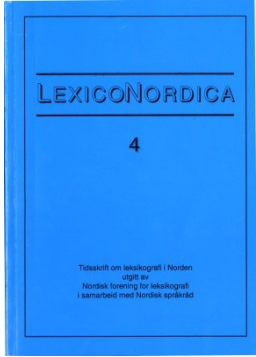


LexicoNordica

Titel:	En ny typ av produkt – en recensent med gammal typ av ammunition. Svar till Aino Kärnä	
Forfatter:	Risto Haarala og Marja Lehtinen	
Kilde:	LexicoNordica 4, 1997, s. 205-211	
URL:	http://ojs.statsbiblioteket.dk/index.php/lexn/issue/archive	

© LexicoNordica og forfatterne

Betingelser for brug af denne artikel

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

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

Søgbarhed

Artiklerne i de ældre LexicoNordica (1-16) er skannet og OCR-behandlet. OCR står for 'optical character recognition' og kan ved tegngenkendelse konvertere et billede til tekst. Dermed kan man søge i teksten. Imidlertid kan der opstå fejl i tegngenkendelsen, og når man søger på fx navne, skal man være forberedt på at søgningen ikke er 100 % pålidelig.

Risto Haarala & Marja Lehtinen

En ny typ av produkt – en recensent med gammal typ av ammunition

Svar till Aino Kärnä

Som huvudredaktörer för CD-Perussanakirja och Suomen kielen perussanakirja vill vi komma med några kommentarer till Aino Kärnäs recension av CD-Perussanakirja vars manuskript vi har haft tillfälle att läsa.

Recensionen följer i stor utsträckning ett gammalt känt mönster: ord som saknas i ordboken räknas upp och definitioner som enligt recensenten är felaktiga påtalas. Vi tänker inte här ta ställning till om ordet *avajaisrieha* '(livad) invigningsfest' borde vara med eller inte eller om definitionen av *kateviljely* 'täckodling' är adekvat nu och i framtiden, utan vi nöjer oss med att ta upp några frågor som är betydelsefulla för lexikografisk praxis och för denna nya typ av ordbok.

Suomen kielen perussanakirja ('Finsk basordbok') I–III, 1990–1994 (se Krista Varantolas recension i *LexicoNordica* 1/1994), publicerades i elektronisk version i februari 1997 under namnet *CD-Perussanakirja*. Innehållet daterades upp i någon mån: ca 2.000 nya ordartiklar lades in och ca 3.000 artiklar modifierades eller kompletterades. De återstående omkring 95.000 artiklarna är oförändrade till sitt innehåll. Produktens nyhetsvärde ligger alltså främst i formen, inte så mycket i innehållet.

Aino Kärnä verkar att i hög grad ha närmat sig den elektroniska ordboken ur "pappersperspektiv", som om hon hade bedömt en tryckt ordbok. På annat sätt kan vi inte förstå vissa av hennes kommentarer om produktens egenskaper. Hon har tydligen missat flera av sökmöjligheterna och hon verkar inte att ha brytt sig om att läsa den tryckta bruksanvisningen eller guiden i själva programmet. Att en vanlig användare inte gör det kan man överse med – en elektronisk ordbok skall gå att använda, åtminstone någorlunda, utan bruksanvisning. Däremot väntar man sig att en recensent skall sätta sig ordentligt in i programmets olika delfunktioner. En viktig infallsvinkel saknas dessutom i recensionen – en relatering till motsvarande produkter ute i världen. Det borde också ha varit tämligen självklart för en recensent att man, när man gör en elektronisk version av en ordbok eller ett uppslagsverk, måste lägga stor vikt vid tekniska lösningar,

sökmöjligheter, snabbhet osv., annars är produkten tämligen oduglig, hur enastående själva innehållet än är.

Mot denna bakgrund verkar Kärnäs vårdslösa påstående om att en bra cd-romprodukt inte åstadkoms genom att innehållet överförs från papper till laser (med vilket hon förmodligen syftar på CD-Perussanakirja) tyda på en fullständig okunnighet om vilka svårigheterna är när det gäller att utveckla en elektronisk produkt. Att datera upp innehållet, att skriva nya artiklar och att bearbeta gamla, har varit bara en liten bråkdel av allt det arbete som har lagts ner på cd-romversionen. Största delen av tiden har gått åt till att fila på den logiska strukturen hos ordbokstexten. Att konvertera filerna till SGML-standard krävde mycket mer än en knapptryckning; noggrann programmering och manuell granskning av en mängd olika saker. Det har gjorts ett stort antal ändringar i själva framställningssättet för att artikelinnehållet skall kunna utnyttjas optimalt i den elektroniska produkten. Uppslagsorden skrivs t.ex. ut i exempelfraserna i CD-Perussanakirja och görs inte som kryptiska förkortningar som i den tryckta boken. Till det vanskligaste i utvecklingsarbetet har hört att ta fram den funktion som ger böjningsparadigmen för samtliga uppslagsord och att ta hänsyn till böjningen i ordbokstexten. Mer om det längre fram.

