

JÄMFÖRELSE AV WEBBTIDNINGSTEXTERS KVANTIFIERBARA LÄSBARHET I  
DAGENS NYHETER OCH HUFVUDSTADSBLADET

Ilpo Partanen

Pro gradu-avhandling i svenska språket  
Institutionen för språk  
Jyväskylä universitet  
Hösten 2006

## ABSTRAKT

### JYVÄSKYLÄ UNIVERSITET

Humanistiska fakulteten	Svenska språket, Institutionen för språk
Författare: Ilpo Partanen	
Titel: Jämförelse av webbtidningstexters kvantifierbara läsbarhet i Dagens Nyheter och Hufvudstadsbladet	
Ämne: svenska språket	Typ av avhandling: pro gradu
Årtal: hösten 2006	Sidoantal: 63 + 7
<p>I denna undersökning redogjordes för den kvantifierbara läsbarheten i webbtidningarna Dagens Nyheter och Hufvudstadsbladet. I metodikdelen avhandlades problematiken bakom läsbarhetsundersökningar (kvalitativ och kvantitativ) samt definierades och redogjordes för den centralaste teorin – C.H. Björnssons <i>läsbarhetsindex</i>, <i>LIX</i> – och dess tillämpning på läsbarhetsmätning i denna undersökning. Därtill behandlades webben (WWW) och dess grundläggande struktur som bas för webbtidningar kort samt den statistisk-matematiska teoretiska bakgrunden och dess tillämpning på signifikansprövning i denna avhandling utreddes.</p> <p>Som undersökningmaterial användes sammanlagt 168 webbtidningstexter (från år 2005) ur tidningarna Dagens Nyheter och Hufvudstadsbladet, varav 84 (42 x 2) var förstasidor och resten bestod av 84 HBL-insidesartiklar. Materialet, dvs. texterna, konverterades till Word-form, LIX beräknades och tabellerades i Excel och slutligen prövades resultaten med lämpliga datorprogram. Därefter presenterades de slutliga resultaten med respektive statistiska uppgifter och figurer. Dessutom värderades LIX-resultat enligt Björnssons LIX-tolk. Största delen av det analyserade förstasidesmaterialet kunde kategoriseras som ”lätt” enligt tolken. När det gäller HBL-insidesartiklarna kunde svårighetsgraden för avdelningarna <i>Ekonomi</i> och <i>Ledare</i> även klassificeras som ”medelmåttig”. I största allmänhet förblev den vanligaste värderingen emellertid ”lätt” också för HBL-insidesmaterialet på grund av variationen, dvs. låga LIX-värden vs höga, inom olika avdelningar.</p> <p>På grund av signifikansprövningsresultaten kan skillnaderna för den kvantifierbara läsbarheten för DN och HBL inte sägas vara följdenligt statistiskt signifikanta. Fast det finns signifikanta skillnader inom och mellan jämförelsegrupperna (t.ex. veckor, veckodagar, insidesavdelningar) går resultaten inte att generaliseras. Resultaten ger dock flera intressanta uppgifter om skillnaderna i läsbarheten i de två svenskspråkiga morgontidningarna.</p>	
Uppslagsord: läsbarhet, läsbarhetsindex, tidningstext, webb, signifikansprövning	
Bibliotek/Förvaringsplats: Aalto bibliotek	
Övriga uppgifter:	

# INNEHÅLL

1 INLEDNING.....	5
2 TIDNINGSTEXT – NÅGRA DEFINITIONER OCH DISKUSSION.....	6
2.1 Definition av tidningstext och begreppet nyhetsartikel .....	6
2.2 Vanlig tidning vs webbsida som tidning.....	7
3 OM TVÅ OLIKA TEXTANALYSMETODER OCH LÄSBARHET .....	8
3.1 Problem med tillämpning av forskningsmetoder.....	8
3.2 Läsbarhet.....	9
3.2.1 Olika definitioner av begreppet läsbarhet .....	9
3.2.2 Olika läsbarhetskriterier .....	10
3.2.3 Definition av LIX .....	11
3.2.4 Kritik – det positiva och negativa med LIX.....	11
3.2.5 Björnssons LIX-tolk .....	12
4 OM WEBBTEXTER.....	13
4.1 Om webbtexTERS struktur.....	13
4.2 Hypertext och (hyper-)länkar.....	13
4.3 Exempel på den dynamiska strukturen .....	14
5 OM SPRÅKSTATISTIK OCH DESS TILLÄMPNING.....	15
5.1 Språkstatistik och några grundläggande begrepp .....	15
5.2 Om två olika språkliga nivåer .....	16
5.3 Statistik-matematiska begrepp.....	16
5.4 Om t-test som testtyp .....	17
5.4.1 Hypotesprövning, signifikansnivåer samt konfidensintervall i t-test.....	17
5.4.2 Syftet med att tillämpa t-test .....	18
5.4.3 Om koden samt t-testkörningar för programmet R .....	19
5.4.4 Redogörelse för boxplotdiagram .....	20
5.5 Om utomstående expertis med analysdel.....	21
6 MATERIAL OCH METOD.....	21
6.1 Utgångspunkt för arbetet .....	21
6.2 Kort om materialets tillgänglighet .....	22
6.3 Hypoteser och forskningsproblem .....	23
6.4 Om forskningsmaterial och -metoder .....	24
7 RESULTAT OCH DERAS TOLKNING .....	28
7.1 Allmänt .....	28
7.2 Om LIX- och statistikvärden .....	29
7.3 Om boxplotfigurerna samt placering av diagramvärdena.....	30
7.4 Spridning av LIX-värdena .....	31
7.4.1 Alla LIX-värden för förstasidorna i DN och HBL.....	31
7.4.2 LIX-värden för förstasidorna i DN för olika veckor .....	33
7.4.3 LIX-värden för förstasidesgrupperna DN vecka 2 och DN vecka 5 .....	35
7.4.4 LIX-värden för förstasidorna i DN för olika veckodagar.....	37

7.4.5 LIX-värden för förstasidesgrupperna DN-måndag och DN-tisdag.....	39
7.4.6 LIX-värden för förstasidesgrupperna DN-lördag och DN-söndag.....	40
7.4.7 LIX-värden för förstasidorna i HBL för olika veckor.....	42
7.4.8 LIX-värden för förstasidorna i HBL för olika veckodagar .....	44
7.4.9 LIX-värden för HBL-insidorna .....	46
7.5 Statistisk signifikans – signifikansresultat för alla jämförelsegrupper .....	48
7.5.1 Alla förstasidor i DN och HBL .....	49
7.5.2 Förstasidor: DN – veckor .....	49
7.5.3 Förstasidor: DN – veckodagar.....	50
7.5.4 Förstasidor: HBL – veckor .....	50
7.5.5 Förstasidor: HBL – veckodagar .....	50
7.5.6 Insidor: HBL.....	51
7.5.7 Förstasidor: DN vs HBL – veckor.....	51
7.5.8 Förstasidor: DN vs HBL – veckodagar .....	52
7.6 Allmänna LIX-resultat.....	52
7.6.1 Förstasidesresultat .....	53
7.6.2 HBL-insidesresultat.....	53
7.6.3 Jämförelse med andra undersökningsresultat.....	54
7.6.4 Slutsatser .....	56
8 SAMMANFATTNING.....	58
LITTERATUR .....	61
BILAGOR .....	64

## 1 INLEDNING

Vanlig tidningstext liknar boktext. Texten i papperstidningar består av lineära textdelar som är skrivna från vänster till höger. Traditionella tidningar är tryckta på tidningspapper som innehåller texten i kompakt och lätthanterlig form.

Tidningstexter som ska läsas på en bildskärm är litet annorlunda. Att läsa denna slags texter kan kräva kunskaper om datorer, Internet och texternas arrangemang på World Wide Web. Dessutom möter de som läser elektroniska texter, inklusive webbtidningar, nya komponenter som saknas i de vanliga pappersbladen: hypertext och dynamiska textstrukturer. Webbtexternas fundamentala egenskaper med hänsyn till webbsidan som tidning kommer att granskas närmare i fjärde kapitlet.

Webbtidningarna har blivit ett alternativt och beaktansvärt medium vid sidan av de traditionella papperstidningarna. Dessutom verkar tendensen vara att elektronisk läsning, inklusive webbtidningar, blir mer och mer populär. Till största delen kan pappersstidningarnas webbversioner betraktas som elektroniska versioner som i stort sett har samma texter och informationsinnehåll – bara i elektronisk form. Det finns emellertid flera intressanta faktorer som påverkar texternas språkliga, strukturella och typografiska tillgänglighet. Alla dessa faktorer kan dock inte behandlas eller analyseras i denna avhandling utan jag först och främst kommer att fokusera på att mäta den kvantifierbara läsbarheten hos webbtidningstext samt utreda eventuella statistiska skillnader mellan de analyserade texterna.

Denna undersökning syftar till att först definiera tidningstext som medium, redogöra för problematiken bakom läsbarhetsundersökningar (kvalitativ och kvantitativ) samt definiera och redogöra för den centralaste teorin (C. H. Björnssons LIX) och dess tillämpning på läsbarhetsmätning – i synnerhet i denna undersökning. Därtill ska webben (WWW) och dess grundläggande struktur som bas för webbtidningar behandlas samt den statistisk-matematiska teoretiska bakgrunden och dess tillämpning på signifikansprövning i denna avhandling utredas.

I huvudanalysdelen av denna avhandling, görs en jämförande datorstödd läsbarhetsanalys mellan några utvalda texter i två morgontidningar, *Dagens Nyheter* och *Hufvudsstadsbladet*.

Som material ska jag använda ett antal utvalda förstasidor (DN & HBL) och insidor (endast HBL) som webbupplagor ur dessa tidningar. Att definiera läsbarhet, avhandla metodik bakom dess tillämpade statistiska granskning samt att analysera mitt utvalda webbtexmaterial är de viktigaste utgångspunkterna i min avhandling. Huvudintresset ligger vid den kvantifierbara läsbarheten samt de statistiskt jämförda skillnaderna (beräknade med statistikdatorprogram) mellan förstasidorna av de ovannämnda webbtidningarna. Därtill ska även insidor (sju olika textavdelningar) i HBL analyseras för att kartlägga eventuella skillnader mellan dessa.

Hela webbtexmaterialet ska först förvandlas till en kvantifierbar form, dvs. mätas (Grammatifix), tabelleras (Excel) och sedan prövas statistiskt (statistikprogrammet R) med hjälp av lämpliga datorprogram. Därefter ska resultaten presenteras med lämpliga statistiska uppgifter och tolkas enligt Björnssons värderingsmetod (LIX-tolk) med hjälp av eventuella figurer samt hänvisningar till det bifogade materialet.

## 2 TIDNINGSTEXT – NÅGRA DEFINITIONER OCH DISKUSSION

### 2.1 Definition av tidningstext och begreppet nyhetsartikel

Enligt Westman (1974, 11) kan språket i tidningstexten anses vara s.k. *bruksprosa* som betyder att språket används i en praktisk kommunikativ funktion. När tidningar betraktas som facktext kan tidningstexten kategoriseras i gruppen som vänder sig till publiken på samhällsnivån. Dessa texter sprider sig genom kanaler som kan nå en bred mottagargrupp. (Melander 1987, 106.) Melin och Lange (2000, 140) definierar begreppet *nyhetsartikel* som ett medel vars funktion är att ”berätta om ett skeende och ge en här- och nu-upplevelse”. De påpekar också att nyhetsartikeln skiljer sig från övriga texttyper genom att dra sig ifrån ett klart ställningstagande. De texter som jag kommer att analysera är nyhetsartiklar skrivna i två *morgontidningar*, den svenska *Dagens Nyheter* och det finlandssvenska *Hufvudstadsbladet*. Eftersom tidningar innehåller olika slags artiklar kan det vara besvärligt att ge en allmängiltig beskrivning av morgontidningstext. Strand (1984, 129) refererar en beskrivning av morgontidningar som Börje Dahlqvist, den dåvarande redaktionschefen på DN ger. Dahlqvist jämför morgontidning med kvällstidning som medium.

Vi skall bara påpeka att vi som yrkesmän uppfattar en tämligen bestämd skillnad mellan morgon- och kvällspress i det här landet, en skillnad som går att bestämma i redigeringstekniska termer. Morgontidningar kan i allt väsentlig sägas vara redovisningstidningar. De har ambitionen att vara en krönika i stort och smått över ett dygns händelser. Det ger dem deras karaktär. Kvällstidningarna däremot har inte givit sig ut för att ha den. De tar upp och beskriver punkthändelser som de anser vara god läsning. De har inte tagit på sig uppgiften att presentera en fortlöpande och balanserad nyhetsinformation. Det behöver inte göra dem till mindre bra tidningar, men annorlunda. (Dahlqvist 1968, 96 refererad i Strand 1984, 129.)

Att definiera tidningstext rent allmänt är inte enkelt. Trots det är det nödvändigt att först definiera mediet som utgör utgångspunkten till de kommande läsbarhetsanalyserna.

## 2.2 Vanlig tidning vs webbsida som tidning

World Wide Web är ett nätverk som består av de datorer som är kopplade till Internet och som stöder WWW. Dessa datorer är tillgängliga till alla som läser och rör sig på WWW, m.a.o. navigerar, på och genom olika webbsidor. En webbsida kan t.ex. vara en webbtidningssida som konstruerats enligt den motsvarande papperstidningens design – bara i elektronisk form. Att en webbtidning ska vara lika läsbar och lättillgänglig samt lätt att använda som den alternativa pappersversionen är ingalunda en lätt uppgift för den digitala skribenten.

En tryckt tidning och en webbtidning avviker drastiskt från varandra när det gäller design. Fast webbsidor ofta anses vara närbesläktade med tryckta tidningssidor och vanliga tidningar finns det en hel del olikheter. En tidningstext och en text på webben är två olika medier som varken kan jämföras eller behandlas på samma sätt. (Englund & Guldbland 1999, 19–20.)

Den första stora skillnaden utgör själva fysiska formen: papperstidning liknar boken – den har konkreta sidor. Dessutom finns det en grundläggande skillnad mellan en vanlig tidningssida och en webbtidningssida. Den vanliga tidningen är en fast och färdig produkt som har en fastställd storlek och layout. Webbsidornas layout är däremot inte fast, utan den beror på användarens skärm och dess egenskaper och anpassar sig automatiskt enligt dem. (Englund & Guldbland 1999, 20–21.)

Webbsidans dynamiska struktur kan göra webbtidningen till en oändlig helhet vars avsedda tillgänglighet inte går att kontrolleras eller planeras på förhand. På grund av detta kan textens längd och omfattning fluktuera mellan några rader och flera tusen rader. Producentens arbetsskicklighet spelar en nyckelroll när det gäller att skapa en välfungerande, dvs. en lättillgänglig och lagom lång webbtidning. (Englund & Guldbrand 20–21.) Dessutom påpekar Englund och Guldbrand (1999, 21) att användarna uppskattar webbsidorna på grund av känsla medan den tryckta pappertidningen analyseras efter dess visuella egenskaper.

### 3 OM TVÅ OLIKA TEXTANALYSMETODER OCH LÄSBARHET

#### 3.1 Problem med tillämpning av forskningsmetoder

Från och med att jag satt igång min undersökningsprocess har jag övervägt olika alternativa sätt att förverkliga en analys av några webbmorgontidningstexter. Efter ett omfattande tankearbete och granskning av litteratur har jag bestämt att jag endast ska använda kvantitativa metoder fast dessa metoder har blivit kritiserade i någon mån. Näslund (1987, 129) skriver att det under 80-talet skedde en metodisk förändring från mer stilistiskt riktade analyser till mer grammatiska (kvantitativa och objektiva). Det finns dock några hindrande faktorer som försvårar kvantitativ tidningstextanalys (se avsnitt 3.2.4).

Det verkar emellertid som om det finns massor av stilanalytiska och kvalitativa analyser av olika slags brukstexter men mycket få kvantitativa undersökningar om t.ex. tidningstexter. Enligt Näslund (1987, 128) har det funnits intresse för språket i massmedierna, inklusive tidningar, under de senaste årtiondena, men språkvetenskapliga undersökningar på detta område är relativt sällsynta.

Jag har valt att utföra en anpassad, läsbarhetsmässigt och även statistiskt pålitlig säranalys av några utvalda texter i webbtidningarna *Dagens Nyheter* och *Hufvudstadsbladet*. Vad tidningstexternas språkliga form beträffar är texterna i webbtidningarna på nätet oftast desamma som i deras pappersversioner. Därför finns det inte någon anledning till att analysera innehållet i respektive pappersupplaga utan utföra analyserna på ett modernt sätt.



Jag vill reda upp den nutida morgontidningstextens (DN & HBL) språkliga läsbarhet genom några kvantitativa mätningar. Strand (1984, 25) påpekar att den kvantitativa LIX-metoden kan användas för att utreda textens grammatiska komplexitet samt som ett mått för själva läsbarheten. Denna LIX-metod som jag kommer att tillämpa i min undersökning ska avhandlas i det följande avsnittet (3.2). Det är logiskt att analysera språkliga egenskaper såsom ord- och meningslängd kvantitativt. Den grafiska och typografiska tillgängligheten går däremot inte att kvantifieras. Då är det lämpligt att använda den kvalitativt deskriptiva forskningsmetoden.

## 3.2 Läsbarhet

### 3.2.1 Olika definitioner av begreppet läsbarhet

Hur ska man definiera begreppet *läsbarhet*? Det finns ju tiotals faktorer som tillsammans gör en text svår eller lättläst. Språket är den främsta förutsättningen till att en viss informationssinnehåll kan förmedlas. Björnsson (1968, 13) beskriver den språkliga läsbarheten på följande sätt:

Läsbarheten är summan av sådana språkliga egenskaper hos en text, vilka gör den mer eller mindre svårtillgänglig för läsaren.

Denna definition omfattar först och främst den språkliga läsbarheten. Historiskt sett har den egentliga språkliga läsbarheten börjats undersöka först efter att man forskat i textens yttre egenskaper, t.ex. typografin. (Björnsson 1968, 14.) Förutom det språkliga påverkar också kvalitativa egenskaper såsom stil, ämne och innehåll textens tillgänglighet.

För att hålla isär den språkliga och den icke-språkliga tillgängligheten använder Björnsson (1968, 14) en vidare definition, *läslighet*, som betyder textens tillgänglighetsgrad som beror på de typografiska egenskaperna. Det handlar om t.ex. teckensnitt och layout. Denna gren är mycket omfattande och besvärlig att avgränsa. (Björnsson 1968, 13–15). Björnsson fokuserar dock sin definition på *läsligheten*, dvs. den typografiska läsbarheten. Han poängterar att faktorer som t.ex. sidantal, sidstorlek, ordantal per sida, osv. i hans läslighetsundersökning ska lämnas utanför granskningen. Han anser att det grafiska hos texter som inte förekommer i

böcker, t.ex. tidningstexter, blir besvärligt att pröva om detta inte görs. (Björnsson 1968, 178.) På grund av detta utesluter Björnsson läslighetsfaktorerna förutom *bokstavsstorlek, radavstånd och radlängd* utanför läslighetsanalysen.

*Läsbarheten* och *läsligheten* kan man också definiera som textens språkliga- och icke-språkliga tillgänglighet. För att kunna ge ett helhetsintryck angående textens tillgänglighet är det förnuftigt att inkludera både de språkliga och de yttre komponenterna i texten och analysera den som en språklig och typografisk helhet.

För att kunna fokusera på den språkliga läsbarheten i min avhandling ska jag inte utsträcka analyserna till att omfatta textens innehållsliga, stilistiska egenskaper förutom redogörelsen för webbsidornas grundläggande principer. Stilen som faktor hos den kommande läsbarhetsanalysdelen avgränsar jag utanför mitt undersökningsområde.

### 3.2.2 Olika läsbarhetskriterier

Björnsson (1968, 20) påpekar att när man först börjar mäta språklig läsbarhet, måste man skapa något slags *kriterium* för att få basis för undersökningen. Exempelvis kan ett bestämt urval texter som *svårighetsgraderas* utgöra ett kriterium som man kan använda som referens när textens språkliga läsbarhet ska mätas. Kriteriet, m.a.o. svårighetsgraden, kan exempelvis skaffas genom att ett antal texter bedöms av experter eller andra människor.

Ett annat kriterium för läsbarhet kan också vara *lästiden*. Att skapa kriteriet på grund av lästiden krävs det också försökspersoner som provläser texterna. (Björnsson 1968, 20–22.)

Förutom svårighetsgradering och lästiden kan t.o.m. *läsförståelsen* utgöra kriteriet. Det kan dock vara svårt att pröva i vilken grad försökspersonerna har förstått texterna. (Björnsson 1968, 21.) Jag anser att både läsförståelsen som kriteriemått och läsvärde som ett sätt att avgöra textens läsbarhet medför några subjektiva faktorer som kan försvaga reliabiliteten i en läsbarhetsundersökning. Därför ska jag inte använda läsförståelse (detsamma gäller läsvärde som läsbarhetsdefinition) som kriterium i min undersökning.

