

15 konferensen
om lexikografi i Norden

4-7 juni 2019

Abstrakt för
plenarföreläsningarna



Anna Helga Hannesdóttir

Göteborgs universitet, Institutionen för svenska språket
anna.hannesdottir@svenska.gu.se

Den tvåspråkiga lexikografins utmaningar förr och nu

Den svenska lexikografins etablering sammanföll med den mest intensiva perioden av standardisering som svenska språket genomgått. Inte bara ortografin och grammatiken behövde regleras, också svenskans ordförråd var i stort behov av utveckling och förnyelse. Dessa utmaningar antog de tidiga lexikograferna i olika utsträckning och på olika sätt. Vid 1800-talets mitt når den tvåspråkiga svenska lexikografins utveckling ett avgörande läge när beskrivningen av svenskan som källspråk respektive målspråk sammanförs i den första enspråkiga svenska ordboken: Dalins Ordbok öfver svenska språket. Följderna för den tvåspråkiga lexikografins fortsatta utveckling dryftas kort.

Idag har datatekniken och det digitala publikationsformatet medfört stora förändringar för lexikografen. Språkteknologins allt viktigare roll inom lexikografen innebär både enorma möjligheter och stora teoretiska och metodiska utmaningar. Inte minst gäller detta den tvåspråkiga lexikografen när nya användningsområden öppnas. Bland annat kan den tvåspråkiga lexikografiska beskrivningen nu läggas till grund för kontrastiva undersökningar av helt andra slag än den traditionella tvåspråkiga ordboken har kunnat erbjuda. För att utnyttja den flexibilitet som det digitala användargränssnittet möjliggör krävs en översyn av den teoretiska grund som lades under det sena 1900-talet. Exempelvis medför uppluckringen av dikotomin målspråk – källspråk helt ny teoretisk och metodologisk situation.

I mitt föredrag kommer jag dels att ge en kort översikt över den svenska lexikografins historiska utveckling, med fokus på de tvåspråkiga ordböckerna och deras utmaningar under olika perioder, dels att beskriva läget för den tvåspråkiga lexikografen idag.



Iztok Kosem

Jožef Stefan Institute (ELEXIS project)
Centre for Language Resources and Technologies,
University of Ljubljana
iztok.kosem@ijs.si

Responsive dictionaries: bringing dictionaries (and lexicographers) and users closer together

A recent study on the use of (monolingual) dictionaries across Europe (Kosem et al., forthcoming) has shown that up-to-date content is among the top three dictionary features valued by users. In fact, in countries where existing dictionaries are rather outdated, the users rank up-to-date content as high as reliability, sometimes even higher. The message is clear – users want dictionaries that reflect contemporary language and that react to the changes in language as quickly as possible. The advances in corpus tools and related lexicographic tools (e.g. tools for neologism detection) have enabled lexicographers to address these user demands rather successfully; however, this is mainly the case for languages with good existing dictionary infrastructures, such as English, German etc.

However, in countries such as Slovenia one of the main problems is the outdatedness of dictionaries, and even the non-existence of or poor accessibility to certain, more specialised dictionaries. This means that these dictionaries need to be made from scratch, which would traditionally also mean years (even decades) of waiting for the users. In order to address this issue, the team at the Centre for Language Resources and Technologies at the University of Ljubljana has tested and implemented several state-of-the-art approaches into the lexicographic process, such as automatic extraction of lexical data from the corpus (Gantar et al. 2016, Kosem et al. 2013), and crowdsourcing (Čibej et al. 2015). Furthermore, we have used semi-automatic methods for the creation of language resources such as Thesaurus of Modern Slovene (Krek et al. 2017a).

The implementation of the methods mentioned above had one important consequence for the dictionary-making process; namely, the data obtained with (semi-)automatic approaches was already of such quality that we decided to present it to the users. Two dictionaries, Thesaurus of Modern Slovene (Krek et al. 2017b) and Collocations Dictionary of Modern Slovene (Kosem et al. 2018), were published in this manner. We named them responsive dictionaries as they are able to quickly “respond” to changes, either because of completed lexicographic work and/or availability of new data (e.g. corpora). Another form of responsiveness is related to the users, as we offered them the opportunity to be involved in the dictionary-making process via voting on particular types of information, suggesting new content, or via specific crowdsourcing tasks.

In my talk, I will present both dictionaries, paying particular attention to the lexicographic process behind both dictionaries, and the contributions and feedback received from the users. In addition, as the dictionaries share certain types of data (collocations, examples), I will show how they were integrated into a large database where all the information from the two dictionaries, as well as from other resources (including crowdsourcing tasks), is combined. I will also discuss some of the solutions that had to be implemented in the dictionary interface to provide the users with the necessary information about the reliability of dictionary data, and to enable their participation in a straight forward manner. I will argue that the adoption of such a model of dictionary publication brings several benefits to both users and lexicographers. In conclusion, I will present future plans, including the implementation of various methods that are being developed within the ELEXIS (European Lexicographic Infrastructure) project.

References

- Čibej, Jaka, Fišer, Darja, Kosem, Iztok. 2015. The role of crowdsourcing in lexicography. In: Kosem, Iztok et al. (eds.), *Electronic lexicography in the 21st century: linking lexical data in the digital age*. Ljubljana: Trojina, Institute for Applied Slovene Studies; Brighton: Lexical Computing, pp. 70-83. https://elex.link/elex2015/proceedings/eLex_2015_05_Cibej+Fiser+Kosem.pdf
- Gantar, Polona, Kosem, Iztok, Krek, Simon. 2016. Discovering automated lexicography = the case of Slovene lexical database. *International journal of lexicography*, 29/2, pp. 200-225. <https://academic.oup.com/ijl/article/29/2/200/2413284/Discovering-Automated-Lexicography-The-Case-of-the?guestAccessKey=95f18766-f10f-4994-a6fa-448cf75ac55e>
- Kosem et al. (eds.). 2017. *Kolokacije 1.0: Collocations Dictionary of Modern Slovene*. viri.cjvt.si/kolokacije
- Kosem, Iztok, Gantar, Polona, Krek, Simon. 2013. Avtomatizacija leksikografskih postopkov. In: Erjavec, Tomaž, Žganec Gros, Jerneja (eds.). *Jezikovne tehnologije, Slovenščina 2.0, 1/2*. Ljubljana: Trojina, zavod za uporabno slovenistiko, pp. 139-164. http://www.trojina.org/slovenscina2.0/arhiv/2013/2/Slo2.0_2013_2_07.pdf
- Kosem, Iztok, Lew, Robert, Müller-Spitzer, Caroline, Riberio-Silveira, Maria, and Wolfer, Sascha. forthcoming. The image of the monolingual dictionary across Europe. Results of the European survey of dictionary use and culture. *International Journal of Lexicography*.
- Krek et al. (eds.). 2017b. *Sopomenke 1.0: Thesaurus of Modern Slovene*. viri.cjvt.si/sopomenke
- Krek, Simon, Laskowski, Cyprian, Robnik-Šikonja, Marko. 2017a. From translation equivalents to synonyms: creation of a Slovene thesaurus using word co-occurrence network analysis. In: Kosem, Iztok et al. (eds.), *Proceedings of eLex 2017: Lexicography from Scratch*, 19-21 September 2017, Leiden, Netherlands. <https://elex.link/elex2017/wp-content/uploads/2017/09/paper05.pdf>



Nina Martola

Institutet för de inhemska språken
Helsingfors
nina.martola@sprakinstitutet.fi

Den tvåspråkiga ordboken – ett absurt åtagande

När man på 1990-talet tog del av vad de stora namnen inom lexikografin sade eller skrev kände man sig lätt uppgiven i vårt hörn av världen. För de som arbetade med ordböcker där engelska är det ena språket och spanska, franska, tyska eller italienska var det andra var det en självklarhet att man skulle göra olika ordböcker för olika målgrupper. Det fanns ”aktiva” och det fanns ”passiva” och det fanns learners’ dictionaries och det fanns översättningsordböcker osv. i båda riktningarna och för båda språkgrupperna.

När man arbetade med ordböcker mellan finska och svenska fanns det i princip tre typer, en fickordbok, en mellanstor för skolbruk och en storordbok för avancerade användare och alla var tänkta för finskspråkiga användare. Ordböcker gjorda i Finland riktas mot den finska marknaden. Den svensk-finska ordboken eller ordboksdelen var i princip en ”passiv” ordbok, eller en avkodningsordbok med en modernare term. Den finsk-svenska riktningen var tänkt för användare som skulle producera text på det främmande språket svenska. I Sverige producerades inga ordböcker mellan finska och svenska riktade till svenskspråkiga. (Svensk-finska Lexin var till för finska invandrare i Sverige.)

Målgruppskrocken var hela tiden närvarande i arbetet. Förlaget hade den stora marknaden för ögonen, dvs. finska användare, medan redaktionen sneglade mycket på en annan grupp, nämligen översättarna från finska till svenska. I det tvåspråkiga Finland, där så mycket officiell text översätts, var och är de en ytterst viktig målgrupp. De har i regel svenska som modersmål eller åtminstone svenska som ett ytterst starkt språk. Att försöka tillgodose olika nivåer på målgrupperna blev en stor utmaning.

I mitt föredrag kommer jag att ta upp de omöjligheter och absurditeter som redaktionen blev tvungen att tackla, dels på grund av ordbokens art, dels utifrån de kommentarer och reaktioner vi fick från olika håll. Föredraget kommer visserligen att utgå från förhållandena på 1990-talet, men problematiken har fått ny aktualitet i och med att många ordböcker nu läggs ut fritt på nätet och på så vis ska kunna betjäna en synnerligen brokig användargrupp.

Abstrakt för
sektionsföredrag

Webbaserad crowdsourcing inom SignWiki-ordboksarbete

SignWiki-ordböckerna i finskt och finlandssvenskt teckenspråk upprättades 2013. Ordböckerna baserar sig på att teckenspråkiga kan medverka genom att själva, eller via administratören, lägga till tecken till sidorna. Eftersom teckenaktörerna har varierande språkbakgrund får man fram språkets lexikaliska och artikulatoriska variation på ett naturligt sätt.

Båda sidornas material produceras med samma metoder (workshoppar, intervjuer, excerpering av tecken ur videomaterial) men sidorna skiljer sig innehållsmässigt p.g.a. olika behov. De finska sidorna har över 4500 teckensidor och innehåller mycket material ur specialområden. De finlandssvenska sidorna består av ca 900 teckensidor varav de flesta hör till baslexikon. Därtill har de finlandssvenska sidorna över 100 korta berättelser.

SignWiki fungerar som en plattform för publicering av material, men inte som en plats för bearbetning av materialet före publiceringen t.ex. i form av diskussion. Därför kommer Finlands Dövas Förbund, som administratör, att ta i bruk ett webbfrågeformulär i Surveypal under våren 2019. Administratören kan samla in tecken via formuläret och språkbrukarna kan ge kommentarer till tecken. Språkbrukarna kan även ge respons på teckenspråk eftersom man kan banda in video via formuläret som gjorts tekniskt enkelt för användaren. I vår presentation framför vi administratörens och språkanvändarnas erfarenheter om denna nya arbetsmetod.

Länkar

Suomalaisen viittomakielen wikisanakirja: <https://finsl.signwiki.org>
Wikiordbok i finlandssvenskt teckenspråk: <https://finssl.signw>

Enspråklig norsk ordbok for fremmedspråklige

I den norske synonym- og definisjonsordboken for fremmedspråklige innvandrere og flyktninger i Norge fokuserer jeg på områder hvor vi avviker fra Bokmålsordboken i lys av den intenderte brukergruppens behov. Områder som ordtilfang, ordklassemarkering, uttale og prinsipper for synonymi og ordforklaringer er sentrale. Ordboken bygger på Norsk-Vietnamesisk ordbok fra 1983 med flere omganger med oppdateringer. Ordtilfanget er spesielt tilpasset fremmedspråklige, særlig innvandrere og flyktninger som har fått grunnleggende norskopplæring og som skal innpasses i det norske arbeidsliv.

Siden 2011 har revisjonene vært drevet av professor emerita Marianne Haslev Skånland og undertegnede uten tilskudd fra offentlige eller private kilder. I foredraget vil jeg illustrere hvordan den intenderte brukergruppen legger føringer på ordboksstrukturen. Når det gjelder ordtilfang går vi lenger enn Bokmålsordboken på leksikalske felt som er spesielt sentrale for fremmedspråklige, som ord som regulerer arbeidsmarkedet, juridisk terminologi og medisinsk terminologi blant annet.

Eit dilemma om lemma

Noko alle ordbokredaksjonar må ta stilling til, er kva kriterium som skal avgjere om noko er eit lemma eller ikkje. Svensén (2009) nemner desse kriteria:

etymologi: ord med same staving og sams opphav blir behandla i same artikkel

semantikk: ord med same staving og med tydingar som høyrer saman, blir behandla i same artikkel

grammatikk: ord med same staving blir behandla i same artikkel om dei òg har same grammatiske trekk

Alle desse kan problematiserast, og til vanleg legg ein noko vekt på alle tre. Eit kriterium Svensén ikkje nemner, er **normstatus:** Ord som rettskrivingsnorma definerer som valfrie former, blir gjerne førte som eitt lemma med to eller fleire variantar. Slik ser brukaren lett kva variantar som er tillatne.

Rettskrivingsvariantar finn vi vel i alle språk (t.d. britisk engelsk: *realise/realize*, dansk: *medie/medium*, svensk: *cybernetik/kybernetik*), men bokmål og nynorsk har spesielt mange variantar som følgje av den norske språkhistoria. Nokre døme er bokmål heim/hjem og mjølk/melk, nynorsk først/fyrst og tenke/tenkje.

Vi skal sjå nærare på korleis skiljet mellom variantformer og lemma har vore praktisert i norske ordbøker. Dei siste åra har det vore ein tendens til at fleire tidlegare variantformer blir skilde ut som eigne lemma. I utgreiinga om nynorskrettskrivinga frå 2012 blei det lagt til grunn at ord skal reknast som separate lemma når dei har ulik semantikk, etymologi eller bøyning. Vi vil drøfte konsekvensar av dette, og vi meiner det er særleg vanskeleg å operasjonalisere ulik etymologi (norsk eller dansk) og ulik bøyning som kriterium for å gjere variantformer om til ulike lemma.

At former med ulik semantikk blir rekna som eigne lemma, er i utgangspunktet sjølvsagt, men for ein del ord er det ein tendens til at éi form dominerer i visse tydingar. Da er spørsmålet kor sterk tendensen skal vere før ein slår fast at det er tale om lemma med ulik semantikk. Vi skisserer korleis vi meiner grensedraginga bør gjerast, og vi vil sjå dette i lys av visse prinsipp i norsk språkpolitikk.

Online Data on Icelandic Inflection: Descriptive or prescriptive: “Why, for whom, by whom, and how?”

The Database of Modern Icelandic Inflection (DMII, Beygingarlýsing íslensks nútímamáls) contains over 280,000 paradigms, with almost 6 million inflectional forms. The project’s original purpose was to create data for use in language technology, and as such, the data must be descriptive rather than prescriptive, as it has to include words and inflectional forms appearing in modern Icelandic texts, irrespective of standardization, spelling, or adherence to standards of correctness or language purity, etc. The result is a large dataset with extensive information on vocabulary and inflectional variants, where usage is sometimes restricted by grammatical or idiomatic context, as well as by style, genre, etc.

