

Lätta och svåra texter – Läsbarhet

Av Margareta Westman

Det finns mycken krånglig svensk prosa och det görs många ansträngningar att komma till rätta med den. Alltför invecklat språk är ett problem bl.a. i skolböckerna, och detta har sporrat pedagoger och statistiker att försöka få fram en metod att mäta svårighetsgraden hos olika texter.

I en bok från förlaget Liber som heter *Läsbarhet* presenterar C.H. Björnsson en idé om hur man ska mäta språklig tillgänglighet. Här beskriver han utförligt det s.k. läsbarhetsindex som han redan tidigare har publicerat i mer kortfattade redogörelser. Det som avses med läsbarhet är den egenskap eller summa av egenskaper som gör språket i en text svår- eller lättläst. Det är en ganska renodlad mätning av teckensystemet som han syftar till, en mätning där textens framställningssätt och innehåll inte ska spela in.

Läsbarhetsindexet, *lix*, räknas ut på så sätt att det genomsnittliga antalet ord per mening i en text adderas med procenttalet för långa ord. Med långa ord menas ord med mer än sex bokstäver. Med mening avses den ordräcka som står mellan stor bokstav och stor bokstav (frånsett egennamn o.d.). Författaren och hans arbetsgrupp inriktade sig på att prova fram ett sätt att mäta språklig svårighetsgrad som var så enkelt att vem som helst kan använda det. (Förmågan att räkna blev viktigare än förmågan att läsa.)

Tanken att meningarnas längd har samband med deras komplexitet torde väl ingen vilja bestrida. Men denna meningsräkning blir mycket mekanisk och man undrar om det inte finns bättre mått. Björnsson diskuterar också meningsbegreppet men tycks ha den uppfattningen att en mening är eller borde vara samma sak som ett huvudsatskomplex, d.v.s.

antingen en enda huvudsats eller en huvudsats med bisatser. Han talar nämligen om meningsskiljande semikolon, tankstreck och komman (!). Å andra sidan anser han att meningar utan predikat inte är några "riktiga" meningar. För Björnssons lix används dock som sagt stor bokstav som meningsgräns. Emellertid är det möjligt att just en mätning av de finita huvudsatsverbens frekvens i ordmassan (ev. relaterad till interpunktionen) bättre skulle spegla hur invecklat språket i en text är, eftersom man där skulle registrera skillnaden i lästlängd mellan meningar som består av uppbyggda huvudsatser och periodiskt uppbyggda meningar. Man sluppe också att uppmuntra reklamtextinterpunktion. Med punkt. Mellan praktiskt taget. Varje satsdel.

Att satsbyggnaden kan vara mer eller mindre komplicerad utan att meningarna egentligen blir längre är ingen statistisk sanning men likafullt sant. Det är alltid svårare för en läsare att läsa och uppfatta en mening som på något sätt avviker från de vanligaste satsbyggnadsmönstren. Verbets delar kan vara åtskilda, subjekt och predikat stå långt ifrån varandra. Flerordiga bestämningar som står före sitt huvudord är i allmänhet mera svårlästa än de som står efter. Vissa konstruktioner är så ovanliga att bara deras förekomst försvårar läsningen. Inget av detta märks när man mäter med lix. Hellre än att bara räkna orden borde man försöka att på något sätt väga de skilda satsledens tyngd mot varandra.

Att Björnsson och hans medarbetare som andra faktor stannade för ordlängden beror på resultaten från flera undersökningar. Ord med mer än sex bokstäver tycks få ökad frekvens ju svårare texten blir, medan ord med färre bokstäver däremot avtar i frekvens i svårare texter. Detta vore alltså en statistisk sanning. De redovisade undersökningarna över just detta är inte särskilt omfattande (ca 55 000 ord) men det är ju ändå möjligt att det förhåller sig så. En följd av detta blir då att en berättelse i imperfekt skulle anses något svårare än en i presens, eftersom så många verb böjs enligt mönstret "kallar - kallade". Namn räknas också efter antal

bokstäver, varje gång de uppträder. En historia om Andersson, Pettersson och Lundström skulle vara svårläsare än en om Olson, Berg och Lund.

