


# NORDISKE STUDIER I LEKSIKOGRAFI

Titel:	Hur blir en traditionell, tryckt ordbok en webbordbok?	
Forfatter:	Ilse Cantell & Caroline Sandström	
Kilde:	Nordiska Studier i Lexikografi 11, 2012, s. 157-168 Rapport från Konferens om lexicografi i Norden, Lund 24.-27. maj 2011	
URL:	<a href="http://ojs.statsbiblioteket.dk/index.php/nsil/issue/archive">http://ojs.statsbiblioteket.dk/index.php/nsil/issue/archive</a>	

© Nordisk forening for leksikografi

## Betingelser for brug af denne artikel

Denne artikel er omfattet af ophavsretsloven, og der må citeres fra den. Følgende betingelser skal dog være opfyldt:

- Citatet skal være i overensstemmelse med „god skik“
- Der må kun citeres „i det omfang, som betinges af formålet“
- Ophavsmanden til teksten skal krediteres, og kilden skal angives, jf. ovenstående bibliografiske oplysninger.

## Søgbarhed

Artiklerne i de ældre Nordiske studier i lexicografi (1-5) er skannet og OCR-behandlet. OCR står for 'optical character recognition' og kan ved tegngenkendelse konvertere et billede til tekst. Dermed kan man søge i teksten. Imidlertid kan der opstå fejl i tegngenkendelsen, og når man søger på fx navne, skal man være forberedt på at søgningen ikke er 100 % pålidelig.

# Hur blir en traditionell, tryckt ordbok en webbordbok?

*Ilse Cantell & Caroline Sandström*

The history of *Ordbok över Finlands svenska folkmål* (Dictionary of Swedish Dialects in Finland) began in the 1920s when the collecting of dialect data was initiated. The editing process started at the beginning of the 1960s, and four volumes were published between 1982 and 2007.

When the digitizing process began in 2010 the first two volumes (*A–Hu*), which only existed on paper, had to be scanned. They were converted into XML format using the structure indicators in the dictionary articles. The manuscripts for the next two volumes (*Hy–Och*) were compiled in a word processing program. Today the dictionary is compiled in a dictionary editing program. In the future, new articles will primarily be published online. A pilot consisting of the interval (*I–K*) will appear in 2013.

In our presentation we discuss difficulties which arise when a traditional printed dictionary is transformed into an online publication. We focus on features that can be used to parse the dictionary text into structured text files. Important questions are e.g.: What structure indicators are typographical? When do we have to pay attention to the content of the text to get proper tagging?

Our target is an electronic dictionary with flexible search possibilities that fully use all elements of the articles. We aim at serving both members of the ordinary dialect-speaking public and researchers in linguistics.

## 1. Projektet *Ordbok över Finlands svenska folkmål*

Bakgrunden till *Ordbok över Finlands svenska folkmål* (FO) sträcker sig ända till 1920–30-talen<sup>1</sup>. Vid denna tidpunkt hade dialekterna redan dokumenterats i några decennier och upplysningar om de svenska dialekterna i Finland hade publicerats i bl.a. Rietz (1862–67), Vendell (1904–07) och Wessman (1925–

---

<sup>1</sup> Ahlbäck (1976: VIII–XIX) och Slotte (1981: 45–64) redogör för bakgrunden till FO, källorna och redigeringsprinciperna för ordboken. Olav Ahlbäck var huvudredaktör (1960–1981) och han efterträddes av Peter Slotte (1982–2009). Sedan 2009 är Caroline Sandström huvudredaktör för FO.

32). År 1927 inrättades Folkmålskommissionen<sup>2</sup> som beviljades statliga anslag och i samband med detta väcktes tanken om att öka insamlingen med sikte på en heltäckande dialektordbok. Ett omfattande, systematiskt insamlingsarbete av dialektala ord och uttryck inleddes i samarbete med Svenska litteratursällskapet i Finland (SLS). Under slutet av 1940-talet påbörjades arbetet på ett ordregister över det insamlade dialektmaterialet kompletterat med andra källor. Resultatet blev ett register som i dag består av omkring en miljon ordsedlar.