The online DMII is used by the general public as a reference, but the descriptive nature of the work makes it difficult to use for people without native or near-native knowledge of Icelandic. The topic of this presentation is the “for whom, by whom, why and how?” of a new, smaller prescriptive version, *The DMII Core*, to be completed in 2019. The DMII Core data will be made available via an API, free of charge. It will be suitable for publication of a prescriptive dictionary of inflection by third party developers, and for use in the creation of teaching material.

As the DMII proper, the project is carried out at The Árni Magnússon Institute for Icelandic Studies. The vocabulary of just over 60,000 headwords, is based on *The Modern Icelandic Dictionary (Íslensk nútímamálsorðabók)*, with additions from *The Icelandic Gigaword Corpus* (2017). All variants in the DMII Core vocabulary have been ranked according to correctness, style, genre, etc., and the DMII Core only contains prescribed forms. The presentation will describe of the choice of vocabulary, and the ranking system of correctness in the DMII Core, with reference to practical problems involved in the juxtaposition of prescription and description in the complex and fluid morphological system of Icelandic.

Websites

The Database of Modern Icelandic Inflection (DMII) [Beygingarlýsing íslensks nútímamáls, BÍN], ed. Kristín Bjarnadóttir. The Árni Magnússon Institute for Icelandic Studies. bin.arnastofnun.is

The Modern Icelandic Dictionary [Íslensk nútímamálsorðabók], eds. Þórdís Úlfarsdóttir & Halldóra Jónsdóttir. The Árni Magnússon Institute for Icelandic Studies. málið.is

The Icelandic Gigaword Corpus [Rísamálbeildin], Steinþór Steingrímsson et al. The Árni Magnússon Institute for Icelandic Studies. malheildir.arnastofnun.is.

Användning av en korpus och en enspråkig ordbok som bas för den nya estnisk-finska ordboken

Den nya estnisk-finska ordboken publicerades i januari 2019 av Eesti Keele Instituut och Institutet för de inhemska språken. Den är den första elektroniska estnisk-finska ordboken. De gamla ordböckerna har redan blivit delvis föråldrade och är gjorda av redaktörer som har källspråket estniska som modersmål, däremot är de flesta i den nya ordbokens redaktion finskspråkiga.

Estniskan har blivit ett allt viktigare språk i Finland, för de estnisktalande är den näst största invandrargruppen i vårt land. Den elektroniska och fritt tillgängliga ordboken med ett responsivt gränssnitt tjänar också pendlare och turister. Ordboken hade cirka 39 000 uppslagsord när den publicerades och den uppdateras fortlöpande.

Den nya estnisk-finska ordboken baserar sig på den nya estniska definitionsordboken som är korpusbaserad. Korpusen med 1,3 miljarder textord baserar sig för sin del på estniska texter samlade från internet. De mest frekventa 35 000 orden har delats i sju frekvensgrupper och de första grupperna finns med som helhet i den tvåspråkiga ordboken. Ordboken innehåller gott om exempel tagna direkt från korpusen. Användning av korpusen och den enspråkiga definitionsordboken som bas gör å ena sidan redaktionsarbetet snabbare, å andra sidan skapar det olika problem, som diskuteras i föredraget.

Att hantera grammatisk förändring i en deskriptiv ordbok

Medan *Svenska Akademiens ordlista* (SAOL) brukar beskrivas som normativ, är *Svensk ordbok utgiven av Svenska Akademien* (SO) tänkt att vara mer deskriptiv, och ett skäl till detta kan t.ex. vara att SO därigenom kompletterar SAOL. I anslutning härtill kan man fundera på om SO är, och ska vara, lika deskriptiv i alla informationskategorier: Gäller deskriptiviteten bara på ”SAOL-området”, dvs. främst i fråga om stavning och böjning, eller gäller den hela ordboksartikeln?

Vi vill diskutera hur SO behandlar den *grammatiska* informationskategorin, särskilt den som redovisas i den avdelning av ordboksartikeln som behandlar konstruktionsuppgifter. Här verkar det nämligen saknas normativa verk. När konstruktionsordboken *Svensk handordbok* utgavs 1966 var syftet tydligt: man ville ”varna för felaktigt eller mindre lämpligt språkbruk”. Den nu över 50 år gamla ordboken (räknat från första uppl.) har dock ingen självklar normativ arvtagare. *Svensket språkbruk. Ordbok över konstruktioner och fraser* (2003) är uttalat deskriptiv, och i SO, utg. 2009, sägs inget om huruvida konstruktionsuppgifterna ska läsas ur ett annat perspektiv än ordbokens övergripande deskriptiva.

Det är svårt att komma ifrån att det som tas upp i en ordbok (särskilt i en med Svenska Akademien som avsändare) uppfattas som ”godkänt” av en auktoritet. Att Sveriges officiella språkvårdsorgan Språkrådet i ett antal fall hänvisar till SO:s konstruktionsuppgifter när man besvarar frågor av fraskaraktär tyder på att konstruktionsuppgifterna knappast utgör något undantag.

Efter en kort bakgrund till konstruktionsuppgifternas plats i ordboken, diskuterar vi vem som ska svara för normeringen av ordförbindelser (om nu inte SO helt och fullt tar på sig denna roll), samt för vem, och varför, olika sorters konstruktionsuppgifter ska inkluderas i ordboken. Vi exemplifierar med några uttrycksätt som har börjat bli frekventa men som ännu inte har inkluderats i ordboken, bl.a. *visa på ngt* och dess sentida syskon *indikera på ngt*.

Erik Bäckerud & Pär Nilsson

Redaktionen för SAOB

Emma Sköldberg

Redaktionerna för SAOL och SO, Institutionen för svenska språket, Göteborgs universitet

Så används Svenska Akademiens ordböcker på nätet – till frågan om barometerglaset

I mars 2016 lanserades den nya webbplatsen för Svenska Akademiens ordbok, <www.saob.se>. I september 2017 kom den gemensamma ordboksportalen <svenska.se> där möjlighet finns att söka i tre ordböcker på en gång, SAOL 14, SO och SAOB. Sedan lanseringarna har antalet sökningar stadigt ökat, men hur utnyttjar användarna egentligen de digitala ordböckerna? I föredraget presenteras och analyseras statistik över hur webbplatserna används.

Bla. presenteras vilka de vanligaste sökorden är och vilka sökningar som inte leder till träff (s.k. nollträffar). Skiljer sig mönstret åt mellan svenska.se och saob.se? Vad kan man dra för slutsatser av detta? Sökningar på sådana ord som *ben* och *cringe* kommer att uppmärksammas – och varför ligger *barometerglas* i topp i SAOB:s loggar över sökord? I föredraget diskuteras även statistik över när och var sökningarna sker och ifrån vilken enhet de oftast utförs: mobil eller dator? Dessutom diskuteras ett antal frågor som kommit till ordboksredaktionen från ordböckernas användare. Exempel ges på hur de insamlade uppgifterna kan användas vid en kommande uppdatering av webbordböckerna.

Karin Cavallin

Svenska Akademiens ordboksredaktion, Lund

Ngram i SAOB

Syftet med denna studie är att identifiera och sedermera analysera återkommande formler i SAOB:s definitionstext.

I digitaliseringsarbetet med SAOB pågår uppmärkning (taggning) av olika ordbokskomponenter (Bäckerud, 2016, s. 102). Under 2019 planerar vi att märka upp bland annat språkprov och definitionstext.

Inom lexikografen har ngram-verktyg (exv. *Ngram Statistics Package* (Banerjee et al. 2003) och *The Sketch Engine* (Kilgarriff et al. 2004) länge använts för att identifiera språkliga samförekomster i korpusar. Dyliga verktyg bör även kunna tillämpas för rent lexikologiska syften, på ordbokstext i sig. När definitionstexten i SAOB väl är identifierad, och därmed enklare att söka i, har vi möjlighet att göra ngram-sökningar på enbart definitionstext.

SAOB:s definitioner innehåller bland annat ett antal återkommande flerordsuttryck. Dessa flerordsuttryck är svåra att räkna upp på rak arm. Flerordsuttryck i SAOB skiljer sig från flerordsuttryck vanligt skriv- och talspråk genom att de kan vara ovanligt långa, och ofta är lite formelartade.

Nedan ser vi ett *dekapentagram*, 15 ord i följd som förekommer i definitionerna för HÅLL-SKJUTS (publ. 1932) och RESERV-SKJUTS (publ. 1957). Att återanvända definitionsuttryck är ett väldigt vanligt förfaringssätt i SAOB. Det återstår att se hur ofta långa, icke-unika ordkedjor förekommer, och hur ofta de är meningsfulla att generalisera över. Hela detta uttryck nedan är inte särskilt återanvändbart, men frasen ”ä. övergående i bet.:" är ett väldigt frekvent flerordsuttryck i SAOB. ...ä. konkret, om själva ekipaget (l. hästen l. hästarna); ä. övergående i bet.: skyldighet att...

Genom att studera ngram i SAOB kan vi få en överblick över vilka definitionsuttryck som faktiskt används. Detta är särskilt angeläget inför den kommande revideringen, då bland annat definitionsspråket i SAOB ska förenhetligas.

Litteratur

Ordböcker

SAOB (1889-) Ordbok över svenska språket utgiven av Svenska Akademien. 1-. Lund, Sverige.

Annan litteratur

Banerjee, Satanjeev and Ted Pedersen. 2003. The Design, Implementation, and Use of the n-gram Statistic Package. I *Proceedings of the Fourth International Conference on Intelligent Text Processing and Computational Linguistics*, s. 370–381. Mexico City, Mexiko.

Bäckerud, Erik. 2016. Indexeringen av SAOB. I *Nordiska Studier i Leksikografi* 13, s. 101-109. Köpenhamn, Danmark.

Kilgarriff, Pavel Rychly, Pavel Smrz och David Tugwell. Sketch Engine. *Proc. Euralex*, s. 105-116. Lorient, Frankrike.

Ulla-Maija Forsberg

Institutet för de inhemska språken, Helsingfors

En etymologisk slangordbok – för vem och varför

Etymologiska ordböcker anses generellt vara riktade mot specialister, lingvister, språkforskare, i kontrast till tvåspråkiga eller enspråkiga allmänordböcker, som en större mängd av språkanvändare behöver.

Ändå fick jag en förfrågan om att sammanställa en etymologisk ordbok över en s.k. Stadin slangi, dvs den gamla helsingforsslangen. Det var en grupp personer som till största delen inte alls var språkforskare utan konstnärer, låtskrivare och representanter för andra Helsingforsbor som är intresserade av den här språkformen som en del av den helsingforsiska identiteten och stadens historia.

I mitt föredrag skall jag behandla två centrala frågor som jag under mitt arbete måste ta ställning till: val av uppslagsord och formulering av själva etymologierna eller förklaringarna. Valet av uppslagsord gäller först och främst att välja vilka ord som överhuvudtaget skall komma med men också hur man hanterar den stora fonetiska variation som de enskilda orden uppvisar. Det senare problemet är inte särskilt svårt med tanke på det kommande elektroniska formatet; allting ryms med och går att hitta. Svårare är att slangen fortfarande lever: nya ord uppstår och det är svårt att veta vilka som kommer att leva vidare och vilka som kommer att ersättas med nya inom en kort tid. Därför kommer huvudvikten att ligga på det äldsta, från svenskan eller ryskan lånade ordförrådet. En väsentlig del av slangorden är nämligen variationer på vanliga finska ord eller nya betydelser för gamla ord, dvs närmast metaforen.

Med tanke på den andra frågan – varför? – lönar det sig att tänka på vad läsaren vill veta. Ska man försöka hitta den autentiska, mera vetenskapligt exakta vägen orden vandrat (konkret: vad är rollen för slangen ”på finska” eller ”på svenska”, som i stort sett har använt samma omformulerade ord) eller ska man nöja sig med ett lånordsoriginal i traditionell stil (lån från det svenska ordet x eller det ryska ordet y)?

Å dela ord i bitar – lingvistisk informasjon uttrykt ved leddeling i ordbøker

Dette innlegget vil drøfta spørsmålet om føremonar og ulempar ved to ulike leddelingsmetodar som har relevans for leksikografi. Metode 1 er markering av hovudfuge, noko som føreset semantisk analyse av ord med meir enn to ledd. Metode 2 er kalla flat leddeling, og inneber at alle ledd og fugeelement blir skilde med leddelingsmerke.

«Ord» forstått som ‘skriven teiknstreng med mellomrom eller skiljeteikn før og etter’ kan vera lange i nordiske språk. Orddelingsreglar finst i norsk rettskriving for korleis ord kan delast ved linjeslutt. Stavinga er grunneininga i dette systemet, og motivasjonen er først og fremst grafisk og estetisk. Leddeling av oppslagsord i ordbøker har eit anna føremål. Leddeling gjeld særleg samansette ord, med ei gråsone mot avleiingsselement. Leddelingsmerket skal syna ordstruktur, og dermed gje indirekte opplysning om opphav, uttale og samanheng med andre ord. Hovudfuge i samansette ord er merkt i Norsk Ordbok, og vart standard i Bokmålsordboka og Nynorskordboka ved siste oppgradering (2012 – 2015). I den underliggjande leddanalysen er fugeelementet markert. Indeksregisteret for Språksamlingane, Metaordboka, bruker eit meir utførleg system, nemleg merking av alle ledd i samansette ord, som t.d. i «skatt*e*ut*jamning*s*midel». Denne måten vart kalla «flat leddeling», av di han ikkje syner noko semantisk hierarki. Metoden vart vald for å spara tid. Om eit ord har meir enn to ledd, og berre ei leddeling skal markerast, må val av hovudfuge kvile på ei semantisk vurdering, som igjen føreset kontroll mot bruk i kontekst. Dersom alle ledd blir markerte, fell kontrollbehovet bort.

Flat leddeling gjev eit anna inntrykk av ordlaginga i norsk enn deling ved hovudfuge. Funksjonsorda (preposisjonar, adverb, pronomen) får ei meir synleg rolle, likeins korteverbalsubstantiv og fugeelement. I siste instans kjem verbfrasen si leksikogenetiske rolle tydelege fram. Spørsmålet om leddelingsmetode vil bli vurdert også i høve til maskinell ordanalyse.

Oddrun Grønvik & Johan Myking

Universitetet i Bergen

Stig Jarle Helset

Høgskulen i Volda

Nettbrukaren i fokus: digitalisering og revisjon av Norsk Ordbok

Arbeidet med Norsk Ordbok starta i 1929 ved Universitetet i Oslo og vart avslutta i 2015. Det Norske Samlaget har gjeve ut verket i 12 band via prosjekt NO2014 (2002–2015). Verket inneheld om lag 330 000 oppslagsord. Universitetet i Bergen (UiB) har i samarbeid med Høgskulen i Volda (HVO) teke ansvaret for å gjere ferdig den digitale utgåva av ordboksverket og revidera alfabetstrekket a–h. (Sjå særskilt abstrakt for posterpresentasjon.)

NO–AH blir eit nytt leksikografisk prosjekt som må retta seg mot nettbrukaren og derfor byggjast på til dels nye faglege prinsipp:

1) NO2014 vart planlagd som papiordbok, og plasspresset i papirutgjevingar var ein realitet. Det hadde følgjer mellom anna for handsaminga av målføretilfanget i formdelen av ordartiklane, som er samantrengde og kryptiske. Korleis bør informasjon om målføre presenterast i ei vitenskapleg nettordbok, slik at utbyttet blir optimalt for heile spennet av ordboksbrukarar? Denne utfordringa er NO–AH så godt som aleine om.

