i svensk-italiensk del. Det finnes oppslagsord fra forskjellige stilnivåer og noen ord og uttrykk som er litterære, poetiske eller alderdommelige. Det brukes stilmarkører og fagmarkører for å markere stilnivå og fagspråk, for eksempel *med.* for medisinsk språk og *vulg.* for vulgært språk. De samme forkortelsene brukes både i papirutgaven og i den elektroniske utgaven.

Det er stort sett ikke problematisk å danne ekvivalensrelasjoner der lemmaene betegner etablerte begreper i begge språkene (jf. Hannesdóttir 2012). De største utfordringene når det gjelder å finne ekvivalenter, ligger gjerne i de begrepene som ikke nødvendigvis finnes i begge språksamfunnene. Det italienske substantivet *mammismo* 'overdreven morskjærlighet for voksne barn' og det svenske verbet *vabba* 'være hjemme fra jobben med sykt barn' er eksempler på lemmaer hvor det ikke finnes direkte ekvivalenter. Lemmaer av denne typen oversettes gjerne med hele setninger, noe som for eksempel gjelder det sistnevnte verbet, som oversettes med *rimanere a casa per accudire un bambino malato*.

Et annet eksempel er det svenske substantivet *smörgåsbord*, som i ordboken er oversatt med *buffet svedese semifreddo*. Dette er

kanskje ikke en helt presis oversettelse, og vil kunne bli misforstått av italienske brukere. *Semifreddo* brukes vanligvis om desserter og synes ikke å passe akkurat her. Dersom poenget skal være at et *smörgåsbord* består av både varme og kalde retter, hadde det kanskje vært bedre med *buffet con piatti caldi e freddi* eller noe lignende, eller bare *buffet tipico svedese*.

Substantivet *källsortering* er et eksempel på et lemma som omtaler et begrep som finnes og er utbredt både i Sverige og i Italia, og der det i utgangspunktet ikke burde være vanskelig å finne en ekvivalent. I ordboken oversettes det med *selezione dei rifiuti*, men her burde man også ha tatt med *raccolta differenziata*, som vel er den vanligste måten å omtale avfallssortering av denne typen på i Italia. Verbet *källsortera* oversettes med *fare la raccolta differenziata [dei rifiuti]* og *fare la differenziata*, og da hadde det vel vært logisk om substantivet *källsortering* hadde blitt oversatt med *raccolta differenziata*. En liten svakhet er det at dette uttrykket ikke står i italiensk-svensk del, verken under *raccolta* eller *differenziato*.

Det finnes flere slike eksempler på manglende samsvar mellom italiensk-svensk og svensk-italiensk del. Substantivet *seriekrock* 'kjedekollisjon' oversettes med *collisione a catena* i svensk-italiensk del, men her hadde *tamponamento a catena* vært en bedre oversettelse. I italiensk-svensk del står *tamponamento a catena* oppført med oversettelsen *seriekrock*, og det er vanskelig å forstå hvorfor dette ikke har kommet med i svensk-italiensk del.

4.2. Egennavn, stedsnavn og forkortelser

Ordboken inneholder mange navn på steder, både nasjoner, regioner og byer. Den inneholder også navn på politiske partier, som *Forza Italia* og *Lega Nord*, men ikke *Partito Democratico* eller *Movimento 5 Stelle* (Femstjernersbevegelsen), og heller ikke adjektivet *pentastellato* ('som gjelder femstjernersbevegelsen'), som brukes mye i de italienske mediene, er med. Forkortelser for fagforeninger

som *CGIL*, *CISL* og *UIL* står i ordboken, og her blir det forklart hva slags fagforeninger det er snakk om. Fra det svenske samfunnslivet er *Landsorganisationen*, *Sverigedemokraterna*, *Socialstyrelsen* og mange andre tatt med, og for disse oppslagsordene er det gitt forklaringer på italiensk som vil være til stor nytte for italienske brukere.

4.3. Neologismer

Ordboken har blitt beriket med mange neologismer, både svenske og italienske. Vi finner *nanoteknik*, *nätfiske*, *nätmobbning*, *rullgardinsmeny* og mange andre nyord i språket, men ikke for eksempel *fartkamera*, *hybridbil*, *transperson* eller *curlingförälder*, som alle står i SAOL14 (2015) og SO (2009). At vi heller ikke finner *wrap*, *tweet* eller *selfie*, er ikke nødvendigvis noe stort problem verken for svenske eller italienske brukere. Det kjønnsnøytrale pronomenet *hen*, som SAOL tok med i sin utgave fra 2015, er heller ikke med, og det er ikke utenkelig at noen italienske brukere av ordboken etter hvert vil trenge hjelp til å forstå dette pronomenet.

4.4. Dialektord og regionale ord

Det er også tatt med noen få dialektord og regionale ord i ordboken, som for eksempel adverbet *mo* 'nå', som markeres som *dialektalt*. Dette adverbet brukes over store deler av Italia, og *regionalt* hadde vært en bedre markør her. Substantivet *anguria* 'vannmelon' er derimot ikke markert verken som dialektalt eller regionalt, til tross for at dette er et ord som for det meste bare brukes i Nord-Italia. Når *mo*, som brukes i Sør-Italia, markeres som dialektalt, kan det virke som om ordboken definerer norditaliensk som standardspråk og søritaliensk som avvik fra standardspråket. Det beste hadde vært å markere både *mo* og *anguria* som regionale ord, noe de fleste italienske ordbøker gjør.

Ellers kan det nevnes at substantivet *terrone* oversettes med *skämtsamt öknamn på syditalienare*, en oversettelse som ikke er helt dekkende. Dette substantivet kan være både nedsettende og diskriminerende, og dette burde ha kommet tydeligere frem i oversettelsen.

5. Elektronisk versjon

Den elektroniske versjonen er i utgangspunktet en digitalisert papirordbok (jf. Svensén 2004:518 om typer av elektroniske ordbøker). Den har som nevnt det samme innholdet som papirversjonen, men med flere klare forbedringer med tanke på lesbarheten.

Når det gjelder forkortelser, er plassbesparing fremdeles noe som det blir tatt hensyn til i papirordbøker, mens det i elektroniske ordbøker kanskje ikke bør være nødvendig å ha forkortelser i det hele tatt.

Den elektroniske utgaven er i det store og hele ryddig og oversiktlig, og enkel å bruke. Den aller største nyvinningen ligger i den funksjonen som gjør det mulig å høre uttalen, ikke bare av enkeltord, men også av hele setninger. Denne funksjonen er forbilledlig, og vil også være til stor hjelp for italienske brukere som skal lære seg svensk.

Søkemulighetene dekker flere behov, og etter hvert som man blir kjent med søkefunksjonene, er ordboken lett å bruke.

Man kan ikke søke i både italiensk-svensk og svensk-italiensk del samtidig, og det hadde i noen tilfeller vært en fordel om det var mulig. Hvis man skriver inn et svensk ord i italiensk-svensk del, får man ingen treff, og på samme måte gir ikke søk på italienske ord treff i den svensk-italienske delen.

Et søk på en verbform som ikke selv er et lemma, gir treff på verbets infinitivform, og man får også opp alle eksemplene som inneholder dette verbet i en eller annen form. Hvis man for ek-

sempel søker etter *bevo* 'jeg drikker', kommer man direkte til verbet *bere* 'drikke', men en lenke til verbenes fullstendige bøyningsparadigmer finnes ikke.

Den elektroniske utgaven er altså mye mer enn bare en publisering av en papirordbok på nett, og den har en rekke nyttige funksjoner som gjør den til en moderne og brukervennlig elektronisk ordbok, samtidig som den naturlig nok, som alle andre elektroniske ordbøker, kan og bør få enda flere funksjoner som etter hvert vil gjøre den enda mer brukervennlig. I utgangspunktet handler det i stor grad om å løsrive seg fra det gamle papirformatet og å tenke nytt og annerledes når fremtidens elektroniske ordbøker skal utformes.

Disse problemstillingene har leksikografien vært opptatt av i mange år allerede (jf. Atkins 1996), og det finnes i det hele tatt svært mange muligheter for å forbedre brukervennligheten i de elektroniske ordbøkene. Man kan begynne med å gjøre så mange ord som mulig klikkbare. Ikke bare lemmaene, men også ekvivalentene kan inneholde lenker til andre artikler, og lemmaene kan ha bøyningsparadigmer som popup-vinduer. I tillegg kan de grammatiske forklaringene inneholde lenker. Det vil si at hvis man for eksempel klikker på *adverb*, kan man få opp en forklaring på hva et adverb er, og hvordan det brukes, med eksempler på begge språkene.

6. Konklusjon

NE:s stora italienska ordbok er en stor og omfattende ordbok som blir stående som en av de største og viktigste ordbøkene mellom et skandinavisk språk og italiensk.

Den siste utgaven av ordboken har blitt betydelig utvidet sammenlignet med de to foregående utgavene, og den fremstår i dag i sin helhet som en moderne og brukervennlig ordbok med en struktur som er ryddig og oversiktlig.

Innholdet i den elektroniske utgaven er i utgangspunktet identisk med innholdet i papirutgaven, men den elektroniske utgaven har i tillegg enkelte funksjoner som gjør at den har tatt et skritt bort fra sitt papiropphav. Dersom denne utviklingen fortsetter, vil den snart stå på egne ben som et selvstendig produkt.

Hvor lenge papirutgaven av denne og andre ordbøker vil fortsette å revideres, vet vi ikke. Det er ikke usannsynlig at de store papirordbøkenes tid vil være over om ikke altfor mange år, og noen revolusjonerende nyheter når det gjelder ordbøker i papirutgave, er det vel ingen som venter på. Det vi derimot forventer, er at forlagene vil arbeide mye med å forbedre og videreutvikle de digitale ordbøkene i tiden som kommer, noe som også er helt nødvendig for forlagenes egen del med tanke på konkurransen fra gratistjenestene som i dag finnes på nettet.

Litteratur

Ordbøker

Dansk-Italiensk Ordbog (2008). Carla Avanza Juul Madsen, Giovanni Mafera & Henning Juul Madsen. København: Gyldendal.

Italiensk blå ordbok (2010). Kolbjørn Blücher, Geir Lima, Diana Haakonsen & Magnus Ulleland. Oslo: Kunnskapsforlaget.

Italiensk-Dansk Ordbog (1999). Knud Andersen & Giovanni Mafera. København: Gyldendal.

Norstedts italienska ordbok (1998). Sören Johanson & Tomaso Tomba. Stockholm: Norstedts.

Norstedts italiensk-svenska ordbok (1994). Sören Johanson. Stockholm: Norstedts.

Norstedts svensk-italienska ordbok (1994). Tomaso Tomba. Stockholm: Norstedts.

SAOL14 = *Svenska Akademiens ordlista* 14:e upplagan. <www.svenska.se> (mai 2019).

SO = *Svensk ordbok utgiven av Svenska Akademien* (2009). <www.svenska.se> (mai 2019).

Annen litteratur

Atkins, B.T. Sue (1996): Bilingual dictionaries. Past, present and future. I: Martin Gellerstam, Jerker Järborg, Sven-Göran Malmgren, Kerstin Norén, Lena Rogström & Catarina Röjder Papmehl (red.): *Euralex '96 Proceedings*. Göteborg, 515–546.

Hannesdóttir, Anna Helga (2012): Ekvivalensrelationer i tvåspråkig lexikografi. I: *LexicoNordica* 19, 39–58.

Svensén, Bo (2004): *Handbok i lexikografi. Ordböcker och ordboksarbete i teori och praktik*. Stockholm: Norstedts.

Geir Lima
Sankt-Peterburgskij gosudarstvennyj universitet
Kafedra skandinavskoj i niderlandskoj filologii
Universitetskaja naberezjnaja 11, aud. 165
199034 Sankt-Peterburg
Russland
geir.lima@gmail.com

Kunnskapsforlagets *Stor norsk-polsk ordbok* – ett informationsspäckat monumentalverk

Morgan Nilsson

Ole Michael Selberg: *Stor norsk-polsk ordbok, Wielki słownik norwesko-polski*. Kunnskapsforlaget 2017. 1503 sidor. Pris c:a 498 NOK.

1. Bakgrund och karakteristik

Stor norsk-polsk ordbok (vidare SNPO) är ett imponerande verk som det enligt förordet tagit över 20 år att producera för den tidigare lektorn i polska vid universitetet i Oslo, Ole Michael Selberg. Idén väcktes hos författaren redan på 1980-talet i samband med granskningen av en mindre norsk-polsk invandrarordbok som dessvärre aldrig kom att publiceras. Större delen av SNPO är författad av huvudredaktören själv, men sedan 2008 har också Gunnar Arneson deltagit i arbetet med cirka en sjättedel av ordboken (bokstäverna M, O, V–Å). Hela verket har fortlöpande korrekturlästs av Adela Skrentni-Olsen. Författaren har också på egen hand svarat för layouten genom att sätta hela verket i TeX med hjälp av egenutvecklade makron.

Ordboken är mycket ambitiöst utformad. Intentionen är enligt författaren, som säger sig inspirerad av Berkovs (1987, 2003) idéer, att den ska vara bidirektionell, alltså i lika hög grad avsedd för polska som för norska användare. Texten har en hög densitet. Mycket information ges med subtila medel, bl.a. genom användning av över 50 specialtecken, 70 grammatiska förkortningar och 160 andra förkortningar, varav många av de senare avser ämnesområden och stilnivåer. En svårighet är dock att samtliga grafiska tecken, till

skillnad från förkortningarna, inte finns samlade i en gemensam lista, utan vissa tecken presenteras bara i den löpande texten i den inledande användarhandledningen, och de kan därför vara lite svåra att hitta när man stöter på dem i ordboken.

I den följande recensionen kommer fokus att ligga på ordbokens omfång i avsnitt 2, artiklarnas utformning i avsnitt 3, anvisningar om uttal och grammatik i avsnitt 4 och det ordförråd som återfinns i ordboken i avsnitt 5. I avsnitt 6 ges en avslutande sammanfattning.

2. Omfång

SNPO är med sina 1503 sidor och över 72 000 uppslagsord den utan tvekan största polska ordboken till eller från ett skandinaviskt språk. Kubitskys (1995) *Polsk-svensk ordbok* är med sina 737 sidor och 37 000 lemman endast hälften så stor och inte på långt när så rik på detaljer eller författad med lika stor akribi (jfr Dura, Malmgren & Toporowska Gronostaj 1996). De båda existerande polsk-danska ordböckerna (Sørensen & Widding 1989, Bostrup, Jarosz & Stręk 2014) är ännu mindre, 350 resp. 515 sidor. Ett nordiskt verk som ligger SNPO:s omfång nära är däremot Kudzinowskis (1988) finsk-polska ordbok på 1316 sidor med 90 000 uppslagsord, men inte heller detta verk har samma detaljrikedom och bidirektionella ambition som SNPO.

Enligt användarhandledningen omfattar SNPO moderat bokmål och riksmål. Riksmålsformer som faller utanför den officiella rättskrivningen markeras med en upphöjd trekant, t.ex. **Hvitehavet**△ → **Kvitsjøen**. Vid vissa enkla ord anges alternativa former, t.ex. ”**sju** a. **syv**” (a. = polska *albo* ’eller’) liksom ”**syv** a. **sju**” följda av två i princip identiska artiklar, men den polska användaren får ingen information om skillnaden mellan formerna. Båda dessa artiklar följs i sin tur av över 40 sammansättningar där talet utgör förled.