### 3.2.3 Definition av LIX

Björnssons läsbarhetsundersökning går först ut på svårighetsgradering av böcker som används som skolböcker inklusive läse-, läro-, fiktions- och sakböcker (Björnsson 1968, 32). I sin undersökning använder han skolböckerna som utgångspunkt för att definiera de *läsbarhetsfaktorer* som ska användas i skolboktexternas svårighetsgradering (Björnsson 1968, 23).

Till dessa faktorer hör två centrala läsbarhetsfaktorer: 1. *långa ord* och 2. *meningslängd* som i Björnssons detaljundersökningar och olika provmätningar visar sig vara de mest avgörande faktorerna (Björnsson 1968, 35, 38, 44, 202.) Björnsson (1968, 47) påpekar att dessa långa ord är ord som består av mer än sex bokstäver. Enligt honom står meningslängden för den genomsnittliga meningslängden mätt i antalet ord. (Björnsson 1968, 46.)

På grund av dessa två introducerar Björnsson (1968, 44–46) *läsbarhetsformeln* som utgörs av summan av meningslängden och procenttalet långa ord ( $lix = \text{meningslängden} + \text{långa ord}$ ). Som resultat ger formeln *läsbarhetsindex*, förkortat: LIX, som kan användas för att definiera textens svårighetsgrad.

Förr i tiden (innan datorer blivit allmänna) har man beräknat LIX för hand. Detta kräver naturligtvis betydligt mer arbete. I min pro gradu-avhandling kommer jag dock att utnyttja ett datorprogram, Grammatifix, med vilket LIX kan beräknas snabbt och relativt pålitligt. Detta slags program gör textens LIX-beräkningar betydligt enklare och snabbare. Tillämpning av programmet förklaras närmare i kapitel sex.

### 3.2.4 Kritik – det positiva och negativa med LIX

Jag anser att läsbarhetsindexet LIX är en lämplig mätningsskåp för läsbarhetsmätningar i min undersökning. Melin och Lange (2000, 51) skriver: ”Den egentligen enda textegenskap som upprepade gånger prövats vetenskapligt är läsbarhet.” Björnssons LIX är den främsta kvantitativa metoden som kan användas för läsbarhetsmätning.

Platzack (1973, 11) anser Björnssons läsbarhetsformel vara ett effektivt verktyg när man ska gradera en text på ett relativt pålitligt sätt på skalan från lätt till svår på en kort tid. Larsson (1976, 54) konstaterar att den genomsnittliga meningslängden kan fungera som ett mått på textens svårighetsgradering när det gäller tidningstext. Han betonar dock att långa meningar inte automatiskt gör en text svårtläst.

Sigurd (1991, 158) påpekar att läsbarhetsindex inte kan användas för att bestämma textens innehåll, inklusive disposition och presentationssätt, utan det kan användas endast på det yttre hos texter.

Näslund påminner om att LIX har kritiserats t.ex. av Gunnarsson (1982) för dess ytliga mätmetod som inte tar hänsyn till t.ex. semantiska eller kommunikationens situationella faktorer – exempelvis läsarens förkunskaper i ämnet. LIX mäter endast textens syntaktiska nivå. (Näslund 1987, 137–138.)

### 3.2.5 Björnssons LIX-tolk

I min undersökning ska jag använda Björnssons (1968, 89) LIX-tolk, dvs. värdegränser som kan användas för att värdera textens svårighetsgrad med en numeral och verbal värderingsskala. Siffrorna samt beskrivningarna anger en LIX-värdering med tio enheters mellanrum från mycket lätt till mycket svår:

20 = Mycket lätt	50 = Svår
30 = Lätt	60 = Mycket svår
40 = Medelmåttig	

Denna värdering av LIX kommer att användas då resultaten i resultatkapitlet ska tolkas. Där kommer LIX-tolken att integreras som en del av tabellerna.

## 4 OM WEBBTEXTER

### 4.1 Om webbtexeters struktur

Det är nästan omöjligt att sammanfatta Internet och dess grundläggande komponenter och principer. Internet består av oräkneliga komponenter som knyts ihop till ett komplext nätverk. Jag ska inte behandla Internet (webb) och dess grundläggande faktorer i minsta detalj utan redogöra för den aktuella grunden på vilken en webbtidning kan operera. En webbsida kan fungera som basis för ett flertal olika elektroniska texter. Att kunna förstå elektroniska texter och informationsstrukturernas konstruktion måste man bekanta sig med Internet och dess grundläggande principer.

Enligt Heimbürger et al. (1990, 19–20) består ett vanligt elektroniskt textdokument av samma delkomponenter som ett vanligt pappersdokument. Både pappers- och webbtexter har komponenter som kapitel, stycken, satser och ord. När det gäller tidningstext är också själva textinnehållet ofta detsamma i pappers- och webbversionen (Englund & Guldbrand 2002, 7). Englund och Guldbrand (1999, 49) betonar dock att den linjära strukturen som boken har lämpar sig dåligt för elektroniska texter.

Heimbürger et al. (1990, 19) skriver att den s.k. normala texten – det gäller också tidningstexten – kan kallas för en *sekventiell* eller *lineär* text. Denna slags text består av sekventiella textkomponenter som utgör en helhet. En sådan lineär text är dubbelriktad, dvs. läsningen kan ske i två riktningar, fram- och baklänges. Webbtexterna och webbtidningarna innehåller emellertid en tredje dimension som nås genom *hypertext* och *länkar*.

### 4.2 Hypertext och (hyper-)länkar

*Hypertext* innehåller *hyperlänkar*. Heimbürger et al. (1990, 19–20) betonar att det handlar om *hypertext* då textens innehåll har frigjorts från textens framställningsstruktur. Då kan man använda prefixet *hyper-* framför orden *text* och *länk*. Den som läser en elektronisk text kan exempelvis först läsa en del av en större texthelhet och sedan fördjupa sina kunskaper om ämnet ifråga genom att navigera via en *hyperlänk* till en ny och mer omfattande textdel.

Hypertext är en icke-lineär helhet som består av individuella textdelar. Textdelarna nås via hyperlänkar som möjliggör tillgång till övriga relevanta textdelar. Genom att navigera och läsa enskilda textdelar som ger information om ämnet kan läsaren konstruera olika framställningar av innehållet. Textdelarna som är anslutna till varandra på grund av saksammanhanget kan kallas för *informationsceller* (jfr engl. *nodes*). I sin tur kallas förbindelserna mellan informationscellerna för *länkar*. (Heimbürger et al. 1990, 19–20.)

#### 4.3 Exempel på den dynamiska strukturen

Länkar fungerar som vägledare för att läsaren ska kunna hitta vidare information eller t.o.m. fördjupa sina kunskaper i ett bestämt ämne. Själva webbläsningen innebär att man navigerar genom länkar som utgör alternativa färdiga stigar för den som söker information på webben. (Kacmar et al. 1988/McAleese 1989 refererad i Heimbürger et al. 1990, 76.)

I följande exempel illustreras skillnaden mellan en vanlig *länk* som inte befinner sig i själva textdelen, dvs. som en del inne i en sats (exempel 1) och en *hyperlänk* som är integrerad i texten (exempel 2).

Exempel 1 Detta är ett exempel på en vanlig länk.

Länken befinner sig utanför själva textdelen.

[klicka här!](#)

Exempel 2 Detta är ett exempel på en hyperlänk som

är integrerad i själva texten.

Båda typer av länkar finns i webbtidningar (DN & HBL) fast vanliga *länkar* som i exempel 1 verkar vara allmännare än *hyperlänkar*. Genom att man använder vissa typer av länkar man kontrollera läsningsprocessen. Englund och Guldbbrand (1999, 21, 49) påpekar att genom hypertextnavigeringen genom hyperlänkar möjliggör s.k. *cirkulärt läsande*. Detta innebär att läsaren inte behöver läsa texten från början till slut utan läsningen kan ske cirkulärt enligt läsarens intressen.

Enligt Bolter (1991, 27) kan vanliga länkar (som i exempel 1) betraktas som moderna versioner av fotnoter. När det handlar om processen där läsaren hämtar information elektroniskt genom länken på skärmen sker datainhämtning automatiskt i stället för att läsaren ska vända blad manuellt. I en vanlig tryckt text navigeras genom att läsaren vänder blad. (Englund & Guldbrand 1999, 21).

Textstrukturen på webben är alltså inte enbart tvådimensionell såsom i papperstidningar. Elektronisk text kan innehålla s.k. *hypertext* som medför att läsaren kan navigera från en textdel (eller webbsida) till en annan. Det är svårt att hålla isär s.k. lineär text och icke-lineär text. Dessutom finns det tiotals olika definitioner för hypertext.

## 5 OM SPRÅKSTATISTIK OCH DESS TILLÄMPNING

### 5.1 Språkstatistik och några grundläggande begrepp

*Statistik* som ett vetenskapligt begrepp syftar på något som är mer eller mindre sannolikt. När statistiken används för att upptäcka det väsentliga och intressanta i forskningsmaterialet och för att presentera resultaten i siffror kan man tala om *deskriptiv statistik*. Det insamlade empiriska materialet som ska användas som utgångspunkt för undersökningen kallas för *rådata*. (Befring 1994, 98.) Då den deskriptiva statistiken tillämpas på språkliga företeelser kan man tala om *deskriptiv språkstatistik*.

Metoderna är hjälpmedel med vilka man samlar in det viktiga ur det empiriska materialet (Befring 1994, 98). Enligt Holme och Solvang (1991, 12) kan begreppet *metod* definieras som följer:

En metod är alltså ett redskap, ett sätt att lösa problem och komma fram till ny kunskap. Allt som kan bidra till att uppnå dessa mål är en metod.

Om att utföra en statistisk undersökning skriver Stukát (1993, 7) följande: ”Den är bara en uppsättning metoder att *samla in, beskriva* och *analysera data*.”.

Begreppet *variabel* betyder ett fenomen eller en egenskap som kan ha mer än ett värde. En adekvat variabel ska definiera olikheter inom det studerade urvalet på ett meningsfullt sätt. (Befring 1994, 45–46.) Variabler som exempelvis ordlängd och ordfrekvens kan kvantifieras och mätas. Detta sker genom att man mäter de fenomen som man vill forska med hjälp av språkstatistiska metoder. Enligt Befring (1994, 56) definieras *mätning* ”som en process där fenomen, egenskaper eller händelser tillskrivs siffervärden enligt bestämda regler”.

## 5.2 Om två olika språkliga nivåer

För att skapa en teoretisk grund för en undersökning ska alla viktiga företeelser och begrepp avhandlas. I det följande behandlas kort två av de viktigaste kvantifierbara (med dator mätbara) variablerna (språkliga nivåerna) i min undersökning.

Sigurd (1991, 154) poängterar att meningslängden är det vanligaste måttet när det gäller *syntaktisk statistik*. Melin och Lange (2000, 16) skriver att till textanalysens *syntaktiska nivå* hör alla faktorer som kan förknippas med sats- och meningsbyggnad, meningslängd, meningarnas struktur och komplexitet. Till den *lexikala nivån* hör sådana faktorer som berör ordförråd (Melin & Lange 2000, 14). Läsbarhetsindex, LIX, består av två variabler; andelen långa ord och den genomsnittliga meningslängden (Björnsson 1968, 46, Strand 1984, 16). Därför är Björnssons LIX en analysmetod som textanalytiskt sett fungerar både på den syntaktiska och den lexikala nivån. I denna pro gradu-avhandling utgör Björnssons *läsbarhetsindex* en väsentlig del av de kvantitativa metoderna med vilka webbtidningstexternas språkliga läsbarhet analyseras.

## 5.3 Statistik-matematiska begrepp

De centralaste begreppen som jag kommer att använda i resultatdelen (även i tabellerna) är det aritmetiska *medelvärdet* ( $M$ ), *medianen* ( $M_d$ ) och *standardavvikelsen* ( $sd$ ). Det aritmetiska medelvärdet kan definieras som summan av ett antal värden dividerad med antalet värden. Medianen kan sägas vara det värde på variabeln som delar materialet på mitten när individerna ordnas efter poäng. Ifall det är fråga om ett material med udda antal blir medianen det mittersta talet, dvs. det värde som den mittersta individen har. Handlar det däremot om ett



jämnt antal ligger medianen mitt emellan de två mittersta individerna. Standardavvikelsen innebär att varje poängvärdes avvikelse från det aritmetiska medelvärdet tas hänsyn till. Standardavvikelsen anses vara mycket mer rättvisande och exakt som spridningsmått än variationsvidden. Variationsvidden anger bara skillnaden mellan det högsta och lägsta värdet i ett material. (Stukát 1993, 32–33, 39–40, Internet 2, Internet 3.)

#### 5.4 Om t-test som testtyp

Ursprungligen skapades *t-testet* av W.S. Gosset som skrev under namnet ”Student”. Därför kallas testet ibland för ”Student’s t-test”. Man använder t-testet för att utreda om skillnader i medelvärde kan förklaras antingen av en slump eller en alternativhypotes. Testet kan användas när det är fråga om enskilda urval, korrelerade grupper eller oberoende grupper. I min undersökning används t-testet för de sistnämnda. (Internet 1, Internet 8.) Enligt Stukát (1993, 73) förekommer t-testning inom språkvetenskaplig forskning t.ex. när man vill utröna om det finns någon signifikant skillnad mellan två grupper aritmetiska medelvärden. Därtill poängterar Butler (1985, 84) att t-testet ska användas då samplet är litet.

##### 5.4.1 Hypotesprövning, signifikansnivåer samt konfidensintervall i t-test

Statistiska signifikanstest syftar till att visa om de observerade skillnaderna mellan resultaten antingen är slump- eller icke-slumpmässiga. Startsituationen i en den statistiska signifikansprövningen innebär att en *nollhypotes* och *mothypotes* samt slumpnivån dvs. signifikansnivån (exempelvis 5 %) bestäms. Därtill utförs själva beräkningarna t.ex. med dator varefter resultaten utvärderas enligt beslutsregeln och de uträknade resultaten. Om värdet är mindre eller lika med det uträknade värdet så förkastas nollhypotesen och mothypotesen accepteras. Om värdet däremot blir större så accepteras nollhypotesen och man kan konstatera enligt den ifrågavarande signifikansnivån att det inte finns någon signifikant skillnad mellan grupperna. (Butler 1985, 69, Internet 1.)

Som redan konstaterats i det föregående avsnittet klassificeras t-test som medelvärde-test. Enligt Stukát (1993, 74) används alltså en nollhypotes och alternativhypotes (mothypotes) i testet:

$H_0$  (nollhypotes) = ”Ingen skillnad mellan grupperna. Skillnaden beror på slumpen.”

$H_1$  (alternativhypotes) = ”Det finns en signifikant skillnad mellan de två grupperna.”

Med testets signifikansnivå avses den största risken för misstag som forskaren är villig att acceptera då nollhypotesen förkastas. Denna risk mäts med p-värde (engl. probability). I språkvetenskap används vanligen 5 % (0,05) signifikansnivå. För att skapa en ännu precisare uppfattning om skillnaderna mellan de analyserade grupperna kan 1 % (0,01) och 0,1 % (0,001) signifikansnivåerna användas. Dessa diskuteras även i kapitel sex. (Butler 1985, 71.) Konfidensintervallen används för att få veta hur mycket medelvärdet avviker från det sanna medelvärdet. Om ett medelvärde ligger innanför eller utanför 95 % konfidensintervall ger också signifikansprövning med 5 % signifikansnivå samma resultat. Om man istället väljer t.ex. 99 % konfidensintervall gäller samma regel för 1 % signifikansnivå. (Internet 7.)

#### 5.4.2 Syftet med att tillämpa t-test

Med hjälp av ett antal parvis körda t-test har jag jämfört 84 förstasidor. Dessutom har 84 artiklar från sju olika huvudavdelningar på insidorna i tidningen *Hufvudstadsbladet* analyserats. På så sätt har bildats partester som ger information om LIX-värdena mellan två olika medelvärdegrupper (t.ex. DN-lördag–DN-söndag). Detta innebär att varje förstasida respektive insidesartikel systematiskt jämförts med varandra för att kunna skapa en pålitlig bild av värdena som täcker hela den kalkylerade läsbarheten som gäller innehållet i *Dagens Nyheter* (endast förstasidor) och *Hufvudstadsbladet* (förstasidor samt insidor). På grund av t-testen har ett antal Excel-tabeller konstruerats i vilka jämförelseparen sorterats enligt allmänt kända statistiska signifikansnivåer (se även kapitel sex, avsnitt 6.4). Eftersom materialet innehåller t.ex. Sex stycken förstasidor från olika veckor (exempelvis 1 söndag/vecka) har vissa veckodagars förstasidor behandlats som värdegrupper bestående av t.ex. Alla DN-söndagsupplagor, dvs. Som genomsnitten av dessa. Samma parvisa t-test-metod och

gruppering har tillämpats även på insidorna i hbl (t.ex. Sport vs utrikes). På basis av alla värden inom varje enskild värdegrupp i förstasides- och insidesmaterialet har spridningen för dessa beräknats. Dessa värdegruppers LIX-spridning illustreras då i resultatkapitlet med respektive boxplotdiagram. (se t.ex. Resultatkapitlet, figur 2.)

#### 5.4.3 Om koden samt t-testkörningar för programmet R

Själva t-testen i min undersökning har förverkligats genom att kalkylera ett "Welch Two Sample t-test" som är en typ av det tidigare nämnda "Student's t-test" (Internet 8), exempelvis för vektorerna "lixDN" (=alla DN-förstasidor som medelvärdegrupp) och "lixHBL" (=alla HBL-förstasidor som medelvärdegrupp). Båda grupperna innehåller 42 successiva observationer (kompleta förstasidor) av LIX-värdena på de ifrågavarande tidningarnas förstasidor. Här redogörs för följande termer som syftar till programmet R och dess funktioner:

förstasidor	funktionen, i vilken de kalkylerade värdena placeras
<-	placeringsoperator
t.test	det kalkylerade testet (detta är också en funktion i programmet R)
lixDN	LIX-värden på DN:s 42 förstasidor
lixHBL	LIX-värden på HBL:s 42 förstasidor, varav test (=funktionernas argument) kalkyleras

För att utföra ett t-test (Student's, typ Welch Two Sample) och sedan placera resultaten i funktionen 'förstasidor' ska en följande instruktion matas in i programmet R:

```
förstasidor<-t.test(lixDN,lixHBL),
```

Alla t-test i denna undersökning har utförts på samma sätt, dvs. genom att variera namnet samt det kalkylerade föremålet och argument på funktionen 'förstasidor' så att det i vart och ett fall motsvarar det ifrågavarande forskningsobjektet. När det handlar om att jämföra insidor i HBL med varandra används till exempel ett följande argument:

ekonomiutrikes<-t.test(HBLEkonomi,HBLutrikes),

de kalkylerade testen och deras värden matas ut och man får följande data (exempelvis: medelvärdegrupper av förstasidorna i DN och HBL)

Welch Two Sample t-test

data: lixDN and lixHBL

t = -4.2363, df = 81.858, p-value = 5.916e-05

alternative hypothesis: true difference in means is not equal to 0

95 percent confidence interval:

-2.939198 -1.060802

sample estimates:

mean of x mean of y

31.04762 33.04762

Resultat, dvs. **t-värde = - 4.2363**

Det finns 81.858 tillgängliga frihetsgrader (df), dvs. ungefär 82 stycken. Testets **p-värde är 0.00005916 ≤ 0.05**, där 0.05 är gränsvärdet för ett statistiskt signifikant test, dvs. skillnaden är statistiskt signifikant. 95 % konfidensintervall är (-2.94 , -1.06). T-testresultat – 4.2363 lämnas utanför detta värdeområde. Utav 95 % konfidensintervallen (eller med 5 % signifikansnivå) och p-värdet tillsammans kan man konstatera att nollhypotesen bekräftas.

#### 5.4.4 Redogörelse för boxplotdiagram

Boxplotdiagrammet består av följande delar: (se boxplotdiagrammen i avsnitten 7.4.1–7.4.9)

- *Maximivärde* inom den ifrågavarande gruppen
- Den övre kanten på lådan kallas för *den tredje kvartilen* (antalet värden är mindre än 75 % av hela värdeområdet inom ett boxplotdiagram)
- Linjen som korsar genom lådan horisontellt är det mittersta talet inom hela värdegruppen (inom det ifrågavarande boxplotdiagrammet) och heter *medianen*
- Den nedre kanten på lådan heter *den första kvartilen* (antalet värden är mindre än 25 % av hela värdeområdet inom ett boxplotdiagram)
- *Minimivärde* inom den ifrågavarande gruppen

Ett boxplotdiagram består av *medianen*, den övre/lägre (*första/tredje*) *kvartilen* samt *minimi/maximivärden*. Lådan i ett boxplotdiagram (eller låddiagram) omfattar 50 % av de observerade värdena. Ett boxplotdiagram kan även ha så kallade ”uteliggare” (engl. ”outliers”), dvs. enstaka extremvärden. Dessa uteliggare presenteras antingen som enstaka asterisker eller ringar i ett boxplotdiagram. Ifall ett värde i ett boxplot ligger från lådan mer än 1,5 gånger kvartilavståndet mellan de yttre kvartilerna betraktas det som en uteliggare. Ett värde som ligger mer än 3 gånger kvartilavståndet från lådan betraktas som en avlägsen uteliggare. När det är fråga om en uteliggare markeras den i ett boxplot med en asterisk (\*). En avlägsen uteliggare betecknas däremot med en ring (o). Ett boxplotdiagram är mycket användbart när man jämför mätvärden från flera grupper eller tidpunkter. (Internet 4–6.)