Ordartikel – boksida

Att det syns bara en artikel i taget i artikelfönstret tycker recensenten inte är någon bra lösning utan saknar "boksida" där omgivande artiklar ger stöd. En sådan önskan är förståelig. Ofta skulle alfabetiskt närstående ordartiklar ge ett visst stöd för den artikel som för tillfället finns uppslagen i fönstret. Men inte heller en tryckt ordbok är ju fullständig härvidlag. På grund av den finska stadieväxlingen hamnar t.ex. verbet *kaataa* 'hälla; fälla m.m.' och dess frekventativa avledning *kaadella* långt från varandra och kan inte kontrolleras på samma sida även om man hade önskat det. Är det på något vis naturligare att bläddra i en tryckt ordbok än att ty sig till "extra mushantering", som Kärnä uttrycker det? Eller borde användaren hellre vänja sig vid att använda ett nytt medium på ett nytt sätt?

Nämnas bör också att artikelfönstret innehåller två knappar med vars hjälp man kan bläddra i texten både i alfabetisk ordning och i träffordning. Dessutom finns det ett särskilt fönster för bläddring framåt och bakåt i alfabetet till de omgivande ordboksartiklarna. Detta fönster tycks ha undgått Kärnä, eller så har hon inte ansett det vara värt att nämna. Lösningen att bara en artikel åt gången visas anser vi vara väl

motiverad. När vi utvecklade prototypen testade vi nämligen en version som bara visade "hela boksidor" och det var den egenskap som fick överlägset mest negativ kritik från användarna. Det är frustrerande om en sökning resulterar i 180 träffar och användaren blir tvungen att bläddra sig genom "boksidor" för att komma åt träffarna när han i stället kan få besked om vilka ordartiklar som innehåller träffar. Vi har inte sett en enda elektronisk ordbok eller uppslagsbok som skulle visa "hela sidor".

Böjningen

Kärnä påstår på två olika ställen i sin recension att de koder som anger uppslagsordens böjning inte skulle förklaras någonstans i CD-Perussanakirja, och att de därför blir helt obegripliga för användaren. Sifferindexen finns dock förklarade i den tryckta bruksanvisningen på sidorna 17 och 18 och dessutom finns modellords- och stadieväxlings-tabellerna som vanliga hjälppiler i Windows.

Nytt i CD-Perussanakirja är att användaren om han så önskar kan be om böjningen för vilket uppslagsord som helst och alltså inte behöver sluta sig till formerna med hjälp av modellord. Böjningsparadigmen är det ingen som har suttit och skrivit in, utan programmet producerar böjningsformerna enligt olika matematiska regler. Med hjälp av maskeringstecken kan användaren få böjningen för flera ord på en gång: en sökning efter böjningen för "*ismi" ger böjningen för alla ord som slutar på *ismi* och är av främmande ursprung. Av paradigmen framgår det direkt om ändelsevokalen skall vara *ä* eller *a*. Valet mellan främre och bakre vokal är inte alltid självklart ens för finskspråkiga.

Kärnä efterlyser också uppgifter om huruvida ett ord som är oböjligt. Den uppgiften ges faktiskt i CD-Perussanakirja: om användaren ber om böjningsparadigmet för ett ord som saknar böjning ger programmet upplysningen att ordet i fråga är oböjligt. Också problematisk adjektivkomparation ger CD-Perussanakirja upplysningar om (se t.ex. uppslagsord som *pitkä* 'lång', *lyhyt* 'kort' och *kiva* 'trevlig; kul'.)