2) NO–AH blir heildigital. Redaktørane får tilgjenge til det digitale etterslepet etter NO2014, som t.d. fulldigitalisering av målføresynopsisen og Norsk dialektatlas, og nettilgjenge til dei eldre ordsamlingane (1600–1850). Noverande band 1 er tynt når det gjeld dei eldste kjeldene. På dette punktet er viktige ting gjort etter at NO2014 vart avslutta – særleg at Aasen 1873 og Ross 1895 no ligg i ordbokshotellet og er søkbare i seg sjølve og via Metaordboka.

3) Kjeldekritikk og kjelde kvalitet kjem høgare på saklista no som enorme korpus byggjer seg opp på alle kantar. I tida før NO 2014 var parolen å utnytta alt tilfang til siste komma, under NO2014 vart ordsamlingane brukte både som redigeringsgrunnlag og til å proporsjonera ordboka, no er overvekta av nytt avistilfang eit problem på grunn av mengda.

Forbindelser og folketro i Ømålsordbogen

Dialektordbogen Ømålsordbogen har undertitlen ”en sproglig-saglig ordbog”. Udtrykket *saglig*, der har sammenhæng med forskningsretningen *Wörter und Sachen*, skal markere at ordbogen lægger stor vægt på at beskrive den materielle fisker- og bondekultur som ordene er indlejret i. Selv om det ikke er direkte fremhævet i titlen, behandler ordbogen imidlertid også den åndelige og litterære folkekultur ret grundigt.

Størst vægt er lagt på de korte folkelige genrer med (mere eller mindre) fast form som ordsprog, eder, rim osv., og ordbogen opererer med en ret detaljeret inddeling både set i forhold til andre dialektordbøger og i forhold til hvordan kollokationer, idiomer ol. er markeret i *Ømålsordbogen*. Mens markeringen ”i forb” kan bruges om udtryk af så forskellige typer som *flyve op*, *flyve og fare* og *flyve i flint*, er der adskillige markeringer til at kategorisere de folkloristiske, sætningsformede forbindelser med, fx ”i talemåde”, ”i ordsprog”, ”i eder”, ”i folkemedicin”, ”i helbredelsesformler”, ”i trylleformularer”, ”i varsler”, ”i vejrregler”, ”i leveregler”.

Den detaljerede opdeling er informativ og medvirker til at skabe et overblik over stoffet i lange artikler. Det vil jeg give en række eksempler på i mit foredrag, men jeg vil også diskutere nogle af de afgrænsningsproblemer den detaljerede beskrivelse medfører for redaktørerne og de ulemper den kan have for brugeren. I diskussionen inddrager jeg også andre dialektordbøgers praksis, både den typiske i moderne i dialektordbøger og Feilbergs særegne i *Bidrag til en ordbog over jyske Almuesmål* (1886–1914), der næsten i lige så høj grad er et folkloristisk opslagsværk som en ordbog.

Inger Schoonderbeek Hansen

Peter Skautrup Centret for Jysk Dialektforskning

Institut for Kommunikation og Kultur

Aarhus Universitet, Arts

Jysk Ordbog – af hvem, til hvem og hvorfor?

Jysk Ordbog blev grundlagt i 1932 af professor Peter Skautrup, som ønskede, på samme måde som Udvalg for Folkemaal i København havde gjort siden 1908, at 'redde hvad der reddes kunne' af det dialektale stof i Jylland, før det var for sent og al dialekt var forsvundet (jf. Gudiksen & Hovmark 2009, Schoonderbeek Hansen 2011). Jysk Ordbog endte ad adskillige mellemregninger på www.jyskordbog.dk, hvor bogstav A-K(æ) er publiceret, resten ligger fremme som 'stubbe'.

Jysk Ordbog har gennem tiderne været redigeret af sprogforskere med et praktisk leksikografisk fokus og kulturhistorisk interesse (jf. Svensén 2004:3). Kilderne er mangfoldige: ordbogskartoteket på 3 mio. sider består af fx skønlitteratur på jysk dialekt (fx Jeppe Aakjær); litteratur om jyske folkeminder og -kultur (fx Evald Tang Kristensens samlinger). Desuden en stor spørgelistesamling (der giver valide input til ords udbredelse), og en stor båndsamling indsamlet i 1970'erne-1980'erne sammen med kollegerne fra Københavns Universitet (<http://www.jyskordbog.dk/jyskordbog/ordbog.html>). Ordbogens tyngdepunkt ligger i perioden 1870-1920, dvs. før industrialiseringen, og den kan derfor karakteriseres som en synkronisk historisk ordbog (Svensén 2004:28). Ca. 47 % af ordbogen er færdigredigeret (ca. 18.500 artikler); resten ligger som 'artikelstubbe' på portalen; heraf forventes yderligere 20.000 artikler at blive færdigredigeret (jf. <http://www.jyskordbog.dk/jyskordbog/ordbog.html>).

Jysk Ordbogs brugere er såvel fagfolk, fx sprogforskere, som ikke-fagfolk, dvs. brugere med en almen interesse for jysk bondekultur. Det kan give visse udfordringer, men når ordbogen beskriver en sprogtilstand der ikke længere svarer til den brugerne kan forholde sig til, hvad får de så ud af den? Hvilke præmisser arbejder synkrone historiske (dialekt) ordbøger under, hvilke overvejelser gør de? Hvorfor er de i det hele taget til? Oplægget handler dermed ikke kun om Jysk Ordbogs *af hvem, til hvem og hvorfor*, men også Jysk Ordbog *før, nu og i fremtiden*.

Hvordan har efterspørgslen efter tematiske dansk-svenske og svensk-danske netordlister forandret sig efter at majoriteten af brugerne nu søger ord via smarttelefoner?

De sidste 2–3 år er der sket uventede store forandringer i antal eksponeringer af visse af mine tematiske dansk-svenske og svensk-danske ordlister på nettet. I samme tidsrum kan det konstateres, at smarttelefonen er gået forbi den stationære computer og har positioneret sig som den vigtigste platform for ordsøgerne til ordlisterne, og som ordbogsudvikler spørger man naturligvis sig selv, hvordan den nye dominerende søgeplatform påvirker udnyttelsen af ordlisterne, og hvordan man kan karakterisere de nye trender.

I foredraget præsenteres og analyseres de relativt hurtige forandringer, der er set de sidste år med hensyn til efterspørgslen efter netordlisterne i Sverige og Danmark (og som tilsyneladende korrelerer med den stigende anvendelse af smarttelefoner), samt de prognoser, der kan opstilles på nuværende tidspunkt. Der foreligger i dag 13 års data for efterspørgslen efter ordlisterne i begge lande, hvilket giver et særdeles godt analyse- og prædiktionsgrundlag samt en ekstra god mulighed for at karakterisere de forandringer, der er sket, og sætte dette i relation til, hvordan smarttelefoner kan have påvirket benyttelsen af ordbøger samt hvilken rolle de kan spille i fremtiden.

Analysearbejdet har fokus på tidsserieanalyse og statistisk modellering, som tilbyder en effektiv metode til at gennemføre analysen på en velstruktureret og forsvarlig måde. Metoden indebærer, at man analyserer tidsserier med det formål at beskrive karakteristiske træk hos data på en modelbaseret måde. Tidsvariationens formaliserede sæson- og trendkomponenter giver mulighed for at beskrive grad af afvigelse fra det allerede kendte og at forudsige den fremtidige udvikling, enten ved en matematisk ekstrapolering af variationsmønstre i historiske data eller ved hjælp af andre tidsserier, som man tror kan påvirke. Dette forventes kunne føre til en bedre forståelse af det, der sker ved overgangen til en ny søgeplatform.

Elina Heikkilä & Jarkko Kauppinen

Institutet för de inhemska språken, Helsingfors

Nedsättande ord i finländska nätordböcker – vilka, hur och varför?

I vårt föredrag kommer vi att jämföra hur nedsättande ord behandlas i finländska nätordböcker. Varför väljer man att överhuvudtaget ta med dem i ordböcker och hur väljs de ut? Finns det någon skillnad mellan ordböcker som baserar sig på någon viss korpus eller i övrigt begränsat material och ordböcker som inte gör det? Hur kan man söka nedsättande ord i olika ordböcker?

Vi koncentrerar oss på de nätordböcker som redigeras på Institutet för de inhemska språken. Bland dessa ordböcker – i dagens läge nio stycken – finns det en- och tvåspråkiga ordböcker som täcker såväl modernt standardspråk med dess stilvariation som äldre språkformer, både dialekter och skriftspråk.

Oberoende på om det är fråga om deskriptiva eller normativa ordböcker är språkbrukarna benägna att tolka dem normativt. Från ordboksredaktörens synvinkel täcker en komplett beskrivning av lexikonet i ett visst språk eller en viss språkform naturligtvis också nedsättande ord och andra slags tabuord. Språkbrukaren kan för sin del använda ordböcker för olika ändamål: hen kan försäkra sig om ett ords betydelse och stilvärde eller kolla vilka ord som går att använda i ett visst sammanhang och vilka som borde undvikas.

I dagens moderna språk sker stilistiska förändringar tämligen snabbt, och det måste man beakta vid uppdatering av nätordböcker. Det som tidigare har uppfattats som skämtsamt eller neutralt, kan i dagens språkbruk uppfattas som nedsättande eller ålderdomligt.

Den tidigaste finska ordboken som innehåller ett stort antal nedsättande ord och uttryck är Christfrid Gananders *Nytt Finskt Lexicon*. Ganander färdigställde manuskriptet år 1787 men fick aldrig se sin ordbok publicerad. Hans försök till en komplett beskrivning av finska språkets ordförråd innehåller ca 300 nedsättande ord. Om manuskriptet hade tryckts utan strykningar på 1700-talet hade dåtida domare haft en ypperlig hjälprea när de skulle döma skymfare och häcklare till bötesstraff.

Om sverigefinskt terminologi- och ordlistearbete

Framför allt på 1960- och 1970-talet flyttade många från Finland till Sverige. Flera av dem kunde ingen svenska och behövde därför information på finska i sina kontakter med myndigheter.

Viss form av sverigefinskt terminologiarbete förekom redan på 1960-talet när tolkar och översättare i sitt arbete skapade finska ekvivalenter till svenska samhällstermer. Arbetet organiserades när Sverigefinska språknämnden bildades 1975. År 2006 införlivades Sverigefinska språknämndens arbete i Språkrådet som är en avdelning vid Institutet för språk och folkminnen.

Ordlistor för den sverigefinska minoriteten

Skapandet av ordlistor i olika specialområden har alltid varit en mycket viktig uppgift för Sverigefinska språknämnden och Språkrådets finska verksamhet. Mest angeläget har det varit att ta fram finsk terminologi för de centrala samhällsområdena skola, arbetsmarknad och det sociala området.

Sverigefinska språknämnden gav ut ordlistor redan på 1980-talet, och Språkrådet har under senare år publicerat bland annat en svensk-finsk socialordlista (2011), en svensk-finsk skolordlista (2014) och en svensk-finsk omsorgsordlista (2018).

Varför sverigefinsk terminologi?

I framtagandet av ordlistor utgår vi alltid från den sverigefinska befolkningens behov av terminologi på finska. Utgångspunkten i vårt arbete är dock svenska: uppslagsorden är på svenska. De svenska och finska samhällena är lika men inte identiska: i båda länderna förekommer terminologi som inte används i det andra landet. Typiska svenska exempel är *landsting* och *lärarygghet* – de företeelserna finns inte i Finland och därmed inte i finskan. Därför har vi språkvårdare i finska tagit fram en finsk term för dessa. Ofta är de finska ekvivalenterna översättningslån men ibland ställer språkens strukturella skillnader till det, och den finska ekvivalenten behöver skapas på ett annat sätt. Även innebörden av begreppen bakom termer kan skilja sig åt, som i fallet *förskola*.

Ett annorlunda exempel på vårt termarbete är vår senaste ordlista. Där är målgruppen undantagsvis de som bara kan lite finska eller inte alls. Ordlistan är tänkt för dem som arbetar inom vården och ibland måste använda finska, trots att de inte kan språket, för att kunna kommunicera med vårdtagare som har förlorat det svenska språket.

Om detta finskspråkiga terminologiarbete i Sverige vill vi gärna berätta på den 15 konferensen om lexikografi i Norden.

Tarja Heinonen

Institutet för de inhemska språken, Helsingfors

Is *alien* Finnish?

This paper deals with the question of what kind of new loanwords are added, or should be added, as headwords to the contemporary Finnish dictionary *Kielitoimiston sanakirja* (available online at <https://www.kielitoimistonsanakirja.fi/>). Generally speaking, the same principles apply to all vocabulary: a word or expression should occur with sufficient frequency in Finnish texts and its form or meaning should be at least partly unpredictable. In addition, if an item is a noun, adjective or verb, it should have inflectional forms. Selection based on these criteria explains why the word *alien* – though alien-looking to Finnish eyes – has made its way into the Finnish dictionary. However, as with any general dictionary of standard language, there are also other issues involved. To begin with, the *Kielitoimiston sanakirja* does not even attempt to include informal registers or specific domains very thoroughly. This inevitably begs the question: where to draw the line?

Especially with respect to loanwords, the dictionary also has a role in codifying the forms to be used. Sometimes a loanword slips into the Finnish phonotactic and inflectional systems without any conflict (e.g., words like *lassi* ‘a yogurt beverage’), but sometimes there are several competing forms, e.g., *appi*, *äppi*, *apsi*, and *appsi* for the English word *app*, and it is not clear which forms should be considered, especially if none of them is regarded as standard. In the case of the word for an ‘app’, the native noun *sovellus* derived from the verb *soveltaa* ‘to apply’ is preferred over forms like *appi* and *apsi*. Preference is shown in the *Kielitoimiston sanakirja* in such a way that there are no cross-references from the standard form to the colloquial forms, but vice versa. In the case of the word *alien*, there are no orthographic variants to the headword, but the variation in pronunciation (/alien/, /eilien/, or /eiliæn/) gives rise to extra variation in inflectional forms (for instance, either *alienia* or *alieniä* in the partitive singular).

Louise Holmer

Göteborgs universitet, Institutionen för svenska språket

Klickande, skrivande och onödigt vattendrickande. Verbalsubstantiv på *-ande* i ordböcker och korpusar.

Svenskans potential att bilda nya ord genom avledning och sammansättning är välkänd. En ständigt aktuell fråga för ordboksredaktioner är vilka av dessa avledningar och sammansättningar som ska tas upp som självständiga uppslagsord i en ordbok och vilka som kan noteras på annat ställe i artikeltexten, alternativt bortses från helt.

I det här föredraget står avledningar av typen neutrala verbalsubstantiv i fokus, t.ex. (*ett*) *klickande* och (*ett*) *skrivande*. Den aktuella ordgruppen på *-ande* har intressant nog beskrivits som ”en nästan värdelös ordgrupp” i förtexten till SAOL 7 (1900), bl.a. med argumentet att verbalnomenet bildas regelbundet till de flesta verb. I ordböcker behandlas annars avledningar med *-ande* sparsamt – om ens alls – i exempelvis förorden, och det är inte heller givet att de upptas som självständiga uppslagsord i lemmalistan. I lexikologiska sammanhang har avledningstypen beskrivits ingående av bl.a. Loman (1962), Söderbergh (1971) och Josefsson (2005). Någon större systematisk genomgång av avledningstypens förekomst i moderna svenska ordböcker har dock inte gjorts tidigare.

I föredraget redovisas och problematiseras hur några olika svenska, enspråkiga ordböcker skiljer sig åt i fråga om hur neutrala verbalsubstantiv på *-ande* behandlas. Resultatet från den lexikografiska undersökningen relateras också till avledningarnas förekomst i skriftbruket i form av olika texttyper från de senaste 50 åren.