### 5.5 Om utomstående expertis med analysdel

På grund av sin kvantitativa karaktär har min forskning och mina analyser fått en statistisk-matematisk karaktär. Därför har jag valt att använda erfarenhet och expertis av Jukka Airola, filosofi magister i statistik, som vänligen har bidragit med sina omfattande kunskaper inom denna forskningsgren. Följande analyser är kontrollerade och utförda i samarbete med honom.

## 6 MATERIAL OCH METOD

### 6.1 Utgångspunkt för arbetet

Utgångspunkten för forskningsarbetet i min undersökning är att samla in och analysera ett begränsat antal webbtidningstexter. Det ska ske inom de tidsramar och resurser som går att använda för att få tillstånd en tillräckligt omfattande och också statistiskt reliabelt analysarbete för pro gradu-avhandlingen.

En av de centralaste problemen med att forska i webbtidningarna är deras växlande tillgänglighet och det att de ofta inte är gratis att använda. Därför kommer jag att basera de kommande läsbarhetsmätningarna främst på texterna på tidningarnas DN och HBL förstasidor. För att kunna framställa en mer reliabel och allmängiltig bild av den

genomsnittliga läsbarheten har valt jag att även analysera ett antal insidestexter i tidningen HBL. Tillgängligheten till denna webbtidning var bra (då materialet samlades in) och detta gör det möjligt att sammanfatta en tillräckligt omfattande bild av läsbarheten i det verkliga innehållet i en svenskspråkig webbtidning. Genom en systematisk jämförelse mellan LIX-värden, kalkylerade med hjälp av Grammatifix kan läsbarheten mätas någorlunda pålitligt.

Nuförtiden finns det knappast någon som vill eller har tillräckliga kunskaper för att kunna utföra detta slags projekt genom att kalkylera LIX för hand. Det skulle innebära en omåttlig mängd av beräkningsarbete och tabellering av data. Därför är den enda motiverade möjligheten att använda datorn som beräkningsinstrument och hjälpmedel i analysen för att få resultaten i sifferform.

LIX-metoden visade sig vara den enda rätta och relevanta enstaka kvantitativa metoden som ger tillräckligt med ny information om de valda webbtidningarnas (DN och HBL) allmänna läsbarhet. Informationen bearbetas genom att utföra valida test (t-test) som ger även statistiskt signifikanta resultat. Det är fråga om en datorstödd analys med statistisk matematik som tillämpas på ett humanistiskt forskningsproblem.

## 6.2 Kort om materialets tillgänglighet

Eftersom det valda forskningsmaterialet är begränsat är det nödvändigt att använda det tillgängliga materialet på ett effektivt sätt. Till exempel finns det på DN:s webbsida endast två tillgängliga kompletta webbupplagor. Dessa DN-exemplar är tillgängliga endast som ett par pdf-webbtidningar som går att nerladdas gratis utan avgiftsbelagd registrering. Insidorna i DN är alltså tillgängliga bara som inbyggda delar av kompletta (och avgiftsbelagda) pdf-upplagor. Detta omöjliggör parvisa analyser av insidorna (t.ex. Inrikes) mellan tidningarna DN och HBL. På basis av detta beslöt jag att utföra denna innehållsliga analys endast för HBL:s insidor. Förstasidorna i DN-webbupplagan (på DN:s webbsida) är däremot tillgängliga som enskilda filer och används som material i min undersökning (se även avsnitt 6.4). Största delen av analyserna baserar sig på båda tidningarnas förstasidor.

### 6.3 Hypoteser och forskningsproblem

Huvudsyftet med denna undersökning är att klarlägga all märkvärdig beräknelig information och kunskaper om möjliga företeelser bakom läsbarhet som kommer fram i de anpassade Grammatifix-mätningarna samt signifikansjämförelserna. Meningen är att ge en pålitlig helhetsbild av den nutida svenskspråkiga tidningstextens läsbarhet genom att utföra en relativt omfattande analys av förstasidor och insidor.

I min undersökning försöker jag att utreda om det finns skillnader i läsbarheten mellan webbupplagorna av morgontidningarna DN och HBL. Läsbarheten mäts med ett lämpligt datorprogram (Grammatifix) och dess resultat jämförs som t-test (statistikprogrammet R) för att redogöra för eventuell statistisk signifikans mellan LIX-resultaten. Resultaten avhandlas i resultatkapitlet samt nio utvalda företeelser illustreras med boxplotdiagram och tolkas med statistiska uppgifter såsom medelvärde, standardavvikelse (värdenas spridning) samt median. Dessutom har jag avsikt för att tolka resultaten enligt Björnssons (1968, 89–90) värdering av LIX (se även avsnitt 3.2.5). Slutligen ska mina LIX-resultat jämföras med några övriga LIX-undersökningsresultat.

Mina ursprungliga hypoteser baserar sig på en noggrann förgranskning av både förstasides- och insidesmaterialet i DN och HBL:

1. Det finns inte några signifikanta skillnader i läsbarheten varken mellan olika veckodagar eller veckor på förstasidorna i DN och HBL när läsbarheten i tidningarna jämförs inom respektive tidningar.
2. Det finns skillnader i läsbarheten mellan förstasidorna i DN och HBL.
3. Det finns några signifikanta skillnader i läsbarheten inom insidorna mellan olika avdelningar (t.ex. Inrikes vs Sport) i HBL.

På basis av de ovannämnda hypoteserna och övriga förantaganden om textmaterialets läsbarhet i DN och HBL har jag formulerat huvudproblemen i min undersökning:

1. Vad är de genomsnittliga LIX-värden, standardavvikelsen (som spridningsmått) samt medianen i webbtidningstexterna (förstasidorna: DN/HBL och insidorna: HBL)?
2. Finns det statistiskt signifikanta skillnader i läsbarheten mellan förstasidorna på olika veckodagar eller veckor inom och mellan tidningarna DN och HBL?
3. Finns det några statistiskt signifikanta skillnader i läsbarheten inom insidorna mellan olika avdelningar i HBL?

Björnsson (1968, 71) betonar att det inte är rimligt att generalisera utifrån provningsresultatet i fråga om korta texter. Därtill hävdar Björnsson att LIX vid korta texter antingen kan bli mycket hög eller låg. I min undersökning har de analyserade texterna enligt Björnssons (1968, 70) råd prövats helt (gäller texter mindre än 2000 ord) och textlängden överskrider inte 1000 ord/text varken i förstasides- eller insidesmaterialet. Idealfallet skulle ha varit om alla texter varit lika långa och sålunda skulle en osäker faktor i mätningen varit utesluten. I praktiken är situationen annorlunda då texterna alltid varierar mycket i längd i olika tidningsnummer. På basis av detta kan det nog förväntas att de genomsnittliga LIX-resultaten i min undersökning inte direkt blir desamma när de jämförs med t.ex. Björnssons (1968) resultat. Vad insidesmaterialet beträffar är det fråga om tidningen HBL i min undersökning men DN i t.ex. Strands (1984) studier. Även skillnaden mellan tidningarnas språk, dvs. HBL (finlandssvenska) och DN (rikssvenska) kan påverka de genomsnittliga LIX-värdena. Med hänsyn till dessa fakta försöker jag att komma på/få fram resultat ur materialet som i någon mån kan betraktas som generella tendenser i läsbarheten.

#### 6.4 Om forskningsmaterial och -metoder

Undersökningen inleddes med en noggrann granskning av några svenskspråkiga webbtidningar. Den ursprungliga gruppen omfattade morgontidningarna *Dagens Nyheter*, *Svenska Dagbladet* och *Hufvudstadsbladet* samt dagstidningarna *Aftonbladet* och *Expressen*. Genom att bekanta mig med alla dessa tidningar som online-webbupplagor valde jag två morgontidningar från denna grupp. De två tidningarna skulle representera materialet för de kommande LIX-mätningarna.



Nästa steg i den förberedande processen innan själva mätningarna innebar att jag skaffade ett tillämpligt mätinstrument för att kunna mäta texternas läsbarhet (LIX). Programmet *Grammatifix* som integreras med Microsoft Word är ett lämpligt, valid och också relativt reliabelt verktyg när det gäller att kalkylera LIX på elektroniska texter. Därför valde jag det som analysmedel. Nyare versioner av Microsoft Word än Word 97 har ett identiskt delprogram som är färdigt integrerat med t.ex. Word 2000. Jag använder Word 97 och då kräver processen *Grammatifix* för beräkningen av LIX.

Efter att ha testat programmet *Grammatifix* med ett par slumpmässigt valda webbtidningstexter, dvs. att programmet skulle ge förväntade och relevanta LIX-värden inom Microsoft Word 97, gällde det att välja texter från sådana webbtidningar som skulle vara de mest tillgängliga innehållsmässigt. För att kunna fokusera denna undersökning på de texter som är tillgängliga åtminstone i två olika morgontidningar bestämde jag mig att avgränsa mätningarna till morgontidningarna *Dagens Nyheter* och *Hufvudstadsbladet*.

DN erbjöd några enskilda insidor varken i html-form eller i pdf-form. Därför bestämde jag mig för att endast utnyttja förstasidorna från tidningen DN. Dessutom samlade jag in lika många förstasidor från HBL i html-form. Allt förstasidesmaterial i både DN och HBL utgörs av två tidsligt oavbrutna helheter från den 7 februari till den 20 mars 2005. Sammanlagt omfattar förstasidesmaterialet i DN och HBL 84 förstasidor, dvs. 42 förstasidor per tidning bestående av sex fullständiga veckor, från respektive webbupplagor (se bilaga 1). Detta antal är valt med hänsyn till allmänna statistiska validitets- och reliabilitetsgränser. I praktiken betydde detta att vänta tills en ny vecka med nya DN-förstasidor på sju veckodagar kom ut och sedan stod till buds som pdf-filer i respektive länkar på DN:s hemsida. Sedan handlade det om att ladda ner var och en förstasidesfil från måndag till söndag tills det eftersökta antalet förstasidor nåddes.

Med *Hufvudstadsbladet* var den motsvarande insamlingsprocessen litet annorlunda. Innan Hbl:s hemsida uppdaterades fanns det ett mycket omfattande arkiv där alla förstasidor, artiklar, insidor (t.ex. Politik, Inrikes, Ekonomi) stod till buds i normal html-form, dvs. som vanliga webbsidor.

Den manuella konverteringen av texterna var en arbetsfas där jag förvandlade de valda texterna (både DN och HBL) till Word-form. Detta gjorde jag genom att använda kopiera och

klistra in-metoden mellan den ursprungliga texten och ett tomt Word-dokument. Dessa gjorde jag noggrant och systematiskt för att kunna utesluta oönskade mänskliga fel – till exempel att lämna bort några delar av de beräknade texterna av misstag. Den enda skillnaden mellan DN och HBL var att förstasidestexterna i DN-materialet som erbjöds som pdf-filer måste först öppnas i programmet Adobe Acrobat Reader. Sedan använde jag samma metod som med HBL. Den kopierade texten klistrade jag i ett tomt Word-dokument. Sedan beräknade jag LIX med Grammatifix.

Dessutom valde jag sju avdelningar från HBL:s insidor: *Ekonomi, Inrikes, Kultur, Ledare, Politik, Sport* och *Utrikes*. Från var och en avdelning valde jag 12 stycken texter, alltså sammanlagt 84 artiklar från olika upplagor. Syftet med insidesmaterialet från HBL är att använda sju olika ämneskretsar för att utsträcka LIX-analyserna över hela huvudsakliga ämnesinnehållet i HBL.

Samplet var relativt litet och det enda riktiga målet var att detta på förhand valda antalet artiklar nåddes. Dessa 84 valda artiklar från respektive webbtidningar (5 webbupplagor) härstammar från tidsperioden från den 5 oktober 2004 till den 19 mars 2005. Då vissa enskilda upplagor av de på varandra följande exemplar inte hade artiklar i alla sju avdelningar måste en ersättande artikel användas. En ungefärligt likadan artikel från en tidsligt nära upplaga fick ersätta den saknande enskilda avdelningen. Antalet vissa insidesartiklar (t.ex. inrikes) från olika dagars upplagor är osystematiskt valt och har gjorts utan någon preliminär bedömning av texternas innehåll.

Då insamlings- och konverteringsprocessen samt LIX-beräkningarna med Grammatifix av både förstasidorna och insidorna var slutförda analyserades LIX-resultaten med hjälp av statistikprogrammet R. Dessa statistiska beräkningar utfördes som parvisa t-test varefter resultaten tabellerades i Microsoft Excel 97. Meningen med att tabellera de beräknade t-test-resultaten var att kunna utreda skillnaderna i läsbarheten (i LIX) genom att utvärdera resultaten enligt de allmänna statistiska 0,1 %, 1 %, 5 % samt 10 % signifikansnivåerna. Om dessa nivåer (i Excel-materialet) talas med respektive språkliga uttryck *mycket signifikant* (0,1 %), *signifikant* (1 %), *nästan signifikant* (5 %) samt *icke-signifikant* (10 %). Dessa fyra allmänt kända signifikansnivåerna har använts för att kunna beskriva resultaten i närmare detalj dvs. med fyra olika signifikansgrader istället för att bara konstatera resultaten vara antingen signifikanta eller icke-signifikanta. Man bör dock lägga märke till att då resultaten i

min undersökning granskas och tolkas enligt 5 % signifikansnivå fungerar denna gräns som vågbrytare mellan ett signifikant vs icke-signifikant resultat även om själva Excel-siffermaterialet utvärderats enligt en mer detaljerad graderingsskala med fyra signifikansnivåer. Då uttrycks resultaten från icke-signifikant till mycket signifikant samt respektive siffror men det avgörande gränsvärdet är 0,05 (se även avsnitt 5.4.4).

Själva jämförelseparen (i t-testen) består av ett antal beräkningar mellan veckodagar inom och mellan de respektive tidningarna DN och HBL. Med ett likadant mönster har dessutom veckor jämförts (se bilagor 4–7). När det gäller insidesartiklarna i *Hufvudstadsbladet* behandlas dessa som jämförelsepar så att alla blir jämförda med varandra (se bilaga 3). Alla resultat av dessa jämförande t-test (p-värden), dvs. utdata från statistikprogrammet R, ordnas och tabelleras i Microsoft Excel 97 enligt de ovannämnda fyra signifikansnivåerna. På så sätt möjliggör det beräknade siffermaterialet även utvärderingen av p-värden enligt fem procents signifikansnivå. Av de allmänna statistiska signifikansnivåerna är fem procents (0,05) signifikansnivå dock den vanligaste inom bl.a. språkvetenskap och därför använder jag den som utgångspunkt i signifikansjämförelserna (se avsnitt 5.4.4). Dessutom utförs T-testberäkningarna med statistikprogrammet R med programmets färdiga funktioner och på detta sätt är det också möjligt att skriva ut motsvarande boxplotdiagram med samma program.

Huvudsyftet med att tolka det bifogade Excel-siffermaterialet är att behandla det väsentligaste, dvs. statistiskt signifikanta resultat (enligt 5 % signifikansnivå) som även jämförs och illustreras som respektive boxplotdiagram. Dessa signifikanta värden markeras i Excel-materialet med fetstil och 5 %-signifikansgränsen framställs med en dubbel våglinje i varje jämförelsegrupp (se t.ex. bilaga 5). De fetstilmarkerade fallen fungerar som en viktig utgångspunkt för de nio illustrerade (boxplotdiagrammen) företeelser som redogörs för i resultatkapitlet. Slutligen tolkar jag de tabellerade resultaten genom att jag redogör för var och en av de utvalda nio fallen med hjälp av respektive boxplotdiagram inklusive tabellerade statistiska uppgifter (t-testets p-värde, medelvärde, standardavvikelse och median) samt granskar LIX-värden enligt gränsvärden i Björnssons LIX-tolk (se avsnitt 3.2.5).

## 7 RESULTAT OCH DERAS TOLKNING

### 7.1 Allmänt

Detta kapitel utgör den viktigaste delen i min undersökning. I resultatkapitlet ska alla de viktigaste företeelserna som analyserna (t-testen) framkallat rapporteras. Naturligtvis behövs alla otaliga individuella t-testresultat inte redogöras för som enstaka fall. Istället försöker jag att behandla alla signifikanta och även några icke-signifikanta resultat antingen med hjälp av en tabell samt respektive figur eller genom att redogöra för och sedan hänvisa till den ifrågavarande delen i det ytterst viktiga bifogade Excel-materialet (bilagorna 1–7).

I avsnitt 7.2 diskuteras och redogörs för LIX- samt statistiksiffrorna som ska användas i rapportering av resultaten. I avsnitt 7.3 utredas boxplotfiguren inklusive boxplotdiagrammet och värdenas placering på dessa kort. En närmare förklaring för boxplotdiagrammet och dess uppbyggnad förklaras i avsnitt 5.4.4. Den centralaste delen av hela resultatkapitlet - även hela undersökningen - är avsnitt 7.4 samt 7.4.1 – 7.4.9. Denna helhet omfattar alla nio företeelser som kommer att beskrivas med hjälp av av tabellerna 1–9, respektive figurerna 1–9 (boxplotdiagrammen) med förklaringar och kommentarer om respektive t-testresultat (p-värden) samt hänvisningar till Excel-materialet. De av de ovannämnda nio företeelser som inte är jämförelser av två (med två diagram) värdegrupper, dvs. 7.4.2, 7.4.4, 7.4.7, 7.4.8 samt 7.4.9 ska behandlas i avsnitt 7.5 angående statistisk information där det även anges procentandelen statistiskt signifikanta resultat (enligt 5 % signifikansnivå) av varje jämförelsegrupp. Dessutom anges signifikansen enligt de fyra allmänna signifikansnivåerna. På detta sätt blir allt signifikant Excel-material, dvs. alla nödvändiga parvisa t-test, noterade (se även följande avsnitt).

Dessutom tolkas/kommenteras värdenas spridning i figurerna efter ögonmått med hjälp av boxplotdiagrammen och på basis av respektive tabellsiffror. Resultaten av statistisk signifikansprövning samt de resultat som inte är illustrerade och beskrivna med figurer och tabeller (DN vs HBL veckor/veckodagar) kommer att behandlas i avsnitt 7.4. Även några övriga allmänna resultat diskuteras i det avsnittet. Dessutom diskuteras de allmänna LIX-resultaten och de genomsnittliga LIX-medelvärden jämförs (förstasidona/insidorna) med

tidigare undersökningsresultat i 7.5. Slutligen granskas de erhållna resultaten i relation till mina hypoteser och forskningsproblem (i kapitel 7) och eventuella slutsatser ska dras.

## 7.2 Om LIX- och statistikvärden

I det beräknade Grammatifix-materialet (i Excel) anges LIX-värden alltid som heltal då de tabelleras. För att kunna förgranska de tabellerade värden mångsidigt innehåller de även eventuella avrundade medelvärden (se bilagor 1 och 2). Fast Björnssons LIX-skala baserar sig på heltal från 20 till 60 anges värden som gäller rapportering av genomsnittet (medelvärdena) och spridningen (medianen och standardavvikelsen) som decimalvärden i min undersökning. Motivering till detta är att jag inte velat avrunda någonting innan resultaten blivit slutliga. Alla värden, dvs. medelvärden, standardavvikelsena samt medianerna är direkt tabellerade från statistikprogrammet R:s beräkningsresultat.

Då det däremot handlar om att tolka värden enligt LIX-skalan (LIX-tolken) avrundas LIX-värden till närmaste heltal från respektive medianen. I praktiken har detta betydelse endast i gränsfall (se tabell 9, HBL-ledare) där exempelvis lätt blir medelmåttig då medianen – LIX-värderingen görs på grund av medianvärdet – 39,5 avrundas till närmaste heltal, 40. Detta är också det enda fallet i numerala resultaten i min undersökning där medianvärdet ligger på gränsen enligt Björnssons LIX-tolk. I boxplotdiagrammen (skalan till vänster i en boxplotfigur) är LIX-tal alltid heltal när det gäller att tolka LIX-resultaten – själva värden som boxplotgraferna illustrerar betraktas som decimalsiffror eftersom statistikprogrammet R, med vilken boxplotten är utskrivna, inte baserar sig på heltal eller en heltalsskala.