Här ges dock bara den ena formen av förledet, men all information ges alltså på två olika ställen i ordboken, vilket kan tyckas gå stick emot den annars dominerande sparsamheten med ordboksutrymmet.

3. Artiklarnas form

Artiklarna i SNPO är mycket informationstäta. De är alfabetiskt ordnade tecken för tecken, dvs. utan hänsyn till blanksteg i flerordsenheter. Layouten är i huvudsak traditionell, enkel och tydlig, men vissa innovativa tillämpningar av typografin är subtila, bl.a. att kursivera en bokstav i ett ord som en fonetisk eller fonologisk markering (se avsnitt 4.1 och 4.2). Den viktigaste innovationen är dock de relativt många grafiska tecken som används i olika funktioner. En del av dessa olika tecken kommer att diskuteras i de avsnitt nedan där de aktualiseras.

Artiklarnas generella struktur är att det fetstilta uppslagsordet följs av numrering av homografer (med romerska siffror), information om uttal (i hakparentes), ordklassangivelse (utom vid substantiv), böjningsändelser, numrering av delbetydelser (med arabiska siffror), markör, polska ekvivalenter (med markör och morfologisk information) och exempelfraser. De flesta av dessa komponenter exemplifieras av följande två artiklar:

fronton [frånŋtåŋ´] *-en, -er* arkit. fronton ⟨-u⟩

frọs`sen adj *-t, frosne* **1** (stiv av frost) zmarznięty [r-z]; (dypfryst) mrożony, zamrożony; (frostskadet) przemarznięty [r-z]; *frossen jord* zmarzniętna gleba; *frosne grønnsaker* mrożone warzywa ≀ mrożonki warzywne **2** (plaget av kulde) *jeg er frossen på føttene* nogi mi marzną ≀ mam zmarznięte nogi; *være frossen av seg* być wrażliwym na zimno ≀ być zmarzlakiem [r-z] F

Med [r-z] anges att *rz* undantagsmässigt inte är en digraf, utan uttalas som två ljud. Tecknet ≀ står mellan alternativa översättningar av exempelfraser och F står för 'familjärt'.

I slutet av artiklarna anges efter symbolen ◊ eventuella idiom, t.ex. **holde** ◊ ***holde seg for god til*** *noe*. Sedan ges efter ▷ hänvisning till andra artiklar där uppslagsordet förekommer, t.ex. **holde** ▷ AVSTAND, KJEFT, ØYE samt 30 andra ord. Efter ↔ ges ibland en hänvisning till ett ord där ytterligare information kan hittas om olika analoga konstruktioner, t.ex. **fjerde** ↔ **femte**. För verb ges efter □ fasta fraser med partiklar, t.ex. **holde** □ **~ av, ~ frem, ~ igjen, ~ på, ~ unna** samt 20 andra förbindelser. Avslutningsvis förekommer ibland efter tecknet ≈ en varning för falska vänner, t.ex. "**diva´n** *-en -er* kanapa [...] (≈ po. dywan: 'teppe')".

Vissa artiklar omfattar ord med mer än en ordklasstillhörighet. Dessa åtskiljs då med bokstäver i en fyrkant, vilket bl.a. förekommer vid namn på språk som förs till samma artikel som motsvarande adjektiv, t.ex. "**norsk** A et *-en* [...] B adj"[1]. Detsamma gäller ord som kan vara både adjektiv och adverb, t.ex. *mulig*, liksom sådana som kan vara både preposition och adverb eller konjunktion, t.ex. *før*, men också verb som används både transitivt och intransitivt eller reflexivt, t.ex. **kas`te** [...] A intr [...] B tr [...] C refl".

4. Uttal och grammatik

Som bidirektionell ordbok ger SNPO omfattande information om uttal och grammatik för båda de ingående språken. Framför allt ger den långt mer information om uttal, böjning och konstruktioner för målspråksekvivalenterna än vad som är vanligt förekommande i tvåspråkiga ordböcker, men även information om de norska uppslagsorden är mycket utförlig. Den notation som används kan delvis kräva att användaren upparbetar en viss vana

1 "et *-en*" ska rimligen tolkas *et norsk, norsken*. Jfr även avsnitt 6.

genom att inledningsvis konsultera användarhandledningen regelbundet. Detta lönar sig dock rikligen på längre sikt, då alla de detaljer som anges utgör en riktig guldgruva, men sättet detaljerna anges på är inte alltid sådant att det tillåter en intuitiv tolkning. Ambitionen har enligt författaren varit att informationen om uttal och grammatik ska vara symmetrisk, dvs. oförutsägbarheter ska redovisas i lika hög grad för båda språken.

Eftersom olika företeelser inom uttal och grammatik är förutsägbara i olika hög grad i de två språken blir inte mängden information om uttal och böjning symmetrisk. Norskans uttal är mindre förutsägbart än polskans, medan norskans morfologi är mera förutsägbar än polskans. Därför finns det en stor mängd information om uttalet av de norska uppslagsorden, medan det finns ytterst lite information om polskt uttal. Å andra sidan finns det en hel del information om polsk morfologi.

Syntaktisk information ges i första hand som en naturlig del av olika exempelfraser. Framför allt handlar det om bruket av prepositioner i norskan och de olika kasusformerna i polskan.

4.1. Uttal

I förordet sägs att användaren förväntas känna till de grundläggande regler som styr sydöstnorskt liksom standardpolskt uttal. Uttalet anges bara då det inte kan utläsas ur ortografin. När det gäller det norska uttalet blir detaljrikedomen därför ganska stor och i själva verket krävs nog relativt omfattande kunskaper om norskans uttalsregler för att kunna tillgodogöra sig alla detaljer. Accenttecken anger betoning, t.ex. ”**åt´ti**” (accent 1) resp. ”**åt`te**” (accent 2), och vokallängd genom att accenten placeras omedelbart efter vokalen om den är lång, t.ex. ”**å`te**”, annars efter följande konsonant. Enstaviga ord lämnas dock omarkerade utom då kort vokal följs av enkel konsonant och då lång vokal följs av flera konsonanter. Bibetoningens placering i sammansättningar

framgår genom att sammansättningsfogen anges med en medelhög punkt före det sista ledet, t.ex. ”**Mid´del·havet**”, men ”**Mid´-delhavs·klima**”. En upphöjd punkt placeras efter lång vokal med bibetoning, t.ex. ”**u´t·re˙dning**”, samt efter stavelse med bibetoning på kort vokal om den inte faller på första stavelsen i slutledet, t.ex. ”**kir`ke·minis˙ter**”.

Oförutsägbara vokalkvaliteter anges med en punkt under vokalen så att *ụ* uttalas som *o*, medan *ọ* uttalas som å, t.ex. ”**blọnd**”. Stum konsonant är i uppslagsorden inte fetstilt, utan i stället kursiverad, t.ex. ”**brø***d* ”. I böjningsformer anges stum konsonant med en punkt under eller över bokstaven, t.ex. ”**gjø`re** v *gjør, gjo`rḍe*”.

Ytterligare ett antal systematiska principer tillämpas i ordboken för att koda olika mindre centrala uttalsföreteelser i norskan direkt i uppslagsordets notation, medan mera osystematiska uttalsavvikelser anges explicit i hakparentes med norskanpassad fonetisk skrift, t.ex. ”**døgn** [døyn]”. Ord vars uttal uppvisar analogi med ett föregående uppslagsord markeras med [↑], t.ex. ”**bọb´sleigh** [-slei] [...] **bọb´sleigh·kjører** [↑]”.

På en punkt förvånas man dock av att en regelbundenhet i norskt uttal noteras vid varje uppslagsord, nämligen det tonlösa uttalet av konsonant före ett *s* som utgör fog i sammansättningar. Detta anges om och om igen, bl.a. vid alla ord i tre hela kolumner från och med ”**liv´s·aften** [-fs-]”, liksom vid alla ord som börjar med *dags-, havs-, skogs-* etc.

För polskans del är uttalet nästan helt förutsägbart baserat på ortografin. Mycket få ord uppvisar oregelbundenheter. Sådana undantag anges i huvudsak med fonetisk skrift i hakparentes, t.ex. ”**fry`se** [...] marznąć [r-z]”, då *rz* i polskan normalt uttalas [ʒ].

Ett enda hjälptecken används direkt i ortografin. I polskan ligger betoningen med nästan fullständig regelbundenhet på ordets näst sista stavelse, och tecknet ˽ används för att ange att två ord i betoningshänseende uttalas som ett, t.ex. ”dobra˽noc” (’god natt’) som alltså betonas på *a*.

Något förvånande är det att undantagsmässig betoning på tredje stavelsen från ordets slut däremot inte anges i ordboken, utan förväntas vara känd. Böjningsformer med sådan betoning är mycket riktigt fullt förutsägbara, men den mindre grupp av lånord som i polskan har bibehållit sådan källspråksbetoning kan vara svåra att förutse.

4.2. Morfologi

Vid de norska uppslagsorden ges för substantiv ändelsen i bestämd form singular, liksom ändelsen i obestämd form plural i de fall då ordet används i plural, t.ex. ”**bok** *-en/-a, bø´ker*”. Substantivens genus anges alltså bara indirekt genom den bestämda artikeln. För verben ges preteritum och verbalsubstantiv, t.ex. ”**li`-me** III v *-te; -ing*”, och för adjektiven ges obestämd form i neutrum singular, samt i mera komplicerade fall även bestämd form/plural, t.ex. ”**gam`mel** adj *-t, gamle*”. Uttal av böjda former anges bara då det inte är förutsägbart.

Information om helt regelbunden böjning av norska uppslagsord anges i förvånansvärt stor utsträckning. Vid substantiv, adjektiv och verb anges i princip alltid böjningen. Vid substantiven fyller detta en funktion, nämligen att ange ordklass, genus och huruvida ordet används i plural, men vid regelbundna adjektiv och verb verkar information om regelbunden böjning överflödig i ett annars så komprimerat ordboksformat.

För de polska ekvivalenterna anges icke förutsägbart genus hos substantiv. Vanligast är behovet vid maskuliner som slutar på *-a* liksom femininer och neutrer som slutar på konsonant, t.ex. ”**ting** [...] rzecz f”. Vidare noteras indelningen av maskulina substantiv i animata och inanimata i de fall då det finns motsättningar mellan ett ords semantik och dess grammatiska klassificering, t.ex. ”**sigarett´** [...] papieros m^a”, då detta ord böjs som ett animat maskulint substantiv. Detta är en ganska komplicerad del av substantivböj-

ningen som påverkar formerna i ackusativ singular och nominativ plural. En hel del stilistisk variation förekommer dock, vilket gör det än svårare att fånga alla detaljer i ordboken, men med fyra olika specialkategorier blir indelningen ändå mer nyanserad än vad som annars är brukligt i polska ordböcker. Att i en ordbok ge fullständig information om all variation i en grammatisk kategori är knappast genomförbart, så här måste SNPO:s system anses väl avvägt.

Precis som de norska uppslagsorden är de polska ekvivalenterna försedda med information om böjningsformer som inte är förutsägbara. I all synnerhet handlar det om ändelsen i genitiv vid många substantiv, t.ex. ”program ⟨-u⟩”. När det gäller övriga kasus, där förutsägbarheten är betydligt högre, anges aktuellt kasus med versal i ett speciellt typsnitt, t.ex. ”**kir`ke** [...] kościół ⟨ℒ -ciele⟩”, vilket anger lokativformen *kościele*.

Substantiv som vid nolländelse har ett epentetiskt *e* mellan de sista konsonanterna i stammen markeras med asterisk, t.ex. ”**ma`ske** [...] maska*”, vilket alltså innebär att genitiv plural, som saknar ändelse, har formen *masek*. En rad andra vokalväxlingar av regelbunden men inte förutsägbar karaktär anges med ett kursivt vokaltecken i den polska ekvivalenten, t.ex. ”**eik** [...] d*ą*b ⟨-u⟩”, vilket ger genitivformen *dębu*.

Ett litet antal mycket oregelbundna polska substantiv markeras med tecknet ∝ och hela deras böjning listas i början av ordboken. Något förvånande omfattar listan bara en halvsida och återstoden av sidan är blank. Man frågar sig vilka kriterier som har avgjort vilka ord som hamnat i denna lista. T.ex. har ett litet antal ord en nominativ singular som helt avviker från stammen i alla andra former, t.ex. ”**dåp** -*en* chrzest ⟨chrztu⟩”. Ingenting i användarhandledningen avslöjar att en sådan genitivform också ska tolkas som information om stammen för alla övriga former. Med tanke på hur få sådana ord det finns i polskan hade det kanske varit klokt att inkludera dem i den inledande listan.

Polska feminina personbeteckningar som är avledda av motsvarande maskulina substantiv anges bara i form av det feminina avledningssuffixet i en speciell hakparentes av typen ⌈ ⌉, t.ex. ”**sku`e·spiller** *-en, -e* aktor ⌈-ka*⌉”, vilket är en mycket praktisk lösning. Icke förutsägbara komparativer anges också i en hakparentes, i detta fall av typen ⌊ ⌋, t.ex. ”**lett** [...] lekki ⌊lżejszy⌋”.

Förvånansvärt nog ger ordboken ingen som helst information om de polska verbens böjning. Det är mycket märkligt med tanke på den högra graden av detalj för alla andra kategorier. Samtidigt är detta ett besvärligt kapitel i den polska morfologin, och det måste därför betraktas som en allvarlig brist att just denna information har utelämnats. Ingen motivering ges heller i användarhandledningen. Det stora flertalet polska verb böjs på ett helt regelbundet och förutsägbart sätt, och det är bara vid en mindre del av verben som det hade funnits behov av att ange icke förutsägbara former. Vissa mycket vanliga och samtidigt mycket oregelbundna verb vore det allra bäst att lista i inledningen och märka med tecknet ∝ på samma sätt som ordboken gör med ett antal substantiv.

Förtjänstfullt är å andra sidan att verbens två aspektformer, som är typiska för de slaviska språken, systematiskt anges för de polska verben. De båda formerna är i förhållande till varandra i det närmaste oförutsägbara. Först ges den perfektiva (fullbordade) formen och sedan, efter ett snedstreck, den imperfektiva (pågående), t.ex. ”**an´·klage** [...] oskarż-yć/ać”, vilket ska utläsas som ”oskarżyć pf, oskarżać ipf”.

Transitivitet noteras däremot explicit endast vid det fåtal verb där samma form kan användas både som transitiv och intransitiv. Detta är dock ett mindre bekymmer eftersom transitiviteten i de allra flesta fall framgår av olika exempel vid de verb som saknar explicit markör.

4.3. Syntax

Syntaktiska konstruktioner redovisas i de allra flesta fall i form av exempelfraser. Särskilt vanliga är sådana exempel vid verb. I polskan redovisas rektionen genom en versal i särskilt typsnitt som anger kasus. Denna motsvarar som regel ett indefinit pronomen i det norska exemplet, t.ex. ”*ha hånd om noe* [...] opiekować się *I*”, där *I* betecknar ett polskt objekt i kasus instrumentalis som motsvarighet till norskans ’om noe’.

Rektion som kan utelämnas anges i vinkelparentes, t.ex. ”**skyl`dig** [...] winny, winien m pred ⟨*i noe G*⟩”, vilket innebär att den norska prepositionsfrasen liksom den motsvarande polska nominalfrasen i genitiv är fakultativ.

Mera sällan ges även annan information av syntaktisk karaktär, t.ex. noteringen m pred i föregående exempel, vilket ska tolkas så att den speciella formen *winien* används i predikativ ställning i maskulinum.

