

Lättläst stil

Språklig variation och likhet i tolv lättlästa litterära prosaverk

Malin Bornhöft

Institutionen för svenska och flerspråkighet
Examensarbete 30 hp
Examensämne: Svenska
Masterprogram i språkvetenskap inriktning svenska
Vårterminen 2016
Handledare: Anna Vogel
English title: Easy-to-read Style



Stockholms
universitet

Lättläst stil

Språklig variation och likhet i tolv lättlästa litterära prosaverk

Malin Bornhöft

Sammanfattning

Den här uppsatsen studerar variation och likhet i och mellan tolv verk för att identifiera vad som är möjligt och omöjligt i texterna. Materialet är tolv nyutgivna verk på tre svårighetsnivåer från Myndigheten för tillgängliga mediers lättlästförlag. Verken är litterära prosaverk. De analyseras stilistiskt både kvantitativt och kvalitativt utifrån tre områden: lexikon och syntax, rytm och interpunktion samt färg och illustration.

Resultatet visar att förlagets svårighetsnivåer inte återspeglas när det gäller variation och likhet. Verken klustrar ihop sig mot mitten och uppvisar ett normalfördelningsmönster, där vissa verk intar extrempositioner. Förutom extremverken tycks stilen uppvisa stor återhållsamhet, förutom för enstaka analysenheter per verk. Vilken analysenhet det gäller varierar mellan verken. Det beror antagligen på att verken är hårt normerade. En slutsats att vissa språkliga enheter är enkla att normera användning av, medan andra är stilistiska val öppna för författarna. En annan är att lättlästa verk innehåller variation som inte alltid går att förutsäga.

Nyckelord

Lättläst, stilistik, korpuslingvistik, Myndigheten för tillgängliga medier, interpunktion.

Innehållsförteckning

1 Inledning	1
1.1 Syfte, frågeställningar och hypoteser	1
1.2 Uppsatsens disposition	1
2 Bakgrund	3
2.1 Lättläst	3
2.1.1 Behöver ämnet beforskas?	4
2.2 Term: normering	4
3 Stilistik	6
3.1 Statistik	6
3.1.1 Att mäta svårighetsgrad	7
3.2 Termer: variation & likhet	8
4 Material	10
4.1 De tre nivåerna	10
4.2 De tolv verken	11
4.2.1 Nivå 1-verken	12
4.2.2 Nivå 2-verken	13
4.2.3 Nivå 3-verken	14
4.3 Korpuslingvistik	15
4.4 Metod för att hantera materialet	16
4.5 Reguljära uttryck	16
5 Lexikon och syntax	17
5.1 Analysenheter och tillvägagångssätt	17
5.1.1 Standardiserad typ/löpord (STT)	17
5.1.2 Grafiska meningar	18
5.1.3 Antal bokstäver	19
5.1.4 Antal löpord och genomsnittlig ordlängd	19
5.1.5 Antal sidor	20
5.1.6 Pronomen	20
5.1.7 Långa ord	21
5.1.8 Egennamn	21
5.1.9 Konjunktioner	22
5.2 Resultat	24
5.2.1 Kvantitativa resultat	24
5.2.2 Lexikon- & syntaxvärde (LSV)	29
5.2.3 Stilistiska kommentarer	31

