

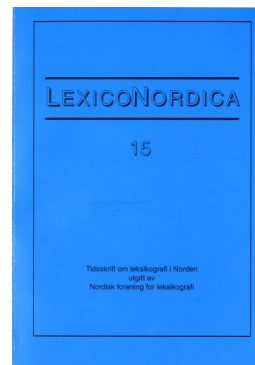
LexicoNordica

Titel: Användarna bakom loggfilerna – redovisning av en webbenkät i Lexin online Svenska ord

Forfatter: Ann-Kristin Hult

Kilde: LexicoNordica 15, 2008, s. 73-91

URL: <http://ojs.statsbiblioteket.dk/index.php/lexn/issue/archive>



© LexicoNordica og forfatterne

Betingelser for brug af denne artikel

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

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

Søgbarhed

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

Ann-Kristin Hult

Användarna bakom loggfilerna – redovisning av en webbenkät i Lexin online Svenska ord

The Internet media gives access to a large number of actual users in more or less real user situations. One year ago I studied the log files from one of the free electronic learner dictionaries in Lexin online (dictionaries for immigrants); the monolingual *Svenska ord*. The log file analysis has now been followed up by a web questionnaire about the users of the very same dictionary. This article presents the result of the questionnaire where users were asked, firstly, basic questions of age, gender, native language etc. and, secondly, questions concerning their searches in the dictionary. Also, the article shortly discusses the conditions for future studies of dictionary users and dictionary usage.

1. Inledning

Med hjälp av loggfiler kan man få reda på vad ordboksanvändare faktiskt slår upp i en ordbok. I jämförelse med traditionella metoder, där informanterna är relativt få och testsituationen mer eller mindre konstruerad, får man via loggfiler tillgång till ett mycket stort antal uppslagningar av verkliga användare. Under våren 2007 genomförde jag en loggfilsstudie som utgick ifrån icke framgångsrika sökningar i den fritt tillgängliga Internetordboken Lexin online Svenska ord (Hult under utg.). I artikeln konstateras bland annat att användarna tenderar att upprepa sina sökningar när de inte får träff i ordboken genom att modifiera sökordet på olika sätt. Jag drar också slutsatsen att användarna förväntar sig mycket av ordbokens lemmauppsättning med tanke på den typ av ord de söker efter och det mycket höga antalet uppslagningar per månad (se vidare Hult under utg.).

Loggfiler avslöjar dock inget om användarnas ålder, kön, sysselsättning etc. De ger heller ingen information om vilka behov användarna vill tillfredsställa eller vad de, efter ordbokskonsultationen, gör med den information de får. De slutsatser som man kan dra av en loggfilsanalys är därför begränsade utan kompletterande studier med hjälp av andra metoder. Under hösten 2007 följdes sålunda loggfilsstudien upp av en enkätstudie på Lexins webbsida för Svenska ord (<<http://lexin2.nada.kth.se/sve-sve.html>>). Syftet med webbenkäten var att få ökad kunskap om användarna bakom loggfilerna. Ett mål var också att få en indikation på hur villiga användarna

var att svara på enkäten och således testa metodens genomförbarhet (inför en kommande större undersökning). Webbenkätstudien är från den aspekten att betrakta som en pilotstudie.

I denna artikel fokuserar jag främst på resultatet från webbenkätstudien men kommer även att beröra tidigare studier av loggfilerna i Svenska ord. Artikeln är disponerad på följande sätt: I avsnitt 2 presenteras kort Lexin online Svenska ord, webbenkäten samt diskuteras urvalet av informanter. Avsnitt 3 ger en bakgrund till användarforskningen och avsnitt 4 redogör för studier av loggfilerna i Svenska ord. Resultatet från webbenkäten redovisas sedan i avsnitt 5. I avsnitt 6 diskuteras slutligen framtida forskning om ordboksanvändning.

2. Material och metod

2.1. Lexin och Svenska ord

Namnet Lexin står för Lexikon för invandrarundervisning. Lexin onlines tryckta förlaga, Lexin-serien, består för det första av en enspråkig svensk ordbok, Svenska ord. Den har tagits fram av Myndigheten för skolutveckling tillsammans med Språkdata under ledning av Martin Gellerstam (se vidare Gellerstam 1999). För det andra ingår ett tjugotal tvåspråkiga lexikon mellan stora minoritetsspråk i Sverige samt engelska och svenska. Svenska ord har fungerat som underlag vid framställning av de tvåspråkiga delarna av ordboksserien och har ca 28 500 uppslagsord. Ordurvalet är gjort med tanke på invandrare som ska lära sig svenska och få tillgång till vanliga ord samt samhällsord som är viktiga för att man ska kunna orientera sig i det svenska samhället. Svenska ord ska också kunna användas som skolordbok. Den fritt tillgängliga Internetversionen av Lexin-serien, Lexin online, har funnits sedan 1994 och i dagsläget är sexton lexikon disponibla, däribland Svenska ord som står i fokus här.