I en del fall ges också syntaktisk information i form av en förklaring, t.ex. ”**med** [...] **9** (om redskap, middel: gjengis med instrumentalis) *skrive med penn* pisać piórem”.

Den syntaktiska informationen är alltså inte lika systematiskt komprimerad, t.ex. genom kodning, som den fonetiska och morfologiska.

4.4. Övrigt

Med tecknet + anges ’också’, dvs. en alternativ möjlighet, t.ex. ”**diva´n** [...] tapczan ⟨+ -u⟩”, vilket innebär att det polska substantivet förutom genitiv på *-a*, vilket ordboken räknar som normalfallet för maskulina substantiv, även kan ha genitiv på *u*. Här hade man dock kunnat önska en notation som visade på frekvensen hos de båda formerna, då genitiv på *-u* i den polska nationalkorpusen (NKJP) liksom i ordboksförlaget PWN:s korpus (KPWN) är betydligt vanligare än formen på *-a*.

5. Ordförråd

Ordbokens omfång och kompakta, informationstäta utformning har givit gott om utrymme för långtgående indelningar i olika delbetydelser med talrika ekvivalenter och exempelfraser (jfr **frossen** i avsnitt 3). Tydlighet i mångfalden upprätthålls i berömvärt hög grad genom användaranpassad disambiguerande information för talare av båda språken, helt i linje med ordbokens bidirektionella intention.

Förutom det centrala ordförrådet i modern norska innehåller ordboken enligt användarhandledningen även vissa äldre, litterära, regionala och dialektala ord. En del produktiva prefix och suffix utgör också självständiga artiklar, t.ex. ”**u-**” och ”**-bar**”. Oregelbundna böjningsformer finns som uppslagsord med hänvisning till det uppslagsord som utgör grundformen, t.ex. ”**er** [ær] pres. → **være**”. En del egennamn finns med ifall polskan och norskan har olika former, t.ex. ”**Pari´s** II geogr. Paryż”, liksom en del vanliga förkortningar, t.ex. ”**LO**”. Sammansättningar som är mycket vanliga, ogenomskinliga eller svåröversatta har också tagits med. Däremot nämns det inte i inledningen vilka principer som styrt valet av fackspråkliga inslag i ordboken.

I det följande granskas närmare några av de kategorier som nämns i användarhandledningen, nämligen sammansatta ord, förkortningar, namn, stilistiskt markerade ord, men också flerordsuttryck, encyklopedisk information och terminologi.

5.1. Sammansatta ord och flerordsuttryck

Enligt användarhandledningen har en del sammansättningar tagits med om de är mycket vanliga, ogenomskinliga eller svåröversatta. Frågan är dock om dessa kriterier verkligen täcker in alla sammansättningar som återfinns i ordboken. Det stämmer att

t.ex. *tomprat* inte är översättbart i sina beståndsdelar och betydelsen är i viss mån ogenomskinlig. Vidare måste man naturligtvis betrakta *tomflaske* och *tomrom* som mycket vanliga, men frågan är om *tomkasse* och *tomsekk* verkligen är så vanliga och/eller så svåröversatta att de klart motiverar en plats i ordboken.

Som ett litet stickprov valdes ett trettiotal någorlunda vanliga sammansättningar ut i dagsaktuell norsk nyhetstext. De allra flesta återfanns i ordboken. De enda som saknades var det relativt nya ordet *klimakrise* och de ganska genomskinliga *redningsbil, russejente* och *ammeku*, medan ord som *valgflesk, nedrustning, voldtektsforsøk, arverekkefølge, alenemor, pøsregn, småbruker, skippertak, knusktørr, vinterføre* och *vogntog* alla fanns med i ordboken.

När det gäller flerordsuttryck är utpräglade egennamn som skrivs med stor begynnelsebokstav i många fall upptagna som egna uppslagsord, t.ex. ”**Rø`de Hal`*v*måne**” och ”**Rø`de Kọr´s**”, medan andra kan vara svårare att hitta. T.ex. återfinns *Den røde plass* bland exempelfraserna inuti artikeln om uppslagsordet **rød**, medan *røde hunder* återfinns som en delbetydelse i artikeln ”**hun***d -en -er* **1** zool. pies [...] **2** med. *røde* [rø`de] *hunder* różyczka **3** astron. *Store hund* Wielki Pies; *Lille hund* Mały Pies”. Detta förklaras i användarhandledningen på följande sätt: ”Fraseologiske uttryck og idiomer er som regel plassert i artikkelen om ordgruppens minst frekvente element.” Semantiskt ligger dock betydelserna **2** och **3** så långt från betydelse **1** att placeringen kan upplevas som långsökt, och ett något större antal flerordsuttryck uppställda som självständiga uppslagsord hade sannolikt bättre svarat mot användarnas intuitiva sökstrategier.

5.2. Förkortningar

Ordboken innehåller många förkortningar som uppslagsord, både allmänspråkliga som *bl.a., ff., f.eks., m.m.*, terminologiska som *dl, mm, aids, DNA,* och förkortningar av namn på t.ex. institutioner,

organisationer, och politiska partier som *FN, USA, LO, OL*. Dessutom förekommer en del avledningar och sammansättningar som utgår från förkortningar, t.ex. *DNA-profil* och *NATO-land*, men märkligt nog saknas *NATO*.

Som ett litet stickprov samlades ett antal förkortningar på dagsaktuella norska nyhetssidor. Av dessa återfanns de politiska partierna *Ap, Frp, H, KrF* och *SV* i ordboken, men inte *MDG* (*Miljøpartiet De Grønne*). Vidare fanns *AS, NRK, NSB, NTB* och *SSB* men inte *VG* (för tidningen *Verdens Gang)*.

5.3. Namn

En stor mängd egennamn återfinns i ordboken. Enligt användarhandledningen har de tagits med i fall då polskan och norskan har olika former, t.ex. ”**Stọck`·họlm** Sztokholm ⟨-u⟩”. Man finner en lång rad geografiska namn med annorlunda polska former, t.ex. *Arktis, Norge, Venezia*, liksom namn på kända gestalter, t.ex. *Shakespeare, Helan og Halvan, Nasse Nøff, Pippi Langstrømpe, Rødhette*, artistiska verk, t.ex. *Svanesjøen, Tolvskillingsoperaen*, byggnader, t.ex. *Eiffeltårnet, Brandenburger Tor, Klippedomen*, och institutioner, t.ex. *Barnevernet, Nobelkomiteen*, men också en del generiska egennamn, bl.a. ”**Ag´nes** Agnieszka”, ”**E`va** Ewa”, ”**Nikola´i** Mikołaj”, ”**Ur´sula** [r-s] Urszula”.

När det gäller geografiska namn hittar man dock även en del fall där det inte föreligger någon skillnad mellan språken annat än i uttalet, t.ex. ”**Euro´pa** [æu-] Europa”, ”**Grena´da** Grenada” eller ”**Soma´lia** Somalia”.

5.4. Encyklopedisk information

Ord för en del begrepp som inte förekommer i Polen tas upp och förklaras i ordboken. Sådan information är snarast att betrakta som encyklopedisk, men fyller en viktig uppgift i en så omfattande

och detaljrik ordbok. Bl.a. återfinns uppslagsordet *Medelsvensson* och här görs en jämförelse med den genomsnittlige *Kowalski* i Polen. Andra exempel är *hardangersøm, haugfolk, lutefisk* och *målsak*, vars innebörd också förklaras.

En annan typ av encyklopedisk information ges vid alla namn på växter och djur, där den latinska beteckningen systematiskt ges i en parentes före den polska ekvivalenten, t.ex. ”**hvi´t·løk** *-en, -er (Allium sativum)* czosnek”.

5.5. Ordförrådets stilistiska variation

Enligt SNPO:s inledning innehåller ordboken förutom det centrala ordförrådet i modern norska även vissa äldre, litterära, regionala och dialektala ord. En rad stilmarkörer tillämpas: B för barnspråk, t.ex. ”**vọv`·vọv·**”; D för dialektalt, t.ex. ”**ful** [...] **3** D (vanskelig) [...] **4** D (sint) [...] **5** D (brunstig)”; F för familjärt, t.ex. ”**mikk**”; † för föråldrat, t.ex. ”**ful** [...] **2** † (frastøtende)”; ℒ för litterärt, t.ex. ”**fọrnøy´e** v *-et* ℒ (more)”; N för neutralt (vilket bara verkar användas vid ekvivalent till ett stilistiskt markerat uppslagsord), t.ex. ”**bọld** adj *-t* poet. [...] dzielny N”; P för folkligt, t.ex. ”**mutter**”; R för regionalt, t.ex. ”**ọl´der** [...] R → **or**”; S för slang, t.ex. ”**ọl´dis**”; V för vulgärt, t.ex. ”**pu`le**”; samt ! för inkorrekt enligt en del språkbrukare, t.ex. ”**fø`re** II [...] **12** (ha til salgs) prowadzić! 𝒜, handlować ℐ”. Flera av dessa markörer förekommer också inom parentes, t.ex. (D) för att ange att ordet är delvis dialektalt etc., t.ex. ”**hål** adj *-t* (D)” eller ”**knul`le** v *-et* (V)”

Det komplexa bruket av olika tecken kan dock leda till tolkningssvårigheter, t.ex. ”**høy`·stål** *-et* D=”, där det trots upprepade genomläsningar av användarhandledningen inte gick att hitta en rimlig uttydning av D=. Möjligen kan man här misstänka ett rent tryckfel.

Som ett litet stickprov valdes de fyra vanligaste slanguttryck som norska ungdomar enligt slang.no uppgett för *gutt* och *jente*,

nämligen *type, kis, fyr, kar* respektive *rype, kjerring, berte, dame*. I SNPO återfinns ”**ty`pe** [...] **3** F (person, fyr) facet, gość, typ”, ”**kis** [I] [k-] *en, -er* S facet, typ”, ”**fyr** [I] *-en, -er* F (mann) facet, gość”, ”**ry`pe** [...] **2** S (jente) cizia, kociak”, ”**ber`te** [I] *-a, -er* S (ung kvinne) babka, cizia”. Vid ordet *kar* saknas, som i NAOB, stilmarkör vid uppslagsordet, men i stället ges den vid de polska ekvivalenterna: ”**kar** [I] *-en -er* mężczyzna[∝] m , człowiek[∝] [...]; (yngre) chłopak m[p-] F; (fyr) facet F”. Vid *kjerring* saknas den aktuella betydelsenyansen: ”**kjerring** [...] **2** F (gift kvinne) mężatka*; (hustru) żona”. För *dame* återfinns bara helt neutrala betydelsenyanser.

5.6. Terminologi

Fackspråk och terminologi diskuteras inte explicit i den inledande användarhandledningen, men man kan lätt konstatera att ordboken innehåller ett rikt urval termer från en mängd olika områden. Förkortningslistan innehåller så gott som 80 olika förkortningar för olika fackområden, vilket är imponerande. Som ett stickprov valdes de 73 termer som även undersöks i Nilsson (2019). Dessa termer hör till grundskolans mera grundläggande begrepp inom matematik, geometri, naturvetenskap och grammatik, bl.a. *multiplikand*, *rettvinklet*, *virvelløs* och *singularis*. Samtliga termer återfinns i SNPO förutom en, nämligen begreppet *fast stoff* (alternativt *faststoff* enligt SNL), vilket bara återfinns i den längre frasen ”*faste stoffers fysikk*” under uppslagsordet ”**stọff**” samt i den längre sammansättningen ”**fas´tstọff-fysikk·**”. Över lag får man intrycket att en mycket stor mängd termer har inkluderats i SNPO, även relativt ovanliga som t.ex. *løsøreregister, mangelsbeføyelse* eller *musettetrekkspill*.

6. Sammantaget intryck

Det finns väldigt mycket gott att säga om Selbergs *Stor norsk-polsk ordbok*, vilket bör ha framgått tydligt av denna recension. Ordbokens främsta styrka är dess omfång. Informationsrikedomen är beundransvärd, både när det gäller urvalet av uppslagsord och mängden information om varje ord, såväl metainformation som ekvivalenter och exempel. Den noggrannhet varmed hela verket är författat är enastående, och några rena felaktigheter har över huvud taget inte kunnat iakttas. Kort sagt är SNPO ett verk av tydligt akademisk rang. Den blir ett mycket gott arbetsredskap för såväl avancerade studenter som för professionella översättare och tolkar.

Vad man skulle önska sig för framtiden är att SNPO tillgängliggörs i en elektronisk utgåva som app för smarta telefoner. Det formatet skulle också möjliggöra löpande uppdateringar med nya ord liksom korrigeringar och andra tillägg. Därefter skulle man naturligtvis önska sig en motsvarande ordbok i den motsatta riktningen, dvs. från polska till norska, men på den punkten måste man nog ge sig till tåls ganska många år med tanke på den tid det har tagit att producera den föreliggande ordboken.

När det gäller brister har några entydiga tryckfel inte kunnat iakttas, men i ett par fall finns en misstanke om tryckfel. I synnerhet vid ord för språk verkar kursiveringen ha missats för det som verkar vara den obestämda artikeln, t.ex. ”**norsk** A et *-en*”, i alla fall om man jämför med användarhandledningen där obestämd artikel är kursiverad. En inkonsekvens skulle dock i så fall vara att obestämd artikel sägs ange att ordet inte förekommer i bestämd form, men vid det aktuella uppslagsordet verkar det i stället handla om att redovisa att ordet kan ha annat genus i obestämd form än det har i bestämd form, vilket inte klargörs i användarhandledningen.

En annan mindre miss är att den polska förkortningen *np.* (= *na pzrykład* 'till exempel') saknas i förkortningslistan i användarhandledningen. Den norska motsvarigheten *f.eks.* finns däremot med.

Enligt listan används vidare "a" för *adjektiv*, medan "adj" står för *adjektivisk*. Av allt att döma verkar dock denna distinktion inte tillämpas i praktiken, då alla de norska uppslagsord som är adjektiv är försedda med markören "adj". Detta kan föranleda viss tvekan inför vad som avses i fall som t.ex. "**hertugin`ne** *-en, -er* księżna adj", där ju ekvivalenten inte är ett adjektiv, utan ett substantiv med adjektivisk böjning.

Artiklarnas kompakta information kan i en del fall vara svår att tolka utan att konsultera förordet, då en del företeelser är mindre vanliga och en del notationer är mindre intuitiva. Sådana mindre frekventa kodningar hade kanske tjänat på ett mera explicit angivelsesätt. De allra flesta kodningarna är dock mycket effektiva och ofta även ganska intuitiva, så på det hela taget är ordboken förvånansvärt lätt att använda. Systemet har utan tvivel gjort det möjligt att inkludera en enorm mängd information utan att ta så mycket utrymme i anspråk, och den som inte är specifikt intresserad av en viss typ av information kommer sannolikt inte att tänka så mycket på specialtecken och förkortningar, utan nöja sig med de givna ekvivalenterna.

Avslutningsvis vill jag varmt gratulera författaren till ett enastående gott och noggrant arbete som kommer att utgöra ett fantastiskt arbetsredskap för väldigt många under lång tid framöver, men för att denna fantastiska ordbok verkligen ska nå ut till dagens användare hoppas nog många med mig på att snart få tillgång till en digital version.

Litteratur

Ordböcker

Berkov, Valerij Pavlovič (1987): *Russisk-norsk ordbok, Russko-norvežskij slovar'.* Moskva: Russkij jazyk.

Berkov, Valerij (2003): *Stor norsk-russisk ordbok, Bol'šoj norvežsko-russkij slovar'.* Oslo: Kunnskapsforlaget.