Redigeringsarbetet på dialektordboken inleddes i början av 1960-talet. Ordbokens redaktion överfördes till Forskningscentralen för de inhemska språken<sup>3</sup> då forskningscentralen grundades. Första häftet av FO kom ut 1976, följt av ytterligare fem häften med tre till fyra års mellanrum. År 1982 publicerades det första hela bandet (*A–E*) och det andra bandet (*F–Hu*) stod färdigt 1991. Utgivningen i häften frångicks under början av 1990-talet och tredje (2000) och fjärde banden (2007) – intervallet (*Hy–Kjys*) respektive (*Kjysa–Och*) – har utkommit som hela volymer. Från och med fjärde bandet är SLS förläggare för ordboken.

FO är en typisk *flergenerationsordbok* (Vikør 1999) med en lång tillkomst-historia. Kännetecknande för dessa ordböcker är att varierande redigerings- och trycktekniker har avlöst varandra under årens lopp, vilket bl.a. kan ha konsekvenser vid överföringen av ordböckerna till digitala databaser.

De första häftena av FO redigerades manuellt och sattes vid tryckningen med datorstyrd fotosättning. Under åren 1993–2005 redigerades FO med ett ordbehandlingsprogram och 2006 övergick man till ett redigeringsprogram. Publiceringen av nytt material sker under 2010–20-talen primärt på webben, men enligt överenskommelse med förlaget SLS kommer FO också i framtiden att tryckas. Den färdiga ordboken beräknas bestå av omkring 125 000 artiklar fördelade på sju band.

Utgångsläget för den digitalisering som behövs för webbordboken är varierande för de olika avsnitten av ordboken. De två första banden har skannats och genomgått textigenkänning (OCR) innan de kunnat konverteras till XML-kod varefter de ännu måste kontrolleras i flera omgångar vid överföringsprocessen. Tredje och fjärde banden redigerades i ett ordbehandlingsprogram och har konverterats och kontrollerats vid överföringen till XML-kod. Eftersom den pågå-

<sup>2</sup> Folkmålskommissionen, vars officiella namn är Kommissionen för undersökning av svenska folkmål och svensk folkkultur i Finland, var en fri sammanslutning av företrädare från olika vetenskapliga kulturinstitutioner.

<sup>3</sup> Forskningscentralen för de inhemska språken, sedan 1.1.2012 Institutet för de inhemska språken, grundades 1976 och är ett statligt finansierat språkinstitut underställt Undervisnings- och kulturministeriet.

ende redigeringen av nya artiklar sker med ett redigeringsprogram i XML-kod kan de nyaste artiklarna från femte bandet framåt tillsammans med artiklarna i de fyra första banden sammanslås till en XML-kodad databas som sedan konverteras till det format användargränssnittet på webben kräver.

## 2. Varför en webbordbok?

Förlagsbranschen både i Finland och på andra håll i Europa har under de senaste åren drabbats av nedskärningar som i synnerhet påverkat utgivningen av stora referensverk och ordböcker (se Laureys 2012). Dessa tendenser märks också i Institutet för de inhemska språkens utgivningspolitik; de egna ekonomiska resurserna för tryckning av pappersordböcker är begränsade och under de senaste åren har det blivit allt svårare att få utomstående finansiering för tryckning.

De stora pågående ordboksprojekten vid institutet har under 2000-talet genomgått flera utvärderingsprocesser. Trots de statliga sparkraven vill institutet hålla fast vid kärnverksamheten i lexikografi och språkvård och även i framtiden satsa på utgivning av kulturbärande historiska ordböcker. De kommer i fortsättningen att publiceras på webben. Framför allt för forskare ger webbordböcker möjlighet till nya ingångar till språkliga data. Webbpubliceringen medför också att dessa ordböcker når ut till nya användare. För FO:s del innebär de nya kraven och övergången till webbpublicering att också de tidigare utgivna banden av ordboken bör göras tillgängliga på webben så att det kompletta ordboksverket kan komma såväl forskare som allmänheten till godo.

## 3. FO:s strukturer och konverteringarna

Man kan uppfatta strukturen i en ordbok på många olika sätt. Den enda struktur som är konkret synlig i en tryckt ordbok är typografin och layouten. Men bakom den finns lexikografens tanke om hela ordbokens och de enskilda artiklarnas struktur.