2.2. Webbenkäten

Enkäten lades upp på Lexin online Svenska ords (hädanefters Svenska ord) webbsida under fyra dygn i slutet av september 2007 (se bilaga 1) och resulterade i 370 inskickade svar. Enkätsvaren sammanställdes automatiskt via det s.k. ACE-systemet av Viggo Kann, tekniskt ansvarig för Lexins

webbsidor.¹ Det har inneburit att jag inte kunnat undersöka varje svar separat, vilket förstås är en stor nackdel vid analysen av enkäterna.

Enkäten innehöll tio slutna frågor och en avslutande öppen fråga där informanterna kunde ge synpunkter på Lexin och Svenska ord. De inledande frågorna behandlade informanternas ålder, kön och sysselsättning och följdes av frågor kring modersmål, utbildning och svenskkunskaper. Därefter spordes informanterna om sina sökningar i Svenska ord.

Fjorton av de 370 enkäterna skickades in blanka. Tabell 1 visar svarsfrekvenser för de slutna frågor i enkäten som endast kunde besvaras med ett kryss, fråga 1–7 samt 10. Tabellen ger alltså uppgift om hur många informanter som svarat på respektive fråga. På fråga 8 tilläts informanterna ange fler än ett kryss och antalet informanter som svarat på denna fråga är således okänt. Samma förhållande gäller för fråga 9, där man bara skulle svara med ett kryss men där ett antal informanter har fyllt i flera. Därför medtas inte dessa två frågor i tabellen.

TABELL 1. Svarsfrekvenser för enkätens slutna frågor

| Fråga | Antal svar |
|---------------------------|------------|
| 1. Ålder | 356 |
| 2. Kön | 354 |
| 3. Sysselsättning | 352 |
| 4. Modersmål | 318 |
| 5. Svenskkunskaper | 196 |
| 6 a. Utbildning i Sverige | 328 |
| 6 b. Utbildning i hemland | 130 |
| 7. Sökfrequens i Lexin | 343 |
| 10. Om nöjd med Lexin | 318 |

Svarsfrekvensen skiftar en del mellan frågorna och det är svårt att svara på varför en del informanter inte har fyllt i hela enkäten. Förklaringen till att fråga 5 och 6b har avsevärt lägre svarsfrekvens är att endast de informanter som hade annat modersmål än svenska ombads svara på dessa två frågor. Frågorna 7–10 samt den sista, öppna frågan skulle bara fyllas i av dem som tidigare använt Svenska ord.

¹ Tack till Viggo Kann, professor i datalogi vid KTH i Stockholm, som har lagt upp webbenkäten på Lexin online Svenska ord. Tack också till Christian Mattsson, projektledare för Lexin, för tillstånd att genomföra webbenkäten.

2.3. Informanturval

Användarna av Svenska ord utgör ett nära nog oändligt urvalsunderlag och att genomföra en totalundersökning, dvs. att få samtliga användare från de dagar enkäten ligger uppe att fylla i enkäten, är rent praktiskt en omöjlig uppgift. Jag måste således göra en stickprovsundersökning. Enkätens informanter utgörs av ett så kallat obundet slumpmässigt urval (Körner & Wahlgren 1998:18). Alla användare av Svenska ord har samma möjlighet att komma med i studien eftersom samtliga av dem erbjuds att delta. Det är dock ett positivt urval eftersom det bara är de som är särskilt intresserade som svarar. Jag har heller ingen, eller mycket liten, kontroll över bortfallet vilket gör det svårt att avgöra hur representativt urvalet är i förhållande till den totala populationen, i det här fallet användarna av Svenska ord. Att jag inte har någon möjlighet att styra urvalet i någon riktning minskar emellertid risken för ett snedvridet sampel och ökar möjligheten att "från begränsade iakttagelser [kunna] dra slutsatser av generell natur" (Hyrenius 1962:9).

På Schæffergårdens symposium diskuterades flera generella problem med användarstudier (se Tarp i denna volym), bland andra:

- a) Informanterna är för få.
- b) Informanterna är inte slumpmässigt valda och utgör därför inte ett representativt urval.
- c) Det är svårt att generalisera resultaten.

I traditionella studier är informanterna ofta studenter eftersom de, för de första, använder ordböcker och, för det andra, är lätta att få tag på. Trots detta är det många studier som baseras på få informanter. Exempelvis undersöker Boogards (1998) 45 holländska gymnasister och Laufer (1992) 57 engelskstuderande på universitetsnivå. Det finns emellertid undantag, exempelvis Atkins & Varantola (1998) i vilken 723 studenter ingår. Problemet med för få informanter är dock överkomligt i studier av Internet-ordböcker som Svenska ord där man kan nå många fler användare. Där har man dessutom större möjlighet att göra ett slumpmässigt urval, vilket alltså ökar möjligheten till generaliserbara resultat. Det som jag i detta fall styr är när webbenkäten läggs upp, hur länge den ska vara tillgänglig och hur många svar jag vill ha. Med tanke på hur många uppslagningar som sker i genomsnitt i Svenska ord under en dag är antalet informanter som svarar på enkäten mycket litet. De 356 enkätsvaren utgör endast 0,5 % av uppslagningarna. Om man istället jämför med antalet unika användare, i form

av IP-adresser (i ett senare material), ökar procentandelen till drygt 2 %. Det absoluta antalet enkätsvar (356) indikerar dock att denna metod är tillämplig.