P-värden (t-testresultaten) anges som direkta beräkningsresultat (programmet R) med respektive decimaler varefter dessa decimalsiffror har värderats enligt allmänt kända fyra 10 %, 5 %, 1 % samt 0,1 % signifikansnivåer (se avsnitten 5.4.1 och 6.4). På basis av R-beräkningarna anges siffrorna i tabellerna 1–9 alltså som decimalvärden. LIX-medelvärdet anges med fem decimaler, standardavvikelsen med sex decimaler och medianen med en decimal (se exempelvis tabell 1).

Vid tabellvärdena kan till begreppet standardavvikelse refereras med förkortning (sd) eller med ordet ”spridning” (se även avsnitt 5.3).

### 7.3 Om boxplotfigurerna samt placering av diagramvärdena

En boxplotfigur inklusive ett eller flera boxplotdiagram kan användas för att illustrera spridningen av en värdegrupp. Jag har tillämpat boxplottdiagram (kan kallas även för låddiagram) med avsikt att ge bästa möjliga illustration av respektive tabellerade statistikvärden. Figurerna har tillämpats enligt de ifrågavarande fallen så att företeelseerna i figurerna ryms på skalan på ett ändamålsenligt sätt. Exaktheten på granskningen hos LIX-skalan varierar mellan 1, 2 och högst 5 LIX-steg (intervallen) i olika figurer. Alla LIX-värden inklusive några eventuella extrema värden ("uteliggarna") syns på värdeområdet i boxplotdiagrammen i figurerna. Uteliggarna i min undersökning är alltid uteliggare med låga värden. När dessa förekommer i ett diagram påverkar de närmast standardavvikelsen, dvs. det hur pass omfattande spridningen för en värdegrupp är. Extrema värden påverkar även medelvärdet samt medianen men i min undersökning (bl.a. på grund av sampelstorleken) har de inte någon stark inverkan på dem.

Medianen syns oftast i boxplotdiagrammen i figurerna som en vågrätt linje som korsar lådan. Det finns också fall där medianen är densamma som första/tredje kvartilen. Då är det bäst att se respektive medianvärdet i tabellen för att få veta var medianen ligger i det diagrammet. Dessutom kan medianen vara det enda värdet som illustrerar en värdegrupp. Om spridningen i gruppen (standardavvikelsen) är noll är det endast medianen som står kvar i diagrammet. Medelvärdet är den enda av tabellvärdena som inte illustreras i ett boxplotdiagram. Orsaken till detta är att medianen anses vara statistiskt mer rättvisande än medelvärdet. Även lådans (innehåller 50 % av observationerna i gruppen) placering, dvs. höjd på stapeln i ett boxplotdiagram ger information om värdenas spridning och placering på skalan.

Dessutom bör man lägga märke till minimi- och maximivärdet samt fallen där en uteliggare ("outlier") utgör minimivärdet i diagrammet. Det finns också fall där första och tredje kvartilen utgör minimivärdet och/eller maximivärdet i ett boxplot, dvs. antingen den övre eller den nedre stapeln saknas fullkomligt. Det är även möjligt att båda staplarna saknas (första/tredje kvartilen presenterar då minimi- samt maximivärde) och diagrammet förefaller som en låda med endast medianen i mitten. De ovan beskrivna avvikande fall med boxplotdiagrammen redogörs för i figurbeskrivningar (i avsnitten 7.4.x) vid behov (se även avsnitt 5.4.4: redogörelsen för boxplotdiagrammet).

## 7.4 Spridning av LIX-värdena

Detta kapitel omfattar nio företeelser som kommer att beskrivas på följande sätt. En helhet börjar med en kort förklaring av den kommande företeelsen. LIX-värdena anges som statistiska siffror i tabellerna 1–9 varefter de illustreras med respektive figurer 1–9 (boxplot) med förklaringar och kommentarer om respektive t-testresultat (p-värden) samt hänvisningar till p-värden i Excel-materialet. Dessutom tolkas värdenas spridning på LIX-skalan i figurerna (i boxploten) först på LIX-skalan efter ögonmått och sedan på basis av den respektive tabellerade standardavvikelsen. Tabellerna samt figurerna bildar alltså nio helheter som tillsammans ska ange de centralaste kunskaperna om de beräknade LIX-resultaten, värdenas spridning och LIX-värdering samt skillnaderna med statistisk signifikansprövning enligt 5 % signifikansnivå.

I avsnitt 7.4.1 jämförs DN:s och HBL:s genomsnittliga LIX-spridning som helheter. Därefter följer 7.4.2 där förstasidornas värden i DN för olika veckor granskas. Sedan tar vi en titt på förstasidornas LIX-medelvärden på veckorna 2 och 5 i DN i avsnitt 7.4.3. Vidare rapporteras det i avsnitt 7.4.4 om alla veckodagars värden för förstasidorna i DN. Skillnaderna i spridningen i två veckodagpar (DN) redogörs för i 7.4.5 samt 7.4.6. Från och med avsnitt 7.4.7 ska värden för tidningen HBL granskas där det först handlar om olika veckors LIX-spridning för förstasidorna. Därtill utredas värden i HBL för olika veckodagar i 7.4.8. Till sist behandlas LIX-spridningen mellan 7 olika ämneskretsar (avdelningar) på insidorna i HBL i avsnitt 7.4.9. Skillnaderna mellan DN:s och HBL:s (veckor/veckodagar) värden kommer att avhandlas i avsnitt 7.5.

### 7.4.1 Alla LIX-värden för förstasidorna i DN och HBL

I detta avsnitt redogörs för spridningen av det genomsnittliga LIX för både DN och HBL angående förstasidors (42 stycken i respektive tidningar) värden som två skilda medelvärdegrupper.

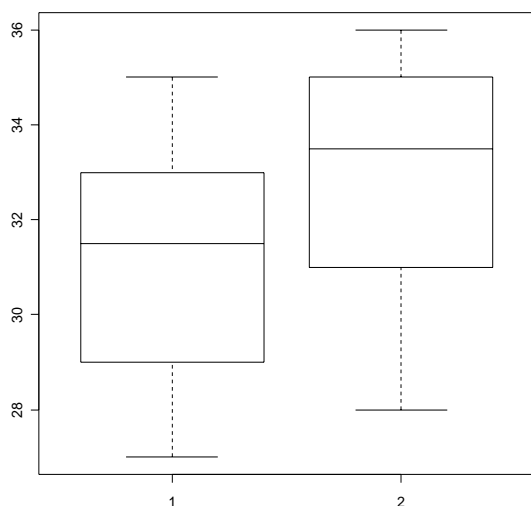
I tabell 1 presenteras LIX-medelvärdena, -spridningen (dvs. standardavvikelsen) -medianen samt -värderingen enligt Björnssons LIX-tolk (se även avsnitt 3.2.5) för alla förstasidor i DN och HBL som två värdegrupper.

*Tabell 1 Medelvärde, standardavvikelse, median samt tolk för alla förstasidors LIX för DN och HBL*

	1=DN	2=HBL
Medelvärde	31.04762	33.04762
Standardavvikelse	2.208105	2.117896
Median	31.5	33.5
LIX-tolk	Lätt	Lätt

Som framgår av tabell 1 finns det två LIX-enheters skillnad mellan medelvärden. På basis av jämnt utdelade LIX-värden (ursprungligen är dessa heltal) är också decimalerna desamma. Detta är dock ett undantag. Standardavvikelsen (också: spridningen, sd) är numeralt, dvs. enligt respektive tabellvärden ungefärligt densamma i båda värdegrupperna. Även medianerna ligger på samma 2 enheters distans från varandra. Enligt Björnssons LIX-tolk kan läsbarheten i båda förstasidesgrupperna (DN och HBL) värderas som "lätt".

Följande figur illustrerar spridningen av alla LIX-värden för förstasidorna i DN och HBL med två boxplotdiagram. Respektive statistiska värden finns i tabell 1. LIX-intervallen i skalan i figur 1 är två enheter.



*Figur 1 Spridningen av alla LIX-värden för förstasidorna i DN och HBL*

I figur 1 kan man se att värden i DN i största allmänhet är lägre än i HBL. Detta syns tydligt också i de motsvarande medelvärdena. I standardfall syns minimi- och maximivärdena i diagrammen som slutpunkter på staplarna. Höjden på lådan illustrerar spridningen bland värdena. Placeringen av lådan på stapeln visar också värdenas spridningsområde. Dessutom



bör även alla de övriga värdena i diagrammen noteras, inte endast värdena på lådans värdeområde. Alla värden på hela diagrammets värdeområde (också utanför lådan) påverkar spridningen i sin helhet. Mellan den övre (tredje kvartilen) och nedre kanten (första kvartilen) i lådan finns de vanligaste värdena (50 % av alla värden) i det ifrågavarande diagrammet. Den horisontella linjen på insidan av lådan betecknar *medianen* hos gruppen, dvs. det mittersta talet inom värdegruppen. På grund av att det finns ett jämnt antal observationer gäller medianen medelvärdet av de två mittersta talen.

Om det istället fanns ett udda antal observationer, skulle medianen direkt vara det mittersta talet. Detta härleder sig direkt från sampelstorleken i undersökningen (som är ett jämnt tal, 42). Det mittersta talet kan inte kalkyleras utav ett jämnt tal. Exempelvis är medianen av talen 1, 2, 3, 4, 5, 6, 7 (7 tal = ojämnt antal) 4, dvs. det mittersta talet i en storleksordnad ordning (här alltså: 4) och *medelvärdet* är 4:  $(1+2+3+4+5+6+7) / 7$ . Av talen 1,2,3,4,5,5 (6 tal= jämnt antal) är *medianen*  $(3+4) / 2 = 3,5$  och *medelvärdet* är 3,333...  $(1+2+3+4+5+5) / 6$ . Således gäller på en allmän nivå: *median*  $\neq$  *medelvärde*!

Efter ögonmått är spridningen (sd), dvs. enligt respektive boxplot lika stort för båda värdegrupperna. Enligt 5 % signifikansnivå är skillnaden mellan alla förstasidor i båda tidningarna statistiskt *signifikant* då t-testets p-värde är 0,00005916. Enligt de allmänna fyra signifikansnivåerna klassificeras skillnaden som *mycket signifikant* (se avsnitt 7.5.1).

#### 7.4.2 LIX-värden för förstasidorna i DN för olika veckor

Detta avsnitt handlar om spridningen av det genomsnittliga LIX för sex olika veckor (samelstorlek: 42 / 7 veckodagar) i DN angående förstasidors värden som sex skilda värdegrupper.

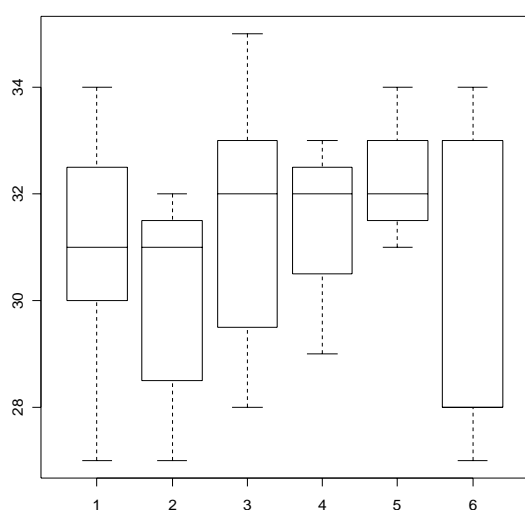
I tabell 2 presenteras LIX-medelvärdena, -spridningen, -medianen samt -värderingen enligt Björnssons LIX-tolk för förstasidorna i DN som sex värdegrupper för olika veckor.

Tabell 2 Medelvärde, standardavvikelse, median samt tolk för LIX i DN för olika veckor (förstasidor)

	1=DN vecka 1	2=DN vecka 2	3=DN vecka 3	4=DN vecka 4
Medelvärde	31.00000	30.00000	31.42857	31.42857
Standardavvikelse	2.380476	2.000000	2.507133	1.511858
Median	31	31	32	32
LIX-tolk	Lätt	Lätt	Lätt	Lätt
	5=DN vecka 5	6=DN vecka 6		
Medelvärde	32.28571	30.14286		
Standardavvikelse	1.112697	3.078342		
Median	32	28		
LIX-tolk	Lätt	Mycket lätt		

Som man kan se av tabell 2 ligger de aritmetiska medelvärden för alla sex veckorna ungefärligt på samma nivå, mellan ca 30–32. Förutom vecka 6 är dessutom medianerna mycket homogena 31–32. Spridningen håller sig ungefärligt mellan dryga 1 och 2,5 men värdegrupp DN vecka 6 avviker från alla övriga fem veckor med sin omfattande spridning (ca 3) och låga median, 28. I värdegrupp DN-vecka 1 är medelvärdet och medianen exakt desamma, 31. På basis av detta låga medianvärde blir även LIX-värderingen annorlunda ("mycket lätt") än med de övriga fem fallen (dvs. veckorna 1–5: "lätt").

Figur 2 illustrerar spridningen av alla LIX-värden för förstasidorna i DN för olika sex veckor med motsvarande boxplotdiagram. Respektive statistiska värden finns i tabell 2. LIX-intervallen i skalan i figur 2 är två enheter.



Figur 2 Spridningen av LIX-värden för förstasidorna i DN för olika veckor

Figur 2 ger vid handen att medianer hos majoriteten av veckorna ligger på samma nivå (31 och 32). Sålunda är grupperna (dvs. veckorna) i största allmänhet likadana. Dessutom ligger lådorna, dvs. majoriteten av samplen, i stort sett på samma höjd på LIX-skalan. Medelvärdet av alla värden i DN är 31.04. Vecka 6 skiljer sig dock från resten av veckorna på grund av sin låga median. Medianen ligger exakt på samma nivå, 28 med första kvartilen i vecka sex diagram, dvs. densamma med första kvartilen. I figur 2 ligger endast diagrammet för vecka 5 i sin helhet ovanför medelvärdet för alla förstasidor för DN (31.04). Resten av veckorna är jämnt spridda ovanför och under detta medelvärde (jfr även med figur 7).

Enligt ögonbedömning är spridningen tydligt störst i diagrammet för vecka sex. Veckorna 1–5 däremot beter sig grafiskt på ett motsvarande sätt som respektive statistiska tabellvärden. Respektive p-värden (t-test) behandlas och diskuteras i avsnitt 7.5.2 (se avsnitt 7.5.2).

#### 7.4.3 LIX-värden för förstasidesgrupperna DN vecka 2 och DN vecka 5

I detta tredje avsnitt utredas två specialfall i DN:s veckojämförelser. Jämförelsen mellan två avvikande värdegrupper vecka 2 och vecka 5 ska ge information om hur LIX-värdena skiljer sig från varandra mellan olika veckor angående värdenas spridning (sd). Detta fall ska fungera som en närmare granskning av olika veckors LIX än i 7.4.2 där alla sex veckor jämförs som skilda grupper.

I tabell 3 presenteras LIX-medelvärdena, -spridningen, -medianen samt -värderingen enligt Björnssons LIX-tolk för förstasidesgrupperna DN-vecka 2 och DN-vecka 5.

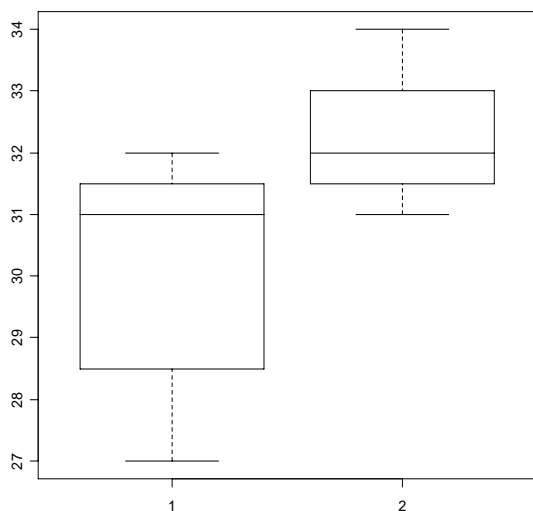
*Tabell 3 Medelvärde, standardavvikelse, median samt tolk för grupperna DN vecka 2 och DN vecka 5*

	1=DN vecka 2	2=DN vecka 5
Medelvärde	30.00000	32.28571
Standardavvikelse	2.000000	1.112697
Median	31	32
LIX-tolk	Lätt	Lätt

Som framgår av tabell 3 finns det en dryg två LIX-enheters skillnad mellan medelvärden för veckorna 2 och 5. Spridningen är avsevärt större för vecka 2 (se figur 3). Mellan medianerna är det endast en skillnad på en LIX-enhet. Enligt LIX-tolken är texten lika lätt hos vecka 2

som hos vecka 5 men de egentliga skillnaderna finns i spridningen dvs. i standardavvikelsen av värdena.

I figur 3 illustreras spridningen av alla LIX-värden för veckorna 2 och 5 i DN med två boxplotdiagram. Respektive statistiska värden finns i tabell 3. LIX-intervallen i skalan i figur 3 är en (1) enhet.



Figur 3 Spridningen av LIX-värden för förstasidesgrupperna DN vecka 2 och DN vecka 5

Det framgår av figur 3 att medelvärdet 31.04 korsar tvärs lådan nära medianen i diagrammet för DN-vecka 2 (medelvärdet är dock inte synligt i figuren) men alla värden i den högra värdegruppen (vecka 5) ligger ovanför 31.04. Huvudsakligen ligger DN:s läsbarhet på LIX-skalan enligt respektive boxplotdiagram mellan dryga 28–31 (boxplot 1) och HBL:s ungefärligt mellan 32–33 (boxplot 2). Detta syns i att boxplotdiagrammen ligger på helt olika höjder i figur 3. DN ligger i stort sett under och HBL helt ovanför medelvärdelinjen. Det intressanta är just divergensen mellan dessa två veckor hos samma tidning. Hypotetiskt förblir LIX relativt statistiskt vid olika tidpunkter inom en och samma tidning. På så sätt är systematiskt upprepande skillnader inom samma tidning under en kort tidsperiod inte förväntade (inte ens för förstasidestext såsom här).

Visuellt bedömt är spridningen (sd) enligt boxplotten betydligt mer omfattande för vecka 2. Enligt 5 % signifikansnivå är skillnaden mellan förstasidesgrupperna DN-vecka 2 och DN-

vecka 5 statistiskt *signifikant* på grund av t-testets p-värde 0,0259. Enligt de fyra allmänna signifikansnivåerna är skillnaden *nästan signifikant* (se även avsnitt 7.5.2).

#### 7.4.4 LIX-värden för förstasidorna i DN för olika veckodagar

I det följande behandlas det genomsnittliga LIX för sju olika veckodagar (samplelstorlek: 42 / 6 veckor) i DN angående förstasidors värden som sju skilda värdegrupper.

I tabell 4 anges LIX-medelvärdena, -spridningen, -medianen samt -värderingen enligt Björnssons LIX-tolk för förstasidorna i DN som sju värdegrupper för olika veckodagar.

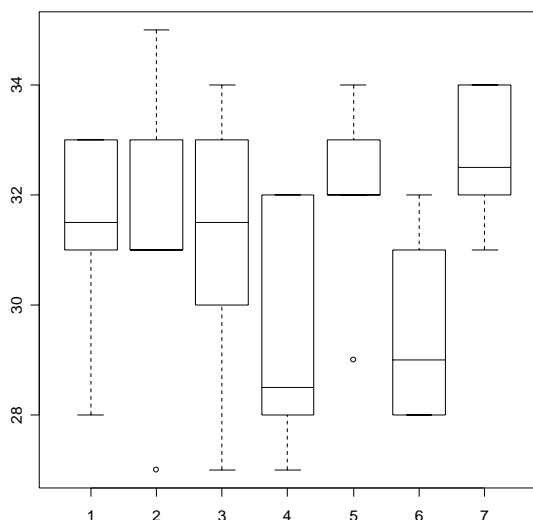
*Tabell 4 Medelvärde, standardavvikelse, median samt tolk för LIX i DN för olika veckodagar (förstasidor)*

	1=DN måndag	2=DN tisdag	3=DN onsdag	4=DN torsdag
Medelvärde	31.33333	31.33333	31.16667	29.33333
Standardavvikelse	1.861899	2.658320	2.639444	2.160247
Median	31.5	31	31.5	28.5
LIX-tolk	Lätt	Lätt	Lätt	Mycket lätt
	5=DN fredag	6=DN lördag	7=DN söndag	
Medelvärde	32.00000	29.50000	32.66667	
Standardavvikelse	1.67332	1.643168	1.211060	
Median	32	29	32.5	
LIX-tolk	Lätt	Mycket lätt	Lätt	

Tabell 4 ger vid handen att de aritmetiska medelvärden för alla sju veckor ligger approximativt på samma nivå, dvs. mellan ca 29–33. Spridningen (sd) för värdegrupperna DN- tisdag (2,6) och DN-onsdag (2,6) är betydligt större än för alla andra veckodagar förutom torsdagen vars spridningsvärde är över två (2,16). Spridningen för måndagen, fredagen, lördagen samt söndagen ligger däremot mellan ca 1,2–1,9. Med undantag av torsdagen och lördagen ligger medianerna över 30. I värdegrupp 5 (DN-fredag) är medelvärdet och medianen exakt desamma, 32. På basis av sina låga medianer blir torsdagen samt lördagen enligt LIX-tolken även värderade som ”mycket lätta” medan de övriga dagarna tolkas som ”lätta”.

