

Natur och digitala, tvåspråkiga ordböcker: En studie av finska som sökspråk i ISLEX

Helga Hilmisdóttir & Marjakaisa Matthíasson

This study deals with the Finnish version of the Icelandic-Scandinavian dictionary ISLEX. The source language of the dictionary is Icelandic and the lemmas that are included represent modern Icelandic language use. The Finnish language has a subordinate status in ISLEX and Finnish words are only entered into the dictionary if they can be used as equivalents for the Icelandic lemmas. However, this is complicated by the fact that ISLEX makes it possible to use both Icelandic and Finnish as search languages. This leads us to the question how representative the Finnish equivalents are for Finnish language use and whether the vocabulary is large enough to decode Finnish texts. This study focuses on words that have to do with nature: plants and animals, the weather and geological and geographical phenomena. The aim is to discuss to what degree Icelandic nature affects the lemma selection in ISLEX and what consequences that has for the user when Finnish is the search language.

1. Inledning

Den isländsk-skandinaviska ordboken ISLEX är en digital ordbok där isländska fungerar som källspråk och danska, finska, färiska, norska (bokmål och nynorska) och svenska fungerar som målspråk. Med andra ord har ordboken utarbetats utifrån det isländska språket och isländska förhållanden, och de skandinaviska språken har en underordnad status från redaktörernas sida (jfr Hannesdóttir 2015:240 om *main language* och *subordinate language*). Digitaliseringen av tvåspråkiga ordböcker har dock medfört nya utmaningar för lexikografer, och den nya tekniken har lett till att etablerade termer som *källspråk* och *målspråk* kan ifrågasättas (Hannesdóttir 2015:236). Digitaliseringen av ordböcker tillåter sökning i både käll- och målspråk, och en islänning kan t.ex. söka finska ord i ISLEX för att avkoda en text på finska. Ur användarens synvinkel fungerar finska då som *sökspråk* och isländska som ett *målspråk* (jfr Hannesdóttir 2015).

På grund av de olika sökmöjligheterna kan man säga att den digitala ordboken har ett bredare syfte än en tryckt ordbok. Ur användarens synvinkel fungerar nämligen den finska versionen av ISLEX både som isländsk-finsk och finsk-isländsk ordbok (jfr Hannesdóttir 2015:240). Bild 1 visar uppslagsordet *vogur* 'vik' när man använder sökspråket isländska och bild 2 för *lahti* om man väljer finska.

Bild 1: Uppslagsordet *vogur*.Bild 2: Det finska sökordet *lahti*.

När användarna söker isländska ord får de fram en artikel som ger information om ordets uttal, böjning och finska ekvivalenter. Som bild 1 visar har uppslagsordet *vogur* två finska ekvivalenter, *lahti* och *poukama*. Om de däremot söker på det finska ordet *lahti* får de fram en lista över alla isländska uppslagsord där sökordet anges som ekvivalent. Ordet förekommer under tre uppslagsord: *flói*, *vík* och *vogur*. Användarna kan sedan klicka sig fram till varje enskilt uppslagsord för att få fram mer information om eventuellt andra ekvivalenter och närmare förklaringar.

Den digitala utvecklingen ställer nya krav på det lexikografiska arbetet. Hanesdóttir (2012; 2015) har påpekat att ordböckernas inre struktur i framtiden kommer att ändras; i digitala ordböcker kommer båda språken att vara jämlika och ge lika bra sökresultat. Men även om ISLEX erbjuder sökning på både isländska och finska har isländskan en tydlig överordnad funktion. Uppslagsorden är valda utifrån isländska förhållanden och beskriver isländsk samhällsordning, kultur, traditioner och natur. Med tanke på detta kan man fråga sig i vilken utsträckning ISLEX fungerar när man använder finska som sökspråk. Ordboken omfattar två helt obesläktade språk som talas i länder som geografiskt sett är ganska olika varandra.

Syftet med denna studie är att undersöka hur bra det finska språket fungerar i ISLEX som sökspråk för naturord, dvs ord som har att göra med biologi, geologi, geografi och meteorologi. I artikeln kommer vi att dryfta följande tre frågor: 1) Finns det några allmänna tendenser angående valet av isländska uppslagsord? 2) Vilka problem kan uppstå när man söker ord som finns både i allmänspråket och som vetenskaplig terminologi? och 3) Hur ska man välja ekvivalenter för ord som har många synonymer, som t.ex. ord för 'storm'? För att svara på dessa frågor

jämför vi ISLEX med *Stora finsk-svenska ordboken* (2004) och gör två nedslag i ISLEX: det isländska uppslagsordet *refur* 'räv' och det finska sökordet *myrsky* 'storm'.