Om det vore så att det fanns en direkt motsvarighet mellan varje lexikografiskt strukturelement<sup>4</sup> och en typografisk lösning, så skulle konverteringen vara enkel. Men så är inte fallet, utan konverteringen måste bygga på en analys av sambandet mellan text och typografi.

---

<sup>4</sup> Ibland används här också för korthetens skull termen *element* om inte risk för sammanblandning med begreppet 'element' i XML-strukturen föreligger.

### 3.1 Den typografiska strukturen

Då FO ska ges ut som webbordbok, ska texten i ordbokens fyra första band överföras till en kodning enligt märkspråket XML. Vid konverteringen finns det inte något annat att utgå ifrån än texten och typografin. Texten i den tryckta FO erbjuder tre typer av fästpunkter som kan användas vid konverteringen till XML-kod. Det kan vara rent layoutmässiga drag, typografiska drag, förekomsten av vissa textsträngar eller av tecken som bara förekommer i vissa delar av ordboksstrukturen. Allt detta ska återges i webbordboken, men ibland kan det inte genomföras på samma sätt som vid tryckningen. Fonter kan t.ex. bete sig annorlunda på skärmen jämfört med hur de ser ut i tryck.

I en webbordbok ter sig frågan om utrymme helt annorlunda än i en tryckt ordbok. På en datorskärm finns det inte lika mycket plats som på en tryckt sida. Å andra sidan kan själva materialet som blir tillgängligt på webben, dvs. databasen, vara nästan oändligt stort.

I FO:s makrostruktur markeras början och slutet av artiklarna av hängande indrag och radbyten. Som typografiska medel att skilja åt elementen i FO:s mikrostruktur har man i den tryckta ordboken utöver rak brödtext använt olika snittvarianter: fet och kursiverad stil, spärning och olika teckengrader.

Dessutom innehåller FO textsträngar som dels står som strukturella markörer mellan olika element i mikrostrukturen, dels bara kan ingå i vissa element. Det kan vara fråga om skiljetecken (semikolon, komma, tankstreck o.d.), siffror och ensamma högerparenteser som i FO används som markörer för de olika betydelsemomenten.

Tecken från landsmålsalfabetet<sup>5</sup> förekommer endast i elementet <uttal>, dvs. i uttalsvarianterna av uppslagsorden och i elementet <fras>, exempelfraserna. Sådana textsträngar kan kodas antingen som elementet <uttal> eller elementet <fras>. Båda dessa har dessutom ett attribut vars två olika värden ”fin” och ”grov” anger vilken ljudbeteckning som använts i källan. Apostrof som anger muljering och omvänd apostrof för partiell tonlöshet är också tecken som endast finns i elementen <uttal> och <fras>.

Spärrad stil har i den tryckta ordboken använts för hänvisningar men också i andra funktioner då man velat ge emfas åt ett ord eller en fras. Med hjälp av hänvisningsmarkörer (=, *Jfr/jfr* eller *se*) ska det vara möjligt att göra skillnad på

<sup>5</sup> Tre tecken som används i ljudbeteckningen i FO och som ingår i landsmålsalfabetet men inte i IPA har på initiativ av FO:s redaktion införts i Unicode-standardens version 5.1.0 i mars 2008. Dessa är tecknet för stockholms-e, tecknet för tonlöst l och o-tecknet med ring nertill. De ingår nu i fältet Latin Extended-C och deras koder är U+2C78 (Latin small letter e with tail), U+2C79 (Latin small letter turned r with tail) och U+2C7A (Latin small letter o with low ring inside).

explicita hänvisningar som automatiskt kan länkas ordboksinternt och ord och fraser som getts emfas genom samma typografiska lösning. Enskilda ord i spärrad stil är implicita hänvisningar, de kan ibland stå i böjd form och kan då inte länkas automatiskt. (Se exempel i artikeln *hydare* nedan.) Fraser i spärrad stil kan ofta inte i sin helhet tolkas som hänvisningar:

lyfta v. [- -] lyfta på (el. uppå) hatten el. mössan [- -]

Fraser av denna typ föregås vanligen av texten *i uttr.* (= i uttrycket). Frasen kan ses som en uppslagsform till efterföljande dialektala exempelfraser och ska då inte länkas till en annan artikel.