Materialet utgörs främst av *Svensk ordbok* (SO 2009), *Bonniers svenska ordbok* (BSO 2010) och SAOB (1898–). Några jämförelser görs också med danska och norska ordböcker. Textmaterialet som illustrerar bruket av avledningarna hämtas från Språkbankens korpussamling Korp och utgörs främst av skönlitteratur, tidningstext och texter från sociala medier.

Nyckelord: verbalsubstantiv, lexikologi, SO, BSO, korpusstudie

Ømålsordbogens brugere – nu og i fremtiden

Ømålsordbogen (ØMO) – en stor videnskabelig dokumentationsordbog over de danske dialekter på Øerne ca. 1750–1950 – har netop gennemgået en omfattende teknisk omstillingsproces med udgivelsen af bind 12 (lindost-march). Processen har imidlertid også givet naturlig anledning til overvejelser over ordbogens brugere nu og i fremtiden, især fordi projektet allerede nu sigter mod en udgivelse på nettet.

På hjemmesiden defineres ØMOs brugere som ”en kreds af fagfolk, først og fremmest sprogforskere, etnologer, historikere og museumsfolk, men den har også fået en kreds af ikkeprofessionelle brugere med interesse for dialekter og landbokultur” (jf. nors. ku.dk). ØMOs kernebrugere er således traditionelt blevet tænkt som fagfolk – eller ”særligt interesserede”. Det er samtidig indlysende at en fremtidig publicering på nettet formodentlig vil give ordbogen ikke kun flere brugere, men også brugere med langt bredere, herunder minimale forkundskaber.

1 En række oplysninger, både sproglige og kulturhistoriske, vil være af interesse for brugere uden specialviden. Hvordan skal man komme disse brugere i møde? Formelle vanskeligheder kan let minimeres (fx opløsning af forkortelser), men mere omfattende tilpasninger kan overvejes i en netudgave, fx en udbygning af de særlige kulturhistoriske afsnit med illustrationer og lydklip (jf. temasiden om brødbagning på dialekt.dk eller projektet ”Kunskapsbank för traditionell småskalig matkultur” på sprakochfolkminnen.se).

2 ØMO rummer yderst værdifulde oplysninger af interesse for andre forskere. Disse brugere skal også tilgodeses i en fremtidig netudgave. Skal bestemte oplysningstyper defineres som ”specialiserede” (fx grammatik), med mulighed for at slå disse oplysninger til/fra afhængig af brugerens behov? Man kan også tænke sig en særlig visning for specialister, med mulighed for søgninger og udnyttelse af databasens oplysningstyper.

Præsentationen vil lægge op til diskussion og inddragelse af erfaringer fra tilsvarende projekter.

Lad brugeren komme til orde Hvordan får ordbogen mest muligt ud af brugerhenvendelser?

I løbet af det seneste tiår har Den Danske Ordbog (DDO) modtaget knap 12.000 e-mails fra ordbogens slutbrugere. Henvendelserne varierer i type, fra kommentarer vedrørende konkrete detaljer i ordbogsopslagene over henvendelser om muligt fagligt samarbejde til klager over generelt sprogligt forfald. Det overvældende antal beskeder gør det umuligt for ordbogens redaktører at give samtlige brugere individuelle, tilbundsående svar, hvorfor redaktionen igennem længere tid har ønsket sig bedre redskaber til at sortere og håndtere henvendelserne (Lorentzen & Trap-Jensen 2016).

Målet med denne undersøgelse er derfor, på baggrund af et udsnit bestående af en hel kalendermåneds mails, at etablere en typologi over de hyppigst forekommende henvendelser. Typologien vil siden danne grundlag for udarbejdelsen af en best practice for hver af de kategorier henvendelserne falder i:

Hvornår giver vi brugeren et længere skriftligt svar? Hvornår kan en henvendelse slet og ret ignoreres? Hvornår bør brugeren dirigeres videre til Dansk Sprognævn, vores egen it-afdeling eller noget helt tredje? Og nok så vigtigt: I hvilket omfang kan processen automatiseres?

I eksisterende studier af slutbrugerens rolle i leksikografien har fokus især været på brugerbaseret eller -genereret indhold samt på undersøgelser og analyser af brugerens oplevelse af onlineordbøgers interface (Meyer & Abel 2017). Her er afsættet et andet: Hvordan kan de daglige henvendelser bidrage positivt til ordbogsarbejdet, og hvordan kan arbejdsprocesserne optimeres med henblik på at få mest muligt ud af brugerhenvendelserne?

Referencer

Lorentzen, Henrik & Lars Trap-Jensen (2016): "User studies, user behaviour and user involvement – evidence and experience from The Danish Dictionary".
< http://www.elxicography.eu/wp-content/uploads/2016/03/Lorentzen_Trap-Jensen_ENeL-2016.pdf >. (set december 2018)

Meyer, Christian M. & Andrea Abel (2017): "User participation in the Internet era".
Kapitel 45 i P. A. Fuertes Olivera (Ed.): The Routledge Handbook of Lexicography .
Abingdon/New York: Routledge. 735-753.

Brugere, anvendelse og udvikling

Ordbog over det norrøne prosasprog (ONP) er et ordbogsprojekt ved Københavns Universitet, som registrerer ordforrådet i oldnordiske prosatekster overleveret i norske og islandske håndskrifter fra ca. 1150 til slutningen af middelalderen. Udgivelsen af trykte ordbogsbind blev indstillet i 2004 og efterfølgende blev en stor del af ordbogsmaterialet gjort tilgængeligt på nettet. Onlineudgaven bliver løbende udvidet med nye ordbogsartikler og udvikling af anvendelsesmuligheder.

Ifølge ONPs redaktionelle principper skal ordbogens materiale genspejle originalkilden så nøjagtigt som muligt, hvilket betyder at citaterne ikke er normaliserede. Denne tilgang til materialet stiller højere krav til brugerne end de fleste ordbøger; samtidigt giver ONP adgang til detaljerede informationer om kildematerialet, som kan anvendes på adskillige måder.

I denne præsentation vil vi fokusere på ONP i lyset af konferencens tema (af hvem, for hvem og hvorfor) og diskutere ordbogens brugergruppe, praktisk anvendelse og udvikling.

Vi giver et kort overblik over historikken bag projektet og hvordan ONP adskiller sig fra sine forgængere. Vi diskuterer hvilke forudsætninger brugerne skal have for at kunne orientere sig i ordbogens materiale og drage fuld nytte af det.

Vi analyserer den feedback vi får fra brugerne og hvad den kan fortælle os om anvendelsen af ONP. Vi diskuterer evt. kritikpunkter og hvordan man kan indarbejde dem ved videre udvikling af projektet. Endvidere giver vi konkrete eksempler på hvordan ONP er blevet benyttet i forskningssammenhæng til forskellige formål.

Til sidst vil vi diskutere udviklingen af ONP som digital ressource og hvordan man kan gøre den mere brugervenlig. Her vil vi vise nogle tiltag for at forbedre grænsefladen og hvordan man kunne muligvis udvide brugergruppen med tilkobling til andre databaser og adgang til forskellige digitale hjælpemidler.

Omdannelsen af en flersproget til en monolingval ordbog

I nogle år har den flersprogede skandinaviske ordbog ISLEX været gratis tilgængelig på Árni Magnússon instituttets hjemmeside (se islex.dk). Ordbogen er mellemstor med ca. 50.000 opslagsord og kildesproget er islandsk. Efter åbningen af ISLEX var man straks klar over at materialet egnede sig godt til dannelsen af flere ordbøger, og i fortsættelsen indledtes arbejdet med en islandsk-fransk ordbog (LEXIA).

Man fandt det endvidere oplagt at bruge ISLEX-materialet til en ny islandsk monolingval ordbog, men dengang var der kun én sådan ordbog tilgængelig på webben. Den nye ordbog skulle være en fyldig beskrivelse af det islandske sprog som det tales og skrives nu, og fik navnet Ordbog over moderne islandsk.

Størstedelen af ordforrådet var allerede udarbejdet og defineret, hovedsagelig på grundlag af korpusser. Lemmaerne var blevet inddelt i definitioner, bøjningerne blev givet med links til en database med bøjningsparadigmaer og udtalen af samtlige opslagsord blev indlæst og fandtes på lydfiler. Mange steder var der links mellem opslagsordene og der fulgte illustrationer med et antal ord. I ordbogen fandtes der endvidere mange sprogbrugseksempler og ordforbindelser.

Da ISLEX-databasen blev kopieret for den monolingvale ordbog var der allerede en stor del af arbejdet gjort, og det som især manglede var fyldige ordforklaringer på islandsk, mens de eksisterende midlertidige ordforklaringer havde været tænkt som hjælp for ISLEX-oversætterne. De seneste år har det redaktionelle arbejde ved Ordbog over moderne islandsk især bestået i at udarbejde nye definitioner med islandske brugere som målgruppe. Dette arbejde har medført en grundlig revidering af indholdet i ISLEX og LEXIA da disse tre ordbøger hænger uløseligt sammen og revisionen sker på de tre steder samtidig. Ordbog over moderne islandsk er tilgængelig på webben (islenskordabok.is).

Erfarenheter av nätprojektet för den finska dialektordboken

Den finska dialektordboken *Suomen murteiden sanakirja* redigeras på Institutet för de inhemska språken i Finland. Den är ett storverk som ger en heltäckande beskrivning av ordförrådet i de finska dialekterna, sammanlagt planeras den omfatta 300 000–350 000 uppslagsord.

På 2010-talet övergick utgivningen från att trycka ordboken till att publicera den på nätet. Det första avsnittet som publicerades enbart på nätet omfattade uppslagsorden från *keus* (lockrop till husdjur) till *kyntsöttää* 'stå eller sitta orörlig'; det publicerades den 28 februari 2012.

Samtidigt påbörjades ett projekt för att bearbeta de åtta tidigare tryckta banden av ordboken för utgivning på nätet. Projektet kunde avslutas i maj 2018. I redaktionen fortsätter vi nu med att årligen publicera ungefär 6 000 nya artiklar på nätet; våren 2019 publiceras avsnittet *loha* 'båt i dåligt skick' – *hysmä* 'skarp krök på med(e)'.

Vilka var de viktigaste erfarenheterna av projektet?

- Det lönar sig att utnyttja de möjligheter som nätet erbjuder: en nätdorbok behöver inte vara en exakt kopia av den tryckta ordboken.
- Det är viktigt att kontrollera att funktioner som planerats att ske automatiskt i samtliga fall fungerar exakt som de borde.

I föredraget kommer vi att ge exempel på utmaningar som arbetet med nätdorboken medfört.

Den finska dialektordboken *Suomen murteiden sanakirja* finns på adressen:
<http://kaino.kotus.fi/sms/>

Students' terminology work: what about, whom for, and why? Compiling terminological glossaries as a part of students' translator competence acquisition

The aim of this paper is to present some practical and technological aspects of competence-based translator training at a particular course of terminology work for translators and interpreters at the University of Helsinki. The paper also provides students' point of view: the results of the surveys among students are presented on how the course has helped them to evolve as translators and what benefits and possible disadvantages they see in the course curriculum.

Terminology work is activities concerned with the systematic collection, description, processing and presentation of concepts and their designations (ISO 1087-1:2000). The Terminology Framework of the EU officially recognizes terminology as a core business in the field of translation (TermCoord 2016). In addition to the EU institutions, terminology related skills can be required from translators in governmental organizations and translation companies.

Working life requirements are reflected in the European Master's in Translation (EMT) framework, which is an essential standard in translator training. According to EMT (2017), students are expected to master systems of concepts and terminology as well as to possess technological competence to implement present and future translation technologies within the translation process. Göpferich (2009: 22) also states in her competence model that terminological sub-competence is, among others, an important part of translator competence.

Universities take into consideration working life requirements and offer courses that train skills in terminology work, glossary compilation and terminology management. At the University of Helsinki, these skills are acquired on the course of Translator's and Interpreter's Terminology Work. The paper addresses the issues of how the course is organized and in which way compilation of terminological glossaries enhances the acquisition of translator competences with a focus on practical issues and technological dimension.

References

ISO 1087-1:2000 Terminology work -- Vocabulary -- Part 1: Theory and application. Edition: 1

Göpferich, S. (2009). Towards a model of translation competence and its acquisition: The longitudinal study TransComp. In: Göpferich, S., Jakobsen, A. L. & Mees, I. M. (eds.) Behind the Mind: Methods, Models and Results in Translation Process Research. Copenhagen: Samfundslitteratur, 11-37.

TermCoord 2016 = Terminology in the European Parliament and the EU institutions. Directorate-General for Translation. http://www.termcoord.eu/wpcontent/uploads/2016/12/Terminology_in_the_EP_and_EU_Institutions_2016_RivneUKR_EN.pdf
Accessed 14.12.2018

EMT (2017). European Master's in Translation. Competence Framework. https://ec.europa.eu/info/sites/info/files/emt_competence_fwk_2017_en_web.pdf
Accessed 14.12.2018

Lennart Larsson

Svenska Akademiens ordboksredaktion, Lund

Bruklighetsangivelserna i SAOB

I och med att Svenska Akademiens ordbok (SAOB) snart kommer att vara avslutad i sin första utgåva och en ny reviderad version är i antågande aktualiseras en mängd frågor när det gäller ordbokens innehåll och presentationssätt; det finns utan tvekan en stor förbättringspotential.

Inte minst gäller det bruklighetsangivelserna, ett område där SAOB kännetecknas av en överflödande mångfald av olika (inte sällan mer eller mindre dunkla eller obsoleta) etiketter när det gäller sådant som frekvens, social och regional begränsning, stilnivå och fackspråkighet. I mitt föredrag kommer jag i första hand att diskutera frekvens – det som i SAOB-sammanhang går under benämningen total bruklighetsinskränkning – och de problem som sammanhänger med dess etikettering: det handlar om angivelsernas mångfald, deras beroende av redaktörernas subjektiva bedömningar och – sist men inte minst – deras avhängighet av när artikeln skrevs.

Rasmus Rask som katalysator för finska vetenskapliga ordböcker och finsk grammatik 1818–1819

Den berömde danske språkforskaren Rasmus Rask (1787–1832) besökte Finland 1818, eftersom han var intresserad av finsk-ugriska språk. Rask lärde sig finska under vägledning av Gustav Renvall, som var docent i finska språket vid Åbo akademi. När Rask frågade om det fanns någon aktuell finsk ordbok i universitetsbiblioteket, visades han manuskriptet av Christfrid Gananders "Nytt finskt lexikon", som hade blivit färdig redan 1786–1787.

Rask började organisera publiceringen av Gananders ordbok med hjälp av Renvall och en rik mecenat i S:t Petersburg. Resultatet blev emellertid en annan bok, Gustav Renvalls "Suomalainen Sana-Kirja" (1823–1826). Renvall hade använt Gananders manuskript som underlag men ersatt svenska med tyska i sitt trespråkiga (finsk-latinsk-tyska) verk. Renvalls ordbok är också ganska koncentrerad (antalet lemman är ca 22 000, medan det hos Ganander är över 35 000), trots att Renvall utnyttjade även tidigare ordböcker och senare ordlistor.

Redan Lauri Hakulinen (1967) skisserade de viktigaste skillnaderna mellan Renvalls och Gananders verk, men Birgitta Romppanen (2001) har påpekat att en undersökning av dessa förändringar är ett forskningsprojekt i sig själv. I vårt föredrag presenteras en mer detaljerad bild av skillnaderna mellan Renvalls och Gananders ordartiklar.