2. ISLEX

ISLEX är en isländsk-skandinavisk ordbok som har funnits på nätet sedan november 2011. Projektet är ett samarbete mellan Árni Magnússon institutet för isländska studier och universitet och forskningsinstitutioner i Norden. Arbetet med den finska versionen av ISLEX inleddes i januari 2013 och i januari 2018 hade 75 % av de isländska uppslagsorden översatts till finska (se Hilmisdóttir & Martola 2015 för en introduktion av den finska ordboken).

Det isländska materialet består av ca 50 000 uppslagsord, 30 000 språkprov och 14 000 fasta och halvfasta fraser (Úlfarsdóttir 2013). Uppslagsorden i ISLEX representerar modernt isländskt språkbruk men omfattar även en del vanlig terminologi och poetiska uttryck.

För redaktören är uppslagsorden i ISLEX grupperade enligt betydelsefält som t.ex. GEOGRAFI, GEOLOGI, DJUR och VÄXTER eller mer specifika grupper som VIND eller OVÄDER. En och samma ord kan ingå i flera olika betydelsefält som t.ex. *rok* 'storm' som ingår i grupperna VÄDER, VIND och VINDHASTIGHET. Som redaktör kan man ta fram dessa listor vilket ger intressanta möjligheter till jämförelser av uppslagsord i olika ordböcker.

3. Naturord i ISLEX

Om valet av uppslagsord i ISLEX kan man säga att det har en tydlig isländsk prägel och att redaktörerna har vinnlagt sig om att inkludera uppslagsord som beskriver landets natur. I detta kapitel diskuteras först valet av isländska uppslagsord i ISLEX och vilka betydelsefält man lagt särskild vikt vid (avsnitt 3.1.). Sedan diskuteras vi problem och utmaningar angående vetenskapliga och folkliga namn på djur och växter (3.2.) och avslutar med en diskussion om vind och oväder och möjligheten att överhuvudtaget hitta tillräckligt nyanserade ekvivalenter mellan två språk (3.3.).

3.1. Allmänna iakttagelser angående valet av isländska uppslagsord

I ISLEX har man inkluderat djur och växter som finns i isländsk natur. Många av dessa arter har spridning i hela Norden såsom fågelarterna *hettumávur/naurulokki* 'skrattmåsa' och *tjaldur/meriharakka* 'strandkata', och växter som *gleymmérei/peltolemmikki* 'åkerförgätmigej' och *sóley/niityleinikki* 'smörblomma'. Man kan

även hitta en del däggdjur som finns i Skandinavien men inte på Island såsom *ikorni/orava* 'ekorre', *elgur/hirvi* 'älg' och *hreysiköttur/kärppä* 'hermelin'.

Emellertid har ordboken en del luckor vad gäller fisk och skaldjur som inte finns på Island. Som exempel kan vi nämna uppslagsord som *silakka* 'strömning', *kuha* 'gös', *muikku* 'siklöja', *siika* 'sik', *rapu* 'kräfta' och *made* 'lake' som alla fångas i Finland och ingår i det finländska köket. Här vill vi dock påpeka att inte ens Nordiska rådets *Fiskordbog* anger isländska namn på dessa arter, så det är möjligt att det inte finns några etablerade ekvivalenter alls.¹ Vanliga finska bärsorter fattas också i ISLEX såsom *lakka* 'hjordron', *tyrni* 'havtorn' och *ruusunmarja* 'nypon'. Om det är redaktörernas mål att finska ska fungera som sökspråk för islänningar som vill förstå finska borde även dessa ord finnas i ordboken.