5.3 Sammanfattning av lexikon och syntax	35
6 Rytm och interpunktion	36
6.1 Analysenheter och tillvägagångssätt	36
6.1.1 Utropstecken och frågetecken	37
6.1.2 Kolon.....	37
6.1.3 Kommatecken.....	38
6.1.4 Tre punkter/ellips	38
6.1.5 Tankstreck/talstreck	38
6.1.6 Bindestreck	39
6.1.7 Versalord	39
6.1.8 Asterisker	40
6.1.9 Parenteser, citattecken och semikolon.....	40
6.2 Resultat.....	41
6.2.1 Kvantitativa resultat	5
6.2.2 Rytm- & interpunktionsvärde (RIV).....	48
6.2.3 Stilistiska kommentarer	50
6.3 Sammanfattning av rytm och interpunktion	55
7 Färg och illustration.....	57
7.1 Tillvägagångssätt	57
7.2 Analyskategorier	57
7.2.1 Placering och antal	57
7.2.2 Distans och innehåll.....	58
7.2.3 Färg	59
7.3 Resultat.....	60
7.3.1 Illustrationernas placering och antal	60
7.3.2 Illustrationernas innehåll och distans	60
7.3.3 Färg	61
7.3.4 Färg- och illustrationsvärde (FIV)	62
7.3.5 Stilistiska kommentarer	63
7.4 Sammanfattning av färg och illustration	64
8 Sammanställning av resultat	66
9 Diskussion	68
9.1 Hur hänger form och innehåll ihop?	68
9.2 Hur hårt går det att normera litterär prosa?.....	68
9.3 Vad innebär förlagets nivåer för läsaren?.....	69
9.4 Metodutvärdering.....	70
10 Avslutning.....	71
Primärlitteratur	72
Litteratur	72
Bilagor	1

Bilaga 1: Nivåernas innehåll.....	1
Bilaga 2: Reguljära uttryck	2
Bilaga 3: Lexikon och syntax, absoluta siffror	3
Bilaga 4: STT	4

1 Inledning

Den här uppsatsen handlar om böcker skrivna på så kallad lättläst svenska. De existerar i mötet mellan två världar: den litterära prosans kreativa, fria värld, och språkvårdens hårt normerade värld – i alla fall är det mitt antagande. Lättläst litteratur skiljer sig från annan genom att ha definierade målgrupper och syften. Dessa böcker verkar alltså leva på premisser som liknar de sakprosatexter brukar lyda under. Det borde påverka språket i dem.

Det är svårt att argumentera mot ideal om allas rätt till språk och kultur, men kritik mot lättläst finns det gott om. Språket lyfts ofta. Det är konstigt och annorlunda säger den som läser annan litteratur. Det är ineffektivt och svårtillgängligt kan språkarbetare säga. Så vad är det som händer med språket? Vad menas egentligen med ”språket”? Och är alla lättlästa verk ens lika varandra?

För att svara på frågor om vad som händer med språket jämför jag i den här uppsatsen stilen i tolv nytgivna lättlästa verk. Jag jämför dem utifrån tre områden som på olika sätt rör ”språket”: lexikon och syntax, rytm och interpunktion samt färg och illustrationer.

1.1 Syfte, frågeställningar och hypoteser

Syftet är att jämföra tolv litterära prosaverk med varandra, för att urskilja vad som är möjligt och omöjligt i lättläst stil.

Jag använder mig av tre frågeställningar för att kunna ringa in vad de lättlästa verken har gemensamt och var de skiljer sig åt:

*Frågeställning 1: Vilken variation och likhet finns det i lättlästa litterära prosaverks **lexikon** och **syntax**?*

*Frågeställning 2: Vilken variation och likhet finns det i lättlästa litterära prosaverks **rytm** och **interpunktion**?*

*Frågeställning 3: Vilken variation och likhet finns det i lättlästa litterära prosaverks användning av **färg** och **illustrationer**?*

I min genomgång av de tre analysområdena kommer jag att ställa upp specificerade hypoteser, som alltså ska ses som underordnade och mer specifika än frågeställningarna. Detta sker löpande i uppsatsen. Hypoteserna baserar jag främst på tidigare forskning och förlagets instruktioner för lättlästa texter och för de tre nivåerna. Jag har viss förtrogenhet med lättläst sen tidigare, och har snabbt bekantat mig med alla verken innan jag påbörjade min studie, vilket i vissa fall givit mig en aning om vad som kan förekomma i dem.

1.2 Uppsatsens disposition

Uppsatsen består grovt av tre delar: bakgrund, analys och avslutning.

Först introducerar jag ämnet (här ovan). Nedan beskriver jag kort området språkvård i allmänhet och lättläst i synnerhet, och redogör för hur jag använder termen *normering*. Därefter presenterar jag fältet stilistisk och beskriver hur statistik brukar användas. Jag redogör för hur jag använder termerna

variation och *likhet* som återfinns i mina frågeställningar. Jag redovisar mitt material och beskriver hur jag hanterat det.