Texten i CD-Perussanakirja är analyserad i sin helhet med ett morfologiskt program, vilket gör det möjligt att söka böjda former genom att skriva in ordets grundform som söksträng. Sådana sökningar är exaktare än sökningar med hjälp av maskeringstecken, vilka i och för sig också är möjliga. För analysen används ett modernt språkteknologiskt program som fått erkännande också utanför Finlands gränser. Detta har inte tillfredsställt Kärnä, utan hon hade föredragit en

annan lösning. Med tanke på utländska användare borde enligt hennes åsikt en sökning på verbet *kesiä* 'fjälla; flaga av' ge också partitiv pluralis *kesiä* av ordet *kesä* 'sommar', eftersom formerna är homonyma. Anser Kärnä då också att en sökning på olika former av *eräs* 'en viss' också borde ta med även illativformen *erään* av *erä* 'parti; omgång m.m.' och att en sökning på t.ex. verbet *purra* 'bita' borde leda till formen *purkaa* också i dess funktion som infinitiv och presens indikativ av *purkaa* 'riva' förutom som imperativ pluralis av *purra*? Det finns, försiktigtvis uppskattat, tiotusentals sådana här ambigüosa formpar i ordbokstexten, och den morfologiska analysen har gjorts uttryckligen för att komma ifrån allt sådant "skräp" i sökresultaten. Utöver det har vi naturligtvis velat behålla möjligheten att söka en bestämd söksträng som sådan: det gäller bara att kunna skriva in de former man söker i sökrutan (också andra än grundformen!), så sköter datorn om resten. Då ingår självfallet homonyma former i sökresultatet, som både nybörjare och längre hunna i så fall måste kunna tolka rätt.

De olika sökmöjligheter CD-Perussanakirja tillhandahåller presenterar Kärnä mycket bristfälligt. Vi nöjer oss med att nämna några. Det finns möjlighet till kombinerade sökningar med hjälp av de logiska operatorerna JA och TAI (AND och OR). Enkla instruktioner – teknisk terminologi har undvikits – om hur dessa används finns både i den tryckta bruksanvisningen och i bruksanvisningsfilen med tanke på användare som eventuellt inte har stött på dem, trots att operatorerna ju är vanliga både i elektroniska uppslagsverk och i olika sökprogram på Internet. Hade Aino Kärnä dragit nytta av denna möjlighet skulle hon ha hittat de ordartiklar som innehåller *tuossa tuokiossa* 'på ett ögonblick' i en enda sökning. Genom att bläddra igenom ett sådant sökresultat hittar användaren troligtvis det han söker betydligt snabbare än genom att leta efter motsvarande information i en tryckt ordbok. På motsvarande sätt skulle söksträngen "**lainen,*läinen*" ha gett samma resultat i en enda sökning som Kärnä behövde två separata sökningar för. Det gäller bara att dra nytta av produktens egenskaper till fullo.

Synonymerna

Det påstås ingenstans i CD-Perussanakirjas bruksanvisningar att synonymer till uppslagsorden skulle vara systematiskt angivna, och den kan sålunda inte betraktas som en synonymordbok. Redaktionen har planer på en systematiskt genomförd synonymangivelse, men det är ingenting som går att göra i en handvändning och minst av allt maskinellt. Enligt vårt förmenande kan dock CD-Perussanakirja i viss

mån utnyttjas när man behöver leta efter synonymer. Resultatet av sådana sökningar är ordartiklar som på något vis hör ihop och där man nog kan utläsa intressant information med relativt liten möda.

Recensionen kan förleda läsaren att tro att *aamu* 'morgon', *aamutui-maan* 'genast på morgonen', *alkaen* 'från och med' osv. enligt ordboken skulle vara synonymer till ordet *heti* 'genast', men så är det självfallet inte, utan orden är helt enkelt en lista på de ordartiklar som innehåller ordet *heti*. Enligt Kärnä är det alltför tidskrävande att bläddra igenom sökresultatet av en sökning efter "heti" i hela texten för att få fram det man är ute efter. Vi är av annan åsikt. Programmet hittar de 183 förekomsterna av *heti* på ett par sekunder, och om man utnyttjar träffönstrets KWIC-funktion går det relativt lätt att ögna igenom beläggsraderna. Det tar knappast mer än ett par minuter att välja ut det som intresserar en och man behöver inte läsa igenom hela artiklar. Det verkar dock som om Kärnä hade missat också KWIC-funktionen. En sökning på "heti" begränsad till metaspråket ger t.ex. följande intressanta resultat: "*vastaus tuli kuin apteekin hyllyltä* ('svaret kom på en gång') *vaikeudetta* ('utan svårighet'), *muuta mutkitta* ('utan vidare'), *heti*", "*Tyrmäsi ajatuksen suoralta kädeltä* ('[han] avfärdade tanken direkt') *heti*, *ilman muuta* ('utan vidare'), *oikopäätä* ('på en gång')", "*paikalla heti*, *kohta*, *oitis*, *viipymättä* ('utan dröjsmål')", "*Lähti siinä samassa* ('[han] gick med det samma') *heti*, *kohta*, *oitis*", "*Seisovalta jalalta* ('på stående fot') *heti*, *asiaa tarkemmin harkitseematta t. tarkistamatta* ('utan att fundera närmare'), *suoraan*, *suoraa päätä* ('direkt'), *suoralta kädeltä* ('på rak arm')". Om användaren kombinerar denna information med de synonymer som ges till själva uppslagsordet *heti* har han fått ut mycket mer information av den elektroniska produkten än han hade kunnat få ut av den tryckta boken och på mycket kortare tid.

