

Amatörlexikografiska insatser på sajten *Folkmun.se*

Emma Sköldbberg & Lena Wenner

This paper describes the results of a study of the user-generated web dictionary *Folkmun.se*. The platform was launched in 2007 and contains (December 2016) about 4 600 headwords (lemmas). We analyse how the content of *Folkmun.se* has developed over time with particular focus on ten contributor aliases and how they have participated in the development of the dictionary. Our study displays some recurrent pattern in how different aliases work. We observe similarities in what kind of words they contribute with and the manner in which they define words. Many lemmas have in common that they are not described in traditional lexicons. Depending on contributor the articles differ in quality regarding both form and content.

1. Inledning

Föreliggande artikel redovisar resultaten av en kvalitativ, diakron studie av den användargenererade ordboken *Folkmun.se*. Med ”användargenererad” menas att sajten består av ord och uttryck som skickats in av allmänheten. Vårt fokus ligger på enskilda bidrag och bidragsgivare och deras roll i sajtens utveckling. Vilka bidrar till ordboken, hur tycks dessa personer arbeta och resonera? Går det att skönja mönster i bidragsgivandet? Artikeln inleds med en forskningsbakgrund och därefter presenteras själva ordbokssajten. Därpå redogör vi för förutsättningarna för studien, vårt material och våra resultat. Artikeln avslutas med en sammanfattande diskussion.

2. Kort om användargenererade sajter

Användargenererade sajter är per definition svåra att undersöka eftersom de snarare är pågående processer än statiska produkter. Vidare är de flyktiga; de har en tendens att byta webbadress eller helt enkelt försvinna efter några år (Törnqvist 2015:155). Trots denna (inneboende) problematik har det de senaste åren forskats allt mer på resurser av detta slag. Exempelvis har det gjorts kvantitativa analyser av *Wiktionary* för att studera hur sajten utvecklas. Viktiga frågeställningar är hur *många* personer som ligger bakom en viss sajt, *när* artiklarna revideras och *hur* de revideras (se t.ex. Wolfer & Müller-Spitzer 2016). Tidigare forskningsinsatser till trots vet vi fortfarande mycket lite om själva bidragsgivarna.

Det råder delade meningar om de användargenererade sajternas värde och kvalitet (se t.ex. Hanks 2012:77–82 och Meyer & Gurevych 2012). Exempelvis

viner *Wiktionary* terräng på en rad områden (såväl lingvistiska som språkteknologiska), och användargenererade ordböcker har potential att bli allvarliga konkurrenter till både förlagsbaserade och akademiska ordböcker (Wolfer & Müller-Spitzer 2016:347). Innehållet i dem fyller också en kompletterande funktion i förhållande till traditionella ordböcker t.ex. när det gäller dokumentation av helt nya, och kanske lite mer tillfälliga, ord. Den kritik som framförts har framförallt riktats mot grundantagandet att samarbete mellan många bidragsgivare leder till förbättrat innehåll (se t.ex. Mattus 2009:183). En annan kritisk synpunkt mot användargenererade sajter är att bidragsgivarna, i stället för att samarbeta i syfte att uppnå mer fullständiga beskrivningar, ibland mest tycks vilja lufta sina fördomar (Lew 2011). På många användargenererade sajter som t.ex. Wikipedia förekommer också vandalism och s.k. *edit-wars* i och med att man kan ändra och t.o.m. radera andras artiklar (Mattus 2009:192). Att bidragsgivare kan förhålla sig till befintliga artiklar och komplettera dem är dock en viktig drivkraft. Aaltonen & Seiler (2016) visar exempelvis att bidragsgivare är mer motiverade att bygga vidare på artiklar än att skapa helt nya.

3. Kort om *Folkmun.se*

Sajten *Folkmun.se* drivs av privatpersonen David Eriksson, som startade den år 2007 i syfte att ”dokumentera dialektala uttryck (men också slang och nyord)” (se vidare Sköldbberg & Wenner 2017 med anförda referenser). Som Sköldbberg (2015:232–233) visat finns det redaktionella riktlinjer till dem som vill föra in ett nytt ord. Bidragsgivarna uppmanas bl.a. att förklara ordet och lägga in språkexempel.