Därefter följer tre analyskapitel, som vardera besvarar en av mina frågeställningar. De heter "Lexikon och syntax", "Rytm och interpunktion" samt "Färg och illustration". De är uppbyggda på liknande sätt. Först presenterar jag vad, hur och varför jag undersökt de analysenheter som jag valt ut, och ställer upp hypoteser utifrån tidigare forskning, vilket kan ses som metodavsnitt. Därefter följer resultatavsnittet, där jag besvarar frågeställningen och avgör om hypoteserna stöds eller ej. Resultatavsnittet har en kvantitativ del, en del där verken tilldelas ett sammantaget variationsvärde, och en del där verken kommenteras stilistiskt var för sig. Kapitlet avslutas med en sammanfattning. Analysdelen av uppsatsen avrundas med en sammanställning av alla tre analyskapitlen.

Avslutningen består av ett diskussionskapitel, där också metodutvärderingen återfinns. Uppsatsen avslutas med att jag kort redogör för vad jag anser är uppsatsens största bidrag.

2 Bakgrund

Språkvården i Sverige ägnar sig åt tre områden: språkvalsplanering, ord- och grammatikvård och tal- och textvård (Språkrådets 2014). Den är i hög grad institutionaliserad, i form av Språkrådet, Terminologisentrum och Svenska Akademien (Teleman 2003:55).

Språkvård som fält intresserar mig, och den här uppsatsen kan ses som en fortsättning på två av mina tidigare uppsatser. Språkrådets verksamhet rör till stor del rådgivning av olika slag, och ett tydlighetsideal med konformitetslegitimeringar präglar råden (Bornhöft 2014:23–24). Det hänger ihop med språkvårdens strävan efter demokrati, ett ideal som dock sällan konkretiseras till fullo (Bornhöft 2015:38). Rekommendationer i kombination med demokrati-ideal ledde mig till fältet lättlästa texter, ett område som i mångt och mycket tycks vara centrum för språkvårdens textideal i praktiken.

2.1 Lättläst

Lättlästa texter existerar för att uppfylla bestämmelser om att alla ska kunna ta del av information och kultur. Lättlästa texter skrivs både av kommersiella förlag (vars texter ofta riktar sig till barn) och av Myndigheten för tillgängliga mediers (MTM) avdelning för lättläst. Det finns samhällsinformation, nyhetsrapportering, fack- och skönlitteratur på lättläst svenska.

Lättläst fungerar som en kategori för verk som publiceras med vissa specifika målgrupper och syften i åtanke, snarare än en bedömning av innehållet. Lättlästa texter riktar sig främst till de 13 procent av befolkningen som enligt MTM tillhör någon eller flera av följande grupper: ”personer med dyslexi och med andra läs- och skrivsvårigheter, nysvenskar, personer med intellektuell funktionsnedsättning, personer med afasi, äldre med demens, ovana läsare, barndomsdöva, andra grupper som har svårt med läsningen av olika skäl¹” (MTM u.å.).

Lättläst skrivs enligt vissa principer, som varierar mellan utgivare och över tid. Vad som är lätt att läsa är inte en enkel fråga att besvara för språkvetenskapen, och än finns inga tydliga regler som garanterar förståelse hos mottagaren. I mångt och mycket lyder lättläst under samma principer som klarspråk, det offentliga språk som ska vara vårdat, enkelt och begripligt enligt språklagen från 2009. Råden gäller bland annat informationsdisposition, rubrikmängd, textlängd, ordval och syntax (se till exempel Språkrådets Klarspråkstest).

Ofta beskrivs lättläst som klarspråk fast ännu mer (Rehnberg 2012). Ännu vanligare ord, ännu kortare meningar, ännu tydligare markering av vad som är viktigt i texten. Råden baseras ofta på allmänt förnuft, av typen att alla vet att långa meningar är lätta att tappa bort sig i, och att obegripliga ord försvårar för läsaren. Mycket har belagts med språkvetenskap, främst kognitiva studier, men resultaten i dem har ofta förvanskats till väl enkla råd, vilket Åsa Wengelin diskuterar i artikeln ”Mot en evidensbaserad språkvård?” (2015).