Bostrup, Lise, Józef Jarosz & Katarzyna Stręk (2014): *Dansk-polsk ordbog*. København: Bostrup.

Kubitsky, Jacek (1995): *Słownik polsko-szwedzki; Polsk-svensk ordbok.* Stockholm: Natur och Kultur.

Kudzinowski, Czesław (1988): *Słownik fińsko-polski; Suomalais-puolalainen sanakirja.* Poznań: Uniwersytet im. A. Mickiewicza.

NAOB = *Det Norske Akademis ordbok*. <naob.no> (juli 2019).

Sørensen, Wanda Strange & Lili Widding (1989): *Dansk-polsk ordbog.* Special-pædagogisk forlag.

Annan litteratur

Dura, Ela, Sven-Göran Malmgren & Maria Toporowska Gronostaj (1996): [Recension av] Jacek Kubicki: *Słownik szwedzko-polski/Svensk-polsk ordbok.* Sthlm: Natur och Kultur 1988. Jacek Kubicki: *Słownik polsko-szwedzki/Polsk-svensk ordbok.* Sthlm: Natur och Kultur 1995. I: *LexicoNordica* 3, 229–241.

KPWN = *Korpus Języka Polskiego PWN* [PWN:s korpus över polska språket]. <sjp.pwn.pl/korpus> (juli 2019).

Nilsson, Morgan (2019): Somaliska tvåspråkiga ordböcker i Norden: utbud och kvalitet. I: *LexicoNordica* 26 (denna volym).

NKJP = *Narodowy Korpus Języka Polskiego* [Den polska nationalkorpusen]. <nkjp.pl> (juli 2019).

slang.no = *slang.no.* <slang.no> (juli 2019).
SNL = *Store norske leksikon.* <snl.no> (juli 2019).

Morgan Nilsson
universitetslektor, fil.dr
Institutionen för språk och litteraturer
Göteborgs universitet
Box 200
SE-405 30 Göteborg
morgan.nilsson@gu.se

Samvittighetsfullt revidert *Norsk ordbok*

Carina Nilstun

Helene Urdland Karlsen & Grete Krogstad: *Norsk ordbok – bokmål*. Oslo: Cappelen Damm 2018. 1063 sider. Pris 345 NOK.

1. Innledning

Norsk ordbok – bokmål (heretter *Norsk ordbok*), som ble utgitt på Cappelen Damm i 2018, er 5. utgave av ordboken som utkom første gang i 1996 under navnet *Cappelens ordbok*[1] og med Aagot Landfald som redaktør. 4. utgave kom i 2006, med Aagot Landfald og Kjell M. Paulssen som redaktører. 5. utgave er redigert av Helene Urdland Karlsen og Grete Krogstad og er ifølge forordet en «nær fullstendig revisjon» av forrige utgave. I det følgende kommer jeg til å sammenligne 4. og 5. utgave for å se nærmere på hva denne moderniseringen består i. Ordboken rommer over 35 000 oppslagsord og 40 000 eksempler, og den finnes bare som papirordbok. *Norsk ordbok* er «tilpasset elever i den videregående skolen, men den kan brukes av norskstudenter på alle nivåer, og av alle andre som trenger å slå opp i en norsk ordbok». (Forord)

Cappelen Damm (og forgjengerne) har lang tradisjon for å utgi ordbøker. Utgivelsene har felles utseende, med egen farge knyttet til hver tittel. De naturlige konkurrentene til denne ordboken er *Bokmålsordboka* (heretter BOB) (2005) og *Kunnskapsforlagets store norske ordbok* (heretter KSNO) (2017). Disse finnes begge både på papir og digitalt og har begge et større lemmatall, henholdsvis 65 000 og 80 000. Papirutgavene koster begge 150 kroner mer enn

1 Cappelen Damm ble dannet i 2007 ved at J.W. Cappelen og N.W. Damm & Søn fusjonerte.

Norsk ordbok, men *Bokmålsordbokas* nettversjon er gratis. Forlaget har altså valgt (i alle fall foreløpig) å bare utgi boken på papir, men har gitt den et prismessig fortrinn mot sin konkurrenter.

Anmelderens bakgrunn er naturligvis formende for anmeldelsen. Jeg har leksikografiske studier fra Universitetet i Oslo. Når det gjelder relevant arbeidserfaring, har jeg jobbet som redaktør i *Det Norske Akademis ordbok* (NAOB) i snart ti år. I 2007 skrev jeg halve den norske kjernen til en serie med bilingvale ordbøker der beskrivelsen av kildespråket er bygget ut såpass mye at de også kan brukes som monolingvale ordbøker. Serien ble utviklet av K Dictionaries og ble i Norge i sin tid utgitt av Vega forlag.

2. Mega- og makrostruktur

Norsk ordbok har en megastruktur bestående av Forord, Innholdsfortegnelse, Brukernøkkel, Veiledning, Liste over forkortelser, Lemmaliste, Minigrammatikk samt Språkregler og skriveråd.[2] Lemmalisten er det man kjenner som ordboken, og det som det er naturlig å vie størst oppmerksomhet i en ordboksanmeldelse. Det øvrige – omteksten – trekkes inn der det er relevant. Makrostrukturen er glattalfabetisk, til forskjell fra i fjerdeutgaven, som hadde reder, og det henvises fra varianter på alfabetisk plass. På alfabetisk plass finner man også rammer med språkinformasjon. Disse er gode og nyttige, og er et spor å følge i den videre utviklingen av ordboken.

Lemmatisering er foretatt etter morfologi, ikke etymologi, ifølge veiledningen (s. 10). Det viser seg for eksempel i at adjektivet *slett* behandles som ett oppslag, og at uttrykket *saken er biff* havner sammen med kjøttstykket. Lemmatisering etter etymologi ville

2 Jeg har i denne anmeldelsen ikke vurdert minigrammatikken og Språkregler og skriveråd, men jeg henviser til minigrammatikken ved et par anledninger.

gitt to adjektiv **slett**: ett med betydningen 'flat', av norrønt *sléttr*, og ett med betydning 'dårlig', fra tysk *schlecht*. *Biff* i *saken er biff* er fra svensk og forkortet form av *bifallit* og altså ikke beslektet med kjøttstykket. Det kommer fra fransk *boeuf*. Men i *Norsk ordbok* finnes homografer som **yr₁** ('nedbør') og **yr₂** ('vrimmel'), **målmann₁** ('keeper') og **målmann₂** ('nynorskbruker') og **himle₁** ('se oppover') og **himle₂** ('bygge himling') – og flere – som har felles morfologi, men ulik etymologi eller orddanning. Disse er lemmatisert på samme måte (det vil si som to oppslag) også i BOB, og det er nærliggende å tro at man her har fulgt BOB uten å tenke på teksten i veiledningen. I 4. utgave var disse ordene redigert sammen, altså etter det morfologiske prinsippet. Veiledningsteksten er på dette punkt identisk i 4. og 5. utgave, men altså ikke fyllestgjørende for løsningene som er valgt i 5. utgave. I 5. utgave er også presensformen av verb føyet til. Dette er en fin og pedagogisk utvidelse, men i minigrammatikken står det: «I ordboken nevnes først – som oppslagsform – *infinitiv*, deretter *preteritum* og til slutt *perfektum partisipp*.» (s. 1044). Igjen er omteksten ikke ajourført til nyeste utgave.

Brukernøkkelen på side 9 viser ordboksartiklenes bestanddeler, godt og effektivt formidlet. På de neste sidene følger en veiledning. Denne fyller halvannen side og rommer noen linjer om hvert av punktene Utvalget av ord, Rekkefølgen av oppslagsordene, Betydning, Faste uttrykk, Uttale og Skrivemåte og bøyning – samt en hel del luft. Både BOB og KSNO har langt mer utfyllende brukerveiledning, og de sidene er vel anvendt plass, til tross for at mange nok ikke leser dem.

3. Lemmaene

Antall oppslagsord i *Norsk ordbok* er som nevnt betydelig lavere enn for de to andre ordbøkene jeg sammenligner den med, og fak-

tisk også lavere enn i 4. utgave (se avsnitt 4). Men lemmaseleksjonen synes å være god, det er ikke mange ord jeg har savnet. (Noen er det, for eksempel *sushi*, *mem(e)*, *miks* og *tannfe*, ingenting graverende altså.) Lemma, betydningsnummer og faste uttrykk står i turkis farge, som også er forsidens farge. I 4. utgave kom både nye betydninger og lemmaer i rede uten forutgående linjeskift. I denne utgaven starter ny betydning på ny linje, og det er en viktig forbedring. Det man vinner i brukervennlighet, forsvarer lett den ekstra plassen det tar.

3.1. Ordklasse og bøyning

Ordklassen substantiv er bare implisitt angitt gjennom angivelse av genus og bøyning. Substantiver som ikke har bøyning, merkes *subst*. For verb står det alltid *v* foran bøyningen. Rekken av bøyningsformer har ingen annen ledetekst enn *fl.* for flertall foran flertallsbøyningen av substantiver og adjektiver, og heller ikke brukernøkkelen eller veiledningen navngir formene. Dette er forholdsvis konvensjonalisert i ordbøkene, men i en ordbok som denne, som retter seg mot elever i den videregående skole og dem som ikke har norsk som morsmål, er det ekstra viktig at spørsmål som kan dukke opp, er besvart. Bøyningsangivelsene er sterkt kodifisert og kunne med fordel hatt brukernøkler for ordklassene substantiv, adjektiv og verb. Substantivartikkelen **suss** kan tjene som eksempel på artikler med en tungt leselig bøyningsangivelse: «*m, n* m: *-en* fl. *-er*, *-ene*, n: *-et* fl. -, *-a/-ene*». Minigrammatikken henviser i liten grad til ordboksartiklene (jf. avsnitt 2).

Samtlige substantiver, verb og adjektiver har bøyning – også sammensetninger og henvisningsartikler. Bindestrek er brukt som plassholdersymbol, og der grunnformen er noe redusert i bøyningene, brukes to prikker som plassholdersymbol. Slik kan det se ut: «**landtunge** *f, m ..ga/-n* fl. *-r, -ne*». Ny ordklasseinndeling er tatt i

bruk,[3] men ikke så detaljert at underinndelingen av determinativer (kvantorer, demonstrativer og possessiver) gis.

Etter bøyning følger i utvalgte tilfeller etymologi. Denne består stort sett bare av lånspråket angitt i forkortet form i en parentes. Iblant gis også opprinnelig form, for eksempel ved **sju** («*norr.* sjau») og **kjøre** («*norr.* keyra»). Danskens rolle i bokmålet er her ikke synlig.

3.2. Uttale

Oppslagsord har trykkmarkering i form av en prikk under den trykksterke vokalen. Er trykket på første stavelse, markeres det ikke. Her ville jeg foretrukket at prikken ble plassert også da. Det er mer brukervennlig og også lettere å gjennomføre for redaktørene. Jeg har funnet en god del oppslagsord som mangler denne prikken (**bikinilinje**, **forkrøplet**, **forlike** og **intellektuell** for å nevne noen). Jeg har også funnet en del inkonsekvens (ved **forward** er trykk markert selv om det ligger på første stavelse, og selv om varianter skal ha trykkmarkering etter samme prinsipp som hovedoppslagsformer, er det noen ganger uteglemt, for eksempel ved **forløyse**). For ord med uttale som avviker vesentlig fra skriftbildet, gis i tillegg uttaleopplysning i skarpe klammer og uten bruk av spesialtegn. Denne opplysningen kommer umiddelbart etter lemmaet, og kan se slik ut: **tic** [tikk], **live** [laiv], **serve** [sørve]. Vokalkvantitet gis ved hjelp av enkle og doble konsonanter, men tonem gis ikke. Avvikende vokalkvalitet (for eksempel *o* uttalt som [å]) alene gir stort sett ikke særskilt uttaleinformasjon. **Jobb/jobbe** har ikke uttaleangivelse, mens **lobb/lobbe** [låbb/låbbe] har. I veiledningen spesifiseres det at det særlig er utenlandske ord som har fått uttaleinformasjon. Uttalene som er angitt, er stort sett ukontroversielle og tidssvarende.

3 I 2005 ble ordklasseinndelingen for norsk til bruk i skoleverket justert etter anbefaling fra Språkrådet og Utdanningsdirektoratet (jf. Språkrådet & Utdanningsdirektoratet 2006).

3.3. Fag- og stilmarkørene

Diasystematisk markering er et vanskelig område, det har nok alle leksikografer erfart. Det er vanskelig å finne et sett av gode markører, og det er vanskelig å anvende dem konsekvent. I *Norsk ordbok* brukes fagmarkørene ofte for å angi *domene*, uten at det er snakk om ord som er spesifikke for det aktuelle faget. Dette kan i noen tilfeller være nyttig, men ofte blir det heller komisk. Fagmarkøren *gastron.* (gastronomi) er i all hovedsak brukt om domenet, og oppleves ofte som overflødig. For eksempel har **nattmat** og **kjevle**[4] (verb) denne markøren. Andre eksempler: **enebolig** har markøren *arkit.* (arkitektur), **vennskapsby** og **vennskapskommune** har markøren *geogr.* (geografi), **gå av med pensjon** er merket *økon.* (økonomi) og **legedom** har markøren *med.* (medisin). I det siste eksempelet ville det vært mer på sin plass å si noe om stilnivået; for de andre eksemplene synes markører å være overflødige.

Markøren *tidl.* (tidligere) brukes både om forhold og gjenstander som hører fortiden til (for eksempel **timeglass** (*særl. tidl.*)), og om ord som nå er lite i bruk (for eksempel **endog** (*særl. tidl.*)). Det er uheldig at så ulike opplysninger har felles markering.

Markørene er svært ofte forkortet. Her er det såpass lite å spare at man burde vurdere å bruke fullformene. Noen fagmarkører blir også i overkant faglige når det i så stor grad er snakk om domene heller enn terminologi. Kanskje kunne *mat* vært en bedre betegnelse enn *gastron.*? Faren er muligens forveksling med forkortelsen *mat.* for matematikk (i *Norsk ordbok* er dette forkortet *matem.*).

3.4. Betydningsbeskrivelsen

Betydningsbeskrivelsen består av definisjon og/eller synonym(er) i rett skrift. Eventuelle fag- og stilmarkører er satt foran, og eventuelle eksempler følger bak. Til sist i betydningen følger i noen

4 Svensk: *kavla*, dansk: *rulle (med kagerulle)*.

tilfeller et antonym. Dette introduseres med MOTS. (motsatt). Her kunne man gjerne føyet til aktuelt betydningsnummer i parentes, slik det gjøres i definisjonene.