Användarforskningen kommer även i framtiden att brottas med problemet att uppnå ett representativt urval av informanter. Jag anser trots allt att det finns möjlighet att i alla fall delvis lösa de generella problemen med användarstudier tack vare de tekniska förutsättningar som det elektroniska formatet erbjuder.

3. Bakgrund

Det finns inte något ett-till-ett-förhållande mellan det sätt på vilket ordböcker är avsedda att användas och det sätt som ordböckerna sedan faktiskt brukas på. Även om lexikografen har en bestämd uppfattning om för vilket syfte en viss ordbok ska brukas vet denne i slutändan ändå inte om ordboken används för det ändamål som avses. Svenska Akademiens ordlista, normen för svenska språkets stavning och böjning, som endast definierar en ca femtedel av lemmarna, används t.ex. ofta just för att få att få veta vad ord betyder (jfr Granström 2005). Det är heller inte säkert att det som lexikografen anser är en användarvänlig utformning av ordboken upplevs på det sättet av användarna. Förhållandet mellan användaren och ordboken är således komplext och ”[o]lika typer av ordböcker brukas av olika typer av användare med hjälp av olika typer av strategier i samband med olika typer av språkliga aktiviteter för att söka olika typer av information som behövs i olika typer av situationer” (Svensén 2004:533).

Antalet empiriska studier kring ordboksanvändaren och ordboksanvändning har ökat kraftigt för varje decennium sedan 1960-talet. Det ökade intresset för användaren har lett till ett fokus på hur användaren faktiskt använder ordboken istället för hur lexikografen föreställer sig att användaren brukar den. Det som traditionellt studeras är hur ordböcker används vid produktion och reception och i samband med översättning (se Atkins & Varantola 1998; Tono 2001). Man har i huvudsak inriktat sig på användningen av inlärningsordböcker och tvåspråkiga ordböcker (Svensén 2004:533f). Däremot vet man lite om ordboksanvändning av svenska, enspråkiga allmänordböcker och de elektroniska varianter som finns. De flesta undersökningar har också genomförts med enbart studenter som informanter. Hur olika yrkesgrupper, framför allt de som yrkesmässigt arbetar med språk, eller människor i allmänhet använder ordböcker har man mindre kunskap om (Svensén 2004:534). Även användarnas förmåga att använ-

da ordböcker, deras konsultationsfärdigheter, har varit av stort intresse att få kännedom om (se Bejóint 1981) liksom sociologiska aspekter såsom användarnas attityder till ordböcker och vana att använda dem. Tarp (2006:54f) uppmärksammar också de situationer där ett behov av en ordbok kan uppstå men där personen ifråga ännu inte valt att lösa det ordboksrelaterade problemet med att konsultera en ordbok, s.k. extra-lexikografiska situationer.

Vad gäller användningen av Internetordböcker befinner sig forskningen i en initial fas och än så länge är det svårt att veta vad utvecklingen av elektroniska ordböcker av olika slag betyder för ordboksanvändningen. Tack vare de tekniska förutsättningarna finns det goda chanser att studera hur Internetordböcker faktiskt används. Dels med hjälp av de automatiskt registrerade sökningarna, loggfilerna, dels genom möjligheten att nå användarna genom att t.ex. lägga upp en webbenkät. Användarstudier med hjälp av loggfiler är dock än så länge relativt få (se Rogström 2000; de Schryver & Joffe 2004; Bergenholtz & Johnsen 2005, 2007) och även webbenkätstudier är mig veterligen än så länge ovanliga. Det är emellertid inte ovanligt att användare av Internetordböcker uppmanas att ge feedback rörande t.ex. felaktigheter i ordboken och även föreslå förbättringar.

4. Loggfilerna i Svenska ord

I Hult (under utg.) redogör jag för en analys av sökningarna gjorda i Svenska ord. Totalt bestod loggfilmaterialet av ett dygns sökningar vilket utgjorde uppemot 17 000 uppslagningar. Nästan 90 %, ca 15 000, kunde taggas morfosyntaktiskt och av dessa var 53 % substantiv, 24 % verb och 21 % adjektiv. De övriga uppslagningarna bestod av ca 1 % adverb och den resterande procenten av, i fallande ordning, prepositioner, pronomen, konjunktioner, interjektioner samt räkneord.

I studien koncentrerar jag mig på de icke framgångsrika sökningarna (s.k. icke-träffar) som uppgick till 41 % av det totala materialet (ca 7000 sökningar). Jag delar för de första upp dem i olika typer av sökningar. Det visar sig att en stor mängd icke-träffar utgörs av rättstavade svenska ord som inte ingår i lemmauppsättningen i Svenska ord. Vidare är sökningar på fraser av olika slag och encyklopediska sökningar rätt så frekventa. Dessa resultat visar inte bara att loggfiler kan utnyttjas till att upptäcka lemma-luckor utan lyfter också fram användarnas behov av att kunna utföra fritextsökningar så att de hittar fram till sammansättningar, idiom m.m.