Renvalls ordbok blev aldrig särskilt populär, men den blev en viktig källa och modell för senare, vanligen tvåspråkiga ordböcker. En annan frukt av Rasks korta besök i Åbo, hans egen finska grammatik, fick ett annat öde: verket blev nedgrävt i arkivet, men genom Renvalls "Finsk språklära" (1841) fick Rasks idéer inflytande även på den finska grammatiska traditionen.

Litteratur

Hakulinen, Lauri 1967: Suomen sanakirjoista. I: *Kirjoituksia suomen kielestä*, 83–107. Tietolipas 51. Helsinki: SKS.

Romppanen, Birgitta 2001: *Från målspråk till källspråk. Utvecklingen av den finsk-svenska och den enspråkiga finska ordboken*. Acta Wasaensia n:o 91. Wasa: Universitas Wasaensis.

15 år med Ordbogen.com – fra banebrydende nyt koncept til konventionel ordbogsportal?

I september 2003 lancerede et par it-iværksættere Ordbogen.com, noget så naivt som en online engelskordbog man skulle betale for at benytte. Ud over at forlange penge for produktet så de ovenikøbet stort på traditionelle redaktionsprincipper og målgruppeovervejelser og tilbød én engelskordbog til alle og lod den tage form efter brugernes behov ved at tilføje ord ud fra deres frekvens i søgeloggerne.

Dengang var konceptet banebrydende, men i dag er stort set alle danske ordbogsudgivere til stede online og inddrager som en selvfølge søgeloggerne i deres redaktionelle overvejelser. Og det er ikke det eneste der har ændret sig i de 15 år der er gået siden da: Først måtte vi sande at engelskordbogens produktionslogik ikke uden videre kunne overføres til vores efterfølgende tosprogsordbøger, siden fik vi også ansvaret for ordbøger udarbejdet efter helt andre principper, og senest har vi i en periode måttet arbejde uden brug af de søge-logger der ellers udgjorde ryggraden i hvordan vi forstod os selv som Ordbogen. Er vi nu bare én ordbogsportal blandt mange?

I dette foredrag vil to af Ordbogens redaktører fortælle om ændringer i vores arbejds-metoder gennem tiden, dels for vores tosprogsordbøger, dels for nogle af de danskordbøger vi har overtaget fra Politikens Forlag. For de førstnævnte har vi skullet imødekomme skole-brugernes krav om mere grammatik og mere pædagogisk stringens. For de sidstnævnte har en eksisterende arbejdsdeling mellem ordbøger målrettet hver sin brugergruppe skullet balanceres mod samme brugeres ønsker. På et tidspunkt fik væksten i antallet af ordbøger på Ordbogen.com desuden den uventede bivirkning at søgeloggerne til sidens fælles søgefelt forplumrede så vi til sidst helt måtte se bort fra dem. Det tvang os til at nøjes med hvad vi kunne udlede af vores kontakt med brugerne via mail og chat, men førte også til den indsigt vi vil argumentere for til slut: at graden af direkte brugerkontakt er det der i dag adskiller os fra andre ordbogsportaler.

Gunn Inger Lyse

Revisjonsprosjektet for *Bokmålsordboka* og *Nynorskordboka*

Universitetet i Bergen

Ut med *adamsslekt* og inn med *arveprinsesse*? Leksikografiske metoder i Revisjonsprosjektet for *Bokmålsordboka* og *Nynorskordboka*

I dette innlegget presenteres kildematerialet, verktøy og metoder som brukes og utvikles i det norske Revisjonsprosjektet for *Bokmålsordboka* og *Nynorskordboka*. I dette prosjektet skal leksikografene revidere to eksisterende ordbøker parallelt, én på bokmål og én på nynorsk. Hvordan kan leksikografene jobbe effektivt, klokt og etterrettelig for å peke ut et modernisert og relevant ordtilfang i begge målformene?

På den ene siden bruker vi de vitenskapelige Språksamlingene ved Universitetet i Bergen. På den andre siden bruker vi søkeverktøy for å analysere tekst på ord- og setningsnivå. Vi vil vise hvordan tekstkorpus er en god empirisk kunnskapskilde når det gjelder **lemmatilfang, fraser og flerordsuttrykk, ordbetydninger, den syntaktiske oppførselen til ord og brukseksempler**. Vi vil samtidig vise eksempler hvor korpus potensielt også er misvisende.

Leksikografene bruker verktøy for språkanalyse som er tilgjengelig gjennom språkinfrastrukturen CLARINO (<https://clarin.w.uib.no/>). **Corpuscle** er et verktøy for å søke på ord og fraser. Gjennom revisjonsprosjektet er Corpuscle utvidet til å søke i flere korpus samtidig (til sammen ca. 2,4 milliarder ord). For eksempel kan vi gjøre komplekse søk i Nynorskkorpuset, Leksikografisk Bokmålskorpus og Nasjonalbibliotekets frie tekster samtidig. Søketreffene kan sorteres med frekvens over tid og distribusjon i ulike korpus, hvilket forenkler leksikografens empiriske arbeid betydelig. Infrastrukturen **INESS** lar oss søke etter syntaktiske konstruksjoner som ikke er så lett søkbare i et tradisjonelt tekstkorpus. NorGramBank inneholder syntaktisk analysert materiale for norsk (til sammen ca. 80 mill. ord). Vi vil vise hvordan vi kan hente ut «wordsketch»-lignende informasjon, som de vanligste verbrammene til et verb (*gjifte seg, gjifte bort, gjifte seg til*) eller typiske eksempler på hvordan *at*-setninger styres av preposisjoner (*etter at, uten at*).

Relevante lenker

Revisjonsprosjektet for *Bokmålsordboka* og *Nynorskordboka*: <https://ordbok.uib.no>

Corpuscle (søk på ordnivå): <https://clarino.uib.no/corpuscle>

INESS (søk på setningsnivå): <https://clarino.uib.no/iness>

Carina Nilstun

Det Norske Akademis ordbok, Kunnskapsforlaget, Oslo

Den elektroniske ordbokens innvirkning på den praktisk leksikografi

Datamaskinen revolusjonerte leksikografien da den ble tatt i bruk i ordboksredaksjonene rundt om. Det ble utviklet digitale redigeringsverktøy og elektroniske tekstkorpus. Forandringene gikk i hovedsak på metodikken, produktet var fortsatt en bok.

Hele tiden har den praktiske leksikografien vært sterkt preget av bokformatet, og da særlig av plassmangelen. Ordboken som teksttype har derfor tradisjonelt vært sterkt kondensert, blant annet med utstrakt bruk av forkortelser, både allmenne og verkinterne. I tillegg var teksten ofte satt i liten skrift og med lite luft og få grafiske virkemidler. Norsk Riksmålsordbok (RMO) var en slik ordbok. Godt innhold, men lite brukervennlig i sin utforming. Med ordbøker i elektronisk format åpnet nye muligheter seg, ikke minst ved at man kunne lenke til og fra ord og betydninger.

Mange ordbøker er i dag elektroniske, men de fleste finnes i tillegg i papirutgave og har ofte en slik som utgangspunkt. De er altså fortsatt knyttet til og dels begrenset av bokformatet. Når vi snakker om ordbøker som bare er elektroniske, må disse sees som en egen ordbokstype – de er noe annet enn papirordbøker tilrettelagt for skjerm. Det Norske Akademis ordbok (NAOB) bygger på RMO, men finnes kun som elektronisk nettordbok. Med rendyrkede elektroniske ordbøker som NAOB kan leksikografien og de leksikografiske produktene igjen ta store utviklingssteg. Plassbegrensningen er opphevet, dét gir store muligheter. I den digitale ordboken er det lett å lage terskelen for å komme med innspill lav, og man har mulighet til å registrere brukeradferd. Både kunnskapen om hva brukerne søker på og de konkrete innspillene de kommer med, er med på å påvirke den videre redigering.

Denne ordboktypen er ikke ferdig utforsket og utformet. I overgangen fra det analoge til det digitale finnes det på ordboksfronten fortsatt upløyd mark. Jeg vil i mitt foredrag se på mulighetene, med utgangspunkt i redigeringen av NAOB.

Christian-Emil Ore

Universitetet i Oslo

Å ta Hans Ross på ordet – Norsk Ordbog (1895) og Aasens Norsk Ordbog (1873) som en helhet

Det 19. århundre var en tid for aktiv målreising i Europa. Det innebar at man ønsket å skape et litteratur- og bruksspråk basert på felles leksikalske, grammatiske og syntaktiske drag i et sett av beslektete dialekter. I Norge gjorde lingvisten og leksikografen Ivar Aasen (1813-1896) et grunnleggende arbeid gjennom mangeårige feltstudier. Det resulterte blant annet i en grammatikk (1848, 1864) og Norsk Ordbog (1873) som la grunnlaget for det nynorske skriftspråket.

En noe mindre kjent språkforsker og leksikograf er Hans Ross (1833–1914). Som gymnasiast ble han introdusert til Aasens arbeid, og som språkforsker arbeidet han med dokumentasjon av norsk språk. Hans omfattende feltstudier resulterte i en ordbok med to tillegg utgitt i perioden 1890–1895 og fire ekstra tillegg (1902, 1910, 1912 og 1913). Ross kalte ordboka Norsk Ordbog. Tillæg til «Norsk Ordbog» af Ivar Aasen. Med det beskjedne navnet (ordboka har 10% flere oppslagsord og omlag 3/4 finnes ikke hos Aasen) og det usedvanlig selvkritiske forordet, har nok Ross satt sitt lys under en skjeppe.

Begge ordbøkene ble digitaliserte i Norsk Ordbok 2014-prosjektet. De har siden blitt analysert og gitt en detaljert TEI-koding. I tillegg er oppslagsordene normalisert til moderne nynorsk og artiklene er lenket inn i Metaordboka, som er en felles indeks på leksikalsk enhetsnivå ('lexical item') for 76 ordbøker. Det gjør det mulig å sammenstille og analysere ordbøkene både på oppslagsordnivå og på leksikalsk enhetsnivå. For sistnevnte finnes knappe 11.000 (27%) både hos Aasen og Ross. For disse er artiklene i Ross en utvidelse av artiklene i Aasen.

Ordbøkene er sammenstilt etter edisjonsfilologiske prinsipper. Leseren kan se ordbøkene som separate tekster eller som et hele der en har tatt Ross på ordet og plassert hans artikler som nettopp utvidelser av Aasens artikler og som nye artikler i en Aasen-Ross ordbok.

Mikkel Ekeland Paulsen

Revisjonsprosjektet for Bokmålsordboka og Nynorskordboka
Universitetet i Bergen

Sammensatte ord – et sammensatt problem

Av alle mekanismer i det norske språket er trolig sammensetningsfunksjonen blant de mer produktive og nyordsgenererende vi har. Bare i forrige setning kunne jeg uten større anstrengelser spleise ordene *ny* og *ord* til *nyord*, som jeg i neste operasjon festa til *genererende*. Selv om denne sammensetninga hverken har treff i Norsk aviskorpus eller Leksikografisk bokmålskorpus, er den såpass gjennomiktig rent semantisk at man kan gå ut fra at totalbetydninga gir seg selv hvis man vet betydninga av delene. Det faktum at *nyordsgenererende* er både infrekvent og semantisk transparent, kan en leksikograf bruke som argument for at sammensetninga ikke bør stå i ordboka.

Hvis man inntar et leksikografisk perspektiv på sammensetninger i norsk, melder det seg fort en rekke spørsmål hvor svaret ikke nødvendigvis gir seg selv. Hva kjennetegner en «vanlig» sammensetning? Er sammensetninga *andedam* mer eller mindre sentral i det norske språket enn *andebryst*?

En ordbokredaktørs svar på disse spørsmålene vil med nødvendighet ha stor innvirkning på lemmautvalget i den gjeldende ordboka, i og med at tilfanget av sammensetninger i norsk språkbruk er betydelig, både i typer og forekomster. I det annoterte norske aviskorpuset, Forskning.no-korpuset og Leksikografisk bokmålskorpus kan man samlet finne ca. 550 unike sammensetninger med førsteleddet *terror* alene. Ordbokredaktører er med andre ord tvunget til å operere med et sett kriterier for å luke ut de sammensetningene som av ulike grunner ikke fortjener et oppslag i ordboka.

I dette innlegget vil jeg drøfte ulike semantiske kriterier for å innlemme/diskriminere sammensatte ord i enspråklige allmennordbøker. Hvor gjennomiktige er egentlig såkalte «gjennomiktige» sammensetninger? Og hvordan kan man i leksikografisk øyemed trekke en grense mellom leksikaliserte og ikke-leksikaliserte sammensetninger?

Har bokmålsbrukere *bein* eller *ben* i *nesa* – eller er det *nesen*?

Om idiomatisk variasjon i norske ordbøker og faktisk språkbruk

I dette innlegget presenteres noen funn fra en større undersøkelse av idiomatisk variasjon i og mellom de norske målformene bokmål og nynorsk. I en empirisk delstudie i doktorgradsprosjektet mitt sammenligner jeg hvordan fire norske ordbøker framstiller det norske fraseinventaret med frasebruken som kan observeres i skriftspråkskorpus: Hvor omfattende er de fraseologiske forskjellene mellom bokmål og nynorsk om en legger *Det Norske Akademis Ordbok*, *Bokmålsordboka*, *Nynorskordboka* og *Norsk Ordbok* til grunn, og speiler ordbøkene den variasjonen som er i hver av målformene? Hva kjennetegner den autentiske frasebruken i bokmål og nynorsk i et variasjonsperspektiv? Og er det forskjell på målformene med tanke på om brukerne holder seg innafor den fastsatte normen når de skriver fraser?

De norske skriftnormalene tillater i utgangspunktet stor variasjon, og i mange tilfeller er det heller ingen entydige svar på hva som er riktig og galt. En normal trekker opp klare grenser mellom hva som er i «innafor» og «utafor», men i levende bruksspråk er det alltid ei gråsoner mellom det udiskutabelt riktige og det udiskutabelt gale.

Disse gråsonene støter en ikke minst på i de faste frasene, da den offentlige normeringa ikke omfatter dette nivået av språket. Hva som forårsaker fraseologisk variasjon og til dels normbrudd er delvis ulikt i bokmål og nynorsk, og jeg knytter det til ulike puristiske ideal i de to målformene, jf. George Thomas' (1991) bruk av purismebegrepet. I materialutvalget mitt reflekteres det ved at idiomgruppene i bokmål har en veksling i ortografi og morfologi med ulik stilistisk virkning, mens purismen i nynorsk gjør at det er mer interessant å se på idiomatisk variasjon i leksikon og syntaks.

Referanser

Thomas, George. 1991. *Linguistic purism*, Studies in language and linguistics. London: Longman

Varemerkenavn i ordbøker

Svensén (2004:505) peker på konflikten som kan oppstå mellom leksikografens ambisjon om å beskrive den faktiske språkbruken og varemerkeinnhaverens ønske om å beskytte sine økonomiske interesser. Utgangspunktet for denne typen konflikter er at varemerkeeieren ikke ønsker at varemerket går inn i dagligtalen som en fellesbetegnelse på produkter av samme art (blir degenerert), fordi det på sikt vil kunne føre til at konkurrenter bruker samme navn i sin markedsføring.

Sentrale spørsmål for leksikografen blir da om det påståtte varemerket er brukt i allmennspråket *før* det lanseres som varemerke, og i hvilken grad betegnelsen faktisk har blitt appellativisert i språksamfunnet. Et viktig skille på veien mot appellativisering er om betegnelsen er ført opp i ordbøker eller ikke. Derfor spiller ordbokredaksjoner ei viktig rolle i slike tilfeller.