Definisjonsfeltet dekker mye. Når betydningsbeskrivelsen er i form av en forklaring, har denne nøyaktig samme utseende som de klassiske substituerbare definisjonene. Eksempler på dette er «**frys**$_1$ kortform av fryseboks el. dypfryser» og «**løs** [...] **5** brukt som etterledd for å lage adjektiv som mangler det forleddet nevner». Dette er likevel greit forståelig og løst på samme måte i BOB. Jeg finner mange gode og treffende ordbeskrivelser som jeg tror er svært nyttige for brukerne. Et komplisert ord som *borgerlig* er enkelt og forståelig gjort rede for over seks betydningsnummer, og alltid er det gitt definisjon og/eller synonym(er). Et annet eksempel er *person* i grammatisk betydning. Begge gjengis nedenfor:

borgerlig [...]
1 som ikke er adelig el. kongelig
2 sivil: *b- strafferett*
3 som følger med det å være borger i en stat: *b-e rettigheter*
4 som bygger på konservativ el. liberal ideologi: *b-e partier*
5 som er preget av en konvensjonell levemåte og tankegang: *b- moral; b- tilværelse*
6 ikke-religiøs: *b- konfirmasjon*

person [...]
4 (*språkv.*) kategori innenfor pronomen og (i enkelte språk) verb: *første p- uttrykker den eller de som snakker, altså 'jeg', 'vi'; annen p- uttrykker den eller de det snakkes til, altså 'du', 'dere '; tredje p- uttrykker den eller de det snakkes om, altså 'han', 'hun', 'den', 'det' el. 'de'.*

3.5. Faste uttrykk

Faste uttrykk er samlet til sist i artikkelen, i alfabetisk rekkefølge. Uttrykkene var i 4. utgave presentert sammen med og på samme måte som eksemplene. Disse er nå hentet ut og gitt samme grafiske utforming som lemmaene og med ny linje for hvert uttrykk. Dette er en stor og viktig forbedring. Det er ikke noe henvisningssystem i uttrykksdelen, og et uttrykk er typisk bare gitt ett sted.

3.6. Ordboken og normen

Norsk ordbok gjengir den fulle bokmålsnormen hva gjelder varianter og bøyninger, til forskjell fra KSNO, som er en ordbok over moderat bokmål. Siden BOB så å si er «bærer» av bokmålsnormen, blir den for fasit å regne. *Norsk ordbok* tar så vidt jeg kan se ingen større snarveier her. De tre variantformene av **høyvann** og **skummetmelk** listes opp, så også **høytlønnet** med sine elleve varianter (jf. BOB for å se samtlige former). At BOB blir en slags fasit, gjenspeiles i flere løsninger, for eksempel lemmatiseringen (se avsnitt 2). Dessuten var normeringen av *stein/sten* over en ganske lang periode gjengitt i forenklet versjon i BOB på nett. Denne forenklingen har forplantet seg til *Norsk ordbok*. Videre har ordet *kraft* en rekke betydninger, hvorav én betydning (avkok-betydningen) har avvikende flertallsform. Dette er det redegjort for i papirversjonen av BOB, men ikke i nettversjonen. I *Norsk ordbok* er dette avviket ikke gjengitt.

Men det finnes også eksempler der det er gjort redaksjonelle valg som avviker fra BOBs praksis. Dette gjelder blant annet håndteringen av aksent i ord som *allé* og *komité*. Disse kan i bokmål skrives med eller uten aksent. Dette er i BOB gjengitt som tre paradigmer: helt uten aksent, aksent i alle former og aksent i grunnform, men uten aksent i bøyde former. Her har *Norsk ordbok* valgt å gjengi bare én løsning, nemlig sistnevnte (jf. Språkrådet 2019 for

en beskrivelse av regler og praksis). Dette valget er både pragmatisk og pedagogisk. Et annet eksempel er at BOB er raus med å gi fulle bøyningsparadigmer, uten alltid å skjele til usus, mens *Norsk ordbok* er mer usus-orientert. *Jiu-jitsu* og *jod* står således uten flertallsbøyning i *Norsk ordbok*, mens man i BOB finner bøyningsformene *jiu-jitsuer*, *jiu-jitsuene* og *joder*, *jodene* respektive *jod*, *joda/jodene*.

3.7. Språket som gjengis, og språket som brukes

Bokmålet har en vid norm hvor alle former er sidestilt. En utfordring når man skal lage ordbøker over bokmål, er å finne et godt redaksjonsspråk. Valgfriheten i bokmål gjør det mulig å skille mellom to varianter av denne målformen: moderat bokmål som er riksmålsnært, og radikalt bokmål, som er nynorsknært (jf. oppslaget **bokmål** i SNL). Fri kombinasjon av former gir ikke et godt språk. NAOB er redigert med moderat bokmål, mens det så vidt jeg vet, ikke finnes noen ordbok som er redigert på radikalt bokmål.

Hvordan er bokmålsnormen håndtert i *Norsk ordbok*? For substantiver som har både feminin og maskulin bøyning, oppgis genus konsekvent i rekkefølgen f(emininum), m(askulinum). For adjektiv med endelse *-et(e)* kommer varianten med endelse *-et* konsekvent før den med *-ete*, for eksempel **frynset**, **frynsete**. I verbbøyningen kommer -a-endelsen alltid før -et-endelsen (*kasta/kastet*), men **forelsket** foran **forelska** og *gav* foran *ga*. Det er altså hverken alfabetet, radikalt/moderat bokmål eller frekvens som er styrende prinsipp for rekkefølgen av lemmaformer og bøyningsformer. For substantivet **to**, som har genusmarkering m(askulinum), n(øytrum) (i den rekkefølgen), er begge eksempler i nøytrum. Ved de to homografene av **svarteliste** finner vi formen *ønskte* i den ene og formen *uønsket* i den andre.

Rekkefølgen på bøyningsalternativer synes altså ikke å henge sammen med redaksjonsspråket, og jeg finner heller ikke en tydelig linje i redaksjonsspråket. For en bruker som ikke er stødig i å

skrive et konsekvent bokmål, er det lite hjelp å få. Bokmålets vide norm har historiske årsaker. Det kan ikke ventes at brukerne av *Norsk ordbok* kjenner historien bak ulike formpar, og jeg mener at ordbøker som gjengir den fulle normen, bør gi veiledning også i formvalg. Veiledningen i *Norsk ordbok* sier hverken noe om redaksjonsspråket som er brukt, eller noe om prinsippene for føring av valgfrie former.

4. Revisjonen

For å gå forordets påstand om at dette er en «nær fullstendig revisjon» av verkets 4. utgave, etter i sømmene, har jeg nærlest et strekk rundt slutten av *m* og begynnelsen av *n*. Sammenligner man lemma for lemma, blir det raskt klart at påstanden holder vann; endringene er gjennomgripende. Mange lemmaer er tatt ut – faktisk står det på baksiden av 4. utgave at den inneholder 45 000 oppslagsord, 10 000 flere enn nyeste utgave! Blant annet var Norges kommunenavn med i 4. utgave, men de er nå ute. Lemmaseleksjonen er som tidligere nevnt meget vellykket, og at man har prioritert kvalitet fremfor kvantitet, er et udelt gode. Som moteksempel har vi *Den store norske bokmålsordboka*, som inneholder 105 000 oppslagsord, deriblant ord som ikke eller i forsvinnende liten grad er å finne i norsk, og hvor svært mange artikler ikke har noen ordbeskrivelse knyttet til seg (Rauset 2015:247, 251–252).

Noen eksempler på hva som er tatt ut, og hva som er kommet inn: I starten av bokstaven *n* er artiklene **nabbe** 'bergknatt' og **nabob** '**1** høy embetsmann i India **2** europeer som har tjent seg rik i Østen' tatt ut, og betydningen 'farlig sterk drink' av **nakkeskudd** likeså. Inn har vi fått **N** som forkortelse for *nord*, **N** som symbol for *nano-*, for et tilfeldig tall og for *nitrogen* og *newton*, samt en videre beskrivelse av bokstaven **N/n**, i tillegg til **nachos** og suffikset **-nad**. Slo man i 4. utgave opp på **måltid**, fikk man bøyning og eksempel,

men ingen definisjon. Verbet **måtte** var tilgodesett med bare tre linjer, måned bare med én. Alt dette er nå bøtet på. Av sammensetninger med *nabo* som førsteledd finnes i 4. utgave **naborett** og **nabovinkel**, i 5. utgave **nabokommune** og **nabolag**. Søk på Nasjonalbibliotekets N-gram viser at avgjørelsene har støtte i usus (NB N-gram 2019). Artikkelen **nabo** kan gjengis som eksempel på fornyelse av artiklene.

> 4. utgave: «den som er/bor nær(mest), granne: *gode n-er; langt til nærmeste n-; n-er i verdensrommet*»

> 5. utgave: «**1** person som bor nær ved; granne: *ha gode n-er; få trøbbel med n-en* **2** land som ligger nær ved: *Russland er vår n- i nord*»

I 4. utgave ble ikke fag- og stilmarkørene skilt grafisk fra definisjonen, mens de nå er kursivert og utskilt med parentes. Det samme gjelder uttrykkene (jf. avsnitt 3.5).

5. En moderne ordbok

I forordet til *Norsk ordbok* betones det at den er en moderne ordbok. Dette har redaktørene lyktes med på mange plan. Designet er innbydende og enkelt. Det nye pronomenet *hen* er med, og også det for meg ukjente *hin*, som skal være synonymt. Disse er begge merket med *uoff.* (uoffisielt). Noen andre moderne ord jeg har funnet, er *norsk-pakistaner*, *norsk-somalier*,[5] *menskopp*, *uten filter*, *mojito*, *mocktail* (nytt for meg) og *abaya*. Pronomenet *hun* er flittig brukt i eksemplene, og eksemplene er fri for gammelmodige kjønnsmønstre. Tidvis er eksemplene sågar samfunnsmessig progressive: «**forelskelse** […] **2** person som en er el. har vært forelsket

5 Disse to skulle vært ført uten bindestrek, analogt med *norskamerikaner*.

i: *hun er hennes nyeste f-*». Uttaleangivelsen [sædselle] er – med mindre det er en feil – basert på en uttale som bare finnes blant ungdom og unge voksne (jf. Uri 2016). *Mæle*, *væpne*, *sæl* og andre ord med *æ* som uttales [e], har ingen uttaleangivelse.

4. utgave hadde ikke alltid med ordklasse, og artiklene kunne være i overkant knappe:

4. utgave:
fem 5
femte 5.

5. utgave:
fem *determ.*
1 grunntallet 5
2 (*utd.*) (god) karakter på ungdomstrinnet på grunnskolen
femte *adj.* ordenstall til fem (5.)

Noen momenter som burde vært mer oppdatert, finner man likevel – sånn er det jo når man først leter. *Norsk ordbok* er bare ajourført etter Språkrådets rettskrivningsvedtak frem til 2015 (jf. Rettskrivningsvedtak etter 2012). Av helt nyere hendelser har vi det at fylkene Nord-Trøndelag og Sør-Trøndelag 1. januar 2018 ble slått sammen til Trøndelag. Dette er kommet med på kartet på innsiden av *Norsk ordboks* perm, men ikke i artiklene. Således er **frøyværing** definert som 'person som bor på el. kommer fra Frøya i Sør-Trøndelag'. Litt mer etablert bør det være at *sukkersyke* ikke lenger er en term som brukes i medisinen (jf. for eksempel SNL og BOB). Dermed burde definisjonen stått på **diabetes**, mens **sukkersyke** burde hatt en annen markør enn *med.* Oppslagsordet **fruktsommelig** er markert med *med., tidl.*, og med det får man ikke riktig formidlet det alderdommelige preget dette ordet har.

Uttrykket *løs og ledig* er bare definert som 'ugift'; her kunne det vært sagt en god del mer. At **høneblund** og **mens** (substantiv) er markert som *slang*, oppleves ikke som moderne, men det sier kanskje først og fremst noe om at ordboken mangler en treffende markør for det mer uformelle ordforrådet.

6. Sagt på en annen måte

Inntrykket er at ordboken er gjennomgått fra a til å, og at alle aspekter har blitt grundig revidert. Jeg savner først og fremst en mer omfattende veiledning. Her skulle jeg særlig ville lese mer om bøyningsformene, siden de er vist i sterkt kondensert format. Jeg savner også en tekst om bokmålsnormen og en redegjørelse for ordbokens redaksjonsspråk. Ikke minst ville dette vært en hjelp til brukerne i hvordan de selv kan forholde seg til normen. Listen over forkortelser kunne vært gjennomgått med henblikk på å skrive flere fullt ut. Ofte er det bare to–tre bokstaver å spare.

Punktene som her er tatt opp, er stort sett småplukk, og den sterkeste innsigelsen gjelder altså ordbokens omtekst. Ordbokens indrefilet – lemmalisten – er svært god; det gjelder utforming og utvalg så vel som innhold. Det er rett og slett imponerende hva en redaksjon på to har fått til! Man kan spørre seg om hvor mange ettbinds ordbøker for bokmål vi egentlig trenger. Jeg tenker at så lenge håndverket er godt og verket har en egenart, så er utgivelsen berettiget. Konkurransemessig er det intet fortrinn at den kun foreligger i papir i en tid da den digitale ordboken tar mer og mer over, også i skolen. Likevel: *Norsk ordbok* bør treffe elevene i videregående skole fordi forlag og redaksjon har lyktes i å lage en oppdatert og lett tilgjengelig ordbok, godt tilpasset sin brukergruppe.

Litteratur

Ordbøker

BOB (2005) = Boye Wangensteen (red.): *Bokmålsordboka*. 3. utgave. Oslo: Kunnskapsforlaget. <https://ordbok.uib.no> (mars 2019).

Den store norske bokmålsordboka (2014). Per-Erik Kirkeby. Oslo: Kagge forlag.

KSNO (2017) = Tor Guttu: *Kunnskapsforlagets store norske ordbok*. 3. utgave. Oslo: Kunnskapsforlaget.

NAOB = Tor Guttu et al. (red.): *Det Norske Akademis ordbok*. Oslo: Kunnskapsforlaget og Det Norske Akademi for Språk og Litteratur. <https://www.naob.no> (mars 2019).

Norsk ordbok: bokmål. 4. utgave (2006). Oslo: J.W. Cappelens Forlag.

Annen litteratur

NB Ngram (2019): <https://www.nb.no/sp_tjenester/beta/ngram_1/> (mars 2019).

Rauset, Margunn (2015): Tungtvegande ny bokmålsordbok? I: *LexicoNordica* 22, 243–261.

Rettskrivningsvedtak etter 2012 (2019): <https://www.sprakradet.no/Spraka-vare/Norsk/normering/Rettskrivingsvedtak/rettskrivningsvedtak-etter-2012> (mars 2019).

SNL = Store norske leksikon (2019): <snl.no> (mars 2019).

Språkrådet (2019): <https://www.sprakradet.no/sprakhjelp/Skriveregler/tegn/Aksentteikn> (mars 2019).

Språkrådet & Utdanningsdirektoratet (2006): *Grammatiske termer til bruk i skoleverket. Tilråding fra Norsk språkråd og Utdanningsdirektoratet*. Mai 2005. Oppdatert 17.03.2006. <https://www.sprakradet.no/localfiles/gramterm.pdf> (juli 2019).

Uri, Helene (2016): Sæd – bokstavelig talt. Kronikk i *Aftenposten* 28.02.2016. <https://www.aftenposten.no/meninger/kronikk/i/Ewej/Kronikk-Sad--bokstavelig-talt--Helene-Uri> (mars 2019).

Carina Nilstun
leksikograf, MA
NAOB, Kunnskapsforlaget ANS
Gullhaug Torg 1
NO-0484 Oslo
carina.nilstun@kunnskapsforlaget.no

Eesti-soome sõnaraamat – en lovande ny estnisk-finsk ordbok på nätet

Mihkel Tõnnov

Heikki Hurtta & Margit Langemets (red.): *Eesti-soome sõnaraamat*. Eesti Keele Instituut, Kotimaisten kielten keskus, Tallinn/Helsinki. Publicerad på internet 9.1.2019. <http://www.eki.ee/dict/efi/>.

1. Hundra år av estnisk-finska ordböcker

I januari 2019 offentliggjorde det estniska Institutet för det estniska språket[1] (EKI) i samarbete med det finländska Institutet för de inhemska språken en ny estnisk-finsk ordbok (ESS, *Eesti-soome sõnaraamat*, fi. *Viro–suomi-sanakirja*). Till skillnad från tidigare ordböcker som publicerats av EKI är denna inte en nätversion av ett tidigare tryckt verk utan en helt ny ordbok.