Studien visar också att användarna tenderar att upprepa sina sökningar

när de misslyckas. De modifierar då sökordet på olika sätt. Två av de vanligaste modifieringarna, vilka jag kallar strategier, är att pröva annan stavning och att ta bort eller lägga till en ändelse. Av resultaten att döma verkar användarna ha höga förväntningar på ordbokens innehåll. Benägenheten att upprepa sökningarna kan möjligen tolkas som en positiv egenskap hos brukarna, men denna envishet kan också leda till att användarna börjar ifrågasätta sina språkkunskaper. Det gäller framför allt i de fall där de stavar rätt i sin första sökning och sedan börjar stava fel.

Att Svenska ord fyller en viktig funktion för många användare står också klart i och med det mycket höga antalet uppslagningar per månad (se tabell 2).

TABELL 2. Antal uppslagningar i Svenska ord under maj–juli 2006 respektive maj–juli 2007

| Månad | Antal sökningar 2006 | Antal sökningar 2007 |
|-------|----------------------|----------------------|
| Maj | 410 154 | 627 593 |
| Juni | 221 865 | 300 567 |
| Juli | 200 048 | 231 654 |

Vad gäller uppslagningar i en Internetordbok är det inte alltid fallet att de sker från ordbokens webbsida eller ens av en fysisk person. I den danska nätordboken Faste Vendinger (<<http://www.idiomordbogen.dk/>>) visade det sig att hela 80 % av de sökningar som gjordes var så kallade automatregistrerade sökningar från sökmotorn Google (se Almind i denna volym). Dessa typer av sökningar utförs alltså inte av enskilda användare från ordbokens webbsida. För att en sådan sökning ska loggas i Lexin krävs en koppling mellan webbserverns logg och Lexins egensökningslogg, annat än datum och klockslag (Viggo Kann, muntlig kommunikation). Jag har även tittat närmare på IP-adresserna från ett senare loggfilsmaterial och hitintills inte funnit några sökningar som uppträder på ett i sammanhanget onormalt sätt, t.ex. så att en stor mängd sökningar utförs på exakt samma tidpunkt från samma IP-adress. Dessa är således inte fysiskt möjliga att göra av en enskild användare. Så vitt jag förstår förekommer det alltså inte automatregistrerade sökningar i Lexins loggfiler.

Det finns anledning att vidare undersöka loggfilernas funktion som ett effektivt verktyg inom användarforskningen. Men, som sagt, de säger ingenting om vilka användarna är eller något om motivet bakom deras sökningar. I följande avsnitt redovisas så resultatet från webbenkäten.

5. Resultat av enkätstudien

Följande frågor låg till grund för webbenkäten:

1. Ålder?
2. Kön?
3. Sysselsättning?
4. Modersmål?
5. Om du har annat modersmål än svenska, hur bedömer du dina kunskaper i svenska språket?
6. a) Din högsta utbildning i Sverige?
b) Din högsta utbildning i ditt hemland (om annat än Sverige)?
7. Hur ofta brukar du söka i Lexins svenska ordbok Svenska ord?
8. Vilken typ av information söker du oftast efter?
9. I vilka situationer brukar du använda Svenska ord?
10. Brukar du få tag på den information du söker i Svenska ord?

Enkäten avslutades med en öppen fråga där informanterna kunde ge sina synpunkter på Lexin och Svenska ord.

5.1. Ålder, kön och sysselsättning

Enkäten inleddes alltså med frågor kring ålder, kön och sysselsättning. Tabell 3 visar åldersfördelningen bland informanterna.

TABELL 3. Informanternas åldersfördelning

| Ålder | Antal | Procent |
|---------------|------------|------------|
| Upp till 15 | 39 | 11 |
| 15–19 | 153 | 43 |
| 20–29 | 86 | 24 |
| 30–39 | 36 | 10 |
| 40–9 | 31 | 9 |
| 50–59 | 8 | 2 |
| 60 eller mer | 3 | 1 |
| Totalt | 356 | 100 |

Det allra flesta uppger att de är 15–19 år (43 %). Tillsammans med åldersgruppen 20–29 år utgör de 67 % av informanterna. Fördelningen mellan könen är jämn; 53 % kvinnor respektive 47 % män har svarat på enkäten.

På frågan om sysselsättning anger så många som 75 % att de är ”Studerande”, vilket inte är förvånande med tanke på att som sagt 67 % av in-

formanterna är i åldern 15–29 år. Även gruppen ”Yrkesarbetande” är relativt stor med 17 %. Det skulle vara mycket intressant att få veta mer om vilka dessa användare är och hur deras ordboksanvändning ser ut då studier kring icke-studerandes ordboksanvändning är få. De övriga två svarsalternativen, ”Arbetssökande” och ”Annat” kryssades för av 3 respektive 5 % av informanterna. De informanter som kryssade för alternativet ”Annat” ombads precisera sin sysselsättning. De svarade bland annat: *läkare, personlig assistent, lärare, konsult, översättare, grafisk formgivare, ströjobb, vikariat, sjukskriven, företagare*. Det verkar alltså som om även de som kryssat för ”Yrkesarbetande” har svarat på denna följdfråga.

Som redan nämnts har jag inte kunnat analysera varje svar för sig och kan således inte jämföra enkätsvaren med varandra. I en framtida enkätstudie är detta en förutsättning för att mer ingående kunna kartlägga användargruppens utseende. Hur ser t.ex. relationen mellan ålder, kön och sysselsättning ut? Detsamma gäller naturligtvis samtliga frågor i enkäten.