I innlegget vil vi redegjøre for hvordan varemerkenavn er behandlet i noen store danske, svenske og norske ordbøker. Vi skal ta for oss både nyere og eldre varemerkenavn, også for å se om rettstvister fra nyere tid kan ha påvirket ordbokredaksjonenes måte å behandle varemerkenavn på. Eksempler på nyere tilfeller kan være det norske *vipps* (app fra DNB Bank som gjør det mulig å sende penger raskt og enkelt), som grafisk ligger svært nær (og fonetisk er identisk med) interjeksjonen *vips* (jf. Grønvik 2017), og det spanske *buff* (hodeplagg som kan brukes både som hals og lue), som mangler en lignende ekvivalent i allmennordforrådet. Eksempler på eldre varemerkenavn som med tiden er godt etablert i allmennspråket kan være *musak*, *lego*, *levis* og *tabasko*.

Litteratur

Grønvik, Oddrun. 2017. «Varemerke og allmennspråk». I Karlsen, Knut E., Dagfinn Rødningen og Håvard Tangen (red.). *I teneste for nynorsken. Heidersskrift til Olaf Almenningen 70 år*. Novus forlag, Oslo.

Svensén, Bo. 2004. *Handbok i lexikografi*. Norstedts Akademiska Förlag, Stockholm.

Caroline Sandström & Harriet Eriksson

Institutet för de inhemska språken, Helsingfors

Suffixavledda verb av finskt ursprung i de svenska dialekterna i Finland

Vid redigeringen av artiklar inom det aktuella avsnittet p- i *Ordbok över Finlands svenska folksmål* har det i materialet för ordboken framträtt en grupp flerstaviga suffixavledda verb på *-isa* (ex. *plottisa, popisa*) och *-ista* (ex. *pomista, potista*). Verb av denna typ förekommer även i andra avsnitt av ordboken och hänför sig ofta semantiskt till betydelser som 'pladdra, skrävla; käbbla; muttra; sluddra; bullra; puttra; skramla'. Det verkar huvudsakligen handla om finska lån i de svenska dialekterna i Finland; motsvarande verb saknas nämligen i ordböcker som beskriver sverigesvenska varieteter. För lejonparten av dessa verb har vi genom etymologiska kontroller i finska ordböcker och arkiv kunnat fastställa finska förebilder. I det finlandssvenska dialektmaterialet förekommer dock ett fåtal suffixavledda verb (t.ex. *bjänista, präppista*) som inte har någon entydig finsk motsvarighet. Detta tyder på att suffixen blivit produktiva ordbildningselement som har kunnat ge upphov till nybildningar i de svenska dialekterna.

I vårt bidrag kommer vi utifrån sökningar i nätordboken *Ordbok över Finlands svenska folksmål* (<http://kaino.kotus.fi/fo/>) inom avsnittet A–P att göra en systematisk kartläggning över suffixavledda verb på *-isa* och *-ista*, men även *-itta*, *-otta* och *-utta* med betydelser som hänför sig till olika former av tal, oväsen och återgivning eller härmande av ljud. Vårt syfte är att undersöka om det finns regionala skillnader i fördelningen mellan olika suffix och betydelser hos de suffixavledda verben i ordbokens material, men också att ur ett kontrastivt språkkontaktperspektiv undersöka om suffixen är produktiva i de svenska dialekterna.

«Å nei, det ordet er ikkje lov på nynorsk! Eller ...?» Om parallell redigering av to norske ordbøker

Skal importordet *glamour* uttalast ulikt på nynorsk og bokmål? Og stemmer det at ein ikkje kan vera *respektfull* på nynorsk, slik ein kan det på bokmål, men at det derimot går fint å vera *respektlaus* på nynorsk, men ikkje på bokmål? Det kan sjå slik ut om ein sjekkar oppslagsorda i dei to norske handordbøkene *Bokmålsordboka* og *Nynorskordboka*.

Dei to ordbøkene har tidlegare vore redigerte av to ulike redaksjonar som har jobba delvis uavhengig av kvarandre. Det har ført til at det er ein del skilnader mellom dei to ordbøkene både i lemmautval og i utforminga av ordboksartiklane. Ein del skilnader kjem av reelle skilnader mellom dei to målformene, medan andre er meir tilfeldige.

Frå å vera to papirutgåver der brukarane slo opp i boka for den målforma dei var interesserte i, og stort sett heldt seg til den informasjonen dei fann der, er situasjonen i dag ein heilt annan. Dei to ordbøkene er digitaliserte og gratis tilgjengelege på nettet (ordbok.uib.no). Begge ordbøkene ligg på den same plattformen, og det er mogleg å søkja parallelt i dei, noko mange brukarar nyttar seg av. Når ein får opp ordboksartiklane for nynorsk og bokmål ved sida av kvarandre, er det lett å samanlikna, og skilnadene vert veldig tydelege. Tilbakemeldingane frå brukarane tyder på at dei gjerne legg meir i desse skilnadene enn det som er meint frå redaksjonen, og at dersom dei til dømes finn eit ord berre i den eine ordboka, konkluderer dei med at det ikkje kan brukast i den andre målforma.

Bokmålsordboka og *Nynorskordboka* skal no reviderast gjennom eit femårig prosjekt ved Universitetet i Bergen, og prosjektet er lagt opp slik at begge ordbøkene vert redigerte samstundes av dei same redaktørane. Ein viktig del av oppdraget er å jamna ut dei umotiverte skilnadene mellom ordbøkene. I innlegget vil vi dela erfaringane så langt og drøfta kva fordelar og utfordringar ei slik parallell redigering fører med seg.

Augmented writing: nye muligheder og nye teorier

Vi lever i disruptionens tidsalder og er lige nu vidne til en lang række store omvæltninger og fundamentale skift i f.eks. banksektoren, forsikringsbranchen, advokatbranchen og sundhedssektoren, idet især automatisering og robotisering (RPA) af rutineprægende opgaver samt anvendelse af kunstig intelligens (AI) efterhånden er blevet så udbredt, at der reelt er tale om en omvæltende disruption.

Denne udvikling er også kommet til f.eks. kommunikationsbranchen, hvor der i stadig stigende omfang anvendes kunstig intelligens til at understøtte journalisters research og arbejde, ligesom nye teknologier som f.eks. Augmented Writing nu hjælper forfattere og skribenter på en lang række områder. Og i en tidsalder, hvor kreativt, nuanceret, korrekt og afvekslende sprog er af afgørende værdi for virksomheder, så kan Augmented Writing blive den næste nye store disruption inden for computer-assisted writing.

Formålet med dette bidrag er således at undersøge og diskutere, hvad Augmented Writing er og ikke mindst hvordan leksikografi som en teoretisk og praktisk disciplin kan bidrage til dette felt. Formålet med bidraget er ligeledes også at diskutere en række nye muligheder for leksikografien i dette felt samt endelig også at diskutere en række teoretiske overvejelser om det jeg kalder Lexicographic Augmented Writing.

Bidraget trækker på en empirisk analyse af ni specielt udvalgte og meget forskellige Augmented Writing-tjenester, som i varierende grad trækker på leksikografiske data, leksikalske data, grammatiske data, retskrivningsdata, stilistiske data, plagiatdata, følelsesintensitetsdata etc. etc. De empiriske data omfatter også en opgørelse over, hvordan de ni tjenester integreres i forfatterens eller skribentens daglige arbejde og i hvor høj grad de reelt er totalt integreret i brugerens workflow. Bidraget vil således diskutere, hvad Augmented Writing reelt er og hvordan leksikografien kan bidrage til dette felt. Denne diskussion omfatter en række teoretiske forslag til det jeg kalder Lexicographic Augmented Writing, herunder en diskussion af hvem det er, der producerer moderne leksikografiske værktøjer og hvilke jobfunktioner, de understøtter, og jeg vil på basis af analysen af de empiriske data opstille en række teoretiske overvejelser, som kan bidrage til feltet.

Litteratur

Boost Editor (2019): Boost Linguistics. < <https://www.boostlinguistics.com/> > (januar 2019)

Google News Initiative (2019): Google News Initiative.
< <https://newsinitiative.withgoogle.com/> > (januar 2019)

Grammarly (2019): Grammarly. < <https://app.grammarly.com/> > (januar 2019)

Persado (2019): Persado. < <https://www.persado.com/products/pro-email> > (januar 2019)

Phrasee (2019): Phrasee. < <https://phrasee.co/> > (januar 2019)

ProWritingAid (2019): ProWritingAid. < <https://prowritingaid.com/> > (januar 2019)

TapRecruit (2019): TapRecruit. <<https://taprecruit.co/> > (januar 2019)

Textio (2019): Textio. < <https://textio.com/> > (januar 2019)

WordRake (2019). WordRake: < <https://www.wordrake.com/> > (januar 2019)

Write Assistant (2019): Write Assistant< <https://beta.ordbogen.com/da/writeassistant> > (januar 2019)

Anden litteratur

Latilya W. Butler (2017): *The Impact of Computer-Assisted Writing on Improving Writing Scores for Urban Eighth-Grade Students*. Doctoral dissertation. Nova Southeastern University. Retrieved from NSUWorks, Abraham S. Fischler College of Education

Simonsen, Henrik Køhler (2017): Den leksikografiske forretningsmodel. In: *LexicoNordica* 24, 2017.

Tarp, S., Fisker, K., & Sepstrup, P. (2017). L2 writing assistants and context-aware dictionaries: New challenges to lexicography. *Lexikos*, 27(1), 494-521.

Multimediale ordbøger: hvordan og hvorfor?

Hvis, som det ofte påstås, at et billede siger mere end 1000 ord, hvad kan man så ikke forvente af videoers forklarings- og illustrationskraft? Videoer er i dag uhyre udbredte, og undersøgelser viser da også, at vi ser mere end 500 millioner timer af video på YouTube hver eneste dag. Og samtidig husker brugerne 95% af et budskab i en video, hvorimod de ved at læse en tekst kun husker 10% af budskabet.

Leksikografi er som bekendt i en voldsom transformationsproces, og vi bør i denne proces også se på de enkelte leksikografiske komponenter og datatyper. Måske er det på tide at skifte nogle komponenter ud med andre – eller supplere eksisterende komponenter - og i langt større udstrækning anvende videoer? I vores gennemgang af litteraturen synes undersøgelser af videoer som leksikografisk ressource at være et temmelig uopdyrket område, og vi vil derfor gerne komme med et forskningsbaseret bidrag, der skal rette op på dette forhold.

Formålet med vores bidrag er at undersøge og diskutere, hvorledes leksikografi kan anvende videoer som leksikografisk ressource og i hvilke brugssituationer videoer er særligt anvendelige med henblik på at øge ordbøgernes nytteværdi. I denne forbindelse vil vores bidrag indeholde specifikke overvejelser om, hvordan videoer kan integreres som funktionelle komponenter og hvilke elementer videoer kan eller bør indeholde i forskellige brugssituationer og til forskellige typer brugere. Endelig vil vi overveje og måle fordele og ulemper ved videoer og komme med anbefalinger til hvilke typer ordbøger og ordbogsformater forskellige typer videoer egner sig bedst til.

Vi vil her formidle de foreløbige resultater af vores teoretiske overvejelser samt gennemgå en række modeller for anvendelsen af både passive videoer og (inter)aktive videoer som informative, leksikografiske ressourcer i multimediale ordbøger.

Terje Svardal
Universitet i Bergen

Nynorskordboka og Bokmålsordboka – gammalt og nytt redigeringsgrunnlag

Nynorskordboka (NOB) og Bokmålsordboka (BOB) skal reviderast i perioden 2018 – 2023. Eksisterande artiklar skal gjennomgåast og det vil bli lagt til nytt stoff der det er naudsynt. I dette innlegget skal eg gå inn på tilfanget som låg til grunn for dei første utgåvene av Nynorskordboka (NOB) og Bokmålsordboka (BOB) og sjå på kva ei utviding av tilfanget og bruk av tekstkorpus har å seie for innhaldet og kvaliteten på ordbøkene.

Ásta Svavarsdóttir

Árni Magnússon instituttet for islandske studier / Islands universitet

Ordbøger og sprogstandardisering. Islandske ordbøger i 1800-tallet og deres rolle i standardiseringsprocessen

Ifølge Einar Haugens (1966) model for udviklingen af standardsprog, består standardiserings-processen af fire faser: valg (af norm), kodificering, antagelse (eller implementering) og elaborering. Ordbøger – samt grammatikker, retskrivningsordbøger og stilistiske vejledninger – spiller en vigtig rolle i kodificeringsfasen og desuden kan de også være af betydning for implementering af normen, dvs. “the gradual diffusion and acceptance of the newly created norm across speakers as well as across functions” (Deumert & Vandenbussche 2003:7). Standardisering af det moderne nationalsprog i Island – i almindelighed anset og omtalt som ”sprogreform” – foregik samtidigt med den politiske selvstændighedskamp i det 19. og tidlige 20. århundrede. I samme periode udkom der et antal islandske ordbøger. I indlægget analyseres og diskuteres deres direkte og indirekte indflydelse på den sprogestandard som var under udvikling, især mht. de ordbøger hvor beskrivelsen af islandsk står i fokus. Hvem lavede disse ordbøger, for hvem og hvorfor?

Referencer

Deumert, Ana & Wim Vandenbussche. 2003. Standard languages: Taxonomies and histories. I: Deumert & Vandenbussche (red.), *Germanic Standardizations. Past to present*, s. 1–14. Amsterdam/Philadelphia: John Benjamins.

Haugen, Einar. 1966. Dialect, Language, Nation. *American Anthropologist* 68(4):922–935

Tilblivelsen af fagordbøger og termsamlinger i Island

I foredraget vil jeg gøre rede for min undersøgelse af islandske fagordbøger og termsamlinger. For det første betragtes hvem tager initiativ til at udarbejde fagsprogsressurser. I Island er det almindeligt at såkaldte termudvalg tager sig af termarbejde og mange sådanne udvalg er blevet oprettet i de sidste årtier. Der iagttages hvordan disse udvalg er blevet til, f.eks. undersøges om der findes årsager som førte frem til deres oprettelse; om udvalgene er opstået af sig selv eller de udnævnes formelt af fagforeninger eller institutioner. Ligeledes iagttages termressurser fra individer som også er udbredte og de forholdsvist få eksempler på termressurser på myndighedernes initiativ. Endvidere diskuteres hvordan Árni Magnússon-instituttet for islandske studier involveres i tilblivelsen af fagsprogsressurser, men instituttet er center for termarbejde i Island.

For det andet undersøges hvilke motivationer ligger bag fagordbøgernes tilblivelse. Islandske fagsprogsressurser er af forskellig art, og der kan ligge forskellige motivationer bag dem, f.eks. sprogrensning, sprogrøgt, effektiv sprogbrug, dokumentation af fagområdets termer. Der søges bl.a. svar på disse spørgsmål: Er det let at finde frem til motivationerne bag termressurserne eller knyttes forskellige motivationer sammen, som gør det vanskeligt at finde frem til primær eller sekundær motivation? Kan man mærke at motivationerne har ændret sig gennem tiderne?

Til sidst undersøges der om man kan mærke at fagordbøger eller termsamlinger rettes mod en speciel brugergruppe, og hvordan de svarer til dens specifikke behov.