Följande figur (figur 4) illustrerar spridningen av LIX-värden för förstasidorna i DN för olika veckodagar med hjälp av motsvarande boxplotdiagram.

Respektive statistiska värden finns i tabell 4 ovan. LIX-intervallen i skalan i figur 4 är två enheter.



Figur 4 Spridningen av LIX-värden för förstasidorna i DN för olika veckodagar

Det framgår av figur 4 att finns tre veckodagar av sju som antingen ligger ovanför eller under medelvärdet 31.04. Av dessa tre (fr, lö, sö) ligger fredagen och söndagen ovanför medelvärdet och lördagen i stort sett under det. Dessutom finns det två speciella fall (ti och fr). Dessa enstaka och mycket avvikande värden som ligger som fristående ”ringar” utanför resten av spridningsvärden och värdeområdet kallas – som konstaterats i avsnitt 5.4.4 – för uteliggare. Uteliggarna påverkar spridningen avsevärt då de ligger tillräckligt långt bort från det övriga värdeområdet i värdegrupp. (Se figur 4, tisdag). När det gäller tisdagen i figur 4 ger respektive extremvärde anledning att förvänta standardavvikelsen för denna grupp förändras drastiskt. Alla förstasidornas medelvärde 31.04 korsar dock tvärs lådan i tisdagens diagram trots dess stora spridning (lådan är relativt liten). Om det inte fanns något extrem värde hos tisdagen skulle spridningen hos den dagen ligga ovanför DN-förstasidornas medelvärde (31.04). Hos fredagen har extremvärdet där inte så betydande roll (jämfört med ti) till exempel på grund av att värden förutom uteliggaren är relativt jämnt spridda. Uteliggarna utgör minimivärden för tisdagen (LIX: 27) och fredagen (LIX: 29). Diagrammen för måndagen, torsdagen samt söndagen har inte några övre staplar som i ett ”standardfall” anger respektive maximivärden utan de tredje kvartilerna i dessa boxplot fungerar även som maximivärde inom gruppernas värdeområden. Dessutom har lördagens diagram däremot inte någon stapel för sitt minimivärde utan den första kvartilen har denna funktion (se även avsnitt 7.3).

Då det handlar i figur 4 (och tabell 4) om sju olika fall bland vilka de finns inte några extrema avvikanden i den genomsnittliga spridningen (förutom uteliggarna i tisdagens och fredagens diagram) ska jag fokusera den visuella granskningen av spridningen på minimi- samt maximispridning. Efter ögonmått är söndagens diagram mest kompakt med inga uteliggare. Enligt ögonbedömning verkar diagram 3 (onsdag) ha den mest omfattande spridningen i figur 4. Emellertid är spridningen i sin helhet för onsdagen (boxplot 3) större än för tisdagen (boxplot 2). Detta framgår senast av respektive värden i tabell 4. Extremvärdet i tisdagens värdegrupp (boxplot 2) påverkar dock spridningen så pass kraftigt att det matematiska värdet (sd) för tisdagen blir större än för onsdagen. Respektive p-värden, dvs. statistiska provningsresultat, ska behandlas och diskuteras i avsnitt 7.5.3 (se avsnitt 7.5.3).

#### 7.4.5 LIX-värden för förstasidesgrupperna DN-måndag och DN-tisdag

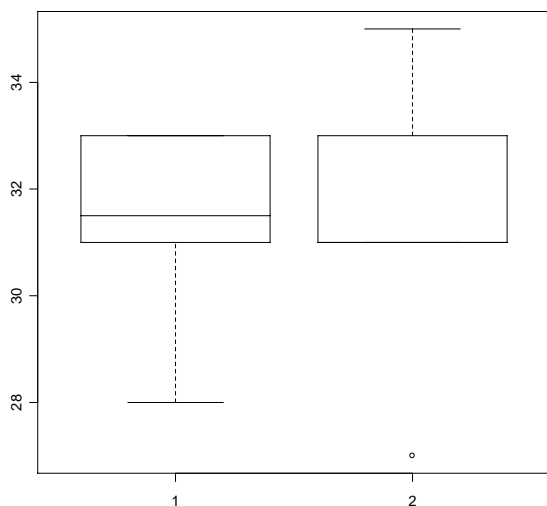
I detta avsnitt behandlas två specialfall i DN:s veckodagjämförelser. Jämförelsen mellan två värdegrupper DN-måndag och DN-tisdag ska ge information om hur LIX-värdena skiljer sig från varandra mellan olika veckor angående värdenas spridning (sd).

I tabell 5 presenteras LIX-medelvärdena, -spridningen, -medianen samt -värderingen enligt Björnssons LIX-tolk för förstasidesgrupperna DN-måndag och DN-tisdag.

*Tabell 5 Medelvärde, standardavvikelse, median samt tolk för grupperna DN måndag och DN tisdag (förstasidor)*

	1=DN måndag	2=DN tisdag
Medelvärde	31.33333	31.33333
Standardavvikelse	1.861899	2.658320
Median	31.5	31
LIX-tolk	Lätt	Lätt

Såsom framgår av tabell 5 är medelvärden exakt desamma men spridningen skiljer sig mycket från varandra. Medianerna är nästan lika. Måndagens median avrundas dock till 32 (inte nödvändigt att ange den avrundade siffran här), så att det blir ett LIX-heltal. Både måndagens och tisdagens värdegrupper värderas som "lätta" enligt LIX-tolken. Figur 5 illustrerar spridningen av alla LIX-värden för måndagarna och tisdagarna som värdegrupper i DN med två boxplotdiagram. Respektive statistiska värden finns i tabell 5. LIX-intervallen i skalan i figur 5 är två enheter.



Figur 5 Spridningen av LIX-värden för förstasidesgrupperna DN måndag och DN tisdag

Figur 5 visar att spridningen i de båda lådorna ligger på samma LIX-värdeområde men är inte lika omfattande för båda fallen. Medianen i det högra diagrammets låda (tisdagen) är det näst minsta värdet förutom uteliggaren. Minimivärdet för DN-tisdag utgörs av uteliggaren nere i diagrammet. Medianen har exakt samma värde (31) som första kvartilen i diagrammet. Dessutom omfattar spridningen hos tisdagen en uteliggare, 27. Värdeområdet från 28 till 31 bildar alltså ett tomrum i diagrammet för tisdagen. Diagrammet för måndagen har inte någon övre stapel som maximivärde utan den tredje kvartilen i boxplottet fungerar på samma sätt som maximivärde. Enligt en hastig ögonbedömning kan spridningen verka vara lika stort i båda värdegrupperna på basis av att lådorna i båda diagrammen har samma höjd. Emellertid avslöjar en närmare granskning av figur 5 att det även finns ett extremt värde nere i tisdagens diagram som gör spridningen för tisdagen betydligt mer omfattande än för måndagen. Enligt 5 % signifikansnivå är skillnaden mellan förstasidesgrupperna DN-måndag och DN-söndag statistiskt *icke-signifikant* på basis av t-testets p-värde 1,000. Skillnaden klassificeras som *icke-signifikant* även enligt de fyra allmänna signifikansnivåerna (se även avsnitt 7.5.3).

#### 7.4.6 LIX-värden för förstasidesgrupperna DN-lördag och DN-söndag

I detta avsnitt ska två specialfall i DN:s veckodagjämförelser granskas. Jämförelsen mellan två värdegrupper DN-måndag och DN-tisdag ska ge information om hur LIX-värdena skiljer sig från varandra mellan olika veckor angående värdenas spridning.



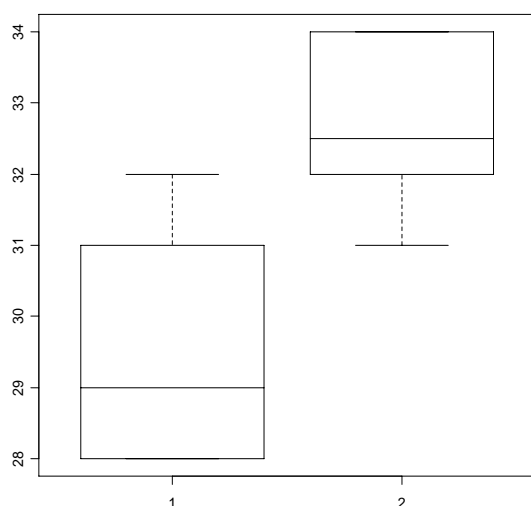
Av tabell 6 framgår LIX-medelvärdena, -spridningen, -medianen samt -värderingen enligt Björnssons LIX-tolk för förstasidesgrupperna DN-måndag och DN-tisdag.

*Tabell 6 Medelvärde, standardavvikelse, median samt tolk för grupperna DN lördag och DN söndag (förstasidor)*

	1=DN lördag	2=DN söndag
Medelvärde	29.50000	32.66667
Standardavvikelse	1.643168	1.211060
Median	29	32.5
LIX-tolk	Mycket lätt	Lätt

Som man kan se av tabell 6 finns det en dryg tre LIX-enheters skillnad mellan medelvärden för grupperna lördagen och söndagen. Mellan spridningsvärden finns ingen stor skillnad fast även medianerna ligger från 3,5 LIX-enheters distans från varandra. Eftersom värderingen görs på basis av medianvärdet (29) förblir den för lördagen ”mycket lätt”. Söndagen värderas däremot ”lätt” då det ligger väl över värde 30.

Figur 6 illustrerar spridningen av LIX-värden för förstasidesgrupperna DN-lördag och DN-söndag med två boxplotdiagram. Intervallen i skalan i figur 6 är en (1) LIX-enhet.



*Figur 6 Spridningen av LIX-värden för förstasidesgrupperna DN lördag och DN söndag*

Figur 6 visar att medelvärdet för alla förstasidor för DN, 31.04, ligger ovanför (om LIX-talet avrundas till 31: exakt på samma höjd med den tredje kvartilen) den tredje kvartilen (övre kanten på lådan) i lördagens boxplotdiagram, dvs. ca 75% ligger under 31.04. Söndagens hela

diagram, dvs. hela värdeområde ligger över medelvärdet. Medelvärdet för alla LIX-värden i DN-förstasidorna fungerar alltså som en slags vågbrytare mellan de två avskilda och invärtes homogena värdegrupperna. Lördagens diagram har inte någon stapel för sitt minimivärde utan den första kvartilen har denna funktion. I söndagens diagram finns däremot ingen stapel för maximivärdet utan den tredje kvartilen illustrerar det.

Visuellt bedömt är spridningen (sd) enligt boxplotten endast litet större för lördagen än för söndagen. Även tabellvärden ligger nära varandra. Enligt 5 % signifikansnivå är skillnaden mellan förstasidesgrupperna DN-lördag och DN-söndag statistiskt *signifikant* på grund av t-testets p-värde 0,004056. Enligt de fyra allmänna signifikansnivåerna är skillnaden *signifikant* (se även avsnitt 7.5.3).

#### 7.4.7 LIX-värden för förstasidorna i HBL för olika veckor

Detta avsnitt handlar om spridningen av det genomsnittliga LIX för sex olika veckor (sampelstorlek: 42 / 7 veckodagar) i HBL angående förstasidors värden som sex skilda värdegrupper.

Tabell 7 anger LIX-medelvärdena, -spridningen, -medianen samt -värderingen enligt Björnssons LIX-tolk för förstasidorna i DN som sex värdegrupper för olika veckor.

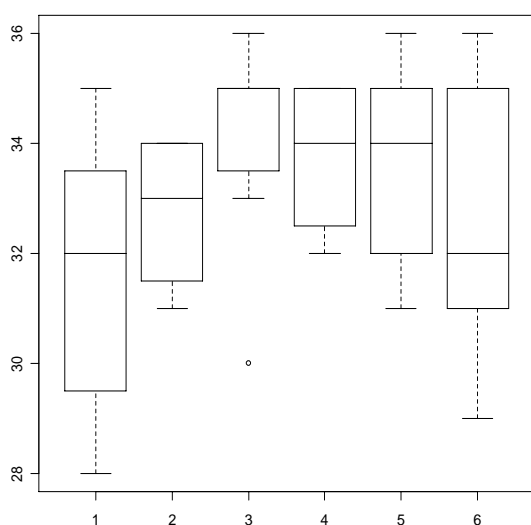
*Tabell 7 Medelvärde, standardavvikelse, median samt tolk för LIX i HBL för olika veckor*

	1=HBL vecka 1	2=HBL vecka 2	3=HBL vecka 3	4=HBL vecka 4
Medelvärde	31.57143	32.71429	34.00000	33.71429
Standardavvikelse	2.636737	1.380131	2.000000	1.380131
Median	32	33	35	34
LIX-tolk	Lätt	Lätt	Lätt	Lätt
	5=HBL vecka 5	6=HBL vecka 6		
Medelvärde	33.57143	32.71429		
Standardavvikelse	1.988060	2.690371		
Median	34	32		
LIX-tolk	Lätt	Lätt		

Som det framgår av tabell 7 ligger de aritmetiska medelvärden för alla sex veckorna ungefärligt på samma nivå, mellan ca 31–34. Spridningen håller sig ungefärligt mellan ca 1,4 och dryg 2,6 och på basis av spridningen (sd) finns det två avvikande värdegrupper (jfr

exempelvis med tabell 2, DN vecka 6) bland de sex veckorna – vecka 1 (ca 2,6) och vecka 6 (ca 2,7). Medianerna ligger med värden från 32–35 relativt nära varandra, dvs. inom 3 LIX-enheter (jfr även tabell 2, medianvärden). På basis av dessa relativt homogena medianvärden blir även LIX-värderingen enligt LIX-tolken densamma för alla veckor, dvs. ”lätt”.

I följande figur (7) illustreras spridningen av alla LIX-värden för förstasidorna i HBL för olika sex veckor med motsvarande boxplotdiagram. Respektive statistiska värden finns i tabell 7. LIX-intervallen i skalan i figur 7 är två enheter.



*Figur 7 Spridningen av LIX-värden för förstasidorna i HBL för olika veckor*

I jämförelse med figur 2 av DN-förstasidorna märker man att lådorna hos i HBL, dvs. i figur 7 ligger på en högre nivå när skalan proportioneras till värdeområdet 30–34 som omfattar de flesta LIX-värdena på hela skalan. Då utelämnas en betydligt mängd av värden ovanför värdet 34 och några värden under 30. På ett motsvarande sätt betar sig värdena på figur 2 där det inte finns mer än ett enskilt värde ovanför 34 och mycket stoff under 30. Det finns dock statistiska skillnader i jämförelsen mellan förstasidorna på DN och HBL (sampelstorlek 84 / 2). Figur 1 föreställer värdena och spridningen i förstasidorna i DN och HBL i genomsnitt och ger en allmän uppfattning om den generella LIX-skillnaden mellan förstasidorna i DN och HBL. Figureerna 2 och 7 avslöjar däremot den verkliga fördelningen och spridningen genom att två skilda helheter granskas som enstaka fall, dvs. värdegrupper för alla veckor för respektive tidningars förstasidor. Diagrammen för vecka 2 samt vecka 4 har inte några övre staplar som respektive maximivärden utan de tredje kvartilerna i dessa boxplot fungerar även som

maximivärden inom gruppernas värdeområden. Därtill har vecka 3 en uteliggare (LIX: 30) och den tredje kvartilen innehåller på samma sätt medianen i boxplottet (LIX: 35). Veckorna 1 och 6 har en mycket mer omfattande spridning än de övriga veckorna fast vecka 3 har en uteliggare som gör det svårt att säga om spridningen på grund av detta extrema värde blir lika stort som för veckorna 1 samt 6. Dessutom har veckorna 3 (som har en uteliggare) och 5 samma spridningsvärden, dvs. 1,988 och 2,000). Respektive p-värden, dvs. statistiska t-test behandlas och diskuteras i avsnitt 7.5.4 (se avsnitt 7.5.4).

#### 7.4.8 LIX-värden för förstasidorna i HBL för olika veckodagar

I det följande behandlas det genomsnittliga LIX för sju olika veckodagar (samplelstorlek: 42 / 6 veckor) i HBL angående förstasidors värden som sju skilda värdegrupper.

Tabell 8 ger vid handen LIX-medelvärdena, -spridningen, -medianen samt -värderingen enligt Björnssons LIX-tolk för förstasidorna i HBL som sju värdegrupper för olika veckodagar.

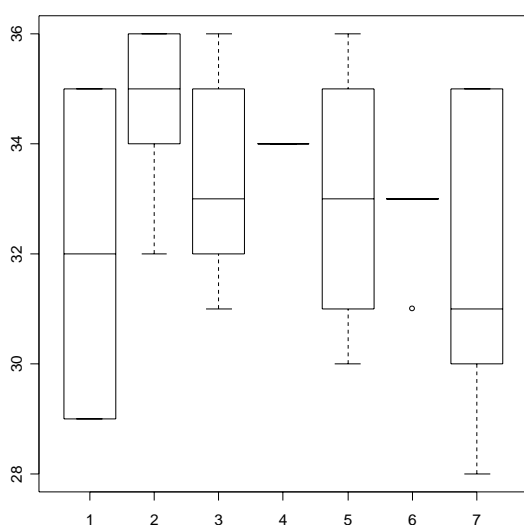
*Tabell 8 Medelvärde, standardavvikelse, median samt tolk för LIX i HBL för olika veckodagar (förstasidor)*

	1=HBL måndag	2=HBL tisdag	3=HBL onsdag	4=HBL torsdag
Medelvärde	32.00000	34.66667	33.33333	34.00000
Standardavvikelse	2.683282	1.505545	1.966384	0.000000
Median	32	35	33	34
LIX-tolk	Lätt	Lätt	Lätt	Lätt
	5=HBL fredag	6=HBL lördag	7=HBL söndag	
Medelvärde	33.00000	32.66667	31.66667	
Standardavvikelse	2.607681	0.8164966	2.804758	
Median	33	33	31	
LIX-tolk	Lätt	Lätt	Lätt	

Såsom man kan se av tabell 8 ligger de aritmetiska medelvärden för alla sex veckorna ungefärligt på samma nivå, mellan ca dryga 31–34. Spridningen varierar drastiskt mellan 0 (torsdagen) och 2,8 (söndagen). Torsdagen kan anses vara en klart avvikande enstaka värdegrupp vars spridning är noll (0). Detta betyder att alla värden inom torsdagens värdegrupp är 34, dvs. alla torsdagsupplagors förstasidor har detta värde (medelvärdet samt medianen) och då blir det naturligtvis inte någon spridning för värdena. Dessutom har lördagen ett mycket lågt spridningsvärde, ca 0,8. Även medianen är exakt samma som

medelvärde. Dessutom ligger även medianerna med värden från 31–35 relativt nära varandra, dvs. inom 4 LIX-enheter (jfr även tabell 2, medianvärden). I värdegrupperna HBL-måndag, HBL-torsdag samt HBL-fredag är medelvärdena och medianerna desamma inom respektive värdegrupper. På basis av relativt homogena medianvärden blir även LIX-värderingen enligt LIX-tolken densamma för alla veckor, dvs. ”lätt”.

I figur 8 illustreras spridningen av LIX-värden för förstasidorna i HBL för olika veckodagar med hjälp av motsvarande boxplotdiagram. Respektive statistiska värden finns i tabell 8 ovan. LIX-intervallen i skalan i figur 8 är två enheter.