5.2. Modersmål, svenskkunskaper och utbildning

Med tanke på att Lexin-ordböckerna från början är lexikon ämnade för invandrarundervisning, och även tydligt marknadsförs som sådana, förvånar det mig att 57 % av de 318 som svarar anger att de har svenska som modersmål. I den tidigare nämnda loggfilsstudien förmodade jag att användarna inte bara utgjordes av andraspråksinlärare, ett antagande som verifieras här. Det finns dock en möjlighet att de som har svenska som modersmål oftare valde att svara på enkäten. Det kan i detta sammanhang tilläggas att jag även har undersökt loggfilerna i Svenska ord från ett fonomorfologiskt perspektiv där jag tittat närmare på typen av felstavningar i materialet (Hult 2008). Syftet var att ta reda på om typen av stavfel kunde förklaras utifrån de svårigheter med svenskt uttal och stavning som personer med annat modersmål än svenska har. Stavfelen visade sig dock vara sådana som många språkbrukare gör, t.ex. i samband med tj- och sj-ljud och dubbelteckning av konsonant. Dessa problem är inte specifika för personer med svenska som andraspråk. Materialet innehöll alltså inte några uppenbara fall där modersmålet verkade inverka på typen av stavfel. Detta bekräftar det resultat som webbenkäten visar, nämligen att en stor del av användarna utgörs av personer med svenska som modersmål eller av personer som har mycket goda kunskaper i svenska.

Förutom svenska är 49 språk representerade bland svaren (inklusive kreol, teckenspråk och det oidentifierade kldayiska). Sammanlagt svarar

136 informanter att de har ett annat modersmål än svenska. De fyra vanligaste modersmålen är arabiska (23), persiska (18), kurdiska (15) och spanska (10). Övriga språk har mellan en och fem representanter (se bilaga 2). Nio informanter anger två eller tre språk. Man kan emellertid fråga sig vad användarna lägger i begreppet modersmål och det finns anledning att fundera kring följande spörsmål: Vad är ett modersmål? Kan man ha fler än ett modersmål? Är *förstaspråk* ett bättre ordval i det här sammanhanget? Frågan kring vad forskare, och inte minst informanterna, uppfattar vara ett modersmål är för omfattande att diskuteras i detta sammanhang men är ändå viktig att lyfta fram.

Påföljande fråga i enkäten gäller hur man bedömer sina kunskaper i svenska om man inte har svenska som modersmål. Sammanställningen visar att 61 % av informanterna bedömer sig ha bra eller mycket bra kunskaper i svenska. Det är dock av flera skäl vanskligt att ställa den här typen av fråga. Hur bedömer man själv sina kunskaper i ett andraspråk? Och hur vet uttolkaren vad varje enskild användare lägger i uttrycken ”Ganska bra” respektive ”Mycket bra”? Frågan är således också svår att utforma svarsalternativ till. Det blir i slutändan ytterst svårt att tolka svaren på ett vettigt sätt. Man skulle således behöva komplettera med både skriftliga och muntliga test för att kunna säga något om detta.

Frågan om utbildning behöver också modifieras i en framtida version. I enkäten framgick inte huruvida utbildningen skulle vara avslutad. Med tanke på att de flesta informanter är i åldern 15–29 år är det kanske oftare icke avslutad utbildning som åsyftas, men det är svårt att uttala sig om.

På frågan om högsta utbildning i Sverige svarar 32 % grundskola, 46 % gymnasium, 4 % folkhögskola och 18 % universitet. De allra flesta har gått eller går således i grundskolan eller gymnasium och en relativt stor andel studerar eller har studerat på universitetet. På frågan om högsta utbildning i hemlandet (om annat än Sverige) svarar 130 informanter. Den siffran stämmer någorlunda väl med de 136 som svarar att de har annat modersmål än svenska (se ovan). Av dessa anger 42 % grundskola, 28 % gymnasium, 5 % folkhögskola och 25 % universitet. De flesta har alltså även här svarat grundskola eller gymnasium men en ansevärd mängd har också någon slags universitetsutbildning.

5.3. Användarna om sina sökningar i Lexin

Som framgår av tabell 2 (se avsnitt 4) utfördes under maj månad 2007 uppemot 628 000 sökningar i Svenska ord (i genomsnitt drygt 20 000/dag). På

frågan om hur ofta informanterna brukar söka i Svenska ord svarar 44 % ”Flera gånger i veckan”, 30 % ”Flera gånger i månaden” och 13 % ”Högst en gång i månaden” respektive ”Detta är första gången”. Detta visar att det inte bara är många uppslagningar i Lexin utan även många unika användare.

De nästföljande frågorna kring informanternas sökningar handlar om vilken typ av information användaren oftast söker efter (8), i vilka situationer de brukar använda Svenska ord (9) samt om de brukar få tag på den information de efterfrågar (10). På frågan om mest eftersökta informationstyp tilläts informanterna att kryssa i flera alternativ. Fördelningen av de totalt 811 kryssen redovisas i tabell 4. Informationstyperna har valts ut efter vilket innehåll som ordboksartiklarna i Svenska ord har.