Normering af ord uden normering – i teori og praksis

Dansk retskrivning er reguleret ved lov, og loven bemyndiger Dansk Sprognævn til at fastlægge retskrivningen. I loven om dansk retskrivning hedder det: “Dansk Sprognævn redigerer og udgiver den officielle danske retskrivningsordbog. Heri offentliggøres den af nævnet fastlagte retskrivning”. Den Danske Ordbog (DDO) indeholder ca. 100.000 opslagsord, mens Retskrivningsordbogen (RO) fra Dansk Sprognævn har ca. 64.000, som ikke alle står i DDO. Det betyder at omkring 40 % af DDO’s ordforråd er uden egentlig officiel normering. Hvilken normeringsmæssig status har disse ord da? Spørgsmålet er dels teoretisk og principielt interessant, dels har det nogle praktiske konsekvenser: Ordbogsredaktionen skal tage stilling til hvilken opslagform og hvilke variantformer der eventuelt skal tages med i beskrivelsen af ordene, og lærere og undervisere skal afgøre hvad de regner for rigtigt og forkert, når de retter elevs og studerendes afleveringer.

I mit bidrag vil jeg diskutere lovgrundlaget og redegøre for redaktionens overvejelser og den praksis det har ført til i DDO. Til diskussionen af lovgrundlaget hører bl.a. en afklaring af hvor langt Sprognævnets bemyndigelse gælder. Omfatter retskrivning fx også bøjning og orddannelser (fugeelementer, afledninger)? Og hvor langt kan man tillade sig at gøre brug af analogier? Når duatlon, biatlon og triatlon er optaget i RO i disse former, kan man da regne skiatlon for en normeret form selvom den ikke står i RO, eller tiltager ordbogsredaktionen sig her en normerende funktion? Er det vigtigt for brugerne at ordbogen formidler hvornår der er tale om officielt normerede former, og hvornår det er ordbogens egne opslagsformer?

Bo-A. Wendt

Svenska Akademiens ordboksredaktion, Lund

Hänvisning eller upprepning i samspelet mellan olika ordboksartiklar

I den tryckta ordboken har det redan av rena kostnadsskäl varit en dygd att ha en så kompakt och utrymmesekonomisk beskrivning som möjligt. Istället för att upprepa information som finns behandlad någon annanstans väljer man i regel att hänvisa till detta andra ställe. Denna beskrivningsekonomi får i *Ordbok över svenska språket* utgiven av Svenska Akademien (SAOB) sägas ha odlats till det yttersta.

När nu planeringen inför en reviderad, digital andra upplaga så smått har inletts, är det därför rätt naturligt att pröva tanken på att i grunden utmana denna princip till förmån för större eller mindre mån av upprepning – mot bakgrund av att utrymme i det digitala formatet inte längre är en begränsande omständighet.

Framförallt handlar det då om tre kategorier med långt driven hänvisningspraxis: etymologier till ord som kan föras till samma indoeuropeiska rot, definitioner som innehåller polysema ord vars betydelse behöver preciseras samt redovisningar av flerordsuttryck (som ju hör hemma i flera artiklar). Att få all information samlad på just det ord och det ställe som man valt att slå upp istället för hänvisningar till andra artiklar (eller ställen) är förstås bekvämt för läsaren. Samtidigt finns det rimligtvis – utöver den fysiska utrymmesbesparingen – en läsbarhetsvinst i att ordbokstexten inte blir alltför överlastad och skrymmande genom att mer fördjupande information hänskjuts till ett annat behandlingsställe.

I mitt föredrag kommer jag att argumentera för denna senare ståndpunkt: att den traditionella utrymmesekonomiska beskrivningen i många fall faktiskt också har uppenbara fördelar för framställningen som sådan, alldeles oberoende av att det fysiska utrymme som sådant är oproblematiskt i den digitala ordboken – och detta särskilt när det digitala formatet gör det mindre ansträngande för läsaren att följa en hänvisning.

Abstrakt för temasektionen om
etymologiska nätresurser

Osmo Nikkiläs finsk-svenska etymologier på nätet

Osmo Nikkilä (1933–2002) var en av Finlands mest utmärkta etymologer och forskare i germansk-finsk språkkontakt samt en av utgivarna av de första två delarna av *Lexikon der älteren germanischen Lehnwörter in den ostseefinnischen Sprachen* (1991–2012). Bland hans arbeten finns även ett stort antal artiklar där han vid olika tillfällen och i väldigt olika sammanhang inte bara behandlar äldre germanska utan även explicit yngre svenska lånord i finska språket. Under sin livstid samlade Nikkilä mycket material om ca 2000 lånord i finskan som han ansåg vara av svenskt ursprung och som hittills inte hade behandlats i etymologiska arbeten. Tillsammans med sina iakttagelser antecknade Nikkilä varje ord på ett pappersark och ordnade dem alfabetiskt med syfte att publicera en kommenterad ordbok över svenska lånord i finska språket.

Det svenska språkets inflytande på finskan har hittills uppmärksammats relativt lite eftersom det redan vid första anblicken verkar uppenbart. Trots detta bör man klassificera svenska lånord i olika skikt på samma sätt som t ex de äldre germanska lånorden, med vars hjälp man nyligen har lyckats datera germanska, finska och samiska ljudutvecklingar inte bara relativt, utan även absolut.

Nikkiläs ovan nämnda material bildar en utmärkt bas för djupgående arbeten. Därför har jag börjat granska etymologierna och publicera dem på i databas på nätet. Presentationen som ska hållas på finska går djupare in på projektets bakgrund, presenterar exempel samt introducerar användningen av databasens resurser.

Suomen murteiden sanaston asteittaista etymologiaa

Käsittelen esitelmässäni viittä suomen murteiden sanaryhmää, jotka ovat osittain venäläisperäisiä mutta joilla on yhteyksiä myös omaperäiseen sanastoon tai muihin lainasanoihin. Niillä on myös ekspressiivisille sanoille tyypillisiä piirteitä.

1) *Ars(s)ina* 'pituusmitta' on venäläislaina, ja *ars(s)inoida* 'suunnitella, miettiä, arvioida' voisi olla siitä muodostettu verbijohdos, mutta verbeissä on paljon äänteellistä vaihtelua, joka liittyy *arvella-* ja *harkita-*verbeihin: *ar(s)seloida*, *ar(s)sikoida*, *ar(s)siloida* ~ *arviloida* ~ *arvikoida* ~ *harsiloida* ~ *harsikoita*.

2) *Hotu* voidaan yhdistää venäjään ainakin merkityksessä 'tie, väylä, reitti; suunta; vauhti, kiire'; *lähteä hotuun* 'matkaan, liikkeelle', *olla hodussa* 'menossa'. Merkityksessä 'pehmeästä, huokoisesta, sisältä ontosta tai lahosta' ja 'huonosta työstä; huonon työn tekijästä' *botu* liittyy selvemmin ekspressiiviseen sanastoon.

3) *Hurtti*, *hurtti* –sanoissa on monenlaista ainesta: 'koira' (muinaisvenäjältä), 'reipas' (ruotsista), 'heitto, vauhti' (*burista* ym). Pejoratiivinen käyttö ja merkitys 'karvainen, pörröinen' voi olla kehittynyt 'koiran' pohjalta, mutta voi olla myös yhteydessä muuta alkuperää oleviin sanoihin, kuten *hurja*.

4) *Hutoi* ~ *hotoi* 'huono, heikko, kelvoton, mitätön' on venäläisperäinen, mutta siinä on myös ekspressiivistä ainesta, kuten *mennä hutoi* 'hukkaan, pilalle, tyhjiin' ja *hutoinen* 'helposti hupeneva, hupa; pehmeä, kuohkea'. Murteissa harvinainen, mutta nykyään puhekielessä yleinen on *buti* 'ohilaukaus, -lyönti tms.'

5) *Jässikkä* 'rasia, laatikko, arkku; kantamus, taakka' on lainautunut venäjältä. Sanassa on paljon äänteellistä variaatiota: *jaassikka* ~ *jaassieka* ~ *jassakka* ~ *jassikka* ~ *jessikka* ~ *jässikkä* ~ *jässäkkä*. Se on sulautunut muuhun sanastoon merkityksessä vankka, tukeva, vanterra' ja 'joukko, ryhmä, rykelmä'.

En etymologisk nätordbok över det äldsta finska ordförrådet

Vårt nätordboksprojekt, som på finska heter Suomen vanhimman sanaston etymologinen verkkosanakirja (<https://sanat.csc.fi/wiki/Etymologiawiki>), har startat vid Helsingfors universitet i början av 2018 med finansiering från Kone-stiftelsen.

Projektet syftar till att etymologisera alla de ord som finskan har ärvt från urfenniskan, de fenniska l. östersjöfinska språkens gemensamma urform, som talades ungefär vid början av vår tideräkning. Databasen innehåller drygt 2000 sökord och baserar sig på Petri Kallios manuskript Itämerensuomalainen yhteissanasto (Gemenfennisk ordlista).

Den viktigaste metodologiska innovationen hos projektet är att branschens alla akademiska forskare får delta öppet i redaktionsarbetet samt publicera och diskutera etymologier på ordbokens wikisidor. Detta innebär bland annat att ordboken uppdateras kontinuerligt, vilket är ett stort steg framåt från det traditionella sättet att utge tryckta etymologiska uppslagsverk. Nätordboken förenar den etymologiska forskningens alla arbetsskeden från första publiceringen genom kollegial granskning, kritik, motkritik och ordboksredigering ända till popularisering. På det viset kan de troliga etymologiernas väg från en forskares hjärna till den etymologiska ordboken förkortas från femtio eller hundra år till ett år eller mindre.

Ordboken består av tre nivåer:

1. All etymologisk information om varje sökord ska få plats i databasen
2. Den etymologiska forskningen av varje sökord i alla tryckta källor sedan den stora finska etymologiska ordboken SSA (1992–2000) gavs i tabellform som bas för den öppna diskussionen om sökordens ursprung
3. Redaktionen ska popularisera alla troliga etymologier till allmänfattliga ordboksartiklar

Nätordbokens wikibas kan i framtiden användas för att utvidga den etymologiska ordboken till yngre finska ordsnitt eller till andra språks lexikon.

Hanne Lauvstad
NAOB, Kunnskapsforlaget

NAOBs etymologier

Det Norske Akademis Ordbok (NAOB) er en stor ettspråklig ordbok for norsk bokmål og riksmål. Ordboken ble publisert digitalt 21.12.2017 og blir løpende oppdatert og vedlikeholdt. Den består av over 200 000 oppslagsord og inneholder flere etymologiske redegjørelser enn noen annen norsk allmennordbok. Foredragsholderen har ledet etymologiarbeidet siden 2006 og vil gi en presentasjon av hvordan dette har vært utført. Det vil bli redegjort for hva slags oppslagsord som har fått etymologier, hva slags opplysninger etymologiene inneholder og hvilke hovedkilder som har vært benyttet. Også dilemmaer knyttet til hvordan etymologiene skulle presenteres, vil bli tatt opp. Dessuten vil foredraget gi et inntrykk av de avgrensninger som har vært gjort og hvilke ordtyper som har budt på særlige problemer. I tillegg vil det bli lagt vekt på hvilke muligheter som ligger i en digital redigeringsform, og det vil bli vist hvordan disse mulighetene er blitt utnyttet i NAOB. Til slutt vil det bli antydnet hvordan NAOB kan komme til nytte som etymologisk ressurs for ordboksbrukere.

Miro Metsämuuronen

Helsingfors universitet

Färöisk lånordsforskning i historisk överblick

Färöiska är ett västnordiskt språk med ca 60 000 talare, de flesta på den Danmark tillhörande ögruppen Färöarna på Nordatlanten. Språkets ordförråd är till huvuddelen västnordiskt, dock med ökad inströmning av danska inslag mot modernare tider. För det icke-skandinaviska som man sporadiskt lagt märke till i färöiskan har källan oftast eftersökts på de brittiska öarna, såväl på keltiskt som på engelskt håll.

Grunden för studiet av lånord i färöiskan lades redan vid skiftet av 1900-talet av den färöiske filologen Jakob Jakobsen, som är mest känd för sitt fältarbete kring det utdöda språket norn i Shetland. Jakobsen, som hade blivit bekant med skotsk gäliska, identifierade en grupp färöiska ord och ortnamnselement för vilka han kunde utpeka ett keltiskt ursprung, såsom drunnur 'Ende, Halestykke af et Kreatur' ← gäl. dronn 'den bageste af Ryggen'. Diskursen som Jakobsen hade startat återupptogs på 1950-talet av prof. Christian Matras som kom att föreslå en rad tillägg till det keltiska språkgodset i färöiskan (t.ex. blak 'kärnmjök' ← gäl. bláthach 'id.'). och hans idéer utvecklades sedermera vidare även av britten W. B. Lockwood. De engelska inflytanden får sin preliminära behandling i Tórður Jóansson's monografi English loanwords in Faroese i 1997. Sammanlagt ett drygt tjugotal färöiska ord har genom tiderna föreslagits en keltisk etymologi. Jóansson listar ut omkring 50 "sannolika" äldre lån från engelskan (lånats före ca 1800) och därtill ett tjugotal "eventuella" lånord från samma period.

I mitt föredrag redogör jag för den färöiska lånordsforskningens historia som ett individdrivet marginalintresse i den bredare kontexten av (väst)nordisk etymologi. Jag tar upp exempel på låneetymologier från olika perioder och diskuterar deras förutsättningar mot den allmänna kontexten av färöiskans ljudhistoriska utveckling. Presentationen anknyter sig till min pågående avhandlingsprojekt om färöisk ljudhistoria, som syftar även på att revidera diskussionen om lånord i färöiskan.

Religiösa kontakter? Mytologiska lånord i sydsamiskan från de skandinaviska språken

I min presentation kommer jag att betrakta de mytologiska lånorden i sydsamiskan som har lånats från de germanska och skandinaviska språken. Religionshistoriker har länge avsett att fenomen i den samiska mytologin har till stor del lånats från den skandinaviska mytologin. Sedan dess har åsikterna ändrats och man medger att den samiska mytologin är oberoende och originell.

Mytologin har ändå fått påverkan från många olika håll. I den här presentationen observerar jag dessa influenser från ett språkhistoriskt perspektiv. Materialet för studiet är det mytologiska lexikonet i sydsamiskan och särskilt de germanska och skandinaviska lånorden i sydsamiskan. Jag granskar hur många av de mytologiska orden har lånats från det skandinaviska hållet och från vilken språk har sydsamiskan fått ord. Jag kommer också att betrakta ordens semantik: hurdana religiösa fenomen refereras av ord i sydsamiskan och har de originala orden i skandinaviskan haft religiösa meningar? Samtidigt betänker jag hur mycket ordens ursprung berättar om religiösa fenomenens ursprung. Har det funnits religiösa kontakter mellan sydsamiska och skandinaviska talare från det språkhistoriska perspektivet, eller hurdana har omständigheterna varit när man lånade mytologiska ord?

I mitt mytologiska ordmaterial finns 75 ord, varav 32 har lånats från det pre- eller urgermanska, urnordiska, fornnordiska eller fornsvenska språket. Några av orden kan förespråka om religiösa kontakter: till exempel, i det sammansatta ordet Hovrengaellies 'åskguden' har den definitiva delen Hovren- lånats från ett fornnordiskt namn Þórr (sv. Tor) för åskguden. Många av orden som har lånats till sydsamiskan har dock ingen religiös mening. Till exempel grundordet gaellies i Hovrengaellies har lånats till sydsamiskan från urnordiskan i en mening av 'make'. Ordmaterialen visar att lånord oftast har utvecklats mytologiska meningar först i sydsamiskan. Det menar att ord har lånats i andra kontexter än religiösa.

Fenniska *litna 'borg, köping', än en gång

Ursprunget för urf. *litna 'borg, köping' (jfr fi. linna, veps. lidn, estn. linn, vot. lidna, sydestn. liin etc.) är omstridd (LägLoS: s.v. linna). Den morfeminternas sekvensen *-tn- är inte förenlig med uralisk fonotax, varför stammen inte är förfennisk. Trovärdiga låneetymologier är föreslagna för andra urfinska lexem med den sekvensen (LägLoS: s.v. vannas, vannoa).