Figur 8 Spridningen av LIX-värden för förstasidorna i HBL för olika veckodagar

Figur 8 visar att måndagens LIX-värden är fullständigt jämnt spridda. Både medianen och medelvärdet är exakt desamma (består av samma linje) i detta specialfall, dvs. 32. Dessutom ligger värdena i diagram 1 symmetriskt runt medelvärdet (29, 32, 35, 32, 35, 29). Hos torsdagen är alla sex (6) värden exakt desamma, dvs. Både medelvärdet och medianen är 34. Eftersom spridningen är noll (0), finns det inte någon låda i den ifrågavarande boxplotgrafan. Vid första ögonkastet verkar lördagen vara identisk med torsdagen men hos denna dag är medelvärdet inte 33 såsom dess median. Påverkan av det extrema värdet sänker medelvärdet till 32,66667. Dessutom är värdena inte helt symmetriska som värdegrupp (30, 31, 35, 35, 31, 36). Det är enbart tisdagens och torsdagens lådor som inte korsar tvärs medelvärdet för alla förstasidor i HBL, dvs. LIX-värdet 33,04 (inte synligt i figuren). Trots detta ligger de på grund av sina medelvärden samt medianer på samma nivå med de andra veckodagarna. Detta

ger anledning till att kunna konstatera att läsbarhetsvärdena för HBL:s olika veckodagar är relativt homogena. Diagrammet för måndagen saknar både den övre och den nedre stapeln och därför utgörs minimi- och maximivärden av den första/tredje kvartilen. Dessutom har diagrammen för tisdagen samt söndagen torsdagen inte några övre staplar utan de tredje kvartilerna i dessa boxplot fungerar även som maximivärde inom gruppernas värdeområden. De grafiskt märkvärdigaste och mest avvikande fallen är dock torsdagen och lördagen. Torsdagens diagram består av endast medianen som på samma sätt innehåller medelvärdet – de båda har alltså LIX-värdet 34. Lördagens boxplot är likadan som torsdagens förutom uteliggaren, dvs. det extrema värdet som utgör minimivärdet inom lördagens värdegrupp. Förutom denna lördags uteliggare (LIX: 31) har alla övriga lördags förstasidor inom lördagens värdegrupp samma värde, 33, och detta blir alltså medianvärdet. Efter ögonmått ser man genast av figur 4 att spridningen för torsdagens värdegrupp är noll. Lördagens spridning är något svårare att bedöma på basis av figuren fast uteliggaren syns tydligt under medianlinjen (LIX-värde: 33) mellan 30 och 32 på skalan. De övriga veckodagarna, dvs. värdegrupperna 1, 2, 3, 5 samt 7 beter sig relativt homogent emellan. Även respektive statistiska tabellvärden (i tabell 8) stöder denna uppfattning. Respektive t-testens p-värden behandlas och diskuteras i avsnitt 7.5.5 (se avsnitt 7.5.5).

#### 7.4.9 LIX-värden för HBL-insidorna

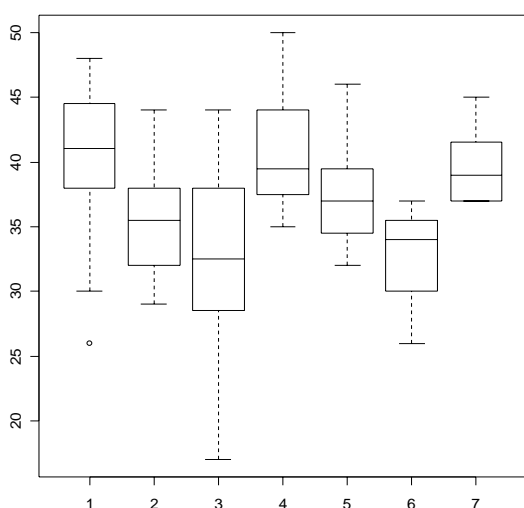
I detta sista avsnitt ska det genomsnittliga LIX för sju olika HBL-insidesavdelningar. I tabell 9 presenteras LIX-medelvärdena, -spridningen, -medianen samt -värderingen enligt Björnssons LIX-tolk för insidorna i DN som sju värdegrupper för olika avdelningar.

*Tabell 9 Medelvärde, standardavvikelse, median samt tolk för LIX för HBL-insidorna*

	1=HBL ekonomi	2=HBL inrikes	3=HBL kultur	4=HBL ledare
Medelvärde	39.91667	35.00000	32.33333	40.75000
Standardavvikelse	6.416716	4.231913	7.854742	4.535216
Median	41	35.5	32.5	39.5
LIX-tolk	Medelmåttig	Lätt	Lätt	Medelmåttig
	5=HBL politik	6=HBL sport	7=HBL utrikes	
Medelvärde	37.33333	32.75000	39.58333	
Standardavvikelse	3.845501	3.744693	2.539088	
Median	37	34	39	
LIX-tolk	Lätt	Lätt	Lätt	

Man kan se av tabell 9 att medelvärden för olika avdelningar varierar från dryga 32 (kultur) till 40 (ledare). Mest spridda är värdena för värdegrupp 3 (Kultur) – värdet är mycket högt, till och med 7,85. Näst störst är spridningen för insidesartiklarna i avdelning HBL-ekonomi, ca 6,4. Den mest homogena värdegruppen enligt spridningen är HBL-utrikes med sitt låga värde, ca 2,5. Värdegruppernas 2, 4, 5, 6 samt 7 spridning ligger relativt stadigt mellan ca 2,5–4,2 i jämförelse med spridningsvärden för grupperna 1 och 3. Medianvärden ligger mellan 32,5 (avrundas dock till 33) och 41. Det enda fallet som klart avviker från de övriga fallen är värdegruppen HBL-ekonomi med sin mycket höga medianvärde, 41. Detta är dock inte det enda fallet där LIX-värderingen enligt LIX-tolken blir även ”medelmåttig” medan majoriteten noteras som ”lätta”. Också värdegrupp 4 (HBL-ledare) värderas enligt tolken som ”medelmåttig” eftersom dess medianvärde avrundas från 39,5 till 40. Då värdet 39,5 är ett gränsfall (det enda i min undersökning) som avrundas uppåt till närmaste fullt tiotal förändras även värderingen för gruppen uppåt på samma sätt.

I följande figur (9) illustreras spridningen av alla LIX-värden för insidorna i HBL för olika sex veckor med motsvarande boxplotdiagram. Respektive statistiska värden finns i tabell 9. LIX-intervallen i skalan i figur 9 är fem enheter.



*Figur 9 Spridningen av LIX-värden för HBL-insidorna*

Det framgår av figur 9 att i tre av de sju diagrammen korsar medelvärdet (icke synlig i figuren) för alla insidor (36,88) respektive lådor. Detta betyder att värdeområdet i fyra fall ligger antingen ovanför eller under medelvärdelinjen vid 36,88. Således ligger Inrikes (2),

Kultur (3) samt Politik (5) genomsnittligt i närheten av medelvärdet (36,88). Lådan i diagram sex (Sport) ligger i sin helhet under den ovannämnda medelvärdelinjen, dvs. åtminstone 50 % av värden för HBL-sport ligger under 36,88. Resten av värden, dvs. De tre avdelningarna som återstår (1, 4 och 7) ligger ovanför samma medelvärde. Alla sju diagram i figuren ligger dock delvis vid medelvärdeområdet. Detta gör det svårt att bedöma samplet antingen jämnt eller ojämnt. Grafiskt sett innehåller figur 9 mestadels standardfall (dvs. där diagrammets alla delar är synliga). Diagram 2–6 har båda staplar (övre/nedre) som anger minimi- samt maximivärden. Den längsta stapeln som tydligt avviker från de övriga staplarna är den nedre stapeln i diagram 3, HBL-kultur. Detta betyder att grupp 3 (Kultur) har mer än ett lågt värde. Emellertid finns det två fall vars boxplot skiljer sig från resten av fallen i figur 9. Diagrammet för värdegrupp HBL-ekonomi har även ett extremt värde (LIX: 26) och det finns inte någon nedre stapel i diagram 7. Då är det – såsom även i några tidigare fall – den första kvartilen som presenterar minimivärdet för respektive värdegrupp.

Efter ögonmått är spridningen störst för värdegrupp 3 (Kultur) medan 6 (Sport) samt 7 (Utrikes) verkar ha en liten spridning på basis av sina kompakta boxplot, dvs. små lådor och inte några långa staplar. De övriga boxplotten (1, 2, 4 samt 5) är ungefärligt lika stora. Om man tolkar spridningen för HBL-insidor antingen på grund av boxplotdiagram i figur 9 eller med hjälp av tabellvärden kan man konstatera att spridningsvärden överhuvudtaget är mer omfattande än i de övriga fallen/figureerna (DN/HBL-förstasidor). Respektive t-testens p-värden behandlas i avsnitt 7.5.6 (se avsnitt 7.5.6).

### 7.5 Statistisk signifikans – signifikansresultat för alla jämförelsegrupper

Såsom konstaterats i början av detta kapitel samt bland beskrivningarna bland de nio presenterade fallen som handlar om en mer omfattande jämförelse än en parvis jämförelse av två värdegrupper (som exempelvis i avsnitt 7.4.1). De avsnitt i vilka det redogörs för parvisa fall har även statistiska uppgifter (enligt både 5 % och de allmänna nivåerna) som sista delen av respektive beskrivning, dvs. efter diskussion om spridningsvärden. De fallen med mer än två värdegrupper och respektive antal boxplotdiagram, dvs. 7.4.2, 7.4.4, 7.4.7, 7.4.8 samt 7.4.9 ska behandlas i detta avsnitt angående statistiska uppgifter. För klarhetens skull anges respektive statistiska informationen även om de parvisa fallen som redan redogjorts för i avsnitten 7.4.1, 7.4.3, 7.4.5 samt 7.4.6. på nytt. De resultaten som inte är illustrerade och



beskrivna med figurer och tabeller (DN vs HBL veckor/veckodagar) kommer att behandlas i detta avsnitt. Alla nödvändiga t-test ska alltså behandlas som delar av sin respektive Excel-jämförelsegrupp. På basis av detta hänvisas också till det bifogade Excel-materialet där t-testen är tabellerade och dessutom de viktigaste fallen (de mest signifikanta) markerade med fetstil.

#### 7.5.1 Alla förstasidor i DN och HBL

Enligt p-värdet (respektive t-test i bilaga 5, nedan) och de fyra allmänna signifikansnivåerna klassificeras skillnaden *mycket signifikant*. För detta första fall finns inte någon jämförelsegrupp (dvs. t-test-tabell) då jämförelsen består endast av ett enskilt parvist t-test. Enligt det ovanstående resultaten skiljer DN:s och HBL:s som två olika tidningar, dvs. som två värdegrupper, fast skillnaden mellan medelvärden 31 (DN) och 33 (HBL) är bara två LIX-enheter. Enligt 5 % signifikansnivå är skillnaden mellan alla förstasidor i båda tidningarna statistiskt *signifikant* då t-testets p-värde är 0,00005916 (se avsnitt 7.4.1 samt även bilaga 1).

#### 7.5.2 Förstasidor: DN – veckor

För att få en allmän bild av denna jämförelsegrupp ska resultaten för respektive t-test (jämförelsepar) i bilaga 7 (den övre tabellen) granskas. Enligt den ifrågasvarande utvärderingstabellen (bilaga 7), dvs. de fyra allmänna signifikansnivåerna klassificeras skillnaden som *nästan signifikant*. Som man ser av tabellen finns det bara ett statistiskt signifikant jämförelsepar. Alla övriga parvisa t-test i tabellen ger resultatet ”icke-signifikant”. Jämförelseparet DN-vecka 3 vs DN-vecka 4 har 1,000 som p-värde vilket betyder att LIX-medelvärden för dessa två grupper är exakt desamma och motsvarande t-test-resultat *icke-signifikant*. Enligt 5 % signifikansnivå är skillnaden mellan förstasidesgrupperna DN-vecka 2 och DN-vecka 5 (illustrerad i figur 7.4.3) statistiskt *signifikant* på grund av t-testets p-värde 0,0259. Som procentandel gör 1 statistiskt signifikant resultat av alla 15 parvisa t-test (enligt 5 % nivå) 7 procent av resultaten i jämförelsegruppen statistiskt signifikanta (se avsnitten 7.4.2 och 7.4.3).

### 7.5.3 Förstasidor: DN – veckodagar

Man får en allmän bild av följande jämförelsegrupp genom att se på respektive t-test i bilaga 4 (den övre tabellen). Som man ser av tabellen finns det fyra (klassificeringen i tabellen är dock enligt de fyra allmänna nivåerna!) statistiskt signifikanta jämförelsepar i alla 21 t-test. Enligt den ifrågavarande utvärderingstabellen (bilaga 4), dvs. de fyra allmänna signifikansnivåerna klassificeras skillnaderna som *signifikant* (1 stycke, illustrerad i figur 7.4.6) och *nästan signifikant* (3 stycken). Skillnaden i alla övriga parvisa t-test i jämförelsegruppen ger resultatet ”icke-signifikant”. Jämförelseparet DN–måndag vs DN–tisdag har 1,000 som p-värde vilket betyder att LIX-medelvärden för dessa två grupper är exakt desamma och t-testresultatet alltså *icke-signifikant* (illustrerad i figur 7.4.5). Enligt 5 % signifikansnivå är skillnaden mellan dessa fyra fetstilmarkerade fall i bilaga 4 statistiskt *signifikant* på grund av respektive p-värden. Som procentandel gör 4 statistiskt signifikanta resultat av alla 21 parvisa t-test (enligt 5 % nivå) 19 procent av resultaten i jämförelsegruppen statistiskt signifikanta (se avsnitten 7.4.4–7.4.6).

### 7.5.4 Förstasidor: HBL – veckor

För att få en uppfattning om följande information (om jämförelseparen för HBL-veckor) ska resultaten för respektive t-test i bilaga 7 (den nedre tabellen) granskas. Enligt den ifrågavarande utvärderingstabellen (bilaga 7), dvs. de fyra allmänna signifikansnivåerna, klassificeras alla skillnader i 15 av 15 jämförelseparen som *icke-signifikanta*. Jämförelseparet HBL–vecka 2 vs HBL–vecka 6 har 1,000 som p-värde vilket betyder att LIX-medelvärden för dessa två grupper är exakt desamma. Dessutom blir skillnaden enligt 5 % naturligtvis också *icke-signifikant* för alla jämförda par samt procentandelen statistiskt signifikanta resultat 0 procent. LIX-spridningen för hela jämförelsegruppen är illustrerad i figur 7.4.7 (se avsnitt 7.4.7).

### 7.5.5 Förstasidor: HBL – veckodagar

Man får en allmän bild av följande jämförelsegrupp genom att se på respektive t-test i bilaga 4 (den nedre tabellen). Som man ser av tabellen finns det två statistiskt signifikanta

jämförelsepar i alla 21 parvisa t-test. Enligt den ifrågavarande utvärderingstabellen (bilaga 4), dvs. de fyra allmänna signifikansnivåerna, klassificeras de två fetstilmarkerade skillnaderna som *nästan signifikanta*. Alla övriga parvisa t-test ger resultatet ”icke-signifikant”. Enligt 5 % signifikansnivå är skillnaden i dessa fyra fetstilmarkerade fall i bilaga 4 statistiskt *signifikant* på grund av respektive p-värden. Som procentandel gör 2 statistiskt signifikanta resultat av alla 21 parvisa t-test (enligt 5 % nivå) 10 procent av resultaten i jämförelsegruppen statistiskt signifikanta. LIX-spridningen för hela jämförelsegruppen är illustrerad i avsnitt 7.4.8 (se avsnitt 7.4.8).

#### 7.5.6 Insidor: HBL

Man får en uppfattning om följande information (om jämförelseparen för HBL-insidor) genom att man granskar resultaten för respektive t-test i bilaga 3. Man får också en allmän bild om LIX-värdena för alla HBL-insidor om man tar en titt på bilaga 2 där alla värden för HBL-insidesmaterial finns tabellerade enligt olika avdelningar, dvs. *Ekonomi, Inrikes*, osv. Det framgår av utvärderingstabellen i bilaga 3 att det finns 9 statistiskt signifikanta jämförelsepar av alla 21 parvisa t-test. Klassificeringen i tabellen – såsom i alla mina Excel-tabeller – baserar sig dock på de fyra allmänna signifikansnivåerna. På basis av dem klassificeras de fetstilmarkerade skillnaderna i bilaga 3 som *mycket signifikanta* (2 st.), *signifikanta* (5 st.) samt *nästan signifikanta* (2 st.). Skillnaden i alla övriga parvisa t-test i jämförelsegruppen ger resultatet ”icke-signifikant”. Enligt 5 % signifikansnivå är skillnaden i dessa nio fetstilmarkerade fall i bilaga 3 statistiskt *signifikant* på grund av respektive p-värden. Som procentandel gör 9 statistiskt signifikanta resultat av alla 21 parvisa t-test (enligt 5 % nivå) 43 procent av resultaten i jämförelsegruppen statistiskt signifikanta (se avsnitt 7.4.9).

#### 7.5.7 Förstasidor: DN vs HBL – veckor

För att få en allmän bild av denna jämförelsegrupp ska resultaten för respektive t-test (jämförelsepar) i tabellen i bilaga 6 granskas. Såsom man ser av bilaga 6 finns det 13 statistiskt signifikanta jämförelsepar i alla 36 parvisa t-test. Klassificeringarna för de fetstilmarkerade (dvs. signifikanta) skillnaderna i bilaga 6 är: *signifikanta* (3 st.) samt *nästan*

*signifikanta* (10 st.). Skillnaden i alla övriga parvisa t-test i jämförelsegruppen ger resultatet ”icke-signifikant”. Enligt 5 % signifikansnivå är skillnaden i dessa 13 markerade fall statistiskt *signifikant* på grund av respektive p-värden. Som procentandel gör 13 statistiskt signifikanta resultat av alla 36 parvisa t-test (enligt 5 % nivå) 36 procent av resultaten i jämförelsegruppen statistiskt signifikanta. Denna jämförelsegrupp är inte illustrerad i avsnitt 7.4.x på grund av dess komplexitet.

#### 7.5.8 Förstasidor: DN vs HBL – veckodagar

Att man ska få en allmän bild av denna jämförelsegrupp ska man se på resultaten för respektive t-test i tabellen i bilaga 5. Som man kan se av bilaga 5 finns det 19 statistiskt signifikanta jämförelsepar av alla 49 parvisa t-test. Klassificeringarna för de fetstilmarkerade (dvs. signifikanta) skillnaderna i bilagan är: *mycket signifikanta* (2 st.), *signifikanta* (6 st.) samt *nästan signifikanta* (11 st.). Skillnaden i alla övriga parvisa t-test i jämförelsegruppen ger resultatet ”icke-signifikant”. Jämförelseparen DN–fredag vs HBL–måndag samt DN–söndag vs HBL–lördag har båda 1,000 som provningsresultat (p-värde) vilket betyder att LIX-medelvärden för dessa två grupper i båda fallen är exakt desamma och motsvarande t-test–resultat absolut *icke-signifikanta*. Enligt 5 % signifikansnivå är skillnaden i dessa 19 markerade fall statistiskt *signifikant* på grund av respektive p-värden. Som procentandel gör 19 statistiskt signifikanta resultat av alla 49 parvisa t-test (enligt 5 % nivå) 39 procent av resultaten i jämförelsegruppen statistiskt signifikanta. Denna jämförelsegrupp är inte (såsom i 7.5.7: DN vs HBL – veckor) illustrerad i avsnitt 7.4.x på grund av dess komplexitet.

#### 7.6 Allmänna LIX-resultat

Om resultaten granskas på en allmännare nivå kan några upprepande/allmänna (inom denna undersökning till någon mån generaliserbara) tendenser bland mina resultat hittas. På basis av att samplet är litet i min undersökning kan dessa slags skillnader mellan förstasidor (exempelvis lördag vs söndag) dock bero på en ren slump även om de beräknade resultaten skulle visa att det finns en statistiskt signifikant skillnad mellan de jämförda grupperna. Emellertid är det intressant att det finns några skillnader som upprepas på ett konsekvent sätt.

### 7.6.1 Förstasidesresultat

När det till exempel gäller LIX-värden för förstasidorna i DN skiljer lördagens värden sig i regel tydligt från söndagens värden. LIX för lördagar på alla sex veckor är lägre än för söndagar (se bilaga 1: värden för DN-förstasidor och bilaga 4, den övre tabellen).

En företeelse som förekommer bland förstasidesmaterialet i Hbl påvisar att alla sex torsdagar har samma LIX-värde, 34. Därtill skiljer torsdagarnas värden från söndagarnas inom båda tidningarna. Inom DN är alla torsdagsvärden mindre än söndagsvärden (se bilaga 1: värden för både DN/HBL-förstasidor).

Om LIX-värdena för DN- samt HBL-förstasidor bedöms som helheter på Björnssons LIX-skala från 20 till 60 kan man konstatera att värdena för båda tidningarna varierar på ett mycket snävt LIX-värdeområde. Hos tidningen DN varierar värdena mellan 27 och 35 – alltså högst åtta (8) LIX-enheter. För HBL ligger det motsvarande området också mellan maximalt 8 enheter, från 28 till 36 och LIX-värderingen enligt Björnssons LIX-tolk dvs. från ”mycket lätt” till ”lätt”. På basis av dessa snäva LIX-områden inom värden för båda tidningarna kan variationen av LIX-värdena sägas vara mycket homogent såväl inom de två avskilda tidningarna som mellan tidningarna som två LIX-värdegrupper (se bilaga 1).