### 3.2 Den lexikografiska strukturen

Den lexikografiska strukturen, framför allt makro- och mikrostrukturen ska vid konverteringen överföras till XML-kod. Enskilda artiklar och nischer med sammansättningar ska analyseras fram och ordboksartiklarnas objektspråkliga och metaspråkliga strukturelement ska få rätt kod.

Artiklarna i FO kan vara antingen enkla artiklar eller nischer bestående av ett simplexord och därtillhörande sammansättningar. Sammansättningarna är dessutom ordnade i grupper beroende på förledens form, t.ex.:

lucka s.

A: luck-byxor

-kapp

-karm

-karvel

-nar

-ring

-skalk

B: lucka-byx

-byxor

C: lucku-byx

-byxor

-ränna

-sträckare

Vid XML-kodningen av elementen i FO:s makrostruktur gäller det att finna gränserna mellan artiklarna, att föra ihop sammansättningarna med simplexordet och att gruppera sammansättningarna enligt förledstyperna som i ordboken markerats med stor bokstav i fetstil.

Artiklarna (för exempel se artikeln *hydare* nedan) består av en formell del där uppslagsordet, dess dialektala uttalsvarianter och grammatiska upplysningar ingår. Betydelsebeskrivningen ges genast efter den formella delen i korta artiklar. I momentindelade artiklar kommer betydelsebeskrivningarna efter respektive siffra. Siffran står i fetstil och efterföljs av en ensam högerparentes. Utbredningen ges också för de olika betydelserna och i samband med betydelserna ges de eventuella exempelfraserna.

**hydare** *hy:dar* Kn, *hy:dar(e)* NL m.3. – 1) rackare, hudavdragare. NL. – 2) man som efter säljägarnas hemkomst från f ä r d e n deltog i arbetet med att h y d a sälarna. Kn.

Förkortningarna som används vid de grammatiska upplysningarna och markörerna för den geografiska utbredningen är konventionaliserade och därför lätta att få rätt kodade vid konverteringen. Dessutom innehåller artiklarna i FO interna och externa hänvisningar, etymologiska uppgifter med angivande av ord i ett långivande språk, t.ex. finska, ryska eller amerikaengelska, och i vissa fall också angivelser av en källa.

### 3.3 XML-kodad strukturerad text

För att ett stort material ska kunna överföras till en databas för utgivning både som tryckt ordbok och som elektronisk ordbok ska materialet kodas. Numera används märkspråket XML.

Den typografiska strukturen är utgångsläget då en traditionell, tryckt ordbok vars manus tagits fram antingen manuellt eller med ett ordbehandlingsprogram ska överföras till det XML-format som är målet för processen. Nedan står ett exempel på en färdig, överförd artikel, en mycket kort sådan. Artikeln *lullog* har överförts från typografiskt formaterad text till FO:s XML-kodade text. XML-kodningens hierarki kan här ses med hjälp av de olika indragningarna. De andra elementen är inkapslingar i det överordnade elementet för artikeln, <art>.

<pre> &lt;art&gt;   &lt;uppsl&gt;lullog&lt;/uppsl&gt;   &lt;text&gt;     &lt;uttal skrift="fin"&gt;lullo(gär) &lt;/uttal&gt;     &lt;gramm&gt;adj.&lt;/gramm&gt;     gungande. vacklande, ostadig.     &lt;ort&gt;Na&lt;/ort&gt;.   &lt;/text&gt; &lt;/art&gt; </pre>	<p><b>lullog</b> <i>lullo(gär)</i> adj. gungande, vacklande, ostadig. Na.</p>
---	---

Figur 1 nedan är ett försök att sammanföra den typografiska och den lexikografiska strukturen, dvs. i detta fall ordbokens mikrostruktur samt strukturen i den XML-kodade texten till en helhetsbild. Taggarna med namn i fetstil får illustrera den lexikografiska strukturen, medan taggarna i vanlig stil representerar elementen i ett XML-kodat FO-dokument. I de två sista kolumnerna till höger står element i den typografiska strukturen vid sina respektive lexikografiska strukturelement och textsträngar som bara kan stå på bestämda ställen i strukturen.