¹ Många av benämningarna på människorna som åsyftas är politiskt laddade av olika skäl, och associeras ofta med negativa attityder. Då och då ändras de på grund av attityder och bestämmelser (från till exempel Socialstyrelsen och intressegrupper). Här citerar jag MTM och kommer annars inte att benämna läsargrupperna.

2.1.1 Behöver ämnet beforskas?

Som språkvetare är vi rika på företeelser att studera, då språk är en så väsentlig del av vad människor gör med varandra. Att välja ämne ur denna enorma mängd innebär alltid vissa svårigheter.

När det gäller lättläst har en rad uppsatser gjorts genom åren som ofta studerar vissa specifika drag av lättlästa texter. En del studerar hur begripliga de egentligen är. En avhandling av Monica Reichenberg (2000) finns på ämnet lättläst språk, och rör elevers inläring. Det mesta som gjorts rör sakprosa. De som inte rör sakprosa tycks oftast handla om jämförelser mellan den bearbetade lättlästa versionen av ett verk och dess original. De studerar skillnader som finns däremellan, och fokuserar på vad som förändras i bearbetningen. Här finns alltså en lucka: vi vet helt enkelt inte så mycket om lättläst litterär prosa förutom i kontrast till deras originaltexter. Jag är övertygad om att den som studerar texterna enskilt och den som jämför dem med andra kommer att finna olika resultat. Dessutom är det först i undersökningar av lättlästa texter utan dylika jämförelser som verk utan original kan studeras.

Att en lucka i vår samlade kunskapsbank finns behöver inte betyda att den måste fyllas eller att det inte finns frågor som är mer aktuella och kräver snabba svar. Att lättläst är ett omdebatterat ämne tyder dock på att det är relevant att beforska. Mycket av kritiken mot lättläst finns sammanfattad i Hanna Sofia Rehnbergs artikel ”Inte klart vad som är lätt” i Språktidningen från 2012, som också finns online. I den framgår att det är många människor som har många åsikter kring lättläst, och att mycket kritik finns. Området beskrivs av Rehnberg som ett ”getingbo”, men det är svårt att hitta åsikterna offentligt, eftersom mycket handlar om ”känslor, aningar och spekulationer” (2012). Och var ska en skribent eller ett förlag i så fall börja om de vill revidera arbetet med texterna? Vi har ju den politiska situationen att MTM ska producera text, och att text alltid ska finnas tillgänglig för medborgarna att ta del av (vilket innebär att bara film till exempel inte räcker).

Utifrån detta kontaktade jag MTM och frågade vad de undrade över och ville ha svar på. Jag hoppades kunna kombinera andras frustration och mina frågor med MTM:s behov, för att på så vis ha en chans att faktiskt göra en undersökning som kan leda till förändring. MTM var mycket positiva till att jag skulle undersöka deras texter, och min kontakt hade frågat kollegor vad de helst ville veta av en student. De var intresserade av att utvärdera de tre nivåer de publicerar verk på, som har olika målgrupper. Nivåerna ska lyda under något olika rekommendationer (mer om dessa i materialkapitlet), och därmed möta olika läsares behov. De undrade vilka skillnader det fanns mellan nivåerna i de texter som faktiskt publicerades. Jag tyckte att ämnet verkade intressant.

Lättläst tycks vara ett fält där många vill mycket, men där det finns viss tröghet att få igenom förändringar. Det verkar också råda delade – och bestämda – meningar om var vi befinner oss idag. Bara det gör ämnet intressant. Jag anser att för att kunna ta konstruktiva steg framåt måste vi veta var vi befinner oss.

2.2 Term: normering

Jag använder mig i uppsatsen av termen *normering*. Det gör jag i en gängse betydelse, där språk alltid regleras och där lättlästa texter regleras explicit.