### 7.6.2 HBL-insidesresultat

LIX-genomsnittet för insidesavdelningarna *Sport* och *Kultur* avskiljer sig från de övriga avdelningarna, dvs. *Ledare*, *Utrikes*, *Ekonomi*, *Inrikes* och *Politik*. Avdelningen *Kultur* har det lägsta LIX-medelvärdet, 32. Dessutom ligger *Sport* mycket nära med värdet 33. På ett motsvarande sätt har *Ledare* den högsta medelvärdet, 41. Därtill ligger både *Ekonomi* och *Utrikes* relativt högt på LIX-skalan med värdet 40. LIX-variationen som gäller avskilda artiklar bland alla sju avdelningar är mycket mera omfattande än för förstasidesresultaten. LIX för HBL-insidesavdelningar varierar mellan det ytterst låga 17 som finns i LIX-värdegruppen *Kultur* och höga 50 i gruppen *Ledare* (LIX-mellanrum t.o.m. 33 enheter). Vid en invärtes granskning av LIX-värdena för olika insidesavdelningar kan man konstatera att de är mycket heterogena speciellt i jämförelse med variationen för förstasidesresultaten i det föregående avsnittet. Då blir även LIX-värderingen för olika avdelningar varierande. Nästan

hela LIX-skalan (förutom värdeområdet 51 till 60) måste tas i bruk då det mycket varierande LIX ska värderas enligt LIX-tolken. Det minsta värdet (LIX: 17) förblir under gränsen för ”mycket lätt” medan det högsta (LIX: 50) värderas som ”svår” (se bilaga 2, LIX-tabellen).

### 7.6.3 Jämförelse med andra undersökningsresultat

I Marjakangas och Vertanens (2003) pro-gradu avhandling forskas i två lättlästa tidningar 8 SIDOR (avsedd för utvecklighetsstörda och personer med lässvårigheter), SESAM (för invandrare och personer med svaga svenskkunskaper) samt *Dagens Nyheter*. Som framgår av tabell 10 baserar sig jämförelsen mellan de två lättlästa tidningarna och DN – såsom materialet också i min undersökning – på ett relativt litet sampel, dvs. antalet analyserade texter. (Marjakangas & Vertanen 2003, 6; 44; 46.)

I följande tabell presenteras/jämförs antalet texter per upplaga samt ordlängd som bokstäver per ord i både Marjakangas och Vertanens (2003, 44) undersökningsresultat för tidningarna DN, 8 SIDOR och SESAM och nedan dessa de motsvarande mättningsresultat i denna undersökning.

*Tabell 10 Antalet texter per tidning och ordlängden som medelvärde (B/O) för DN, 8 SIDOR och SESAM (Marjakangas & Vertanen 2003) samt mina resultat för förstasidor (DN & HBL) samt insidor (HBL)*

	Antalet texter per tidning	Bokstäver per ord (B/O)
DN	49	5,4
8 SIDOR	66	4,6
SESAM	56	5,0
DN (förstasidor)	42	5,1
HBL (förstasidor)	42	5,2
HBL (insidor)	84	5,2

Som man kan se av tabell 10 är antalet texter någorlunda lika omfattande. Det totala antalet texter i Marjakangas och Vertanens pro-gradu är 171 (Marjakangas & Vertanen 2003, 36). I min undersökning är det motsvarande antalet 168. Att antalen ligger så pass nära varandra är en ren slump men gör dock samplen och också resultaten mycket väl jämförbara. Enligt Marjakangas och Vertanens (2003, 44) undersökningsresultat är den genomsnittliga ordlängden för DN 5,4. I mina Grammatifix-mättningsresultat är den respektive siffran något lägre, 5,1. Detta är på grund av att Marjakangas och Vertanens (2003, 44) siffra (5,4) baserar

sig på DN-insidesmätningar medan siffran 5,1 i min undersökning är den genomsnittliga ordlängden för förstasidesmaterialet i DN. På basis av denna jämförelse skulle man kunna anta att texten i allmänhet på förstasidorna i DN (mina forskningsresultat) består av något kortare ord än på insidorna (Marjakangas och Vertanen). I sin tur syns denna skillnad även på respektive LIX-värden i tabell 11. Även värdet för HBL-insidor (mitt resultat) är lägre än respektive värdet för DN-insidestext (Marjakangas & Vertanen 2003, 44). Man bör dock lägga märke till att mätmetoden för LIX är annorlunda i min vs Marjakangas och Vertanens avhandling. Både 8 SIDOR (4,6) och SESAM (5,0) förblir under mina resultat för bokstäver per ord (B/O).

Följande tabell anger minimi-, maximi samt medelvärdet för LIX i både Marjakangas och Vertanens (2003, 46) undersökningsresultat för tidningarna DN, 8 SIDOR och SESAM och nedan dessa de motsvarande mätningsresultat i min/denna undersökning.

*Tabell 11 Minimi- och maximivärden samt LIX- medelvärden för DN, 8 SIDOR och SESAM (Marjakangas & Vertanen 2003) samt mina resultat för förstasidor (DN & HBL) och insidor (HBL)*

	Minimivärde	Maximivärde	LIX-medelvärde
DN	13	60	43
8 SIDOR	12	45	28
SESAM	18	60	35
DN (förstasidor)	27	35	31
HBL (förstasidor)	28	36	33
HBL (insidor)	17	50	37

Som det framgår av tabell 11 ligger minimivärden för LIX i mina undersökningsresultat (DN-förstasidor: 27 och HBL-förstasidor: 28) förutom HBL-insidor (LIX: 17) på en högre nivå än för DN-, 8 SIDOR- samt SESAM-insidor (Marjakangas & Vertanen 2003, 46). När det däremot gäller respektive maximivärden ligger alla av dem på en tydligt högre nivå för Marjakangas och Vertanens resultat (DN-, 8 SIDOR- samt SESAM-insidor). Endast maximivärdet bland HBL-insidesmaterialet når upp till 50. På basis av jämförelserna ovan kan man anta att LIX-variationen för insidestext är mycket mer omfattande än för förstasidestext. Värderingen enligt Björnssons LIX-tolk (se även avsnitt 3.2.5) för insidestext (DN, 8 SIDOR och SESAM samt HBL-insidor) varierar mellan ”mycket lätt” och ”mycket svår”. I min undersökning är den högsta värderingen dock ”svår” (LIX: 50, HBL-insidor). Beträffande värderingen av LIX-medelvärden enligt LIX-tolken ligger endast 8 SIDOR (Marjakangas & Vertanen 2003, 46) i kategorien för ”mycket lätt”, dvs. under gränsvärdet,

för ”lätt” , 30. Alla övriga, dvs. alla mina förstasides- samt insidesresultat (LIX: 31, 33 samt 37) och Marjakangas och Vertanens SESAM (LIX: 35) värderas som ”lätta”. Endast DN-insidestext på basis av LIX-medelvärdet 43 värderas som ”medelmåttig” (Marjakangas & Vertanen 2003, 46).

Strands (1984, 25) medelvärde för läsbahetsindexet i *Dagens Nyheter*-morgontidningsreportage, 43, vilket kan jämföras med Marjakangas och Vertanens (2003, 46) resultat (exakt detsamma: 43) och mina LIX-resultat för insidesmaterialet i HBL. Strands medelvärde verkar dock ganska högt i jämförelse med det totala beräknade medelvärdet för LIX, 37, för alla HBL:s avdelningar. Denna skillnad på 6 LIX-enheter hos medelvärdena avslöjar dock inte det faktum att det också finns höga värden fast medelvärdet ligger relativt lågt. Det högsta värdet i mitt HBL-insidesmaterial var till och med 50. Dessutom finns det flera värden långt över 40 (se bilaga 2).

Mina LIX-mätningar med Grammatifix på förstasidesmaterialet indikerar stadigt relativt låga LIX-medelvärden (DN-förstasidor: 31) jämfört med tidigare undersökningsresultat (DN) av exempelvis Strand (1984, 25). Hans mättningsresultat för den genomsnittliga ordlängden (dvs. bokstäver per ord) för morgontidningsreportage (DN-insidestext), är dock detsamma som mitt medelvärde för HBL-insidor, 5,2 (se även tabell 10 samt bilaga 2).

Tyvärr kan inte några absoluta eller utanför denna undersökning generaliserbara slutsatser dras om den generella LIX-beräknade läsbarheten hos morgontidningarnas innehåll (förstasidor och insidor) antingen på grund av mina och/eller Marjakangas och Vertanens (2003) forskningsresultat. Att kunna generalisera LIX-medelvärden och sedan dra några allmängiltiga slutsatser om den genomsnittliga läsbarheten (alltså LIX-medelvärden) i DN och HBL skulle kräva en omfattande analys – exempelvis en stickprovundersökning – av (webb-) morgontidningar utförd på ett relativt stort sampel (betydligt större än i både min och Marjakangas & Vertanens 2003 undersökningar).

#### 7.6.4 Slutsatser

För att kunna pröva de tre ursprungliga huvudhypoteserna samt forskningsproblemen (som nämnts i avsnitt 6.3) i denna undersökning har jag räknat fram de genomsnittliga LIX-värden,



standardavvikelsen (värdenas spridningsmått) och medianen för det utvalda webbtidningsmaterialet (förstasidorna: DN/HBL och insidorna: HBL). Därtill har materialet blivit jämfört på ett omfattande sätt som jämförelsegrupper (se avsnitten 7.4.x samt bilagorna 3–7).

På basis av mina forskningsresultat kan jag (inom denna undersökning) konstatera följande:

1. Det finns jämförelsegruppvisa signifikanta skillnader mellan förstasidorna för olika veckodagar eller veckor inom och mellan tidningarna DN och HBL (även mycket signifikanta). Det procentuella antalet statistiskt signifikanta resultat inom respektive värdegrupper blir emellertid oftast lågt. På grund av detta faktum kan man enligt min första forskningshypotes konstatera att det finns någon (statistisk) signifikant skillnad i den allmänna (jämförelsegruppvisa) läsbarheten varken mellan olika veckodagar eller veckor på förstasidorna i DN och HBL när läsbarheten i tidningarna jämförs inom och mellan respektive tidningar. Om det istället fanns åtminstone 50 % signifikanta resultat i flera jämförelsegrupper skulle man eventuellt kunna förkasta min hypotes, dvs. konstatera då att det finns skillnader i den allmänna läsbarheten mellan de ovannämnda (se bilagorna 4–7).
2. När det gäller förstasidorna som två hela värdegrupper (DN vs HBL) kan man på basis av signifikansprövningen konstatera att det finns en statistiskt mycket signifikant (enligt 5 %: signifikant) mellan dem. Man kan konstatera att det finns skillnader i den allmänna läsbarheten mellan förstasidorna i DN och HBL fast skillnaden mellan medelvärdena för dessa värdegrupper är bara 2 LIX-enheter. Också min andra forskningshypotes ska sålunda behållas (se avsnitt 6.3 samt bilaga 5, nedanför tabellen).
3. Beträffande HBL-insidesmaterialet kan man konstatera att nästan hälften av skillnaderna (43%) inom jämförelsegruppen är statistiskt signifikanta. Trots detta kan man inte säga att det i allmänhet finns statistiskt signifikanta skillnader i läsbarheten inom insidorna mellan olika avdelningar i HBL (se bilaga 3). Eftersom min tredje ursprungliga forskningshypotes går ut på att det finns några skillnader mellan läsbarheten inom insidorna mellan olika avdelningar (t.ex. *Inrikes* vs *Sport*) i HBL kan hypotesen behållas. Om det däremot handlar om den allmänna (inom jämförelsegruppen i bilaga 3) läsbarheten ska hypotesen förkastas.

## 8 SAMMANFATTNING

Webbtidningstext fungerar på basen av World Wide Web och enligt de grundläggande principerna på Internet, dvs. hypertext, (hyper-)länkar samt cirkulärt – dock också lineärt – läsande. En webbtidning fungerar på en komplicerad bas som bildas genom att använda en webbsida med hypertext och icke-lineära komponenter istället för tryckt text på tidningspapper. Då det var fråga om att analysera tidningstexter på basis av en lämplig teori satte jag avhandlingsprocessen igång med att jag definierade och diskuterade tidningstext samt nyhetsartikel som begrepp varefter två olika forskningsmetoder för läsbarhet presenterades och metoden i denna undersökning fokuserades på den kvantitativa läsbarhetsteorin av Björnsson, dvs. textens läsbarhetsindex, LIX, dess mätning och värdering av svårighetsgraden med LIX-tolken.

Huvudsyftet med denna undersökning var att klarlägga all signifikant kalkylerbar information och kunskaper om möjliga företeelser bakom de läsbarhetsresultat som kommer fram i de anpassade Grammatifix-mätningarna (LIX-beräkning) samt signifikansjämförelserna (signifikansprövning) för det utvalda webbtextmaterialet.

Som undersökningmaterial användes sammanlagt 168 webbtidningstexter ur tidningarna Dagens Nyheter och Hufvudstadsbladet, varav 84 (42 x 2) var förstasidor och resten bestod av 84 HBL-insidesartiklar. Materialet, dvs. texterna, konverterades till Word-form, LIX beräknades samt tabellerades i Excel och slutligen prövades resultaten med lämpliga datorprogram. Därefter presenterades de slutliga resultaten med respektive statistiska uppgifter samt figurer.

Avsikten var att utreda och att ge en bild av den nutida svenskspråkiga webbtidningstextens genomsnittliga kvantifierbara läsbarhet genom att utföra en jämförelseanalys av förstasidor och några insidor. Dessutom avsåg jag att pröva den statistiska signifikansen mellan de beräknade LIX-skillnaderna för att få fram valida och reliabla resultat samt för att kunna besvara mina forskningshypoteser och –frågor på grund av de erhållna resultaten. Fast resultaten granskades relativt mångsidigt och även några upprepande tendenser bland de allmänna LIX-resultaten kunde påpekas kan resultaten inte generaliseras till att omfatta förstasidestext (DN/HBL) eller insidestext (HBL).

Enligt min första forskningshypotes kan man konstatera att det fanns någon (statistisk) signifikant skillnad i den allmänna (jämförelsegruppvisa) läsbarheten varken mellan olika *veckodagar* eller *veckor* på förstasidorna i DN och HBL när läsbarheten i tidningarna jämförs inom och mellan respektive tidningar. När det gällde förstasidorna som två hela värdegrupper (DN vs HBL) kunde man på basis av signifikansprövningen konstatera att det fanns en statistiskt mycket signifikant mellan dem. Min andra forskningshypotes bör sålunda också behållas.

I största allmänhet kunde även min tredje forskningshypotes behållas. Eftersom den gick ut på att det finns några skillnader mellan läsbarheten inom insidorna mellan olika avdelningar (t.ex. *Inrikes* vs *Sport*). Fast nästan hälften av skillnaderna (43 %) inom jämförelsegruppen för HBL-insidestexterna (84 st.) var statistiskt signifikanta kan man inte generalisera att det i allmänhet (inte ens inom denna undersökning) finns statistiskt signifikanta skillnader i läsbarheten inom insidorna mellan olika avdelningar i HBL. För att man överhuvudtaget skulle kunna generalisera LIX-medelvärden och sedan dra några allmängiltiga slutsatser om den genomsnittliga läsbarheten i DN och HBL skulle kräva en mer omfattande analys – exempelvis en stickprovundersökning av (webb-) morgontidningar med ett relativt stort sampel. Exempelvis bygger både min och Marjakangas och Vertanens undersökning (2003) båda på sampel på ca 170 texter. När det gäller insidesundersökning skulle också en jämförande analys av insidesavdelningar (t.ex. som stickprov) mellan tre eller flera tidningar utföras för att kunna få fram vidare generaliserbara LIX-resultat.

I största allmänhet finns det emellertid flera intressanta faktorer som påverkar tidningstexternas språkliga, strukturella och typografiska tillgänglighet. Alla dessa faktorer kunde dock inte behandlas eller analyseras i denna avhandling utan jag först och främst fokuserade på att mäta den kvantifierbara läsbarheten hos webbtidningstext samt utreda eventuella statistiska skillnader mellan de analyserade LIX-resultaten.

Denna undersökning syftade till att först definiera tidningstext som medium, redogöra för problematiken bakom läsbarhetsundersökningar (kvalitativ och kvantitativ) samt definiera och beskriva den centralaste teorin och dess tillämpning på läsbarhetsmätning – i synnerhet i denna undersökning. Därtill behandlades webben och dess grundläggande struktur som bas för webbtidningar och den statistisk-matematiska teoretiska bakgrunden och dess tillämpning

på signifikansprövning i denna avhandling utreddes. På grund av dessa teorier utfördes LIX-analyserna och slutligen statistiska signifikansprövningarna på de erhållna LIX-resultaten.

## LITTERATUR

### PRIMÄRLITTERATUR

Förstasidesmaterialet (84 st.) som webbupplagor: Dagens Nyheter och Hufvudstadsbladet (7.2.– 20.3.2005)

Insidesmaterialet som artiklar (84 st.) i webbupplagor: Hufvudstadsbladet (5.10.2004, 8.12.2004, 14.12.2004, 4.1.2005, 19.3.2005)

### SEKUNDÄRLITTERATUR

Befring, E. (1994). *Forskningsmetodik och statistik*. Studentlitteratur: Lund.

Björnsson, C.H. (1968). *Läsbarhet*. Stockholm: Liber.

Butler, C. (1985). *Statistics in Linguistics*. Oxford: Basil Blackwell.

Dahlqvist, B. (1968). Möten med mottagare. I: Broman, S (red.) *Vårt behov av press, radio och TV. Reflexioner kring en undersökning*. Stockholm, 89–99.

Englund, H. & Guldbrand, K. Personliga och läsarvänliga webbtexter. I: *Språkbruk* 1/2002, 5-7.

Englund, H. & Guldbrand, K. (2001) *Klarspråk på nätet*, andra upplagan., Sundbyberg, Pagina Förlags AB.

Heimbürger, A., Alkula, R. & Kuhanen, T. (1990). *Hyperteksti ja hypermedia*. Valtion teknillinen tutkimuskeskus, tiedotteita.

Holme, I & Solvang, B (1991). *Forskningsmetodik. Om kvalitativa och kvantitativa metoder*. Studentlitteratur: Lund.

- Larsson, K. (1976). Några synpunkter på syntaxen i en svensk landsortstidning. I: *Språket i spalterna*. Ord och stil. Språksamfundets skrifter 8. Lund: Studentlitteratur, 53–78.
- Marjakangas, H. & Vertanen, S. (2003). *Om läsbarhet i lättlästa och vanliga tidningstexter. En jämförelse av 8 SIDOR, SESAM och Dagens Nyheter*. Opublicerad pro gradu-avhandling. Institutionen för språk, svenska språket, Jyväskylä universitet, Jyväskylä.
- Melander, B. (1987). Något om populärvetenskapliga texter och populärvetenskapligt språk. I: Gunnarsson B-L (red.) *Facktext. Ord och stil*. Språkvårdssamfundets skrifter 18. Malmö: Liber, 104–123.
- Melin, L. & Lange, S. (2000). *Att analysera text*. Lund: Studentlitteratur.
- Näslund, H. (1987). Fackjournalistik – en hybrid? I: Gunnarsson B-L (red.) *Facktext. Ord och stil*. Språkvårdssamfundets skrifter 18. Malmö: Liber, 125–144.
- Platzack, C. (1973). *Språket och läsbarheten*. Lund: Skrifter utgivna av Svensk-läraryrskommittén 152.
- Sigurd, Bengt (1991). *Språk och språkforskning*. Lund: Studentlitteratur.
- Strand, H. (1984). *Nusvenskt tidningsspråk. Kvantitativa studier av reportage i morgon-, kvälls- och veckopress*. Stockholms universitet: Institutionen för nordiska språk.
- Stukát, S. (1993). *Statistikens grunder*. Lund: Studentlitteratur.
- Westman, M. (1974). *Bruksprosa*. Lund: Liber Läromedel.