			Typografi	Text
<Mikrostruktur>				
<uppsl>				
</uppsl>				
<homograf>				
</homograf>				
<text>				
	<uttal>		Kursiv/stor rak stil	tecken från lands- målsalfabetet
	</uttal>			
	<ort>		Rak stil	Lista
	</ort>			
	<gramm>		Rak stil	Lista
	</gramm>			
	<ort>		Rak stil	Lista
	</ort>			
	<mom>		Fet stil	tankstreck, siffra, högerpar.
	<fras>		Kursiv/stor rak stil	tecken från lands- målsalfabetet
		<rak_text>	Rak stil	
		</rak_text>		
		<fras_forkl>	Fet stil	
		</fras_forkl>		
		<ort>	Rak stil	Lista
		</ort>		
		<hanv>	Spärrad stil	
			(inkapslingar)	
		</hanv>		
	</fras>			
	</mom>			
</text>				
</Mikrostruktur>				

Figur 1 Den typografiska och lexikografiska strukturen i FO.



## 4. Planering av webbordboken

Inför planeringen av *Webb-FO*<sup>6</sup> har redaktionen haft stöd av språkteknolog- och lexikografkolleger vid Institutet för de inhemska språken. Arbetet har inletts med ett pilotprojekt där funktionerna testas på ett kortare avsnitt ur ordboken.

Kriterierna för en bra webbordbok är varierande. Pastor och Alcina (2010) har jämfört ett antal elektroniska ordböcker utgående från olika kriterier. Kriterier som de diskuterar är ordbokens databas som resurs och andra eventuella resurser som kan nås t.ex. via hyperlänkar. Vidare behandlar de sökningar och hur sökresultatet presenteras. Nedan i avsnitten 4.1 och 4.2 diskuterar vi hur *Webb-FO* kunde byggas upp utifrån dessa kriterier.

FO:s redaktion har bekantat sig med några webbordböcker och tänkt ut en rad prioriterade egenskaper då *Webb-FO* tas fram. Problem som iakttagits hos gratisordböcker på webben är att de innehåller så mycket reklam att det ibland är svårt att skilja mellan innehåll och reklam. Sökfunktionerna i det brokiga sortimentet är sällan särskilt genomtänkta eller sofistikerade.

En god förebild bland webbordböckerna är *EDD Online (English Dialect Dictionary)*. Den föreligger tills vidare bara i en betaversion som forskare kan komma åt att testa. Alla planerade funktioner är inte ännu färdiga, men de finns redan med som knappar eller kryssbara rutor i användargränssnittet. *EDD Online* har en del sådana egenskaper som kunde tillämpas för *Webb-FO*. Den är jämförbar med FO dels därför att den är en dialektordbok och innehåller liknande lexikografiska kategorier som FO, dels därför att den genomgår en motsvarande digitaliseringsprocess.

Samtidigt med utgivningen av *Webb-FO* pågår webbpublicering av den finska dialektordboken *Suomen murteiden sanakirja (SMS)*<sup>7</sup> vid Institutet för de inhemska språken. Artikelstrukturen i SMS och FO är rätt olika, så det är bara en del av arbetet med SMS som kan utnyttjas för FO. Bägge ordböckerna kommer att nås via materialbanken *Kaino* (<http://kaino.kotus.fi/>) och därför koordineras projekten. En annan förebild för FO är webbordboken över karelskan *Karjalan kielen verkkosanakirja*<sup>8</sup> med sina mångsidiga sökmöjligheter. Den publicerades redan i oktober 2009 på *Kaino*.

<sup>6</sup> Arbetsnamn för FO:s webbversion.

<sup>7</sup> Ett pilotavsnitt av SMS publicerades på webben i februari 2012.

<sup>8</sup> En recension av den karelska webbordboken finns i *LexicoNordica* 17 (Cantell 2010).

#### 4.1 Ordboksresurser

Vi har ovan i avsnitten 3.2 och 3.3 beskrivit databasen som resurs för webbordboken. Framför allt är det viktigt att FO:s databas och kodning ska utnyttjas till fullo. Om exaktare kodning behövs för någon funktion som annars inte är möjlig, är redaktionen beredd att satsa på detta.