En tydlig tendens i ISLEX är att redaktörerna varit ambitiösa med att inkludera geologisk terminologi. En söking på GEOLOGI ger 333 isländska uppslagsord. Endast ett fåtal av dessa ord återfinns i *Stora finsk-svenska ordboken* (2004): t.ex. *jäätikkö* 'glaciär', *tulivuorenpurkaus* 'vulkanutbrott'. I ISLEX ingår däremot en rad geologiska termer som är viktiga i ett land som Island. Där kan man nämna ord som *öskulag/tuhkakerros* 'asklager', *dyngja/kilpitulivuori* 'sköldvulkan' och *skriðjökull/vuorenjuurijäätikkö* 'utloppsglaciär'. Andra betydelsefält som bedömts som viktiga i ISLEX är kategorierna VÄDER (390 ord), VIND (113 ord) och SNÖ (58 ord). En mera systematisk jämförelse mellan ISLEX och *Stora finsk-svenska ordboken* (2004) skulle sannolikt visa att det finns vissa betydelsefält som täckts in med mera ambition i den finska ordboken. Man kan t.ex. tänka sig att ord som har med skogen att göra har en större betydelse i en finsk kontext. I ISLEX är kategorin SKOG mycket liten och består av endast 12 ord.

3.2. Isländska och finska rävar

Som tidigare nämnts har redaktörerna för ISLEX strävat efter att inkludera de djur och växter som finns på Island. När det gäller fåglar, fiskar och växter har man tagit med både specifika arter som t.ex. *týsfjóra* 'ängsviol' och den mer allmänna beteckningen *fjóra* 'viol' som omfattar alla blommor av violfamiljen. Uppslagsorden är ofta försedda med bildmaterial, och efter artnamnen kommer den latinska termen inom parentes.

Eftersom ISLEX bygger på isländska förhållanden kan vissa djurbeteckningar tolkas på olika sätt beroende på om användaren är från Island eller Finland. Det bästa exemplet är uppslagsordet *refur* 'räv' som är ett ord i allmänspråket som kan syfta på flera arter. När islänningar använder ordet *refur* syftar de oftast på polarräv eftersom andra arter av rävsläktet inte lever vilt i isländsk natur. På finska heter polarräven (eller fjällräven) däremot *naali* och man gör en tydlig skillnad mellan *naali* 'polarräv' och *kettu* 'räv el. rödräv' som båda lever vilt i Finland. Eftersom användare på Island och Finland sannolikt har olika uppfattningar om

¹ Se <<http://docs.norden.org/faktaof/Tabel-med-latin.htm>> (hämtat i juli 2017).

hur en prototypisk räv ser ut har redaktörerna valt att inte visa bildmaterial för uppslagsordet *refur*.

3.3. Oväder

Isländskan har ett stort ordförråd för att beskriva väder och det kan vara svårt att beskriva alla betydelsenyanser. Dels syftar väderorden ofta på abstrakta företeelser som olika personer kan uppfatta på olika sätt, dels kan det finnas geografisk variation när det gäller vilka ord man använder och hur man definierar dem. Definitionen av väderord blir också mer komplicerad i och med att vissa ord kan användas både som meteorologiska termer och som vardagligt språk med ibland mindre precision. Ett exempel på detta är orden som används för att beskriva vindhastighet. Ordet *rok* är t.ex. mycket frekvent i allmännspråket och syftar på det som vi upplever som stark eller mycket stark vind. Som meteorologisk term används *rok* däremot för en vindstyrka som motsvarar 10 på Beaufortskalan, som betyder att *rok* är lite starkare än *stormur* (9 på Beaufort) och svagare än *ofsaveður* (11 på Beaufort). Därför kan man säga att det vardagliga ordet *rok* har en mycket snävare betydelse inom meteorologin. På grund av detta har man i ISLEX delat upp uppslagsord som *rok* i två olika betydelser: den allmänna och den terminologiska. Tabell 1 visar de meteorologiska termerna på isländska och finska med svensk översättning (se finska termer och tabell i Similä 1981:44–45).

Beaufort skala	Isländsk term	Finsk term och svensk översättning
0	logn	tyyni 'lugnt'
1	andvari	hyvin heikko tuuli 'mycket svag vind'
2	kul	heikko tuuli 'svag vind'
3	gola	kohtalainen tuuli 'måttlig vind'
4	stinningsgola	
5	kaldi	navakka tuuli 'frisk vind'
6	stinningskaldi	
7	allhvasst	kova tuuli 'hård vind'
8	hvassviðri	
9	stormur	myrsky 'storm'
10	rok	
11	ofsaveður	
12	färviðri	hirmumyrsky 'orkan'

Tabell 1: Beaufortskalan på isländska och finska.