Eriksson har gett oss tillgång till en databas där allt material som inkommit till *Folkmun.se* från starten till december 2016 finns samlat.¹ Totalt innehåller databasen 5 523 bidrag. Av dessa är 4 641 poster synliga på webbsidan och de har skickats in av 1 664 unika bidragsgivare. Eriksson är ensam moderator och bestämmer vilka bidrag som publiceras på webben. Han går igenom alla insända och rapporterade ord, men är restriktiv med att stryka eller censurera material. Det motiverar han med att om ett nedsättande ord existerar, är det bra att språkbrukarna kan googla sig fram till definitionen och sedan förhoppningsvis välja att inte använda det. Det händer att synnerligen nedsättande bidrag inte görs publika, men de vanligaste skälen till utebliven publicering är dels att bidragsgivaren inte

¹ Stort tack till David Eriksson som så frikostigt delat med sig av sitt material och svarat på våra frågor. Avseende studiens forskningsetiska aspekter så har bidragsgivarna inte informerats om sin medverkan i studien. Erhållen data är anonymiserad så att den inte går att koppla till fysiska personer. Däremot går det att utifrån anonymiserade IP-uppgifter till exempel utläsa om flera användaridentiteter använder samma dator-uppkoppling. Därigenom menar vi att vi har beaktat befintliga forskningsetiska riktlinjer om informanternas anonymitet och skydd av informanter.

gett sitt medgivande till att artikeln publiceras på sajten (något som görs via mejl), dels att hen av misstag skickat in ett och samma bidrag flera gånger.

Statistik via Google Analytics visar att antalet inskickade ord med något undantag har minskat för varje år sedan sajten startades, men ser man till antalet sajtbesökare har siffran stadigt ökat. Under oktober månad 2007 hade *Folkmun.se* 3 500 besökare. Det kan jämföras med oktober månad 2016 då 12 000 personer besökte sidan.

4. I tio bidragsgivares fotspår

I den här studien har vi valt att närma oss bidragsgivarna genom att undersöka de spår de lämnat efter sig i materialet. Man kan se tydliga likheter mellan den valda metoden och studier som har gjorts av ordboksanvändares sökmönster (jfr t.ex. Almind 2008 och Hult 2017).

I avgränsningssyfte har vi närstuderat artiklar som härrör från tio olika bidragsgivare (se vidare nedan). Just deras signaturer (dvs. bidragsgivarnamn) valdes ut eftersom deras bidrag skiljer sig åt till antal och omfång m.m. De aktuella bidragsgivarna får, inom ramen för vår studie, representera olika typer av användare som är med och utvecklar *Folkmun.se*. Mer specifikt har vi funderat kring deras val av signatur och undersökt hur många bidrag de skickat in, när de skickats in och hur många av dessa som är publika på webbsidan. Naturligtvis har vi också granskat vad det rör sig om för ord och uttryck. Finns det återkommande informationskategorier i artiklarna? Och, inte minst, vilken lexikografisk kvalitet har artiklarna? Notera att den statistik som vi presenterar nedan inbegriper allt material som lagts in av användarna, även sådant som av olika anledningar inte gjorts publikt via sajten.

4.1. Dimbo-Klas

Den första bidragsgivare som vi följt kallar sig Dimbo-Klas. Möjligen är detta en man som bor eller kommer från småorten Dimbo i Västra Götalands län. Om så är fallet kan hen vilja signalera någon form av geografisk hemvist genom sin signatur. Men detta kan vi, med tanke på aktuell metod, förstås inte veta. Dimbo-Klas har endast lagt in fyra artiklar, tre simplex och ett flerordsuttryck, och det har skett under loppet av ca fem minuter en kväll i september 2013. Samtliga artiklar rör etablerade dialektord från ett specifikt angivet område. Ett typiskt exempel är substantivet *glisker* (se figur 1).