Hänvisningarna i ordboken ska vara länkade så att man genom att klicka kommer fram till det ställe som hänvisningen avser. Dessutom kunde det skapas länkar vid hänvisningar till åtminstone två andra ordböcker som finns eller kommer att finnas tillgängliga på webben, nämligen till SAOB och SMS. I framtiden planerar Institutet för de inhemska språken att utveckla en ordboksportal och det vore fint att där kunna länka direkt mellan husets olika ordböcker. Externa länkar är sårbara, eftersom de inte nödvändigtvis är bestående; t.ex. databasen för SAOB kan i framtiden bli omplacerad.

#### 4.2 Sökningar

Webbordboken ska vara mångsidigt sökbar. Det ska vara möjligt att göra (1) enkla sökningar i lemmalistan, (2) avancerade sökningar med olika begränsningar och filter som är relevanta med hänsyn till innehållet och (3) fritextsökningar i hela databasen.

Avancerade sökningar i *Webb-FO* ska kunna begränsas till olika element i artiklarna. Förutom bland uppslagsorden – som kan vara simplexord och/eller sammansättningar – ska det vara möjligt att söka bland uttalsvarianterna och exempelfraserna och dessutom i brödtexten vars viktigaste del är betydelseförklaringarna. Betydelseförklaringarna är dock inte taggade, vilket betyder att sökningen här ger en hel del extra brus.

De avancerade sökningarna ska också kunna begränsas med hjälp av sökfiter enligt beläggort (socken) eller dialektområde, uttalsvariant (t.ex. enligt kvalitet och kvantitet), ordklass, böjningskategori (deklination och konjugation), etymologi (t.ex. finska, ryska, amerikaengelska som långgivande språk) samt enligt olika kombinationer av dessa.

*Webb-FO* ska t.ex. kunna konsulteras för följande sökningar:

- Vilka svaga maskuliner förekommer i dialekten i Tenala? (svar t.ex.: *buske* 1, *hägne*; *hyvel*)
- Vilka sammansatta adjektiv finns vars efterled slutar på *-og*? (svar t.ex.: *hårdmjölkog*, *häll-fattog*, *hämnd-girog*)
- Vilka substantiv ur gruppen femininer f.1/f.2 kan i vissa dialekter ha maskulin böjning m.1/m.2? (svar t.ex.: *al* 1, *dörr*, *jord*, *lagg* 1, *lagg* 2, *not* 1)

- Vilka homografer av ord som slutar på *-a* förekommer som både substantiv (svf.) och verb? (svar t.ex.: *balga*, *håsa*, *mara* med homografer)
- Vilka finska lån finns belagda i norra svenska Östbotten? (svar t.ex.: *kytta* s.1, *käpel* 1, *käpel* 2, *lamp* 1)

Exaktare kodning av innehållet i databasen behövs om *Webb-FO* ska kunna erbjuda sina användare möjlighet att söka synonymer eller antonymer eller s.k. indirekta etymologier, dvs. etymologiska angivelser som finns i den artikel som det hänvisas till men som också avser det ord som det hänvisas ifrån. Sökningar enligt bynamn eller önamn kan inte heller göras utan att exaktare kodning införs.

För att webbordboken ska kunna användas för sökningar som kan ge svar på frågor om t.ex. terminologi, utbredning och vissa typer av förleder eller efterleder i sammansättningar krävs likaså exaktare kodning. Sådana frågor är t.ex. följande: (a) Vilka sällfångstermer förekommer i dialekten på Bergö? (svar t.ex. *krabb* s.2, bet. 2, *kryp-stång*) eller (b) Vilka sammansättningar finns med adjektivisk förled som slutar på *-og*? (svar t.ex.: *hårog-buse*, *hårog-harpa*, *kålog-bisse*, *kåtog-nackad*).

Då artiklar som ingår i sökresultatet visas borde det vara möjligt att tekniskt komprimera i artikelfönstret för en översikt av betydelserna. Också för SMS har en komprimeringsfunktion utarbetats. Artiklarna i SMS visas antingen i sin helhet eller komprimerade så att endast dispositionen, dvs. momenten och de inledande förklaringarna, syns eller i en vy där exemplen uteblir.

## 5. Sammanfattning

Stora vetenskapliga ordböcker som är fritt tillgängliga på webben är ett steg i utvecklingen mot ett mer demokratiskt kunskapssamhälle. En webbordbok har i jämförelse med en tryckt ordbok en helt annan potential att nå nya användargrupper. Det ligger dock ett drygt arbete bakom en webbutgivning.