Som framgår av tabell 1 så har den finska skalan färre vindord än den isländska. Medan den isländska skalan använder 12 olika substantiv och ett adjektiv/adverb för att precisera vindstyrka finns det endast 4 olika ord på finska: *tyyni* 'lugnt', *tuuli*

'vind', *myrsky* 'storm' och sammansättningen *hirmumyrsky* 'orkan'. Notera t.ex. att den korrekta termen för vindstyrka 9, 10 och 11 är *myrsky* på finska medan motsvarande vindstyrka beskrivs med tre olika termer i isländskan: *stormur* (9), *rok* (10) och *ofsaveður* (11). Det finska sökordet *myrsky* ger användaren således tre isländska (terminologiska) alternativ att välja mellan.

När de finska redaktörerna sökte finska ekvivalenter till isländska väderord valde de att fokusera mer på tydlighet än på att få fram en stor variation i ordförrådet. Detta återspeglas i de finska ekvivalenterna i ISLEX. När man söker på **tuuli* och **myrsky* får man fram 61 respektive 35 isländska uppslagsord. Tabell 2 visar sökresultaten för det finska sökordet *myrsky* där sökordet förekommer utan preciserande adjektiv eller förled (sökning utan asterisk). De isländska definitionerna är hämtade från den enspråkiga nätdataboken *Íslensk nútímamálsorðabók* (=ÍNO) som bygger på samma databas som ISLEX.

Ekvivalent	Definition i ÍNO översatt till svenska
aftakarok	mycket stark vind
aftakaveður	väldigt stark vind
foráttuveður	stark vind och nederbörd, i sht hög vindhastighet
hávaðarok	mycket stark vind
hryssingur	storm med snö
ofsarok	väldigt stark vind
ofsaveður	väldigt stark vind
ofsi	svår storm
ofviðri	svår storm som orsakar skada på byggnadsverk
óveður	stark vind med hög vindhastighet, ibl. nederbörd
rok	stark vind
rosi	stark vind med regn, snöblandat regn eller snö
slagveður	storm och mycket regn
slagviðri	storm och mycket regn
stormur	mycket stark vind
veðurhamur	dåligt väder med hög vindhastighet
veðrahamur	dåligt väder med hög vindhastighet

Tabell 2: Uppslagsord som motsvarar det finska sökordet *myrsky* med definitioner hämtade från ÍNO i svensk översättning.

När man talar om storm och oväder påverkas valet av ord av många faktorer: Medför stormen nederbörd av något slag? Vilken årstid rör det sig om? Hur farligt är det att vara ute, hur påverkar stormen miljön och hur plötsligt börjar den?

Även om uppslagsorden i tabell 2 semantiskt sett är ganska lika finns det fina skillnader mellan dem. Man talar t.ex. om *slagviðri* om det både blåser och regnar, medan *rok* enbart syftar på vindhastigheten och *óveður* på oväder i allmänhet. Ordet *hávaðarok* leder däremot tankarna till vår hörupplevelse av mycket stark vind och *ofsi* betyder något extremt eller häftigt. En del ord används mycket frekvent

i isländskt tal- och skriftspråk, i synnerhet *rok* 'stark vind' och *stormur* 'mycket stark vind'. Andra ord som t.ex. *veðurhamur* och *veðrahamur* är mindre frekventa i modern isländska och visar mindre syntaktisk variation. En jämförelse av förekomsten av vindord i den isländska tidningsdatabasen *Tímarit.is* visar att *stormur* har 23 220 belägg medan *veðurhamur* har endast 107 belägg.

I nuläget ordnas listor på uppslagsord som innehåller ett finskt sökord i alfabetisk ordning. I de flesta fall spelar detta inte någon större roll eftersom sökorden oftast ger endast ett eller två uppslagord. När det gäller sökord som *myrsky* 'storm' eller *tuuli* 'vind' blir det dock mer komplicerat. För isländska användare spelar antalet ord inte någon större roll eftersom de snabbt kan gå igenom listorna och hitta de ord som passar kontexten, men för finska användare kan det här vara betydligt svårare. Idealt skulle man kunna programmera ordboken på så sätt att uppslagsorden rangordnas enligt frekvenslistor och därmed göra valet enklare för finskspråkiga användare.