The screenshot shows the Folkmun.se website interface. At the top left is the logo 'folkmun.se' with a sun icon. To the right is an alphabetical menu 'ABCDEFGHIJKLMNOPQRSTUVWXYZAAO' and a search bar with the text 'Sök på folkmun.se'. Below the search bar are links for 'Senaste', 'Lägg till ord', 'Kontakt', and 'Blogg'. The main heading is 'Glisker' in green. Below it are navigation links: '« glirv', 'Överraska mig!', and 'Gloglas »'. The article content includes the title 'Glisker', a quote: '"Vår hund har blivit gammal, han är inte så glisker längre"', a note: 'Västgötska (Redväg) för "på alerten", "på hugget".', and the author: 'av Dimbo-Klas 2013-09-18 - Rapportera'. At the bottom is a link: 'Lägg till din egen definition av Glisker'.

Figur 1: Artikeln *glisker* i Folkmun.se (något beskuren; hämtad 30.11.2017).

Ordboksartikeln innehåller ett lemma, ett språkprov, en dialektangivelse, en geografisk uppgift (Redväg) samt en betydelseangivelse i form av två synonyma uttryck.

4.2. Go4it

Signatur nummer två har valt att kalla sig Go4it. Hen har skickat in totalt fem bidrag en fredagsförmiddag 2013 och samtliga artiklar är publika på sajten. Det rör sig om fyra enkla ord och ett uttryck (*fläng* 'dum', *tölit* 'tråkigt', *töla* 'prata; pratglad kvinna', *Urban* 'namn; person som är på dåligt humör' och *ge-rää* 'ge dig, sluta'). De kan alla sägas vara relativt etablerade, talspråkliga uttryckssätt och enligt Go4it hör de till Karlstadsdialekt. I artiklarna finns förutom lemma och dialektangivelse många andra informationskategorier representerade, t.ex. uppgifter om uttal, betydelsevariation, konnotationer och exempel. Go4it:s bidrag är på så vis relativt rika på information, men vissa av artiklarna rymmer samtidigt ett antal skrivfel och brott mot svenska skrivregler.

4.3. Benson

I materialet återkommer signaturen Benson 61 gånger. Bland Bensons artiklar finns många dubletter; totalt rör det sig om 47 unika bidrag. Typiskt har hen lagt in 3–4 bidrag åt gången, i regel med cirka två minuters mellanrum. Artiklarna, som rör ord som *linka* 'halta el. knäa' och *pyngla* 'liten börda', ger ett synnerligen

enhetligt intryck. Först återfinns ett uppslagsord (simplex) och därefter en enklare betydelseangivelse, oftast i form av en eller två synonymer. Benson använder också konsekvent vissa etablerade förkortningar. De allra flesta orden karakteriseras dessutom som ”blekingska”. Sammantaget vittnar bidragen om såväl goda dialektkunskaper som ett systematiskt arbetssätt från just denne bidragsgivares sida.

4.4. Luscinia

Signaturen Luscinia – ett möjligt kvinnligt förnamn, men också det vetenskapliga namnet på ett släkte trastfåglar – har skickat in ca 50 bidrag mellan 2010 och 2011. Så gott som samtliga av dessa är publika. Typiskt har denna bidragsgivare lagt in ett par ord, gärna på kvällen, för att sedan tystna igen. De inmatade orden tycks, om man studerar dem lite snabbt, ha få gemensamma nämnare. Exempelvis har Luscinia den 16/1 2010 lagt in ord som *snigeltåg* ’långsamt tåg’, *golbög* ’den som samarbetar med polisen’ och *zlatanist* ’anhängare av Zlatan’. Men om man granskar samtliga bidrag framträder ett mönster. Ett flertal av orden anges vara skånska, t.ex. *piggatrå(d)* ’taggtråd’ och *cocabonnar* ’glasögon med tjocka glas’. Andra uppslagsord är semantiskt besläktade med varandra, t.ex. *småglin*, *glin* och *glytt* (som alla betyder ’små barn’). Värt att notera är att hen också bidragit med flera slangord som kopplas till fågelskådning, och valet av alias är därmed kanske inte så märkligt.

4.5. Hetrometrus

Hetrometrus (en signatur som sammanfaller med en skorpionbeteckning) är med i den absoluta starten av sajten under 2007, och är ytterst aktiv under tre månaders tid. Hen lägger in 208 bidrag varav så gott som alla är publika. Artiklarna har överlag enkel struktur, t.ex.: ”Hossor Strumpor Sydsåknskt”.