TABELL 4. Mest sökta informationstyp i Lexin online Svenska ord

| Informationstyp | Antal kryss | Procent |
|--------------------------------|-------------|------------|
| Betydelse | 250 | 31 |
| Stavning | 166 | 21 |
| Exempel | 91 | 11 |
| Böjning | 54 | 7 |
| Uttal | 45 | 5 |
| Stående uttryck | 45 | 5 |
| Sammansättningar/avledning | 44 | 5 |
| Ordklass | 41 | 5 |
| Uttal via ljudfunktion | 31 | 4 |
| Bildtema | 22 | 3 |
| Annat: översättning, synonymer | 22 | 3 |
| Totalt | 811 | 100 |

”Betydelse” hamnar överst i uppställningen och svarar tillsammans med ”Stavning” för 52 % av de totalt 811 kryssen. Det här resultatet styrks i hög grad hos Svensén (2004:545) som jämför olika studiers resultat beträffande mest sökta informationstyp i ordböcker. Under ”Annat” har majoriteten av informanterna svarat att de söker efter synonymer eller att de använder Svenska ord i samband med översättning.

Att många informanter vill ha reda på ords betydelse bekräftas också i följande fråga där nästan hälften (49 %) av informanterna uppger att de använder ordboken för att förstå ett ord i en text de läser (reception). En stor del (35 %) svarar också att de använder Svenska ord för att få hjälp med ett ord när de skriver (produktion). I övrigt svarar 41 informanter (9 %) att de använder ordboken för att få veta hur ett ord uttalas. Denna siffra stämmer

väl med de 45 som kryssat för ”Uttal” i frågan om mest sökta informationstyp (se tabell 4). Resterande 7 % svarar ”Annan situation” och de har preciserat sina svar på olika sätt. Några specificerar vilken situation de befinner sig i: ”Jobbar med TTS utveckling och behöver något att verifiera uttalsfrågor, ordklassfrågor med”, ”till min läxa” eller ”förberedelser inför föreläsningar etc”. Andra specificerar varför de använder Lexin: ”För att få veta hur ett ord böjs”, ”när jag behöver hjälp att förklara ett ord”, ”översätta ord” eller ”För att hitta exakt rätt uttryck, för språklig variation”.

En stor majoritet av informanterna anser att de helt eller delvis brukar få tag på den information de söker efter; 46 % svarar ”Ja” och 50 % svarar ”Delvis”. Detta resultat kan tyckas förvånande med tanke på att andelen misslyckade sökningar i Svenska ord i genomsnitt ligger runt 40 % (se avsnitt 4).

5.4. Användarnas kommentarer kring Lexin och Svenska ord

Ett hundratal informanter har valt att svara på den sista frågan som lyder: ”Har du någon kommentar kring Lexin och Svenska ord? Är det till exempel något du saknar i ordboken?”.

Drygt hälften av kommentarerna handlar om att informanterna saknar ord: dels ord från allmänspråket, dels svåra ord och fack- och ämnesord inom t.ex. biologi, medicin, psykologi, politik och teknik. Ett par informanter specificerar exakt vilka fackord de saknar. Många skriver också längre inlägg om vad de tycker är bra och dåligt, vad som kan förbättras och i vilka situationer de använder Lexin och även, som frågan inbjöd till, mer allmänt om Lexin. Ett tjugotal informanter efterfrågar till exempel fler tvåspråkiga lexikon, bland annat på språk som exempelvis tyska och franska vilka inte är typiska moderna minoritetsspråk i Sverige.

Informanternas omdömen och önskemål om Lexin och Svenska ord förstärker min uppfattning att användarna inte verkar medvetna om vilken typ av ordbok Lexin är eller vilken målgrupp den ursprungligen är anpassad för. I en uppdatering av Internetversionen bör man eventuellt fundera över om man tydligare ska skilja på målgruppen för Lexins tryckta förlaga och Internetversionen och på så sätt anpassa den senare till användarnas önskemål. Internetversionen utgörs idag i stort sett av den tryckta versionen i digitalt format och man har i mycket liten utsträckning utnyttjat de fördelar som det elektroniska formatet erbjuder.

5.5. Sammanfattning

Trots att det inte har varit möjligt att undersöka varje enkätsvar separat har jag ändå fått viss kännedom om vilka användarna bakom loggfilerna är. Majoriteten av dem är studerande i åldern 15–29 år, men andelen högutbildade och yrkesarbetande är också relativt hög. Med tanke på vilka som utgör Lexins åsyftade målgrupp har en överraskande stor andel av informanterna svenska som modersmål. För många av användarna är Lexin förmodligen inte ett lexikon för invandrarundervisning utan en lättillgänglig och framför allt gratis ordbok på Internet. I övrigt anges 49 olika modersmål (arabiska, persiska, kurdiska och spanska uppges flest gånger).

Många av informanterna söker ofta i ordboken och de mest efterfrågade informationstyperna är betydelse och stavning, närmast följt av språkexempel. Svenska ord verkar användas i ungefär lika stor utsträckning för reception och produktion. Så många som 96 % av informanterna brukar alltid eller delvis få tag på den information de söker, att jämföra med de i genomsnitt ca 40 % icke-framgångsrika sökningarna per dag i min tidigare undersökning. I fältet för kommentarer skriver många informanter längre inlägg och motiverar varför de tycker att något är bra eller dåligt och vad som kan bli bättre. De uppvisar ett intresse för ordboken men avslöjar också sin okunskap om vilken typ av ordbok Lexin online är.