Själva beräkningen av ord och långa ord är naturligtvis problematisk. Sålunda upplöses förkortningar och alla siffror och tal läses ut. Begreppet ideogram förefaller författaren helt obekant och därför får man sådana orimligheter som att talet 12 ska räknas som kort ord men 14 som långt. Ett decimalbråk som 36,21 räknas som två långa ord. En konsekvens av detta är att alla tal lika gärna kan skrivas med bokstäver, det har ingen inverkan på läsbarhetsindexet.

Sexhundrasträttioåttahundrafyrtionio skulle vara lika lätt eller svårt att uppfatta som 673 849.

Hit hör också problemet med ordvalet. Vi har tyvärr ännu ingen förteckning över vad man skulle kunna kalla det centrala och gemensamma ordförrådet för alla svenskar. Men så mycket vet vi dock att de vanligaste orden är det lättbegripligaste. Detta gäller både i allmänhet och i snävare sammanhang. T.ex. *ordet förgasare* är kanske inte ett vanligt ord, abstrakt och totalt sett, men det är ett väldigt vanligt ord i beskrivningar av bilmotorer. Det är också det ord man väntar sig som beteckning för företeelsen. Att kalla *förgasare* för *grej* skulle inte göra språket i beskrivningen lättbegripligare. I en skriftlig framställning bygger man dessutom i regel på vad man redan har sagt, ett ord är inte lika svårt varje gång det förekommer.

Metoden att räkna meningslängd och procentandel långa ord kom C.H. Björnsson och hans arbetsgrupp fram till efter att ha provat flera metoder. Man försökte mata olika faktorer, bl.a. ovanliga ord, abstrakta ord, bisatser per mening, personord etc. Emellertid tycks sättet att definiera t.ex. faktorn abstrakta ord rätt märkligt. Man provade och fann så småningom att det skulle räcka att beakta sex suffix, nämligen *-ande*, *-ende*, *-het*, *-ing*, *-tet* och av någon anledning verbsuffixet *-era*. Frekvensen för ord med dessa suffix beräknades, dock endast när

de uppträdde i sin grundform. Och av någon besynnerlig anledning (sneglände mot en tänkt, dåligt programmerad datamaskinsberäkning?) räknade man alla ord som slutade på dessa bokstavskombinationer, t.ex. också ord som *flera*, *stekhet*, *ring*, *sotet*. Knappast någon svenskspråkig person skulle uppfatta de orden som innehållande vad vi kallar suffix.

Jag tar upp saken för att belysa den fyrkantighet som vidlåder många av de språkliga resonemangen i boken, särskilt i kontrast med den elegans varmed erhållna värden attenuationskorrigeras eller omräknas till stanineskala. Men det är också synd att frågan om språkets abstraktionsnivå fuskas bort eftersom den verkligen berör en allvarlig svaghet hos stora delar av vår sakprosa ur begriplighets-synpunkt.

För att få reda på vilka texter som kan anses svåra och vilka som kan anses lätta, och därmed prova pålitligheten hos lix, gjorde man dels en skolboksundersökning med elever och lärare, dels en allmän undersökning. I den senare deltog 26 vuxna personer i olika åldrar och med helt skilda sysselsättningar. De fick bedöma sammanlagt 80 textstycken på vardera 500 ord av olika litteraturarter från böcker för barn över skönlitteratur, dags- och veckopress och saklitteratur till facklitteratur. Deras uppgift var att svårighetsgradera texterna inbördes. Det visade sig då att Björnssons lix i stort sett ordnade texterna i samma ordning som "Medelvensson" gjorde när han tyckte till. Dessvärre gjordes tydligen texturvalet helt på känn. "Valet av texter gjordes inte särskilt systematiskt" (s, 49). Det är svårt att tro att detta begränsade urval kan täcka alla arter av nutida svenskt skriftspråk.