Inledningsvis dokumenterar Hetrometrus mer allmänt kända, vardagliga ord. Därefter går hen över till att i följd lägga in dryga trettioalet uppslagsord som karakteriseras som ”göingska”. (Göinge ligger i nordöstra Skåne.) Därpå matar hen in mer allmänt skånska ord – över 70 uppslagsord efter varandra har etiketten ”skånska”. Avslutningsvis återgår Hetrometrus till mer allmänna svenska slangord och slanguttryck. En närmare granskning av de renodlat skånska orden ger vid handen att de faktiskt är alfabetiskt ordnade. Vidare ger artiklarna ett mycket enhetligt intryck och de har dessutom matats in i mycket snabb takt. Mycket tyder alltså på att Hetrometrus utgått från någon förlaga, men vi har inte kunnat utröna vilken denna i sådana fall skulle kunna vara.

4.6. Mads & Preben

Vidare har vi granskat bidrag skrivna av Mads & Preben. Av vårt material framgår att dessa två män (?) i princip endast varit aktiva vid ett par mycket sena lördagskvällar i september 2007. Sammanlagt ligger de bakom 31 bidrag, men inget av dessa syns på sajten. Rent innehållsmässigt uppvisar deras ordboksartiklar stora likheter; nästan uteslutande rör det sig om grova beteckningar på könsorgan med obscena betydelsebeskrivningar (under lemman som *drängakuk* och *klitta*) eller rasistiska inlägg (under lemman som *habibber*, *blåneger* och *zigenar arsle*). Med tanke på den sena timmen kan det kanske vara två vänner som, i deras tycke, roat sig några ”blöta” kvällar. Man kan dock betrakta deras insatser som en form av vandalisering på webbsidan (se avsnitt 2), i synnerhet som de i ett fall använt sig av Hetrometrus namn och lagt in en allmänt kränkande artikel under denna signatur.

4.7. Richard (J)

Signaturen Richard, alternativt Richard J, har under ca ett års tid skickat in 41 bidrag till *Folkmun.se*. Typiskt har hen matat in 3–4 bidrag i taget och detta på kvällstid. Bland uppslagsorden återfinns *pröjsa* ’betala’, *bull-TV* ’mikrovågsugn’, *dumknollad* ’person som är dum/bortkommen’, *syntetneger* ’person som ligger i solarium’, *lakritstomte* ’afrikan’, *pervo* ’pervers person’, *gambiaskalle* ’negerboll’, *brödkvarn* ’mun’, *PLO-burgare* ’kebab’, *skinkprins* ’homosexuell man’ och *dadeltrampare* ’person från mellanöstern’. Som synes rör det sig mest om (etablerade) slangord och enkla förklaringar, inte sällan med rasistiska och/eller homofoba inslag. Till skillnad från Mads & Prebens bidrag är dessa artiklar synliga på sajten. Richard (J) har bekräftat att hen vill att de ska synas och sajtägaren har också godkänt dem.

4.8. Chooniii

Bidragsgivare nummer åtta har valt att kalla sig Chooniii. Signaturen kan möjligen kopplas till det multietniska ordet *shono/shonne* med betydelsen ’snubbe’ och med tanke på bidragens karaktär är detta inte helt orimligt. Signaturen kan kopplas till åtta bidrag som alla är publika. Uppslagsorden lades in under loppet av en kvart, sent en fredagskväll i juni 2009. Bland dessa återfinns *göt*, *katt* och *yani* med betydelseorna ’rumpa’, ’snygg’ och ’typ’ samt verbet *klippa*, vilket är försett med språkexempel som ska illustrera de tre olika betydelser som anförs, ’stjåla’, ’slå’ och ’åka fast’. Vad gäller innehållet i artiklarna kan vi se att det efter hand blir fylligare – bidragsgivaren tycks bli lite varmare i kläderna.