Ordboken utkom lite över ett år efter att samma utgivare publicerade den finsk-estniska storordboken, *Suomi–viro-suursanakirja*[2] (SVSS) fritt tillgängligt på nätet för att fira Finlands 100 år av självständighet. Även den nya estnisk-finska ordboken skulle enligt förordet till ESS offentliggöras redan till Estlands 100-årsjubileum i februari 2018 – men blev dock nästan ett år försenad och lanserades istället i och med ”året för det estniska språket” som året 2019 blev utsett till av Estlands utbildningsministerium (HTM 2018).

Nyligen var det också 100-årsjubileum för den allra första estnisk-finska ordboken, Lauri Kettunens *Virolais-suomalainen sana-*

1 Eesti Keele Instituut.
2 Med estnisk parallelltitel *Soome-eesti suursõnaraamat*, ursprungligen publicerad år 2003.

kirja som utgavs 1917 i Helsingfors. År 1958 kom den ut i omarbetad utgåva, med en ny tryckning 1965. I Estland publicerades 1991 en ny upplaga på basis av 1917 års utgåva som enligt baksidan av titelbladet var ”delvis uppdaterad till modernt skriftspråk”. Trots denna anmärkning hade boken vid det laget redan begränsad praktisk användning, eftersom bägge språken hade förändrat sig mer än vad en delvis uppdatering förmådde att täcka.

Att Kettunens ordbok hade blivit föråldrad konstaterades redan av Kokla et al. i förordet till deras ordbok *Eesti-soome sõnaraamat* som publicerades 1972. Även denna bok fick flera nytryckningar samt en uppdaterad utgåva år 1993 – som därefter igen utgavs på nytt, senast år 2007 med en varnande anmärkning på omslaget om att det är en oförändrad nyutgåva.

År 2015 följde en ny *Eesti-soome sõnaraamat* sammanställd av Haak et al. som även innehåller nyare ordförråd.

Eftersom såväl SVSS som estnisk-finska ordböcker av de ovannämnda utgivarna nämns som viktiga föregångare i förordet till ESS, ska jag också jämföra hur de olika författarna har valt att lösa diverse utmaningar – både formella och språkliga.

ESS har tagits emot utan större uppmärksamhet. Nästan fem månader efter lansering finns det inga recensioner i estniska tidningar eller språktidskrifter, bara ett par pressmeddelanden (som huvudsakligen består av utdrag ur ordbokens förord) och två artiklar skrivna av redaktionsmedlemmarna själva (dvs. Hurtta 2019 och Kuusk 2019).

Likväl verkar folk ha hittat fram till ordboken. Under de tre första månaderna efter lanseringen gjordes 750 sökningar per dag i ordboken, vilket redaktionen tolkade som bevis på intresse för och behov av resursen (Kuusk 2019). Till slutet av maj hade antalet ökat till 925 per dag, med sökord på såväl finska som på estniska (ESS: Statistika).

Den här recensionen kommer troligen att präglas av min bakgrund som (tidigare) filologistudent och (nuvarande) översättare

och självverkänd språknörd. Bortsett från åtskilliga estniska ordböcker (oftast utgivna av EKI) har jag framför allt använt mig av enspråkiga nätordböcker på de tre skandinaviska språken (mest DDO, BNO, SAOB och SO) men ibland även av finska verk.

2. Storordbokens lillebror

Av förordet till ESS framgår att den i hög grad baseras på SVSS:s databas från 2003, med uppdateringar från andra, nyare ordböcker. Dessutom uppmuntrar båda ordböckerna användaren att prova att göra samma sökning även i den andra ordboken (genom en länk på sidan), så det är logiskt att betrakta dem som närmare relaterade till varandra än till någon av de andra estnisk-finska eller finsk-estniska ordböckerna.

När storordboken utkom i bokform var den enligt förordet riktad till dem som visste vilket finskt uppslagsord de var ute efter, alltså huvudsakligen finnar, men i nätutgåvan kan ordboken lika bra användas av dem som vill hitta en finsk översättning av ett estniskt ord genom att ställa in sökfunktionen på att söka i översättningar. Detsamma gäller faktiskt för alla EKI:s nätordböcker, vars pappersutgåvor bara kunde användas av dem som redan kunde uppslagsorden.

Med tanke på ESS' nära relation till SVSS är det något av en överraskning att den bara innehåller ca 39 000 uppslagsord mot drygt 90 000 i SVSS. Detta gör mig även lite tveksam inför vem den nya ordboken egentligen är ämnad för – för att vara en enkel fickordbok (t.ex. i mobilen) är den alltför innehållsrik och för användning vid exempelvis översättning är den inte innehållsrik nog. För jämförelses skull kan nämnas att Kettunen (1991) har mellan 25 000 och 30 000 uppslagsord,[3] Kokla et al. (1988) har

3 Enligt min beräkning.

mellan 30 000 och 36 000,[4] och Kokla et al. (2007) samt Haak et al. (2015) har ca 45 000 lemman vardera.

Eftersom antalet uppslagsord är så pass mycket mindre än i SVSS är det konstigt att det finns med en del specifika ord (t.ex. *fennomaania* 'fennomani', *tegevusluba* 'koncession', *lekseem* 'lexem', *raadioamatöör* 'radioamatör', *ahvileivapuu* 'apbrödsträd', *fotofoobia* 'ljuskänslighet'). Däremot saknas många ganska vanliga ord (flest adjektiv?) helt: *tüür* 'roder' och *tüürima* 'att styra' (uppförda är dock *tüürimees* 'styrman', *tüürpoord* 'styrbord', *pöördetüür* 'roder'), *tuhm* 'matt, svag, dov' (ordet finns i förklaringar), *kohmetu* 'olustig, illa till mods' (finns i en förklaring), *pöörane* 'galen, tokig' (finns i flera exempelsatser, med olika översättningar), *taibukas* 'fyndig', *võluv* 'charmfull', *korralagedus* 'oordning', *kangekaelne* 'envis' (finns i var sin exempelsats) m.fl.[5]

Dessutom finns det en del ordartiklar som inte är helt färdigredigerade – detta nämns också i förordet och sådana artiklar bär motsvarande varning. Enligt förordet håller redaktionen på att komplettera dem löpande. För det mesta verkar det vara förhållandevis sällsynta sammansättningar och fraser, men det finns även vanligare ord och uttryck som inte är färdiga än (t.ex. *halvama* 'att förlama', *kark* 'krycka', *kerjama* 'att tigga', *kurja vanduma* 'att svära, förbanna', *lohisema* 'att söla, dröja' och *omal ajal* 'i sin tid').

3. Innehåll i en ordartikel

En typisk ordartikel (jfr figur 1) består av själva uppslagsordet (samt eventuell ordgräns ifall det är en sammansättning), en ikon

4 Enligt min beräkning.
5 Exemplen är tagna från ordbokens statistiksida – och för att ett ord ska dyka upp på topplistan måste det ha varit tillräckligt många användare som försökt hitta det.

för att visa böjningsformer, en förkortning för ordklass,[6] förklaring av ordet på estniska och översättning till finska. Därefter följer eventuella bibetydelser samt exempeluttryck och -satser, och ibland även hänvisningar till relaterade ord. Om möjligt, ges det också antonymer till adjektiv. Tyvärr är det dock inte alla uppslagsord som har all denna information (se avsnitt 3.2 och 3.5 nedan).

sõna|raamat M

sõna+raamat

2

sg nominatiiv	sõna+raamat
sg genitiiv	sõna+raamatu
sg partitiiv	sõna+raamatut
sg illatiiv	sõna+raamatusse
pl nominatiiv	sõna+raamatud
pl genitiiv	sõna+raamatute
pl partitiiv	sõna+raamatuid

teatmeteos, mis annab infot teatud keele või valdkonna sõnade kohta

sana|kirja, sanasto

- saksa-eesti sõnaraamat koolidele saksalais-virolainen koulusanakirja
- e-sõnaraamat verkkosanakirja
- slängisõnaraamat slangisanakirja

Figur 1: Ett exempel på en ordartikel.

Nedan ska jag koncentrera mig på några aspekter av ordartiklar som är särskilt välgjorda och några andra som kanske kunde bli bättre med hjälp av konstruktiv kritik.

3.1. Ortografiska val

Även om det estniska skriftspråket ligger förhållandevis nära uttalet är det två aspekter som inte återspeglas exakt i skriften: palataliserat resp. opalataliserat uttal av vissa konsonanter (som kan

6 Ordklass markeras dock inte vid sammansättningar och inte heller vid de flesta verb. Vid sammansatta ord kan användaren vid behov kontrollera ordklassen genom att slå upp efterledet. Vid verb är markeringen däremot helt inkonsekvent.

vara den enda skillnaden mellan två ord) och lång resp. överlång ljudlängd (som oftast bara skiljer ordets böjningsformer och därför kan ses som mindre relevant här).

Som brukligt i ordböcker används standardortografin även i EES, bortsett från markerandet av ordgräns i sammansättningar. Likadant gjorde Kettunen (1991) som också markerade stamändringar. Kokla et al. (2007) och Haak et al. (2015) indikerade palataliserat uttal vid homografi och Kokla et al. (1988) (som var utgiven i Finland och uppenbarligen riktad till finsktalande) markerade det genomgående. I SVSS markeras palatalisation inte; vid homografer tillfogas dock genitivformen som borde klargöra om ordet har palataliserat uttal eller ej.

3.2. Morfologisk information och uttal

Klickar man på M-ikonen (som står för morfologisk information), visas uppslagsordets böjningsformer. För nomina består dessa av de grundläggande kasusformerna: nominativ, genitiv, partitiv och illativ i ental samt nominativ, genitiv och partitiv i flertal. För verb visas indikativ presens singularis 3:e person, indikativ presens impersonal,[7] indikativ imperfekt singularis 1:a och 3:e person, imperativ presens singularis 3:e person, *ma*- och *da*-infinitiv och *tud*-particip.

Morfologisk information ges i ESS bara för de estniska uppslagsorden, ej för finska översättningar. Motsvarande funktion saknas helt i nätutgåvan av SVSS där det bara anges (numerisk) böjningstyp som då länkar till respektive artikel i nätutgåvan av *Kielitoimiston sanakirja* (2018) som visar nominativ, genitiv, partitiv och illativ i singularis samt pluralis. I pappersutgåvan hänvisades läsaren till böjningstabellen i bokens slut. Likadant var det i de tidigare estnisk-finska ordböckerna, med undantag av Kettunen

7 Det estniska verbets impersonalform är i stort sett detsamma som passiv i andra språk.

(1991) och Kokla et al. (1988). Båda nämnde oftast två böjningsformer i själva ordartiklarna (för nomina var det uppslagsordet i nominativ följt av genitivform i ental, för verb *da*-infinitiv följd av indikativ presens singularis 1:a person), Kokla et al. (1988) hänvisade dessutom med typnummer till böjningstabellen.

Som en modern funktion i ESS finns det möjlighet att lyssna på (maskinläst) uttal av vissa böjningsformer, oftast nominativ, genitiv och partitiv singularis för nomina, och indikativ presens singularis 3:e person och infinitivformerna för verb. Även uttalet är något som finns endast för de estniska orden och motsvarande funktion har heller inte blivit tillagd till nätutgåvan av SVSS.

Det verkar dock som om både morfologisk information och uppgifter om uttal har blivit tillagda automatiskt men att det av någon anledning inte fungerat vid homonymer (eller egentligen homografer), varför finskspråkiga läsare inte kan kontrollera direkt hur t.ex. *tee* ’tie; sv. väg’ och *tee* ’tee; sv. te’ böjs eller uttalas. Situationen blir värst vid ord som är homografer men inte homofoner (alltså skrivs lika men uttalas olika – jfr avsnitt 3.1 ovan). Detta gäller huvudsakligen ordpar där det ena ordet innehåller en av de palataliserade konsonanterna /lʲ, nʲ, sʲ, tʲ, d̥ʲ/ och det andra ett opalataliserat /l, n, s, t, d̥/, eftersom dessa inte åtskiljs i den estniska ortografin. Därmed vet man inte utan kontext eller ytterligare böjningsformer om uppslagsord som t.ex. *palk* eller *tall* ska uttalas /pɑlk/ ’palkka; sv. lön’ eller /pɑlʲk/ ’hirsi; sv. bjälke’, /tɑlʲː/ ’talli; sv. stall’ eller /tɑlː/ ’karitsa; sv. lamm’.

För estnisktalande användare framgår uttalet samt böjningen så snart de ser på förklaringen (eller någon böjningsform i eventuella exempelsatser). Den saknade informationen vore nyttig först och främst just för finsktalande som håller på att lära sig estniska (eller av någon annan anledning vill veta hur ett eller annat estniskt ord uttalas och/eller böjs men som inte nödvändigtvis har kunskaper om estniskans uttalsregler).

Märkvärdigt är också att i strid med föregångarna har *ma*-infinitiv valts som verbens uppslagsform. I finsk tradition är det *da*-infinitiv och tidigare har detta följts även i de estnisk-finska ordböckerna, även om man annars använder *ma*-infinitiven som verbens uppslagsform i estniska ordböcker. Det fungerar dock att söka även efter *da*-infinitiven.

3.3. Grundliga förklaringar

Förklaring på estniska ges för alla uppslagsord, även om ordet inte har någon homonym (eller homograf). Detta är annars något som oftast kan ses i enspråkiga ordböcker, och kan kännas lite konstigt här, särskilt för ord som antingen har en mycket abstrakt eller tvärtom väldigt konkret och allmänt känd betydelse. Till exempel förklaras den första betydelsen av ordet för ”jag” (*mina*) i svensk översättning som ’ordet som talaren använder till att peka på sig själv’. Å andra sidan är något slags förklaring absolut nödvändig för homografer och polysema ord, och då kan man lika gärna vara konsekvent och ha med förklaringen även om denna egentligen inte behövs – i alla fall i en nätutgåva. Blir ordboken någon gång utgiven på papper så borde detta kanske tänkas igenom en gång till.

Sådana lite överdrivna förklaringar har förresten redan tidigare blivit kritiserade av Hint & Anttila (2013:285–287) i deras recension av en annan ordbok utgiven av EKI, den estniska etymologiska ordboken (*Eesti etümoloogiasõnaraamat*).

I föregångarna till ESS fanns förklaringar bara där polysemi eller homografi krävde det, och de var då så korta som möjligt.

3.4. Förkortningar

Det används flera förkortningar som markerar ordklass, fackområde, stilnivå eller dylikt, men till skillnad från i många andra nätordböcker (oftast de som är överförda från pappersutgåvor

med minimala ändringar) visas en liten inforuta vid muspekaren när denna hålls över en förkortning. Detsamma sker i nätutgåvan av SVSS. Det finns dock ett fåtal förkortningar i bägge ordböckerna som saknar inforuta, oftast i löpande text eller i hänvisningar till specifika betydelser av ett uppslagsord (t.ex. hänvisas från artikeln *valge* 'vit' till "3. täh" (dvs. *tähendus* '(bi)betydelse') av *must* 'svart' och till "2. täh" av *punane* 'röd').

3.5. Relaterade ord, antonymer och synonymer

Logiken bakom valet av relaterade ord förblir ganska oklar. I de flesta ordartiklar saknas sådana, och där de anges, är de bara ett eller två. Exempelvis hänvisas det från *punane* 'röd' till *punaselt* 'rött (adv.)' och från *saama* 'att få' till *saamine* 'det att få' – men inga relaterade ord anges under t.ex. *must* 'svart; smutsig', *valge* 'vit' eller *andma* 'att ge', även om analogierna egentligen finns.