Lemmaselektionen och "konsekvens"

Enligt Kärnä är det bra att det finns fler kausativa verb som uppslagsord i Perussanakirja än i Nykysuomen sanakirja ('Nufinsk ordbok'), men hon efterlyser ännu fler och räknar upp ett antal avledningar i denna grupp som inte finns med i ordboken. Enligt den tillspetsade slutkommentaren "skapar [detta] ett intryck av att man vid lemmaselektionen inte har använt sig av de medel (ibland tämligen mekaniska sådana) som finns för att täcka hela ordfält, utan vilka ord som kommit med, även sådana som utgör delar av typiska ordfält, har varit beroende av slumpen". Konstaterandet visar att skribenten inte är

insatt i vilken utsträckning det är möjligt och meningsfullt att beskriva finskans oerhört rika avledningsmaterial i en ordbok. Det går exempelvis att bilda nomen agentis-avledningar till i stort sett varje verb med hjälp av suffixet *ja*, *jä*, t.ex. *kaatua* 'falla (omkull)' + *ja* = *kaatuja* 'person som faller (omkull)'. På motsvarande sätt går det att bilda nomen actionis-avledningar till så gott som alla verb med hjälp av suffixet *minen*. Av dessa är det dock bara en liten del som tas med som uppslagsord i finska ordböcker, eftersom de flesta är mekaniska bildningar och inte egentligen skulle tillföra ordboken någon ny innehållslig information (avledningar som avviker till sin betydelse är naturligtvis en sak för sig). Påpekas bör också, att den typ av kausativbildningar som Kärnä efterlyser, sin potentiella produktivitet till trots, är väldigt sällsynta i texter. I en finsk textkorpus på ca 17 miljoner löpord finns det bara sex belägg: *villiinnyttää* 'göra vild' (som finns med i Perussanakirja), fyra förekomster, *pidennyttää* 'låta förlänga', en förekomst, och *suurennuttaa* 'låta förstora', en förekomst. Det som behövs när det gäller att bedöma vilka avledningar som skall tas med är framför allt sunt lexikografförnuft.

Priset

Vad gäller priset konstaterar Kärnä att CD-Perussanakirja kostar ungefär hälften så mycket som den tryckta versionen (de exakta siffrorna är 1.680 vs. 950 mark) men att man för att använda programmet behöver utrustning för minst 10.000 mark. På basis av de registreringskort som har skickats in till oss verkar det som om folk i Finland numera skulle ha synnerligen effektiva datorer, mycket effektivare och dyrare än vad som behövs för CD-Perussanakirja. Folk förnyar med andra ord sin datorutrustning hela tiden på grund av olika program som kräver kapacitet; det vore naivt att tro att någon skulle skaffa sig en 10.000 marks datorutrustning uteslutande för CD-Perussanakirjas skull. Enligt Kärnäs logik borde priset på datorn alltså slås på priset för ett ordbehandlingsprogram eller för ett spel (och i det senare fallet räcker 10.000 mark knappast!). Med samma logik borde priset på bokhyllan också slås på priset för en tryckt ordbok.

Litteratur

Nykysuomen sanakirja (I–VI) ('Nufinsk ordbok') 1951–61. Porvoo: WSOY.

- Suomen kielen perussanakirja*. (I–III) ('Finsk basordbok') 1990–94.
Päätoimittaja ('huvudredaktör'): Risto Haarala. Helsinki: Painatus-
keskus Oy.
- Varantola, Krista 1994: Suomen kielen perussanakirja – Finsk basord-
bok. Rec. i *LexicoNordica* 1, 309–317.