4.9. Tjo fräzzen

Signaturen Tjo fräzzen har lagt in tio bidrag en eftermiddag i juli år 2009 och åtta av dessa visas på sajten. Hen kan sägas stå bakom en ny typ av gruppspråk i ordboken i förhållande till de signaturer som vi behandlat hittills. Bland bidragen finns nämligen flera engelskklingande uttryck och förkortningar varav vissa anges komma från dataspelsvärlden, särskilt *World of Warcraft*. För att visa hur hen kan ha resonerat i samband med bidragsgivandet ges följande exempel: klockan 14.14.04 lägger hen in uppslagsordet *inv ffs* med betydelsen 'när du vill att nån ska invita dig jätte mkt till ett party'. 39 sekunder senare läggs artikeln *ffs* in som anges betyda 'for fuck sake'. Tjo fräzzen verkar alltså ha kommit på att också *ffs* kan tjäna som uppslagsord i *Folkmun.se* och att det kan behöva förklaras.

4.10. Tyke

Slutligen har vi valt att titta närmare på signaturen Tyke som står bakom 49 bidrag (varav 46 är publika). Under loppet av ca 3,5 års tid har hen återkommit med några månaders mellanrum och då matat in ca 10 bidrag åt gången. Bland bidragen finns flera slangartade ord men också ett antal flerordsuttryck. Betydelsebeskrivningarna är ofta långa och inte sällan anför språkprov. Det stora antalet bidrag över en relativt lång tidsperiod gör att det går att urskilja ett mönster. Användaren har främst lagt in semantiskt besläktade ord och uttryck som kan föras till tre övergripande teman. Dessa är: 1) sex och könsorgan (t.ex. *aktertankare*, *polera hjälmen*, *åderpåk*, 2) ointelligens (t.ex. *felknullad*, *bombad*, *larmet går men brandkåren sover* och 3) meningslös, icke framgångsrik, aktivitet (*slakrunka*, *torrjucka* och *fastna i vinkelvolten*). Det är förstås svårt att säga varför just dessa teman återkommer, men kanske har hen associerat sig fram till mer eller mindre synonyma ord och uttryck till dem som hen redan bidragit med.

5. Sammanfattande diskussion

I den här artikeln har vi redogjort för en studie som utgår från drygt 5 500 bidrag som matats in *Folkmun.se*:s webbformulär 2007–2016. Materialet är omfattande och här har vi bara haft möjlighet att göra några nedslag. I avgränsningssyfte närstuderas tio signaturer som representerar olika typer av bidragsgivare. Personerna ifråga har lagt in allt från 4 till 208 artiklar var. Sammanlagt omfattar vår studie närmare 470 inlägg.

Vår undersökning visar att många bidragsgivare tycks vilja uttrycka något genom sitt val av signatur. Vidare kan vi skönja mönster i deras arbetssätt (hur många ord de matar in åt gången, vilket tempo de har) och vad de bidrar med (simplex/flerordsuttryck, semantiskt innehåll, varietet och stilnivå).

Artiklarna i *Folkmun.se* är, både till form och innehåll, av skiftande kvalitet. Bidragsgivarna har i varierande grad följt riktlinjerna. Tydligt är också att både rasistiska och homofoba bidrag är publika på webbsidan. Det bidrar till att sajten innehåller ett flertal beskrivningar av ord och uttryck som inte är beskrivna i mer konventionella ordböcker. Sajten kan därför potentiellt ha en kodifierande verkan jämförbar med mer etablerade ordböcker.

I tidigare forskning framhålls ofta motsättningen mellan utbildade lexikografer och amatörer respektive traditionellt utarbetade ordböcker och användargenrerade. Också i vår titel antyds att insatserna på *Folkmun.se* görs av amatörlexikografer, men utifrån vår studie kan man fråga sig om det verkligen bara är amatörer som bidrar. Många bidrag är systematiskt uppbyggda och visar på kännedom om grammatik, lexikografi, dialekter osv.

Avslutningsvis kan man fundera kring möjliga motiv till att skicka in ordboksartiklar. Några förefaller vilja testa sajten (se t.ex. signaturerna Dimbo-Klas, Go4it och Tjo fräzzen). Andra verkar primärt vilja sprida kunskaper om eller bevara en dialekt (Benson och Hetrometrus) eller sociolekt (Choooniii). Åter andra tycks vilja lyfta fram mer eller mindre spridda slangord (Luscinia, Tyke) eller tänja på gränserna genom att lägga in slangord som uppenbarligen kan väcka anstöt (Richard (J)). Slutligen tycks några mest vilja orsaka förtretligheter för andra eller förstöra (Mads & Preben). Mer omfattande vandalism förekommer dock inte vilket kan bero på att man på *Folkmun.se* inte redigerar varandras artiklar (jfr t.ex. Wikipedia). Vill man förtydliga innehållet i en annans artikel kan man endast lägga in en ny post.