Antonymirelation är markerad mellan adjektiv-, adverb- eller substantivpar såsom *must* 'svart'/ *valge* 'vit', *must* 'smutsig' / *puhas* 'ren', *pikk* 'lång' / *lühike* 'kort', *valjusti* 'högljudd' / *tasa* 'tyst', *harva* 'sällan' / *sageli* 'ofta' – men inte mellan t.ex. *harv* 'sällsynt, ovanlig' / *sage* 'upprepad, vanlig', trots att *harv* är förklarat som *mittesage* 'inte upprepad, ovanlig'. Därtill finns det exempel som *rõõmus* 'glad' som hänvisar till motsatsen *kurb* 'trist' – men denna hänvisar inte tillbaka; motsvarande gäller för *suur* 'vuxen' (en av bibetydelserna) / *laps* 'barn'.

Behandlingen av synonymer verkar också vara något inkonsekvent: dels har de självständiga ordartiklar, dels inte. Kanske är det ett arv från pappersordböckerna – om synonymer skulle ligga nära varandra i en tryckt bok (som t.ex. *telekaamera* och *televisioonikaamera* 'tv-kamera' eller *sageli* och *sagedasti* 'ofta') är de ihopslagna i samma artikel, om inte (som t.ex. *buss* och *autobuss* 'buss' eller *faks* och *telefaks* 'fax') får de egna artiklar. Det faktum att såväl *troll* (1:a betydelsen) som *trollibuss* 'trolleybuss' har en

egen artikel pekar dock på att det är frågan om en ren tillfällighet. I ett elektroniskt verk vore det i alla fall önskvärt att antingen slå samman synonymerna eller hålla isär dem i bägge fallen, då man oavsett måste använda sökfunktionen.

I exemplen ovan kan man i princip argumentera att dessa ord inte har exakt samma stilnivå – att det andra ordet i varje synonympar låter lite gammaldags, även om det inte är försett med motsvarande markering i ESS – men det finns även synonymer med exakt samma betydelse och stilnivå som har egna artiklar (t.ex. *maikelluke* och *piibeleht* 'liljekonvalj' eller *tasapisi* och *pisitasa* 'småningom').

Det är kanske en smakfråga hur man bäst borde behandla synonymer, men om utgivarna tänker hålla ordboken uppdaterad (som det antyds i förordet) vore det nog enklast för dem – och knappast sämre för användarna – om de hade ett konsekvent genomfört system.

I de tidigare estnisk-finska ordböckerna samt i SVSS markerades varken synonymi eller antonymi, så försöket att göra det i ESS kan i alla fall anses som berömvärt.

4. Design, sökfunktioner och användarvänlighet

Gränssnittet för ESS följer samma grundprinciper och generell design som de andra ordböckerna på EKI:s webbsajt. Det finns dock väsentliga skillnader mellan dem när det gäller till exempel vilka delar av en ordartikel man kan söka ifrån eller vilka metatecken som accepteras.

I ESS kan användarna välja att söka efter uppslagsord, utvidgat uppslagsord,[8] grammatisk information, ordklass, stilnivå och fack-

8 Detta kan användas för att hitta fraser eller sammansättningar innehållande ett visst ord eller led som inte nödvändigtvis är ett självständigt uppslagsord i ordboken.

område, finsk översättning, eller vilken som helst text på estniska eller finska samt efter förekomsten av ett visst ord i hela ordboken. Som standard söks det bland uppslagsorden. Ifall inga resultat finns utvidgas sökningsområdet gradvis (om inte användaren har ändrat inställningen som styr detta, jfr figur 2).

Figur 2: ESS' huvudsida med olika inställningsmöjligheter (på finska).

Möjligheten att söka efter ordklass, stilnivå e.d. har dock lite tveksamt värde eftersom sökningen är begränsad till 50 resultat och det inte finns möjlighet att precisera sökningen med ytterligare kriterier. Med tanke på avsnitt 2 ovan är det nog klokast att välja att söka överallt – det är många saknade uppslagsord som faktiskt förekommer i exempelsatser.

I likhet med många andra moderna nätordböcker (däribland DDO, BNO och ordböckerna på portalen svenska.se) visas det en förslagslista när användaren börjar skriva i sökfältet. Som ett ovanligt och lite besynnerligt drag följer listan inställningen av sökområde. Därvid är det alltså möjligt att söka efter en finsk översättning även om man är lite osäker på hur ordet ska stavas (ifall man av någon anledning inte vill söka i SVSS), eller kontrollera till exempel vilka stilbenämningar eller grammatiska beteckningar

som används i ordboken (ifall man av någon anledning inte vill slå upp listan över dessa i användningsinstruktionerna).

Det finns även möjlighet att söka med hjälp av metatecken: * motsvarar noll eller flera vilka som helst tecken, *?* motsvarar ett enkelt tecken, *[abc-f]* motsvarar ett enkelt *a*, *b* eller något tecken i alfabetiskt omfång från *c* till *f*, och *[^abc-f]* motsvarar vad som helst annat än de angivna tecknen eller teckenomfånget. Det verkar dock som om de estniska specialtecknen *š*, *ž*, *õ*, *ä*, *ö*, *ü* inte följer sin plats i det estniska alfabetet, vilket gör att man vid en sökning som *[s-t]okolaad* inte får fram den förväntade artikeln om *šokolaad* ('choklad'). Sannolikt är det tecknens placering i Unicode-standard som gäller – men detta tydliggörs inte på ESS' webbsida.

De delar av en ordartikel som matchar sökordet markeras med gult och ifall träffen återfinns inne i en sats markeras denna med ljusgult. Detsamma gäller även när det söks efter stilbenämning eller grammatisk beteckning.

Lyckligtvis (enligt min åsikt) har EKI inte velat följa trenden att skapa ett förenklat gränssnitt av webbsidan för mobilenheter. Istället förändras bara storleken och placering av webbsidans element så att allt innehåll är läsbart även på smartmobiler som, efter dagens standard, har blygsam skärmupplösning.[9]

5. Sammanfattning

Ordet som kanske bäst karakteriserar ordboken, är "lovande" – grundprinciperna och de tekniska möjligheter som erbjuds är rimliga och kan byggas ut men samtidigt är det klart att det finns rum för ganska väsentliga förbättringar. Först och främst med tanke på saknad morfologisk information och uttalet av homografer,

9 På liknande sätt har det gjorts på BNO och portalen svenska.se. Däremot brukar texten på DDO vara alltför liten om man inte zoomar in manuellt.

då det är just denna del av ordboken som är helt ny jämfört med SVSS. Innehållet kunde gärna utökas med ytterligare några tusen uppslagsord – och förstås borde resterande artiklar färdigredigeras. Andra brister (såsom ett inkonsekvent system för relaterade ord, inkl. antonymer och synonymer) är förhållandevis små och utgör inget hinder för att man snart ska kunna betrakta ordboken som en värdig efterföljare i den hundraåriga traditionen av estnisk-finska ordböcker.

Litteratur

Ordböcker

BNO = *Bokmålsordboka* og *Nynorskordboka*. Universitetet i Bergen i samarbeid med Språkrådet. <https://ordbok.uib.no/> (maj 2019).

DDO = *Den Danske Ordbog*. Det Danske Sprog- og Litteraturselskab. <http://ordnet.dk/ddo> (maj 2019).

Eesti etümoloogiasõnaraamat (2012). Koostanud ja toimetanud Iris Metsmägi, Meeli Sedrik, Sven-Erik Soosaar. Tallinn: Eesti Keele Sihtasutus. <http://www.eki.ee/dict/ety/> (maj 2019).

ESS: Statistika = *Sõnaraamatu päringute statistika*. <http://www.eki.ee/dict/efi/stats.html> (maj 2019).

Haak, Anu, Helga Laanpere, Mari Maasik, Sirje Ootsing & Jaan Õispuu (2015): *Eesti-soome sõnaraamat*. Tallinn: Valgus.

Kettunen, Lauri (1991 [1917]): *Eesti-soome sõnaraamat. Virolais-suomalainen sanakirja*. Tallinn: Roto.

Kielitoimiston sanakirja (2018). Päivitetty 6.6.2018. Helsinki: Kotimaisten kielten keskus. <https://www.kielitoimistonsanakirja.fi/> (maj 2019).

Kokla, Paul, Helga Laanpere, Mart Mäger & Arno Pikamäe (1988 [1971]): *Virolais-suomalainen sanakirja*. Suomalaisen Kirjal-

lisuuden Seuran toimituksia 301. 2. painos. Helsinki: Suomalaisen Kirjallisuuden Seura.

Kokla, Paul, Helga Laanpere, Mart Mäger & Arno Pikamäe (2007 [1993]): *Eesti-soome sõnaraamat*. Neljas trükk. Tallinn: Valgus.

SAOB (1893–) = *Svenska Akademiens ordbok*. A–vret. <https://svenska.se/> (maj 2019).

SO (2009) = *Svensk ordbok utgiven av Svenska Akademien*. Stockholm: Norstedts. <https://svenska.se/> (maj 2019).

svenska.se = <https://svenska.se/> (maj 2019).

SVSS (2003) = *Suomi–viro-suursanakirja*. Päätoimittaja Valdek Pall. Eesti Keele Instituut, Kotimaisten kielten tutkimuskeskus. Tallinn/Helsinki. <http://www.eki.ee/dict/fie/> (maj 2019).

Annan litteratur

Hint, Mati & Raimo Anttila (2013): Hajamärkusi kauaoodatud etümoloogiate puhul. I: *Keel ja Kirjandus* 4, 285–295. <http://kjk.eki.ee/pdf/MHintRAnttila4_13.pdf> (maj 2019).

HTM (2018) = Haridus- ja Teadusministeerium (20.11.2018): *Eesti keele aasta 2019 tähistab eesti keele mainimist riigikeelena sajand tagasi*. <https://www.hm.ee/et/uudised/eesti-keele-aasta-2019-tahistab-eesti-keele-mainimist-riigikeelena-sajand-tagasi> (maj 2019).

Hurtta, Heikki (2019): Esimene Eesti-Soome tunnel käigus. I: *Sirp* 26.04.2019. <https://sirp.ee/s1-artiklid/varamu/esimene-eesti-soome-tunnel-kaigus/> (maj 2019).

Kuusk, Külli (2019): EKI keelekool: kolm kuud eesti-soome sõnaraamatu elust. I: *Postimees* 15.03.2019. <https://arvamus.postimees.ee/6546050/eki-keelekool-kolm-kuud-eesti-soome-sonaraamatu-elust> (maj 2019).

Mihkel Tõnnov
MA
mihkel.tonnov@eesti.ee

MEDDELELSER

Mindeord for Jens Axelsen

Lotte Follin

Et helt liv i ordbøgernes tjeneste

Jens Axelsen blev født 22. februar 1928, og han døde 19. maj 2018. Han var dansk leksikografis "grand old man", der ydede en uvurderlig indsats for ordbogsarbejdet i Danmark gennem sit mere end 50 år lange virke.

Han indledte sin ordbogskarriere som medarbejder på *Dansk-Engelsk Ordbog* ved Hermann Vinterberg og C.A. Bodelsen i 1955. Fra 1964 var han redaktør på Gyldendals røde *Dansk-Engelsk Ordbog* og *Engelsk-Dansk Ordbog* i mange udgaver samt forfatter til *Dansk-Engelsk Undervisning* og Gyldendals *Retskrivningsordbog*. Herudover bidrog han til mange andre af Gyldendals ordbogsudgivelser.

Han interesserede sig også for ordbøgernes historie, både leksikografisk og som dokumentation af samtiden, og det blev til *En rød klassiker* (Axelsen 1995), en levende skildring af dansk-engelsk ordbog gennem tiderne.

Han vedblev med at være hovedforfatter på de engelske ordbøger, indtil han fik undertegnede som redaktør i 2000, hvorefter han gradvis trak sig tilbage fra det daglige arbejde med ordbøgerne. Vi fejrede i 2005 hans 50-års jubilæum på Gyldendal, og *Dansk-Engelsk Ordbog*s 10. udgave fra 2007 var den sidste fra hans hånd.

En fremsynet ordbogsforfatter

Et kort forsøg på at redegøre for Jens Axelsens kvaliteter som ordbogsforfatter vil være utilstrækkeligt, for de var mange, så vi holder os til få eksempler.

Han var en indsigtsfuld ordbogsforfatter med stor didaktisk viden fra sit virke som lærer, og det gav ham en knivskarp fornemmelse for, hvilken viden man kunne forudsætte hos brugeren. Han var en tidlig fortaler for, at ordbøger skulle kunne bruges uden instruktion i erkendelse af, at ingen læser brugervejledningen, og det derfor er afgørende, at brugeren møder den nødvendige information på den forventede plads.

Han var inkluderende, længe før det kom på mode. Brugeren skulle ikke møde uigennemskuelige grammatiske forkortelser, men gode eksempler, der illustrerede ordenes brug.

Eksemplets mester

Jens blev af både brugere og kolleger rost for sine gode eksempler, og det er fuldt fortjent. Leksikografer ved, at eksempler ofte – tilsigtet eller ej – afslører forfatterens idiosynkrasier, og for dem, der ikke har fornøjet sig med at læse alle Jens' eksempler, kan der afsløres to her.

Jens satte stor pris på sin ven og utrættelige informant i England, W. Glyn Jones. Så hvis et eksempel skulle illustrere noget særlig positivt, fandt man ofte en Dr. Jones indblandet.

Omvendt kunne han ikke fordrage ost, så hvis der var brug for at illustrere noget særlig negativt, var det ofte den sagesløse ost, der måtte holde for.

Mennesket Jens Axelsen

Jens var et varmt menneske, altid interesseret i nye bidragydere til ordbøgerne, studentermedhjælperne, og alle os i redaktionen. Han nød at komme til forlagets efterårsreception, hvor han morede sig meget med at spille diva, der forventede, at redaktøren agerede

skaffer og forsynede ham med lækkerier fra buffeten. Så fandt han en god plads i gården, hvor han holdt underholdende hof.

Han ville altid gerne med til den efterfølgende middag og en episode fra en af dem illustrerer både hans facon og store stolthed over ordbøgerne.

Vi havde bænket os som de første ved et 8-mandsbord, et ungt par nærmer sig forsigtigt, og den ene præsenterer sig som debuterende lyriker. Jens griber hans hånd med ordene, "Goddag, jeg hedder Jens Axelsen. Det er mig, der finansierer din udgivelse." Et eventuelt indtryk af arrogance fordampede hurtigt ved Jens' levende interesse for dem og deres litterære forhåbninger. Han var aldrig selvhævdende, men havde, hvad vi måske kan kalde det en ikke ubetydelig fornemmelse af eget værd – en fornemmelse, vi er mange, som deler. Da Jens fyldte 70, blev han da også hyldet af kolleger fra hele branchen med et sjældent fornøjeligt og informativt festskrift, *Elefant, se også myg* (Garde, Jarvad & Thomsen 1998).

Jens efterlader sig en kostelig arv i form af mange ordbøger, der vil bestå som pejlemærke for leksikografisk kvalitet længe endnu. Ud over sit professionelle bidrag vil han blive husket for sin varme og humor. Æret være hans minde.

Litteratur

Axelsen, Jens (1995): *En rød klassiker – Dansk-engelsk ordbog 1845-1995*. København: Gyldendal.