4. Avslutning och diskussion

I denna artikel har vi diskuterat naturord i ISLEX för att ta fasta på vilka problem som kan uppstå när man har en digital ordbok som har två eller flera sökspråk. ISLEX har en stor målgrupp och utgångspunkten är att den ska fungera för både isländsk- och finsktalande användare. I analysen kom vi fram till att ISLEX har en tydlig tonvikt på isländska naturförhållanden. Det finns ovanligt många geologiska termer i ISLEX om man jämför ordboken med *Stora finsk-svenska ordboken* (2004). Även isländska djur och växter är mycket väl representerade. På så sätt kan man säga att ISLEX passar bra för finska användare som vill avkoda isländska texter. Om islänningar vill använda ISLEX för att avkoda finska texter är det risk att de stöter på vissa luckor i ordförrådet.

Som Hannesdóttir (2015) har påpekat har digitaliseringen av ordböcker lett till att användarna kan söka på målspråket. Om tanken är att tillmötesgå dessa nya användare och förbättra sökresultaten för sökspråket finska finns det en del saker som kunde förbättras. Man kunde t.ex. fokusera på ord som har med skogen att göra och tillägga en del fiskar, svampar, växter, bär och djur som har en central roll i finländsk kultur (både terminologi och ord ur allmänspråket). Sådana tillägg är enkla att göra och kräver inte mycket planering. Om man å andra sidan vill tillägga ord för abstrakta företeelser som väderförhållanden skulle det kräva en mer systematisk genomgång av det finska ordförrådet. Finska språket har precis som isländska ett stort och nyanserat ordförråd som beskriver väderförhållanden men eftersom ISLEX har utarbetats från källspråket isländska erbjuder ordboken få väderord på finska: De finska redaktörerna ansåg att det var viktigare att hitta tydliga och beskrivande ekvivalenter än att registrera alla nyanser inom vissa finska betydelsefält.

Slutligen kan man fråga sig om det går att redigera en existerande ordbok på så

sätt att man får lika bra sökresultat på det överordnade och underordnade språket. Hur får man t.ex. in finska sökord som inte har en ekvivalent på isländska? Vår studie visar att trots att man till viss mån kan använda finska som sökspråk i ISLEX kan ordboken inte ersätta en finsk-isländsk ordbok. För att få en webbordbok som ger lika bra sökresultat på båda språken är det viktigt att planera ordboken utifrån båda språken från början.

Källförteckning

Ordböcker

ISLEX. <www.islex.is> (augusti 2017).

Íslensk nútímamálsorðabók. <islenskordabok.arnastofnun.is> (januari 2018, preliminär version).

Järvelä, Tuomas (2008): *Suomi-islanti-suomi sanakirja*. Helsinki: Gaudeamus.

Stora finsk-svenska ordboken (2004). Helsingfors: WSOY.

Övrig litteratur

Hannedóttir, Anna Helga (2012): Ekvivalensrelationer i tvåspråkig lexikografi. I: *Lexico Nordica* 19, 39–58.

Hannedóttir, Anna Helga (2015): What is a Target Language in an Electronic Dictionary? I: *Electronic Lexicography in the 21st century: Linking Lexical Data in the Digital Age*. Ljubljana/Brighton: Trojina, Institute for Applied Slovene Studies/Lexical Computing Ltd, 236–249.

Hilmisdóttir, Helga & Nina Martola (2015): Flerspråkig lexikografi och obesläktade språk – erfarenheter från finska Islex. I: *Nordiske studier i leksikografi 13. Rapport fra 13. konference om leksikografi i Norden København 19.–22. maj 2015*, 143–154

Similä, Artturi & Toivo Vuorela (1981): *Artturi Similän sääkirja*. Porvoo: WSOY. *Timarit.is*. <timarit.is> (augusti 2017).

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

Úlfarsdóttir, Þórdís (2014): ISLEX – a Multilingual Web Dictionary. I: Calzolari et al (red.): *Proceedings of LREC 2014, Ninth International Conference on Language Resources and Evaluation*. <<http://www.lrec-conf.org/proceedings/lrec2014/index.html>> (augusti 2017).

Helga Hilmisdóttir
forskningslektor, fil.dr
Árni Magnússon institútet för isländska
studier
Laugavegur 13
101 Reykjavík
helga.hilmisdottir@arnastofnun.is

Marjakaisa Matthíasson
Redaktör, fil.mag.
marjakaisa@internet.is