## INTERNETKÄLLOR

Internet 1: <http://hem.passagen.se/benthorn/statistik/ttest.htm> (22.3.2006, kl 9)

Internet 2: <http://www.infovoice.se/fou/bok/10000049.htm> (6.5.2006, kl 20.10)

Internet 3: <http://matmin.kevius.com/metrologi.html> (6.5.2006, kl 20.50)

Internet 4: <http://www.netmba.com/statistics/plot/box/> (7.5.2006, kl 19.05)

Internet 5: <http://joyx.joensuu.fi/~ek/tilasto1/boxplot.htm> (7.5.2006, kl 19.20)

Internet 6: <http://sv.wikipedia.org/wiki/L%C3%A5dagram> (7.5.2006, kl 19.30)

Internet 7: <http://www.infovoice.se/fou/bok/statmet/10000047.htm> (8.5.2006, kl 20.25)

Internet 8: <http://fi.wikipedia.org/wiki/T-testi> (8.5.2006, kl 21)

BILAGA 1: Grammatifix-resultat för LIX-värden i DN och HBL - förstasidorna

(B/O står för bokstäver per ord)

Tidning	Datum	Veckod.	LIX	B/O	Tidning	Datum	Veckod.	LIX	B/O
DN	20050207	måndag	31	5,2	HBL	20050207	måndag	29	4,9
DN	20050208	tisdag	31	5,2	HBL	20050208	tisdag	32	5,1
DN	20050209	onsdag	33	5,2	HBL	20050209	onsdag	35	5,3
DN	20050210	torsdag	27	4,9	HBL	20050210	torsdag	34	5,2
DN	20050211	fredag	32	5,2	HBL	20050211	fredag	30	4,9
DN	20050212	lördag	29	5,1	HBL	20050212	lördag	33	5,1
DN	20050213	söndag	34	5,4	HBL	20050213	söndag	28	4,6
DN	20050214	måndag	32	5,3	HBL	20050214	måndag	32	4,7
DN	20050215	tisdag	31	5,1	HBL	20050215	tisdag	34	5,1
DN	20050216	onsdag	27	4,8	HBL	20050216	onsdag	34	5,3
DN	20050217	torsdag	29	4,9	HBL	20050217	torsdag	34	5,4
DN	20050218	fredag	32	5,1	HBL	20050218	fredag	31	4,8
DN	20050219	lördag	28	4,8	HBL	20050219	lördag	33	5,2
DN	20050220	söndag	31	5,2	HBL	20050220	söndag	31	5,0
DN	20050221	måndag	33	5,1	HBL	20050221	måndag	35	5,2
DN	20050222	tisdag	35	5,3	HBL	20050222	tisdag	35	5,3
DN	20050223	onsdag	30	5,0	HBL	20050223	onsdag	36	5,2
DN	20050224	torsdag	28	5,0	HBL	20050224	torsdag	34	5,1
DN	20050225	fredag	33	5,4	HBL	20050225	fredag	35	5,3
DN	20050226	lördag	29	5,1	HBL	20050226	lördag	33	5,2
DN	20050227	söndag	32	5,4	HBL	20050227	söndag	30	4,7
DN	20050228	måndag	33	5,3	HBL	20050228	måndag	32	4,8
DN	20050301	tisdag	31	5,2	HBL	20050301	tisdag	35	5,2
DN	20050302	onsdag	30	5,1	HBL	20050302	onsdag	32	5,1
DN	20050303	torsdag	32	5,2	HBL	20050303	torsdag	34	5,1
DN	20050304	fredag	29	5,2	HBL	20050304	fredag	35	5,2
DN	20050305	lördag	32	5,4	HBL	20050305	lördag	33	4,9
DN	20050306	söndag	33	5,2	HBL	20050306	söndag	35	5,1
DN	20050307	måndag	31	5,1	HBL	20050207	måndag	35	5,3
DN	20050308	tisdag	33	5,3	HBL	20050308	tisdag	36	5,3
DN	20050309	onsdag	33	5,3	HBL	20050309	onsdag	31	5,0
DN	20050310	torsdag	32	5,2	HBL	20050310	torsdag	34	5,2
DN	20050311	fredag	34	5,4	HBL	20050311	fredag	31	5,1
DN	20050312	lördag	31	5,2	HBL	20050312	lördag	33	5,2
DN	20050313	söndag	32	5,1	HBL	20050313	söndag	35	5,2
DN	20050314	måndag	28	5,0	HBL	20050314	måndag	29	5,0
DN	20050315	tisdag	27	5,0	HBL	20050315	tisdag	36	5,4
DN	20050316	onsdag	34	5,4	HBL	20050316	onsdag	32	5,0
DN	20050317	torsdag	28	5,1	HBL	20050317	torsdag	34	5,3
DN	20050318	fredag	32	5,5	HBL	20050318	fredag	36	5,2
DN	20050319	lördag	28	5,0	HBL	20050319	lördag	31	4,8
DN	20050320	söndag	34	5,4	HBL	20050320	söndag	31	4,7
		medelv.	31,04762	5,09			medelv.	33,04762	5,17
		avrundat	31	5,1			avrundat	33	5,2



BILAGA 2: **HBL - Insidesmaterialet** ordnat som enskilda artiklar: LIX samt B/O-värden

	dag												medelvärde	avrundat
	1	2	3	4	5	6	7	8	9	10	11	12		
HBL ekonomi	30	40	39	26	46	41	42	48	44	41	37	45	39,91667	40
HBL inrikes	35	44	37	36	29	36	33	35	40	39	31	31	35,50000	36
HBL kultur	30	38	22	35	29	42	33	44	32	17	28	38	32,33333	32
HBL ledare	38	50	40	39	37	35	45	36	43	38	42	46	40,75000	41
HBL politik	34	37	39	37	37	32	33	40	35	46	37	41	37,33333	37
HBL sport	27	37	35	26	36	32	34	30	30	35	37	34	32,75000	33
HBL utrikes	39	45	39	40	42	42	37	37	39	37	37	41	39,58333	40
	totalt												36,88095	37

Motsvarande B/O-värden för HBL-insidesmaterialet (B/O står för bokstäver per ord)

	dag												medelvärde	avrundat
	1	2	3	4	5	6	7	8	9	10	11	12		
HBL ekonomi	4,8	5,2	5,3	4,8	5,9	5,3	6,1	6,0	5,9	5,2	5,0	5,6	5,43	5,4
HBL inrikes	5,4	5,4	5,4	5,3	4,6	5,2	5,0	5,1	5,7	5,3	4,8	4,6	5,15	5,2
HBL kultur	4,5	4,8	4,1	4,5	4,5	5,5	4,7	5,6	4,9	3,4	4,6	5,1	4,68	4,7
HBL ledare	5,2	6,2	5,4	5,5	5,4	5,0	5,8	5,1	5,7	5,2	5,6	5,7	5,48	5,5
HBL politik	5,0	5,2	5,2	5,2	5,1	4,8	4,9	5,0	5,3	5,9	5,5	5,9	5,25	5,3
HBL sport	4,7	5,4	5,0	4,3	4,9	5,0	4,8	4,6	4,8	5,0	5,1	4,9	4,88	4,9
HBL utrikes	5,1	5,4	5,2	5,3	5,5	5,4	5,3	5,1	5,6	5,2	5,3	5,5	5,33	5,3
	totalt												5,17	5,2

BILAGA 3: HBL - Utvärdering av insidesmaterialet enligt t-testens resultat

Jämförelse mellan **insidorna** enligt p-värde (mest signifikanta först) **inom HBL**

Jämförelsepar	Fel av första slaget	Signifikansnivå	
<b>Sport vs Utrikes</b>	<b>p-värde= 0.0004483</b>	<b>mycket signifikant</b>	
<b>Ledare vs Sport</b>	<b>p-värde= 0.0001154</b>	<b>mycket signifikant</b>	2 st. mycket
<b>Ekonomi vs Sport</b>	<b>p-värde= 0.003696</b>	<b>signifikant</b>	signifikanta skillnader,
<b>Kultur vs Ledare</b>	<b>p-värde= 0.004913</b>	<b>signifikant</b>	5 st. signifikanta skillnader
<b>Politik vs Sport</b>	<b>p-värde= 0.007271</b>	<b>signifikant</b>	samt 2 st. nästan signifikanta
<b>Inrikes vs Ledare</b>	<b>p-värde= 0.007743</b>	<b>signifikant</b>	skillnader
<b>Kultur vs Utrikes</b>	<b>p-värde= 0.009245</b>	<b>signifikant</b>	
<b>Inrikes vs Utrikes</b>	<b>p-värde= 0.01026</b>	<b>nästan signifikant</b>	
<b>Ekonomi vs Kultur</b>	<b>p-värde= 0.01702</b>	<b>nästan signifikant</b>	
Ledare vs Politik	p-värde= 0.05944	icke-signifikant	
Ekonomi vs Inrikes	p-värde= 0.06109	icke-signifikant	totalt 9 / 21 st. signifikanta skillnader
Kultur vs Politik	p-värde= 0.06512	icke-signifikant	enligt 5% signifikansnivå, dvs. 43 %
Inrikes vs Sport	p-värde= 0.1062	icke-signifikant	
Politik vs Utrikes	p-värde= 0.1070	icke-signifikant	
Inrikes vs Kultur	p-värde= 0.2357	icke-signifikant	
Ekonomi vs Politik	p-värde= 0.2471	icke-signifikant	
Inrikes vs Politik	p-värde= 0.2788	icke-signifikant	
Ledare vs Utrikes	p-värde= 0.4473	icke-signifikant	
Ekonomi vs Ledare	p-värde= 0.7172	icke-signifikant	
Ekonomi vs Utrikes	p-värde= 0.8694	icke-signifikant	
Kultur vs Sport	p-värde= 0.8704	icke-signifikant	

BILAGA 4: DN och HBL - Utvärdering av förstasidesmaterialet enligt t-testens resultat

Jämförelse mellan <b>veckodagar</b> enligt p-värde (mest signifikanta först) <b>inom DN</b>			
<b>signifikant</b>	<b>DNlöDNsö</b>	<b>0.004056</b>	
<b>nästan signifikant</b>	<b>DNtoDNsö</b>	<b>0.01118</b>	1 st. signifikanta och
<b>nästan signifikant</b>	<b>DNfrDNlö</b>	<b>0.02600</b>	3 st. nästan signifikanta
<b>nästan signifikant</b>	<b>DNtoDNfr</b>	<b>0.03939</b>	skillnader
icke-signifikant	DNmåDNlö	0.1011	
icke-signifikant	DNmåDNto	0.1172	
icke-signifikant	DNmåDNsö	0.1771	
icke-signifikant	DNtiDNto	0.1844	
icke-signifikant	DNtiDNlö	0.1872	totalt 4 / 21 st. signifikanta
icke-signifikant	DNonDNto	0.2185	skillnader enligt 5 % signifikansnivå,
icke-signifikant	DNonDNlö	0.224	dvs. 19 %
icke-signifikant	DNonDNsö	0.2462	
icke-signifikant	DNtiDNsö	0.3005	
icke-signifikant	DNfrDNsö	0.4493	
icke-signifikant	DNmåDNfr	0.5291	
icke-signifikant	DNonDNfr	0.531	
icke-signifikant	DNtiDNfr	0.6165	
icke-signifikant	DNtoDNlö	0.8836	
icke-signifikant	DNmåDNnon	0.9022	
icke-signifikant	DNtiDNnon	0.9154	
icke-signifikant	DNmåDNti	1.000	

Jämförelse mellan <b>veckodagar</b> enligt p-värde (mest signifikanta först) <b>inom HBL</b>			
<b>nästan signifikant</b>	<b>HBLtoHBLlö</b>	<b>0.01032</b>	
<b>nästan signifikant</b>	<b>HBLtiHBLlö</b>	<b>0.02197</b>	
icke-signifikant	HBLtiHBLsö	0.0512	
icke-signifikant	HBLmåHBLti	0.0671	2 st. nästan signifikanta
icke-signifikant	HBLtoHBLsö	0.09714	skillnader
icke-signifikant	HBLmåHBLto	0.1275	
icke-signifikant	HBLtiHBLfr	0.2122	
icke-signifikant	HBLtiHBLon	0.2186	
icke-signifikant	HBLonHBLsö	0.2639	totalt 2 / 21 st. signifikanta
icke-signifikant	HBLtiHBLto	0.3276	skillnader enligt 5 % signifikansnivå,
icke-signifikant	HBLmåHBLon	0.3514	dvs. 10 %
icke-signifikant	HBLtoHBLfr	0.3907	
icke-signifikant	HBLfrHBLsö	0.4138	
icke-signifikant	HBLlöHBLsö	0.4347	
icke-signifikant	HBLonHBLto	0.4441	
icke-signifikant	HBLonHBLlö	0.4694	
icke-signifikant	HBLmåHBLfr	0.5275	
icke-signifikant	HBLmåHBLlö	0.5819	
icke-signifikant	HBLfrHBLlö	0.7752	
icke-signifikant	HBLonHBLfr	0.808	
icke-signifikant	HBLmåHBLsö	0.8376	

BILAGA 5: DN och HBL - Utvärdering av förstasidesmaterialet enligt t-testens resultat

Jämförelse mellan <b>veckodagar</b> enligt p-värde (mest signifikanta först) <b>DN vs HBL</b>			
<b>mycket signifikant</b>	<b>DNlöHBLti</b>	<b>0.0002101</b>	
<b>mycket signifikant</b>	<b>DNtoHBLti</b>	<b>0.0007974</b>	2 st. mycket signifikanta skillnader
<b>signifikant</b>	<b>DNlöHBLto</b>	<b>0.001114</b>	
<b>signifikant</b>	<b>DNtoHBLto</b>	<b>0.003214</b>	
<b>signifikant</b>	<b>DNlöHBLlö</b>	<b>0.003519</b>	6 st. signifikanta skillnader
<b>signifikant</b>	<b>DNlöHBLon</b>	<b>0.00459</b>	
<b>signifikant</b>	<b>DNmåHBLti</b>	<b>0.00708</b>	11 st. nästan signifikanta skillnader
<b>signifikant</b>	<b>DNtoHBLon</b>	<b>0.007403</b>	
<b>nästan signifikant</b>	<b>DNtoHBLlö</b>	<b>0.01104</b>	
<b>nästan signifikant</b>	<b>DNfrHBLti</b>	<b>0.01595</b>	
<b>nästan signifikant</b>	<b>DNmåHBLto</b>	<b>0.01713</b>	
<b>nästan signifikant</b>	<b>DNonHBLti</b>	<b>0.02261</b>	
<b>nästan signifikant</b>	<b>DNlöHBLfr</b>	<b>0.0227</b>	totalt 19 / 49 st. signifikanta skillnader enligt 5 % signifikansnivå, dvs. 39 %
<b>nästan signifikant</b>	<b>DNtoHBLfr</b>	<b>0.02488</b>	
<b>nästan signifikant</b>	<b>DNtiHBLti</b>	<b>0.02854</b>	
<b>nästan signifikant</b>	<b>DNsöHBLti</b>	<b>0.03056</b>	
<b>nästan signifikant</b>	<b>DNfrHBLto</b>	<b>0.03272</b>	
<b>nästan signifikant</b>	<b>DNsöHBLto</b>	<b>0.04295</b>	
<b>nästan signifikant</b>	<b>DNonHBLto</b>	<b>0.04657</b>	
icke-signifikant	DNtiHBLto	0.05743	
icke-signifikant	DNlöHBLmå	0.08624	
icke-signifikant	DNtoHBLmå	0.08851	
icke-signifikant	DNmåHBLon	0.1006	
icke-signifikant	DNtoHBLsö	0.1395	
icke-signifikant	DNonHBLon	0.1404	
icke-signifikant	DNlöHBLsö	0.1409	
icke-signifikant	DNmåHBLlö	0.1531	
icke-signifikant	DNtiHBLon	0.1718	
icke-signifikant	DNonHBLlö	0.2323	
icke-signifikant	DNmåHBLfr	0.2344	
icke-signifikant	DNfrHBLon	0.2353	
icke-signifikant	DNonHBLfr	0.254	
icke-signifikant	DNtiHBLlö	0.2852	
icke-signifikant	DNtiHBLfr	0.2986	
icke-signifikant	DNfrHBLlö	0.4086	
icke-signifikant	DNsöHBLsö	0.4498	
icke-signifikant	DNfrHBLfr	0.4506	
icke-signifikant	DNsöHBLon	0.4988	
icke-signifikant	DNsöHBLmå	0.5965	
icke-signifikant	DNonHBLmå	0.5995	
icke-signifikant	DNmåHBLmå	0.6292	
icke-signifikant	DNtiHBLmå	0.6747	
icke-signifikant	DNonHBLsö	0.757	
icke-signifikant	DNsöHBLfr	0.7846	
icke-signifikant	DNfrHBLsö	0.8088	
icke-signifikant	DNmåHBLsö	0.814	
icke-signifikant	DNtiHBLsö	0.837	
icke-signifikant	DNfrHBLmå	1.000	
icke-signifikant	DNsöHBLlö	1.000	

Jämförelse mellan **alla förstasidor** som värdegrupper enligt p-värde **DN vs HBL**  
 Skillnaden är **mycket signifikant**  $lixDNlixHBL$   $p= 0.00005916$

BILAGA 6: DN och HBL - Utvärdering av förstasidesmaterialet enligt t-testens resultat

Jämförelse mellan <b>veckor</b> enligt p-värde (mest signifikanta först) <b>DN vs HBL</b>			
<b>signifikant</b>	<b>DNv2HBLv4</b>	<b>0.002062</b>	
<b>signifikant</b>	<b>DNv2HBLv3</b>	<b>0.002813</b>	3 st. signifikanta
<b>signifikant</b>	<b>DNv2HBLv5</b>	<b>0.005772</b>	skillnader
<b>nästan signifikant</b>	<b>DNv4HBLv4</b>	<b>0.01214</b>	
<b>nästan signifikant</b>	<b>DNv2HBLv2</b>	<b>0.01350</b>	
<b>nästan signifikant</b>	<b>DNv6HBLv3</b>	<b>0.01895</b>	10 st. nästan
<b>nästan signifikant</b>	<b>DNv4HBLv3</b>	<b>0.01992</b>	signifikanta skillnader
<b>nästan signifikant</b>	<b>DNv6HBLv4</b>	<b>0.0223</b>	
<b>nästan signifikant</b>	<b>DNv1HBLv3</b>	<b>0.02583</b>	totalt 13 / 36 st. signifikanta
<b>nästan signifikant</b>	<b>DNv1HBLv4</b>	<b>0.02683</b>	skillnader enligt 5 % signifikansnivå,
<b>nästan signifikant</b>	<b>DNv6HBLv5</b>	<b>0.03224</b>	dvs. 36 %
<b>nästan signifikant</b>	<b>DNv4HBLv5</b>	<b>0.04391</b>	
<b>nästan signifikant</b>	<b>DNv1HBLv5</b>	<b>0.04937</b>	
icke-signifikant	DNv2HBLv6	0.05522	
icke-signifikant	DNv5HBLv4	0.05536	
icke-signifikant	DNv3HBLv3	0.05651	
icke-signifikant	DNv3HBLv4	0.06268	
icke-signifikant	DNv6HBLv2	0.0771	
icke-signifikant	DNv5HBLv3	0.07752	
icke-signifikant	DNv3HBLv5	0.1031	
icke-signifikant	DNv6HBLv6	0.1224	
icke-signifikant	DNv4HBLv2	0.1227	
icke-signifikant	DNv1HBLv2	0.1315	
icke-signifikant	DNv5HBLv5	0.1681	
icke-signifikant	DNv1HBLv6	0.2311	
icke-signifikant	DNv2HBLv1	0.2346	
icke-signifikant	DNv3HBLv2	0.2640	
icke-signifikant	DNv4HBLv6	0.2976	
icke-signifikant	DNv6HBLv1	0.3699	
icke-signifikant	DNv3HBLv6	0.3733	
icke-signifikant	DNv5HBLv1	0.5274	
icke-signifikant	DNv5HBLv2	0.535	
icke-signifikant	DNv1HBLv1	0.678	
icke-signifikant	DNv5HBLv6	0.7071	
icke-signifikant	DNv4HBLv1	0.9036	
icke-signifikant	DNv3HBLv1	0.919	

BILAGA 7: DN och HBL - Utvärdering av förstasidesmaterialet enligt t-testens resultat

Jämförelse mellan <b>veckor</b> enligt p-värde (mest signifikanta först) <b>inom DN</b>			
<b>nästan signifikant</b>	<b>DNv2DNv5</b>	<b>0.0259</b>	
icke-signifikant	DNv5DNv6	0.1238	1 st. nästan
icke-signifikant	DNv2DNv4	0.1594	signifikanta skillnader
icke-signifikant	DNv1DNv5	0.2295	
icke-signifikant	DNv4DNv5	0.2523	
icke-signifikant	DNv2DNv3	0.2625	totalt 1 / 15 st. signifikanta
icke-signifikant	DNv4DNv6	0.3479	skillnader enligt 5 %
icke-signifikant	DNv3DNv6	0.409	signifikansnivå, dvs. 7 %
icke-signifikant	DNv1DNv2	0.4119	
icke-signifikant	DNv3DNv5	0.4316	
icke-signifikant	DNv1DNv6	0.5715	
icke-signifikant	DNv1DNv4	0.696	
icke-signifikant	DNv1DNv3	0.7486	
icke-signifikant	DNv2DNv6	0.92	
icke-signifikant	DNv3DNv4	1.000	

Jämförelse mellan <b>veckor</b> enligt p-värde (mest signifikanta först) <b>inom HBL</b>			
icke-signifikant	HBLv1HBLv3	0.0778	
icke-signifikant	HBLv1HBLv4	0.08896	0 st. mycket
icke-signifikant	HBLv1HBLv5	0.1370	signifikanta eller
icke-signifikant	HBLv2HBLv3	0.1900	signifikanta skillnader
icke-signifikant	HBLv2HBLv4	0.2002	
icke-signifikant	HBLv3HBLv6	0.3319	0 st. nästan
icke-signifikant	HBLv1HBLv2	0.336	signifikanta skillnader
icke-signifikant	HBLv2HBLv5	0.3694	
icke-signifikant	HBLv4HBLv6	0.4044	totalt 0 / 15 st. signifikanta
icke-signifikant	HBLv1HBLv6	0.4378	skillnader enligt 5 %
icke-signifikant	HBLv5HBLv6	0.5118	signifikansnivå, dvs. 0 %
icke-signifikant	HBLv3HBLv5	0.6947	
icke-signifikant	HBLv3HBLv4	0.7617	
icke-signifikant	HBLv4HBLv5	0.8788	
icke-signifikant	HBLv2HBLv6	1.000	