T.E. Karsten (1943-44: 280–293) har föreslagit ett lån från tidig urg. *fligna-, en hypotes förkastad i LägLoS (s.v. linna) men senare modifierad och argumenterad av J.Koivulehto (1999: 121; jfr LägLoS s.v. vannoa). Hans argumentation för en substitution med *-tn- håller medan hans rekonstruktion av *-i- (ist f *-e-) i förgm. +pliknó med hänvisning till en slutanmärkning i Kluge & Seebold (2002: s.v. Fleck) är svårförenlig med nordisk vokalism. En förment tidig betydelseförskjutning från 'lapp' till 'plats' > 'köping' > 'borg' kan inte stödas med komparativt material, även om en förskjutning till 'köping' är belagd för en avledd n-stam vlecken långt senare i tyska/nederländska.

På symposiet ifrågasätts premissen för Koivulehtos slutsats: "es scheint kein anderes passende Original zu geben". En studie av etymologin för medelnederländska/-lågtyiska klippe som motsvaras (kanske genom lån) av skandinaviska klippa /klippe ska evaluera möjligheten av ett original *klibna- med betydelsen 'naken stenmassa' (DEO s.v. klippeI.) > 'fästningsberg' > 'borg' > 'köping'.

DEO=Nielsen, N. Å. (1985 [1966]). Dansk etymologisk ordbog: ordenes historie (4. opl.).
Karsten, T. E. (1943–44). Finnar och germaner: Folk- och kulturberöringar från tre årtusenden. Folkmålsstudier, IX–X.

Kluge, F., & Seebold, E. (2002). Etymologisches Wörterbuch der deutschen Sprache (24. Aufl).

Koivulehto, J. (1999). Review of LÄGLOS, Bd. II. Kratylos, 44, 119–122.

LägLoS=Kylstra, A. D., Hahmo, S.-L., Hofstra, T., & Nikkilä, O. (1991–2012). Lexikon der älteren germanischen Lehnwörter in den ostseefinnischen Sprachen (Vol. 1–3).

Poster

Johan Schalin & Juha Kuokkala

Helsingfors universitet

Fenniska *šike-/*šiko- 'svett(as)'

Den medelurfinska e-stammen *šike- 'svett' > fi. "hiki" est. "higi" etc., den parallella o-stammen *šiko- 'svettas(s)' med avledningar bildade på dessa (eller på *šiges-) saknar en accepterad etymologi. Ursprunget kan tillskrivas hög ålder då båda stammarna har efterföljare i hela den fenniska språkfamiljen, med motsvarigheter även i sydestniska (SKES: s.v. hiki). Ett förslag till baltisk etymologi har enhälligt avvisats (Junttila 2015: 201 med nätbilagan "tutkimushistoriallinen taulukkoliite" s.v. hiki).

I en posterpresentation vid temasektionen ska det visas att stammen troligen är lånad från ett germanskt original *sigan- som har följare i svensk dialektalt "sega, sege" m. 'vatten som tränger upp genom marken, källåder' ≈ "sege" m. 'vätska som rinner ur sår eller böld' (SAOB s.v. -sega), med ytterligare motsvarighet i fvn. m. "blóðsegi ≈ blóðsigi" 'blodklump'. Denna deverbala n-stam kan från början antas ha haft en mer allmän betydelse av drypande vätska som sipprar fram och silar ner, vilket kan ses i efterföljare till verbet *seihwan och dess vernerska växelformer (jfr Koivulehto 1999: 50–57), liksom i tillhörande indoeuropeiskt data. En anknytning till fukt i terräng finns också i finskan, nämligen i adjektivet fi. "hikevä" 'fuktig, drypande [jord]'.

Den semantiska problematiken förbunden med att knyta adjektivet "heik(k)o, heikeä" 'svag, mager, dålig' med motsvarighet i votiska "eikko" (SKES: s.v. heikko) till germanska avledningar av samma rot diskuteras också.

Junttila, S. (2015). Tiedon kumuloituminen ja trendit lainasanatutkimuksessa: kantasuomen balttilaislainojen tutkimushistoria. Helsinki. <https://helda.helsinki.fi/handle/10138/158777>

Koivulehto, J. (1999). Verba mutuata: quae vestigia antiquissimi cum Germanis aliisque Indo-Europaeis contactus in linguis Fennicis reliquerint. (K. P. Ruppel, Red.). Helsinki. SKES=Toivonen, Y. H., Itkonen, E., & Joki, A. J. (Red.). (1955-1981). Suomen kielen etymologinen sanakirja I–VII (Vol. 1–7). Helsinki.

Abstrakt för
postersektionen

Oddrun Grønvik, Stig Jarle Helset & Johan Myking

Universitetet i Bergen

Høgskulen i Volda

Prosjekt NO–AH: fulldigitalisering av Norsk Ordbok

Norsk Ordbok (1929–2015)

Norsk Ordbok er det store ordboksverket over norsk talemål og det nynorske skriftspråket. Arbeidet med Norsk Ordbok starta i 1929 ved Universitetet i Oslo og vart avslutta i 2015. Det Norske Samlaget har gjeve ut verket i 12 band via prosjekt NO2014 (2002–2015). Verket inneheld om lag 330 000 oppslagsord.

Prosjekt NO–AH (2019–?)

Universitetet i Bergen (UiB) har i samarbeid med Høgskulen i Volda (HVO) teke ansvaret for å gjere ferdig den digitale utgåva av ordboksverket. I statsbudsjettet for 2019 har Norsk Ordbok fått 2 millionar kroner til oppstart og igangsetjing. UiB og HVO vil også legge inn betydelege eigne ressursar i prosjektet, og har som ambisjon å byggje opp eit permanent fagmiljø.

Norsk Ordbok skal bli digitalt komplett

Berre alfabetstrekket i–å, ca 220 000 oppslagsord, er i dag publisert som nettordbok. Alfabetstrekket a–h er tilgjengeleg for publikum berre på papir. Dette gapet i den elektroniske versjonen av Norsk Ordbok utgjier nesten 40 prosent. Dessutan er det manglar i Norsk Ordbok fordi arbeidet starta så tidleg at mange ord som er i dagleg bruk i dag, ikkje eksisterte då dei første banda i Norsk Ordbok vart publiserte.

Norsk Ordbok og Språksamlingane

Sidan 2016 har UiB hatt ansvaret for Språksamlingane. Dette er den vitskaplege databanken som ligg til grunn for arbeidet med Norsk Ordbok. Samlingane vart flytta frå Universitetet i Oslo til Universitetet i Bergen i 2016 i samarbeid med Høgskulen i Volda.

Norsk Ordbok er digital infrastruktur

Digitaliseringa av Norsk Ordbok skal bidra til å gje det norske samfunnet ein nødvendig digital infrastruktur i møte med automatisering, robotisering og påverknad frå globale engelskspråklege aktørar. Digitaliserte språkbasar av høg kvalitet er ein føresetnad for til dømes talegjenkjenning og god kvalitet på automatiserte svar på norsk.

For whom? End users and lexicographical data on web portals

Introduction

This poster seeks to address one aspect of the overarching research question “for whom?”. Web portals are of increasing relevance from the point of view of the individual end user as they are perceived as convenient tools in providing access at once to a variety of databases, dictionaries and other types of lexicographical data. At the same time, their existence can make lexicography work more complicated than before.

Case

As a case in point we shall present newly gathered information on the Icelandic web portal malid.is: (1) on user statistics; (2) on the results of a user opinion poll carried out in November 2018. Our main results are, firstly, that the web portal has gained great popularity in a relatively short period of time, and, secondly, that while most end users are content, some users tend to miss information that is of relevance, and information from some of the data collections are much more prominent than from others, resulting in somewhat unfortunate imbalance; also better usage instructions are needed, among other things as to the different nature of individual data collections.

Conclusion

Web portals are important intermediary links between editors of dictionaries and other data collections on the one hand and end users of lexicographical information on the other hand. As a consequence, lexicographers may find themselves persuaded to adjust their products – as to technology, layout and content – for the possibility that their data be presented through a secondary medium such as web portals. Editors may find this fact limiting for the presentation of their material. At the same time, web portal editors must seek a reasonable balance in presenting information from different sources, and provide end users with some flexibility to tailor data presentation to individual needs. If this is successful the needs of both groups, i.e. dictionary makers and dictionary users, can be satisfied at once.

Monica von Martens

Göteborgs universitet

Redigeringsmiljön för Stora finsk-svenska ordboken – återbruk, samarbete och vidareutveckling

Postern beskriver hur vi byggde upp en relationsdatabas och ett redigeringsystem för Språkinstitutets Stora finsk-svenska ordbok. genom att tillgodogöra oss erfarenheter dels från huvudredaktören Nina Martolas tidigare arbete med ordboken, dels från liknande projekt vid Göteborgs universitet och hur detta projekt i sin tur har berikat vidareutvecklingen av redaktionsmiljön för Svenska Akademiens lexikaliska databas vid Göteborgs universitet.

Den tvåspråkiga ordboken levererades i form av en XML-fil i ett format optimerat för visning på nätet och uppdraget var att skapa ett redigeringsystem som möjliggjorde uppdatering av ordboken. Redigeringsystemet är avsett att användas endast under ett begränsat antal år, varefter den uppdaterade ordboken skall flyttas hem till Språkinstitutet och integreras med där befintliga system. Fokus låg därför på att snabbt och resurseffektivt få till stånd ett användbart system för en eller ett fåtal expertanvändare. Ett krav var också att det skulle finnas en snabb och smärtfri process för att publicera nya versioner på nätet.

Beskrivningen av den valda datastrukturen och programmiljön kan vara av intresse för andra projekt som står i färd med att utvärdera sina redaktionssystem och publiceringsprocesser. Om intresse finns kan jag ordna en demo/frågestund kring den tekniska implementationen.

A word frequency dictionary of Icelandic child-directed speech

This paper presents the making of a word frequency dictionary of Icelandic child-directed speech. It will be the product of a transcribed corpus of 50 hours of video recordings of two children during their language acquisition. The corpus includes child-directed speech by parents and other adults. The dictionary will be immediately important in research on Icelandic language acquisition.

A variety of changes are underway in Icelandic syntax. Some of them are more lexically rooted than others, such as dative substitution, where accusative subjects are replaced by dative subjects (Yang 2016 and references cited there), and a change from dative-nominative verbs to dative-accusative pattern (Árnadóttir and Sigurðsson 2013). In addition to these, a recent study suggests that children acquire dative subjects that originate as indirect objects in passives earlier than dative subjects that originate as direct objects (Sigurðsson, Nowenstein and Sigurjónsdóttir 2018). That is a surprising result unless it turns out that ditransitives, such as ‘give’, are more frequent than dative-taking monotransitives, such as ‘help’, in the input of children acquiring Icelandic.

Even though a change like dative substitution has been investigated in great detail, linguists have not had a direct access to the input of children acquiring Icelandic. The dictionary that is currently underway, will solve that problem as it will be made available in open access, as well as the transcriptions of the video recordings. It will therefore be an important tool for language acquisition researchers.

References

Árnadóttir, H., and E.F. Sigurðsson. 2013. Case in disguise. B. Fernández and R. Etxepare (eds.): *Variation in Datives: A Micro-Comparative Perspective*. OUP, Oxford.

Sigurðsson, E. F., I. Nowenstein and S. Sigurjónsdóttir. 2018. The acquisition of dative subjects in L1 Icelandic. Paper presented at Galana 8, Bloomington.

Yang, C. 2016. *The Price of Linguistic Productivity*. MIT Press

Mette-Marie Møller Svendsen,
Thomas Troelsgård & Nicolai Hartvig Sørensen
Det Danske Sprog- og Litteraturselskab

Salmesang og superordbog: ordbogslinkning i praksis

Projektet 'Musik og sprog i reformationstidens danske salmesang' er et tværinstitutionelt samarbejde mellem Det Kongelige Bibliotek og Det Danske Sprog- og Litteraturselskab. Som en del af projektet udvikles en superordbog, en database der kommer til at rumme i alt 12 digitaliserede ordbøger og gloselister som er relevante i forhold til sproget i reformationstidens salmebøger og generelt for perioden. Ordbogsværkerne omfatter dansk-dansk, dansk-latin, latin-dansk og latin-dansk-græsk.

På samme måde som den norske metaordbog giver et overblik over samtlige opslag på et ord fra en række forskellige materialer, så giver superordbogen også mulighed for at skabe et indblik i sprogets udvikling ved hjælp af en mængde danske resurser (Ore 2001). Den samlede superordbog vil kunne være til gavn for alle der arbejder med det ældre danske sprog.

Hvad gør man for at de varierende stavemåder for dansk i 1500-tallet er til mindst mulig besvær i linkningsprocessen? Hvilke dele af arbejdet kan automatiseres, og hvor meget skal gennemgås manuelt? Hvordan tages der bedst højde for ordenes etymologi?

Denne poster vil give et indblik i de forskellige praktiske led der indgår i linkningsprocessen, bl.a. hvilke overvejelser der gøres i forhold til udvælgelse af lemmaækvivalenter og hvilke it-værktøjer der inddrages undervejs.

Projektet er støttet af Carlsbergfondet og Velux Fonden og forventes afsluttet medio 2021.

Referencer:

Ore, Christian-Emil (2001). "Metaordboken - et elektronisk rammeværk for Norsk Ordbok?" I: Gellerstam, Martin et al. (ed.). Nordiska studier i leksikografi. Göteborg

Bo-A. Wendt

Svenska Akademiens ordboksredaktion, Lund

SAOB – för vem och hur?

Inom några år kommer första upplagan av Ordbok över svenska språket utgiven av Svenska Akademien (SAOB) att vara färdigställd. Eftersom språket har utvecklats under ordbokens utgivningstid kommer innehållet naturligtvis att behöva kompletteras, rättas och uppdateras. Svenska Akademien har därför fattat beslut om en mycket ambitiös revidering som också innefattar innehållsliga förändringar. Detta betyder att det för första gången på 125 år är möjligt att tänka i nya banor vad gäller ordbokens presentations- och innehåll.

På Svenska Akademiens ordboksredaktion har vi så smått påbörjat planeringsarbetet inför en ny upplaga som med all sannolikhet kommer att publiceras endast i digitalt format. Målet är en optimalt funktionell och användarvänlig ordbok som beskriver den lexikala utvecklingen under de drygt femhundra år som brukar betraktas som den nysvenska perioden.

En central fråga är därvid vem ordboken skall rikta sig till. I förordet till första häftet av SAOB, från 1893, står det ”att ordboken är bestämd ej blott för vetenskapsmännen, utan för hela den kunskapssökande svenska allmänheten”. Vad betyder detta för revideringsarbetet? Hur gör vi SAOB mer tillgänglig och begriplig för fler? Ambitionen är att den nya digitala ordboken skall vara användbar och tillgänglig för en bredare läsekrets utan att ge upp den komplexitet som en vetenskapligt ambitiös och detaljerad beskrivning innebär.

I vår posterpresentation kommer vi att redogöra för olika förslag på hur en ordboksartikel i SAOB kan göras mer överskådlig. Man kan tänka sig olika grafiska lösningar för att göra den nuvarande kompakta och informationstäta ordbokstexten lättillgänglig. Språkexemplens kronologi kan exempelvis framgå tydligare. Vidare bör brödtexten få en moderniserad stavning och mindre ålderdomlig prägel. Ordbokens många förkortningar kan lösas upp och dess många termer förklaras.