Då man planerar en helt ny elektronisk ordbok, kan man börja med att planera vilka sökningar som ska vara möjliga och sedan fatta beslut om databasens innehåll och kodning enligt önskemålen. Då man digitaliserar en existerande ordbok, utgår man från det innehåll som finns, tänker ut vad som kunde vara önskvärt att söka i detta innehåll och försöker sedan åstadkomma en kodning som gör alla de önskade sökningarna möjliga. Detta kan vara en stor utmaning, då ordboken inte ursprungligen planerades och byggdes upp med tanke på en webbutgivning.

Webbordboken planeras så att det blir möjligt att genom sökningar och relevanta begränsningar och sökfilter få fram det centrala i ordbokens innehåll. För FO:s del är exponeringen av variationen i de svenska dialektområdena i Finland viktig. Ordboken innehåller en mångfald av dialektala uttalsformer och belysande exempelfraser; ett digert dialektmaterial som redaktionen för FO vill göra tillgängligt för alla användare.

## LITTERATUR OCH KÄLLOR:

- Ahlbäck, Olav, 1976: Inledning. Ordbok över Finlands svenska folkmål, utg. av Forskningscentralen för de inhemska språken. Helsingfors. S. VII–XXVIII.
- Cantell, Ilse, 2010: Ordbok över karelskan på Internet. I: *LexicoNordica* 17. S. 277–288.
- FO = Ordbok över Finlands svenska folkmål, utg. av Forskningscentralen för de inhemska språken, 1976 ff. Helsingfors.
- Laureys, Godelieve, 2012: Mot en hållbar lexikografi. En balansgång mellan återanvändning och förnyelse. I: *Nordiska studier i lexikografi* 11. (Skrifter utgivna av Nordiska föreningen för lexikografi 12.) Lund
- Pastor, Verónica och Alcina, Amparo, 2010: Search Techniques in Electronic Dictionaries: a Classification for translators. I: *International Journal of Lexicography*, Vol. 23 No. 3. S. 307–354.
- Rietz, Johan Ernst, 1862–1867: *Svenskt dialektlexikon*. Ordbok öfver svenska allmogespråket. Lund.
- SAOB = Ordbok över svenska språket, utg. av Svenska Akademien, 1893 ff. Lund.
- Slotte, Peter, 1981: Ordbok över Finlands svenska folkmål – problem och metoder. I: *De finlandssvenska dialekterna i forskning och funktion*, red. Bengt Loman. Åbo. S. 45–64.
- SMS = Suomen murteiden sanakirja, utg. av Forskningscentralen för de inhemska språken, 1985 ff. Helsingfors.
- Vendell, Herman, 1904–1907: Ordbok över de östsvenska dialekterna. Helsingfors.
- Wessman, V. E. V., 1925–1932: *Samling av ord ur östsvenska folkmål*. Tillägg till H. Vendells Ordbok över de östsvenska dialekterna. Helsingfors.
- Vikør, Lars S., 1999: *Fleirgenerasjonsordbøker og tida*. I: *Nordiska studier i lexikografi* 4. Rapport från Konferensen om lexikografi i Norden Esbo 21–24 maj 1997, red. Peter Slotte et al. Helsingfors. S. 395–405.

## Internethänvisningar

- Aineistopalvelu Kaino (Materialbanken Kaino): <http://kaino.kotus.fi/> (september 2011)
- EDD Online: [http://www.uibk.ac.at/anglistik/projects/speed/startseite\\_edd\\_online.html](http://www.uibk.ac.at/anglistik/projects/speed/startseite_edd_online.html) (april 2011)
- Karjalan kielen verkkosanakirja (Webbordbok över karelskan): [http://kaino.kotus.fi/cgi-bin/kks/kks\\_etusivu.cgi](http://kaino.kotus.fi/cgi-bin/kks/kks_etusivu.cgi) (september 2011)

Suomen murteiden verkkosanakirja (Webbordbok över finska dialekter): <http://freax.kotus.fi/sms/?p=main> (april 2012)

*Ilse Cantell*

Ordbok över Finlands svenska folkmål  
ilse.cantell@kotus.fi

*Caroline Sandström*

Ordbok över Finlands svenska folkmål  
caroline.sandstrom@kotus.fi