I sin glädje över överensstämmelsen diskuterar så författaren hur lix ska användas. På den punkten blir man rätt betänksam. Han menar att alla böcker och alla texter kunde förses med en siffra för lix som ett slags konsumentupplysning. I boken ges också medelvärdestabeller för

lix i olika svåra texter: mycket lätt text - lix 20, lätt text — lix 30, medelmåttig text - lix 40, svår text — lix 50, mycket svår text - lix 60. Visserligen varnas man för att uteslutande förlita sig på lix men samtidigt framhålls hur praktiskt det är för att bedöma texter. Det sägs att lärare skulle ha stor glädje av detta och än mer elever. Författaren menar tydligen på fullt allvar (s. 117) att eleverna i ett skolbibliotek som nu bläddrar i böcker de tänker låna, tittar efter bilder och läser litet här och var på prov, att dessa elever i stället borde få "svart på vitt genom att (man) i boken ange(r) läsbarhetsindex".

C.H. Björnsson polemiserar (s. 96) mot invändningar som uppenbarligen har framförts, mot att lix används för bedömning av skönlitterära texters språkliga tillgänglighet. Men det är inte främst konstverkets helighet frågan gäller utan det är snarare så att lixuppgifter troligen är rätt meningslösa när det gäller kvalitetsskönlitteratur. Av bilagorna framgår också att "Medelvensson" vid bedömningen av de 80 texterna tyckte att ett fotbollsreportage med lix 42 var enklare att läsa än Eyvind Johnsons *Tidens gång* med lix 30. Ett kåseri av Kar de Mumma kom på 19:e plats bland de 80 texterna (från lätt till svårt), medan Sven Lindquists *Hemmaresan* kom på 44:e. Båda hade lix 35. Ser man så vidare i tabellerna i bilagorna över beräkningar av lix för olika skönlitterära verk och jämför siffrorna med uppgifter om lämpliga lixnivåer för olika årskurser i skolan, kommer man till slående resultat. Hur ska en djurintresserad liten skolelev i andra klassen förstå att hon ska välja *Tatte nallebjörn* (barnbok, lix 25) och inte Vesaas *De svarta hastarna* (lix 23), fastän den senare har ett språk som enligt tabellen s. 225 passar just för åttaåringar. På samma sätt skulle Lagerkvists *Bödeln* och Hemingways *Den gamle och havet* passa för årskurs 4 (tioåringar).

Invändningar av den här typen bemöts på s. 115 med hänvisning till att det här är fråga om medeltal. Men det rimmar illa med förslaget

om lixbeteckning på alla texter och böcker. Om lixuppgiften kan vara verkligt missvisande, då anger inte läsbarhetsindexet textens svårighetsgrad. Långa parataktiska meningar kan höja lix utan att höja språkets otillgänglighet, en knutig syntax kanske inte påverkar lix utan bara språkets begriplighet, korta men ovanliga ord kan sänka lix men göra texten svårare.

Men viktigare ändå är att den rent språktechniska sidan inte alltid är lika betydelsefull när det gäller för en läsare att tillägna sig en text.

Författaren påpekar om och om igen att vad man eftersträvar är ett objektivt mått på språkets svårighetsgrad. Risken blir att detta tekniska mätande tilldelas för stor betydelse. Det är ju lätt att tro att en låg lixsiffra garanterar att en text, en bok är mycket lättlast, d.v.s. lätt att tillägna sig. Kan man alls bortse från framställningssättets betydelse? När författaren på s. 16 vill visa att man kan det, säger han bara att "samma" yttre ämne kan behandlas på olika sätt. Jesus i templet kan framställas som en ren barnberättelse eller behandlas i en svår teologisk utredning. Olika framställningssätt således. Framställningssättet är just det språkliga greppet över det ämne som skildras och kan inte skiljas från språket. I halvt medvetande om detta gör författaren också den reservationen att man inte kan mäta poesi med lix.

Eftersom lix enligt författaren inte ska användas ensamt utan bör kombineras med en allmän bedömning, kan man undra vad lixberäkningen egentligen ska tjäna till. I denna bok som vimlar av uttryck som interkorrelation, attenuationskorrigering, validitetskoefficienter, ideal sammanvägning etc. upprepas ideligen att man måste kunna komma på ett sätt att bedöma språksvårighetsgrad som inte kräver insikter i språklära. Det ligger mycket i det. Men kan inte praktiskt taget alla svenskspråkiga personer redan bedöma svensk text i det här avseendet, i varje fall om skillnaderna är stora? Och är skillnaderna inte märkbara, så märks de ju inte. Man kan fråga sig om

någon person i undersökningen avvek mycket kraftigt från de andras åsikt om vilken text som var svårare och vilken som var lättare. Om icke visar det att lixberäkningar är överflödiga. Och om någon verkligen avvek från de andras åsikt skulle lix i hans fall inte vara något läsbarhetsindex.