Garde, Anna, Pia Jarvad & Knud Troels Thomsen (1998): *Elefant, se også myg. Festskrift til Jens Axelsen*. København: Gyldendal.

Lotte Follin
sproglig konsulent; Jens Axelsens redaktør 2000-2007, Gyldendal Ordbøger
Kong Georgs Vej 13
DK-2000 Frederiksberg
lotte.follin@gmail.com

Rapport från styrelsen för Nordiska föreningen för lexikografi

Pär Nilsson

Nordiska föreningen för lexikografi (NFL) arbetar för att utveckla det praktiska och teoretiska ordboksarbetet i Norden och för att upprätta och upprätthålla nordiskt lexikografiskt samarbete. Under 2019 arrangerades det 26:e *LexicoNordica*-symposiet. Föredragen resulterade i nummer 26 av tidskriften *LexicoNordica* (detta nr). Under året genomfördes även den 15:e konferensen om lexikografi i Norden. I samband med konferensen hölls också generalförsamling i NFL. Sedan generalförsamlingen konstituerar styrelsen sig enligt följande:

Pär Nilsson, Sverige, ordförande
Henrik Hovmark, Danmark, vice ordförande
Hanne Lauvstad, Norge, kassör
Kristin Marjun Magnussen, Färöarna, sekreterare
Maria Lehtonen, Finland

Helga Hilmisdóttir, Island, suppleant
Margunn Rauset, Norge, suppleant

Mer om symposiet och konferensen

LexicoNordica-symposiet hölls 10–12 januari 2019 på konferensgården Lysebu i Oslo. Temat var *nordiska ordböcker med två eller flera språk*. Tolv inbjudna föredragshållare närvarade och totalt deltog ett tjugotal personer tillsammans representerande samtliga nord-

iska länder. Symposiet och det efterföljande numret av tidskriften stöttades av Nordplus Nordens språk. För symposiet mottogs även stöd från Fondet för dansk-norsk samarbeid.

Den 4–7 juni avhölls den 15:e konferensen om lexikografi i Norden, i Helsingfors i Finland. Liksom de två senaste konferenserna hade årets upplaga ett övergripande tema, nämligen: *ordböcker – av vem, för vem och varför?* Omkring 110 anmälda besökare deltog i konferensen, och förutom de 3 plenarföredragen hölls 49 enskilda sektionsföredrag fördelade över 3 parallella sektioner. Flera av föredragen hölls inom temat med ämnen som behandlade bl.a. feedback från ordboksanvändare, besöksstatistik och digital utveckling. Förutom sektionsföredragen hölls en sektion med posterpresentationer och dessutom en separat workshop om etymologiska nätresurser. Konferensen arrangerades av Institutet för de inhemska språken i Finland tillsammans med Helsingfors universitet.

Konferensen inleddes med en mottagning i och rundvisning av Institutet för de inhemska språkens fina lokaler på Hagnäskajen i centrala Helsingfors. Dag 2 fortsatte i det stilrena, funktionalistiska Forsthuset (arkitekt Jussi Paatela) på Helsingfors universitet. Efter registrering och inledande välkomsttal höll professor Anna Hannesdóttir från Göteborgs universitet det inledande plenarföredraget. Föredraget, som kan sammanfattas som en historisk exposé över tvåspråkig lexikografi, hade titeln ”Den tvåspråkiga lexikografins utmaningar förr och nu”. Dagen fortsatte med sektionsföredrag fram till kvällen och avslutades med mottagning i universitetets huvudbyggnad och med tal av dekanen vid humanistiska fakulteten på Helsingfors universitet, Pirjo Hiidenmaa.

Iztok Kosem från universitetet i Ljubljana startade dag 3 med sitt plenarföredrag: ”Bringing dictionaries (and lexicographers) closer together”. Föredraget utgjorde en översikt över det digitala lexikografiska tillståndet i Europa. Efter lunch följde en exkursion med guidad visning till Villa Hvitträsk, en nationalromantisk arki-

tektvilla i Kyrkslätts skogar, som har fungerat som bostad och ateljé åt Eliel Saarinen, Herman Gesellius och Armas Lindgren. Dagen avslutades med konferensmiddag på restaurang Haltia i Esbo.

Dag 4 inleddes med NFL:s generalförsamling följt av konferensens sista plenarföredrag ”Den tvåspråkiga ordboken – ett absurt åtagande” av Nina Martola från Institutet för de inhemska språken, Helsingfors. Föredraget fokuserade den tvåspråkiga lexikografins utmaningar med huvudsaklig utgångspunkt i ett enskilt verk, nämligen *Stora finsk-svenska ordboken*. Efter de sista sektionsföredragen avslutades konferensen med tal och avtackning på fredagseftermiddagen.

Konferensens föredrag kommer att publiceras i tryckt form i den kommande konferensrapporten, *Nordiska studier i lexikografi* 15 under 2020. Nytt för årets konferens var att plenarföredragen streamades och tillgängliggjordes digitalt i realtid. Dessa föredrag finns tillgängliga på NFL:s hemsida (<http://nordisk-leksikografi.com/konferenser.html>).

Övrigt

Styrelsen har under 2019 lagt ned mycket tid och kraft på att ta fram en ny hemsida, sedan den tidigare webbplatsen under Nordisk sprogkoordination plötsligt lades ned i samband med att den organisationen upphörde. Förändringen blir för NFL:s del förhoppningsvis till det bättre då den nya hemsidan har bättre funktionalitet och bl.a. tillåter visning av rörligt bild och uppladdning av betydligt större dokument. Den nya hemsidan finns tillgänglig på adressen: <www.nordisk-leksikografi.com> (och <www.nordisk-lexikografi.com>). Liksom tidigare används NFL:s Facebook-sida som kompletterande informationskanal.

Styrelsen arbetar vidare med att både bevara och i högre grad tillgängliggöra NFL:s samtliga publikationer. Därför kommer för-

eningens tidskrifter att arkiveras i Dansk BiblioteksCenters (DBC) arkiv för danska nätpublikationer, som en fast punkt för publikationerna också vid framtida omorganisationer av hemsidan o. dyl. Ett samarbete har också inletts med European Lexicographic Infrastructure, ELEXIS, med avseende på sökverktyget *Elexifinder*, en databas där lexikografiskt relevanta artiklar och texter tillgängliggörs.

Fullständig årsberättelse för föreningens verksamhet under de två senaste åren finns på NFL:s hemsida (<http://nordisk-leksikografi.com/om_nfl.html>).

Tack

Styrelsen tackar alla dem som på olika sätt bidragit till NFL:s arbete under det gångna året. Ett stort tack riktas till konferenskommittén för den 15:e konferensen om lexikografi i Norden, Caroline Sandström, Ulla-Maija Forsberg, Maria Lehtonen, Charlotta af Hällström-Reijonen, Lotta Jalava, Santeri Junttila, Hanna Lehti-Eklund och Klaas Ruppel. NFL tackar också för finansiellt stöd i samband med konferensen; tack till Nordplus, Svenska Litteratursällskapet i Finland och Vetenskapliga samfundets delegation.

De avgående styrelseledamöterna Þórdís Úlfarsdóttir och Caroline Sandström tackas för värdefulla insatser för föreningen. Ett särskilt stort tack riktas till Caroline för driv och hårt arbete i samband med konferensen.

Styrelsen tackar även redaktionen för *LexicoNordica*, landsredaktörerna och de två huvudredaktörerna, Sturla Berg-Olsen och Emma Sköldberg, som återigen lagt ned ett stort arbete på tidskriften. Den avgående huvudredaktören Emma Sköldberg förtjänar ett särskilt stort tack för ovärderligt arbete under många år i föreningens tjänst och för ständig entusiasm. Tack också till Fondet for dansk-norsk samarbeid, och till Nordplus för stöd också

till symposiet. Laurids Kristian Fahl tackas för hjälp med layout och distribution av *LexicoNordica* och Dansk Sprognævn för hjälp med uppladdning av föreningens publikationer. Rikke Hauge tackas igen för sina värdefulla insatser för föreningens administration och ekonomi.

Framåt

NFL blickar framåt och har närmast det 27:e *LexicoNordica*-symposiet i sikte. Detta äger rum i januari 2020 med temat *kontroversiella ord i nordiska ordböcker*.

2021 arrangeras den 16:e konferensen om lexikografi i Norden, och turen har nu kommit till Sverige. Konferensen kommer att bli ett samarrangemang mellan Göteborgs universitet och Svenska Akademiens ordboksredaktioner i Göteborg och Lund. Konferensen går av stapeln i Lund.

Mer information om både symposiet och konferensen kommer inom kort att presenteras på NFL:s hemsida.

Pär Nilsson
ordboksredaktör
Redaktionen för Svenska Akademiens ordbok
Dalbyvägen 3
224 60 Lund
Par.Nilsson@svenskaakademien.se

REDAKSJONELT

1. LexicoNordica **udkommer** hvert år i november. Tidsskriftet indeholder leksikografiske bidrag som er skrevet på et af følgende nordiske sprog: dansk, finsk, færøsk, islandsk, norsk (bokmål eller nynorsk) og svensk. Bidrag på engelsk, fransk eller tysk kan også optages hvis særlige forhold taler for det.

2. **Bidrag** sendes til det medlem af redaktionskomitéen som repræsenterer bidragyderens land:

   - Anna Braasch, Institut for Nordiske Studier og Sprogvidenskab (NorS)/Center for Sprogteknologi, Københavns Universitet, Njalsgade 136, DK-2300 København S. <braasch@hum.ku.dk>.
   - Kjetil Gundersen, Erika Nissens gate 7, NO-0480 Oslo. <gunderkj@yahoo.com>.
   - Lennart Larsson, Institutionen för nordiska språk, Uppsala universitet, Box 527, SE-751 20 Uppsala. <lennart.larsson@nordiska.uu.se>.
   - Harry Lönnroth, Institutionen för språk- och kommunikationsstudier, PB 35, FI-40014 Jyväskylä universitet. <harry.h.lonnroth@jyu.fi>.
   - Ásta Svavarsdóttir, Árni Magnússon-instituttet for islandske studier, Laugavegi 13, IS-101 Reykjavík. <asta.svavarsdottir@arnastofnun.is / asta@hi.is>.

   Seneste tidspunkt for aflevering af bidrag er **den 1. april** hvis artiklen skal kunne trykkes i det nummer af tidsskriftet som udkommer i november samme år. Bidraget indleveres digitalt i både tekstbehandlingsformat og i PDF-format.

3. **Illustrationer** der skal medtages i artiklen, indsættes i manuskriptet og vedlægges som separate grafikfiler, helst i JPG-format og minimum 300 ppi.

4. **Manuskript**: Bidraget forfattes i LexicoNordicas stilark, der kan rekvireres ved henvendelse til redaktionen. Manuskriptet **indledes** med titel på artiklen og forfatterens navn. For tematiske og ikke-tematiske bidrag følger et **abstract** på engelsk på op til 10 linjer og dernæst selve artiklen, som opdeles i kapitler. Bidraget afsluttes med angivelse af forfatterens navn, titel samt post- og e-mailadresse. Bidrag kan normalt have et omfang på højst 20 sider.

5. **Citater**: Kortere citater (op til 3 linjer) bringes som en del af teksten med dobbelte anførselstegn omkring, mens længere citater eller fremhævelser af større vigtighed gives i et afsnit for sig selv **uden** anførselstegn.

6. Vi anbefaler en tilbageholdende brug af **fodnoter**. Evt. nødvendige noter gennemnummereres i teksten med højtstillet angivelse uden parentes.

7. **Litteraturhenvisninger** foretages i teksten efter følgende model:

   som det fremgår af Herbst (2009)
   som det fremgår af Borin & Forsberg (2011:18)
   (se Herbst 2009:158ff.)

   I den løbende tekst angives ikke hele internetadresser, men et forfatternavn eller en angivelse af titlen på internetbidraget, som bruges i litteraturlisten. Her angives internetadresser uden understregning, men omgivet af < >.

8. Særlige angivelser: Vær tilbageholdende med brug af **fede** typer; **sprogeksempler** markeres med kursiv, fx: ordet *ungkarl* har synonymet *alenemand*; **betydninger** af sproglige enheder

angives ved hjælp af enkelte anførselstegn, fx: 'en ugift mand'; dobbelte anførselstegn bruges ved **citater** eller **forbehold**, fx: De er vokset op i de "glade" tressere. Tegnsætningsreglerne, bl.a. for brug af komma, tankestreg, bindestreg (i betydningen 'fra … til'), er forskellige i de nordiske lande, og forfatterne bør naturligvis følge reglerne for det sprog som bruges i artiklen.

9. **Litteraturangivelser**
I tilfælde af en længere litteraturliste kan den inddeles i to dele i lighed med nedenstående eksempel. Hvad angår angivelser som *red.*, *eds.*, *Hrsg.*, anbefales det så vidt muligt at bruge originalsproget. Det vigtigste er dog konsekvens inden for samme liste.

Litteratur

Ordbøger

ALD (1948) = A.S. Hornby, E.V. Gatenby & H. Wakefield: *A Learner's Dictionary of Current English.* London: Oxford University Press.

COBUILD (1987) = *Collins* COBUILD *English Language Dictionary.* Editor in Chief: John Sinclair, Managing Editor: Patrick Hanks. London/Glasgow: Collins.

ELEXIKO = Annette Klosa m.fl.: *elexiko.* Mannheim: Institut für Deutsche Sprache. <www.elexiko.de> (maj 2008).

Italiensk-Dansk Ordbog (1999). Knud Andersen & Giovanni Mafera. København: Gyldendal.

Jarvad, Pia (1999): *Nye Ord. Ordbog over nye ord i dansk 1955-1998*. København: Gyldendal.

Norstedts stora engelska ordbok (2000). Stockholm: Norstedts.

Oxford-Hachette French Dictionary (1994). Oxford: Oxford University Press.

STO = Anna Braasch m.fl. (red.): *Sprogteknologisk Ordbase*. København: Center for Sprogteknologi. <www.cst.dk/cgibin/defisto> (april 2007).

Anden litteratur

Haiman, John (1980): Dictionaries and Encyclopedias. I: *Lingua* 50, 329-357.

Jónsdóttir, Halldóra & Þórdís Úlfarsdóttir (2011): ISLEX – en flersproget nordisk ordbog. I: Birgit Eaker, Lennart Larsson & Anki Mattisson (red.): *Nordiska studier i lexikografi* 11. Lund: Nordisk förening för lexikografi, 353–366.

Lakoff, George & Mark Johnson (1980): *Metaphors we live by*. Chicago/London: The University of Chicago Press.

Mugdan, Joachim (1985a): Grammatik im Wörterbuch: Wortbildung. I: Herbert Ernst Wiegand (Hrsg.): *Studien zur neuhochdeutschen Lexikographie IV*. Hildesheim/Zürich/New York: Olms, 237-308.

Nikula, Kristina (2012): Samspelet mellan text och bild i enspråkigt svenska ordböcker. I: *LexicoNordica* 19 (dette bind).

NRG = Jan Terje Faarlund, Kjell Ivar Vannebo & Svein Lie (1997): *Norsk referansegrammatikk*. Oslo: Universitetsforlaget.

Svenonius, Peter (1996): The Verb-Particle Alternation in the Scandinavian Languages. <http://ling.auf.net/lingBuzz> (september 2011).

Zgusta, Ladislav (1971): *Manual of lexicography*. The Hague: Mouton.

10. LexicoNordica udkommer både som trykt tidsskrift og i en **internetudgave**. Ved indsendelse af et bidrag til redaktionen erklærer forfatterne sig derfor indforstået med både trykt og digital publicering.