En framtida webbenkät behöver modifieras på flera sätt. Dels vad gäller utformning av frågorna så att missförstånd undviks, dels vad gäller frågornas urval och relevans. Dessutom skulle enkäten kunna finnas tillgänglig på andra språk än svenska.

Vad gäller det slumpmässiga informanturvalet kan man inför framtida studier av Lexin diskutera intresset av att undersöka de användare som har svenska som modersmål när Lexin inte är ämnad för denna målgrupp. Man skulle kunna välja ut de informanter som har svenska som andraspråk från det slumpmässiga urvalet och således fokusera på Lexins ursprungliga målgrupp och funktion som ett lexikon för invandrarundervisning.

6. Framtida forskning om ordboksanvändning

Ordboksanvändning är ett eftersatt forskningsområde inom svensk, för att inte säga nordisk, lexikografi. Innan man ger sig i kast med en större studie av brukarna och bruket av svenska ordböcker bör man avsätta ordentligt med tid för att noggrant planlägga sin undersökning. För det första ska man vara klar över vad det är man vill undersöka. För det andra är det i regel

nödvändigt att ta till flera metoder för att nå fram till resultat av värde. Som enda metod har t.ex. loggfiler eller en webbenkät begränsad nytta men tillsammans med andra kan de däremot bidra med värdefull kunskap om användarna och deras ordboksanvändning. Det är för det tredje av största vikt att försäkra sig om ett representativt urval av informanter, loggfiler etc. När det gäller studiet av Internetordböcker är det flera tekniska villkor som måste uppfyllas för att man ska kunna uppnå statistiskt hållbara resultat. Det handlar främst om att ha tillgång till loggfilernas IP-adresser och att rensa bort de automatregistrerade sökningarna från sökmotorn Google, i de fall de existerar. Vad gäller webbenkäter ska svaren kunna analyseras separat och i bästa fall ska de kunna kopplas till respektive enkätsvarares sökningar i ordboken.

Man bör naturligtvis även ta till andra metoder. Till exempel kan man genomföra intervjuer med ett urval av dem som svarat på enkäten. Så kallad "eye tracking", där ögats rörelser över en dataskärm registreras, är också en metod värd att undersöka. Det finns därtill intressanta alternativa metoder som kan användas för att söka reda på de ordboksrelaterade behov som uppstår i extra-lexikografiska situationer (se Tarp i denna volym). I en mer heltäckande studie av bruket av svenska Internetordböcker behöver även materialet utökas. Jämte Lexin online finns t.ex. Svenska Akademiens ordbok (<<http://g3.spraakdata.gu.se/saob/>>) och Nationalencyklopedins ordbok (<<http://www.ne.se>>). Genom att tillämpa flera metoder och undersöka olika typer av ordböcker hoppas jag att jag ska kunna dra mer generella slutsatser om användningen av svenska Internetordböcker.

Litteratur

- Almind, Richard 2008: Søgemønstre i logfiler. I: *LexicoNordica 15*, 33–55.
- Atkins, Sue B. T. & Krista Varantola 1998: Language learners using dictionaries: The final report on the Euralex/AILA research project on dictionary use. I: Atkins, S. B. T. (ed.), *Using dictionaries. Studies of Dictionary Use by Language Learners and Translators*. (Lexicographica Series Maior 88.) Tübingen: Niemeyer, 21–81.
- Bejoint, Henri 1981: The Foreign Student's Use of Monolingual English Dictionaries. A Study of Language Needs and Reference Skills. I: *Applied Linguistics*. Vol. II. No. 3, 207–222.
- Bergenholtz, Henning & Mia Johnsen 2005: Log files as a Tool for Improving Internet Dictionaries. I: *Hermes – Journal of Linguistics 34*, 117–141.