Även om lix skulle ha visat sig vara en mätare som vid flera tillfällen hade graderat språklig tillgänglighet på ungefär samma sätt som "Medelvensson", får man och kan man inte göra en kullerbytta och använda lix som om det vore en analysmetod. Men det gör Björnsson. Lix ska alltså användas inte bara för mätning och gradering av existerande text utan också som "objektivt" påverkningsmedel på skribenter. "Objektivt" tycks här betyda "som är uttryckt i siffror". På s. 136 skriver författaren och hans medarbetare: "Vi till och med tror att de sporadiska framstötarna om att skriva läsbart som numera allt oftare görs blir mer eller mindre verkningslösa om de inte sammankopplas med läsbarhetsprovning och lix." Men för det första så kan inte en statistisk sanning räknas som sanning i alla enskilda fall. För det andra tycks författarna möjligen ha en aning men ingen egentlig kännedom om den stora förändring som har skett och sker inom svensk sakprosa för närvarande. För det tredje är en mera direkt språklig hjälp säkert till större nytta på de områden där man ännu skriver för krångligt. Alltså hellre än att räkna med lix ska skribenter av skilda slag få regler för hur man lägger upp en framställning, skriver mindre invecklade meningar och använder begripligare ord och ett konkretare språk, gärna genom exempel ur autentisk text.

Ett exempel på hur lixberäkning kan påverka en text erbjuder själva boken *Läsbarhet*. Jag räknade efter bokens anvisningar och fick fram ett lix på 44 vilket enligt tabellerna anger texten som medelmåttig (till svar). Men sedan kommer det intressanta: det genomsnittliga antalet ord per mening var 14 vilket enligt tabell tjugo karakteriserar lätt text.

Procentandelen långa ord blev 30 vilket karakteriserar svår text. I stället för att ge akt på sitt ordförråd skulle således en skribent kunna koncentrera sig på att sätta punkt ofta. Förhållandet mellan ord och meningsbyggnad — lix mäter inte någondera delen vidare bra — är inte enkelt kompensatoriskt; man kan inte ta igen på gungorna vad man förlorar på karusellen. Obegripliga ord i enkla meningar gör texten obegriplig. Riktigt snåriga meningar gör texten otillgänglig även om orden är enkla.

Ska vi ändå försöka frambringa en mät- och analysmetod för språklig svårighetsgrad, kan vi knappast hoppas att den kan gälla för annat än sakprosa och nybörjarlitteratur. Och i stället för att bygga den på statistiska beräkningar som grundar sig på otillåtligt grova mått, borde vi kanske ändå ta till vara den insikt *i* språkets byggnad och funktion som faktiskt finns tillgänglig. En någorlunda pålitlig metod måste föregås av omfattande undersökningar av ords och ordformers och satsstrukturers frekvenser, kort sagt av skilda språkmönster. Först på den grunden kan någon meningsfull statistik göras upp. Samtidigt måste nog viss hänsyn tas till framställningssättet; vilken roll spelar uppläggningsvariationer och upprepningar, konkretion och abstraktion för läsbarheten? Språket är ett långt mer komplicerat system, även rent tekniskt, än lix upphovsman har klart för sig.

Margareta Westman var professor i svenska och föreståndare för Svenska språknämnden från 1985 till sin bortgång år 2000.

Denna text publicerades första gången i tidskriften *Språkvård* 1969 och senare i boken *Språkets myller* (Norstedts ordbok). Hennes efterträdare som föreståndare för Språknämnden, Olle Josephson, samlade där ett antal av det hon skrivit.

www.fungerandemedier.se publicerar texten med godkännande av redaktionen för Språkvård och Norstedts ordbok.