För att se hur typiska de tio bidragsgivare som studerats här är, skulle vi i framtiden vilja göra fler närstudier. Vi skulle också vilja testa andra undersökningsmetoder, exempelvis lägga upp en enkät på ordbokssajten. Därigenom kan vi få kontakt med personer som bidrar med ord och få mer säkra svar på frågor som inte går att besvara enbart utifrån nuvarande material. Förutom att vi får veta mer om personen bakom en signatur (ålder, geografisk bakgrund m.m.) kan vi få mer kunskap om vad som driver en person att publicera ord på *Folkmun.se* och vilket syfte hen har med sina bidrag.

Litteratur

Ordböcker

Folkmun.se. <<http://www.folkmun.se>> [september 2017].

Wiktionary (svensk version). <<http://sv.wiktionary.org/wiki/Wiktionary>> [september 2017].

Annan litteratur

- Aaltonen, Alexi & Stephan Seiler (2016): Cumulative growth in user-generated content production: Evidence from Wikipedia. I: *Management Science* 62(7), 2054–2069.
- Almind, Richard (2008): Søgemønstre i logfiler. I: *LexicoNordica* 15, 33–55.
- Hanks, Patrick (2012): Corpus evidence and electronic lexicography. I: Granger, Sylviane & Magali Paquot (eds.): *Electronic Lexicography*. Oxford: Oxford University Press, 57–82.
- Hult, Ann-Kristin (2017): *Ordboksanvändning på nätet. En undersökning av användningen av Lexins svenska lexikon*. (Göteborgsstudier i nordisk språkvetenskap 27.) Göteborg.
- Lew, Robert (2011): Online Dictionaries of English. I: Fuertes-Olivera, Pedro A. & Henning Bergenholtz (eds.): *E-lexicography. The Internet, Digital Initiatives and Lexicography*. London/New York: Continuum, 230–250.
- Mattus, Maria (2009): Wikipedia – Free and Reliable? Aspects of a Collaboratively Shaped Encyclopaedia. I: *Nordicom Review* 30, 183–199.
- Meyer, Christian, M. & Iryna Gurevych (2012): Wiktionary: a new rival for expert-built lexicons? Exploring the possibilities of collaborative lexicography. I: Granger, Sylviane & Magali Paquot (eds.): *Electronic Lexicography*, Oxford: Oxford University Press, 259–291.
- Sköldberg, Emma (2015): Bland runda ord och dialektala uttryck. Om den användargenererade sajten Folkmun.se. I: Sandström, Caroline, Cantell, Ilse, Grönros, Eija-Riitta, Nuolijärvi, Pirkko & Eivor Sommar Dahl (red.): *Perspektiv på lexikografi, grammatik och språkpolitik*. (Instituten för de inhemska språken 39.) Helsingfors, 229–241.
- Sköldberg, Emma & Lena Wenner (2017): Folkmun och Luxikon. En jämförelse mellan två användargenererade ordbokssajter. I: *Språk och stil. Tidskrift för svensk språkforskning* 27, 21–48.
- Törnqvist, Lars (2015): Nordiska dialekt- och slangordböcker på Internet. I: *LexicoNordica* 22, 151–169.
- Wolfer, Sascha & Carolin Müller-Spitzer (2016): How Many People Constitute a Crowd and What Do They Do? Quantitative Analyses of Revisions in the English and German Wiktionary Editions. I: *Lexikos* 26, 347–37.

Emma Sköldberg
lektor, docent
Institutionen för svenska språket
Göteborgs universitet
Box 200
405 30 Göteborg
emma.skoldberg@svenska.gu.se

Lena Wenner
forskningsarkivarie, fil. dr
Instituet för språk och folkminnen
Vallgatan 22
411 16 Göteborg
lena.wenner@spraakochfolkminnen.se