- Bergenholtz, Henning & Mia Johnsen 2007: Log Files Can and Should Be Prepared for a Functionalistic Approach. I: *Lexikos 17*, 1–21.
- Bogaards, Paul 1998: What type of words do language learners look up? I: Atkins, S. B. T. (ed.), *Using dictionaries. Studies of Dictionary Use by Language Learners and Translators*. (Lexicographica Series Maior 88.) Tübingen: Niemeyer, 151–158.
- de Schryver, Gilles & David Joffe 2004: On How Electronic Dictionaries are Really Used. I: Geoffrey, W. & S. Vessier (eds.), *Proceedings of the eleventh EURALEX International Congress. Lorient, France, July 6–10, 2004*. Lorient: UBS, 187–196.
- Granström, Alexandra 2005: (Hur) använder du ordböcker? En undersökning bland svensklärarstuderande och svensklärare. I: *Nordiske Studier i Leksikografi 8. Rapport fra Konference om leksikografi i Norden, Sønderborg 24–28 maj 2005*, 121–128.
- Gellerstam, Martin 1999: LEXIN – lexikon för invandrare. I: *Lexico-Nordica 6*, 3–18.
- Hult, Anki (under utg.): Från ord till handling. En studie i ordboksanvändning på nätet. I: *Nordiska studier i lexicografi. Rapport från 9. Konference om lexicografi i Norden, Akureyri 22.–26. maj 2007*.
- Hult, Ann-Kristin (2008): *Användarnas sökstrategier i en elektronisk ordbok*. Institutionen för svenska språket. Göteborgs universitet (opubl. kursuppsats).
- Hyrenius, Hannes 1962: *Statistiska metoder*. 4 uppl. Stockholm: Almqvist & Wiksell.
- Körner, Svante & Lars Wahlgren 1998: *Statistiska metoder*. Lund: Studentlitteratur.
- Laufer, Batia 1992: Corpus-based versus lexicographer examples in comprehension and production of new words. I: Tommola et al. (eds.), *EURALEX'92 Proceedings I–II. Papers Submitted to the 5th EURALEX International Congress on Lexicography in Tampere, Finland*. Tampere: Tampereen Yliopisto, 71–76.
- Rogström, Lena 2000: Användarstrategier i OSA-databasen. I: Byrman G., H. Lindquist & M. Levin (utg.), *Korpusar i forskning och undervisning. Rapport från ASLA:s höstsymposium, Växjö, 11–12 november 1999*. Uppsala, 255–265.
- Svensén, Bo. 2004. *Handbok i lexicografi: Ordböcker och i teori och praktik*. 2 uppl. Stockholm: Norstedts Akademiska Förlag.
- Tarp, Sven 2006: *Leksikografien i grænselandet mellem viden og ikke-viden. Generel leksikografisk teori med særlig henblik på lærerleksikografi*. Bind 1. Handelshøjskolen i Århus: Center for lexicografi.

- Tarp, Sven 2008: Kan brugerundersøgelser overhovedet afdække brugernes leksikografiske behov? I: *LexicoNordica 15*, 5–32.
- Tono, Yukio 2001: *Research on Dictionary Use in the Context of Foreign Language Learning. Focus on Reading Comprehension*. (Lexicographica. Series Maior 106.) Tübingen: Niemeyer.

Elektroniska referenser

<<http://g3.spraakdata.gu.se/saob/>>
<<http://lexin2.nada.kth.se/sve-sve.html>>
<<http://www.idiomordbogen.dk/>>
<<http://www.ne.se>>

Ann-Kristin Hult
Doktorand
Institutionen för svenska språket
Box 200
SE-405 30 Göteborg
ann-kristin.hult@svenska.gu.se

Bilaga 1. Frågorna som ställdes i webbenkäten

- 1. Ålder:** upp till 15
 15–19
 20–29
 30–39
 40–49
 50–59
 60+

- 2. Kön:** Kvinnligt Manligt

- 3. Sysselsättning:** Studerande
 Yrkesarbetande
 Arbetsökande
 Annat Vad: _____

- 4. Modersmål:** _____

5. Om du har annat modersmål än svenska, hur bedömer du dina kunskaper i svenska språket?

Dåliga Ganska bra Bra Mycket bra

- 6. a) Utbildning i Sverige:** Grundskola (eller motsvarande)
 Gymnasium (eller motsvarande)
 Folkhögskola
 Universitet
- b) Utbildning i ditt hemland
 (om annat än Sverige):** Grundskola (eller motsvarande)
 Gymnasium (eller motsvarande)
 Folkhögskola
 Universitet

7. Hur ofta brukar du söka i Lexins svenska ordbok, Svenska ord?

Detta är första gången
 Högst en gång i månaden
 Flera gånger i månaden
 Flera gånger i veckan

Resten av frågorna gäller dig som har sökt i Svenska ord förut.

8. Vilken typ av information söker du oftast efter? Du kan kryssa i flera alternativ.

Stavning

Uttal

Uttal via ljudfunktion

Böjning

Ordklass

Betydelse

Exempel

Stående uttryck

Sammansättningar/avledningar

Bildtema

Annat Vad: _____

9. I vilken situation brukar du använda Svenska ord?

För att förstå ett ord i en text som du läser

För att få hjälp med ett ord när du skriver en text

För att få veta hur ett ord uttalas

Annat Vad: _____

10. Brukar du få tag på den information du söker efter i Svenska ord?

Ja

Nej

Delvis

Har du någon kommentar kring Lexin och Svenska ord? Är det till exempel något som du saknar i ordboken? Skriv gärna det här!

Tack för din medverkan!

Bilaga 2. Antal användare/modersmål:

| | |
|-----|---|
| 182 | svenska |
| 23 | arabiska |
| 18 | persiska |
| 15 | kurdiska |
| 10 | spanska |
| 5 | engelska, polska |
| 4 | bosniska, finska, thailändska |
| 3 | albanska, franska, turkiska |
| 2 | dari, grekiska, kroatiska, serbiska, syrianska, ungerska |
| 1 | armeniska, azeriska, berbiska, bulgariska, estniska, farsi, filippinska, gujarati, hebreiska, isländska, kantonesiska, kil-daniska, kinesiska, kreol, litauiska, malaysiska, mandarin, nederländska, norska, pashtu, rumänska, ryska, serbokratis-ka, slovenska, somaliska, tagalog, teckenspråk, tjeckiska, turkmeniska, urdu, vietnamesiska |