Jag vill se allt språk som något som skapats och skapas i social förhandling, där systematik uppkommer. Språket är socialt reglerat, och den som uttrycker sig obegripligt kommer bli uppmanad att anpassa sig efter det gemensamma systemet. Utöver det är skriftspråk normerade i hög utsträckning. Svensk språkvård arbetar för att skriftspråket ska normeras på ett sätt som gör det enkelt att använda för både läsare och skribenter (Svenska språknämnden 2005:17). Inom dessa ramar går det att normera mer eller mindre; semiotiska resurser kan ligga nära eller långt borta från konventionaliserade användningar, där kanske experimentell lyrik visar på språkets och normeringens gränser.

Lättlästa texter normeras genom särskilda instruktioner för lättlästa texter som författaren måste förhålla sig till och i hög utsträckning följa, och genom redaktörsarbete. Instruktionerna gäller flera typer av språkliga enheter, som syns i Tabell 1. Instruktionerna fungerar också som en beskrivning av vad förlaget anser att deras lättlästa verk ska innehålla. Lättlästa texter är alltså hårt normerade, och tydliga ideal finns för dem att uppfylla.

Det är intressant att studera lättlästa texter just som normerade texter, eftersom det kan säga något om hur normeringsarbete går till. Hur mycket går det att normera texter, och vad innebär samma instruktioner för olika typer av berättelser?

3 Stilistik

Stilistik formar den breda teoretiska bakgrunden för den här uppsatsen. För att reda ut vad som kan rymmas i fältet vänder jag mig till tre böckers inledningskapitel.

Stilistik är ”läran om stil” enligt Per Lagerholm i verket *Stilistik* (2008:10) vilket i sig självt är tydligt men obegripligt. Hos Lagerholm har *stil* sin allmänspråkliga betydelse men undersöks vetenskapligt. Lagerholm säger att stilistiska undersökningar ofta börjar i en intuition som sedan måste beläggas vetenskapligt (2008:10). Så fungerar ansatsen i den här uppsatsen, där intuitionen säger något om att språket i lättlästa verk är omdiskuterat och kritiserat. Standardverket *Style in Fiction* av Geoffrey Leech och Mick Short har som syfte att visa ”the student of English that examining the language of a literary text can be a means to a fuller understanding and appreciation of the writer’s artistic achievement.” (2007:1). Däri ryms ingen kritik eller granskning av vad som sker med språket, utan fokus tycks vara på positiva omdömen. De exempelverk de går igenom är sedan klassiker vars stil kan antas vara av god kvalitet, men frågan om hur stilistiken används på texter utan hög verkshöjd kvarstår. Stilistiken de förordar tycks därmed ha lite gemensamt med kritiska textstudier och diskursanalytiska ambitioner. Leech & Shorts syfte ligger närmare litteraturvetenskapen än lingvistik, men diskuterar många frågor som den som studerar litterära verk ställs inför. Ur den hämtar jag alltså inte nödvändigtvis metoder, eftersom jag inte vill utvärdera verken som bra eller dåliga.

I antologin *Det skönlitterära språket: tolv texter om stil* är författarna snarare språkvetare än litteraturvetare. Där skriver redaktör Carin Östman i inledningen att antologin vill ”ge exempel på vad vi menar att språkvetenskapen kan tillföra kunskapen om litterärt språk” (2015:7). *Stilistik* betyder här ”studiet av hur språk skapar mening i text – inte bara litterär text – och vilka språkliga konstruktioner som kan användas för att förmedla ett visst innehåll.” (Östman 2015:7). Trots inskottet handlar antologin främst om litterär stil, då språkvetenskapen bör syfta till att studera hela språket, alltså även det i skönlitteratur som sällan är språkvetarens främsta material. Inom språkvetenskapen kan *mening* innebära olika saker, men Östman verkar referera till en vid bemärkelse av att innehåll förmedlas genom texter, och att innehåll och form är oskiljaktiga så att där formen ändras förändras också innehållet. Det är Östmans stilistisk jag håller mig närmast.

Från stilistiken hämtar jag några begrepp för perspektiv, som jag använder i mina kvalitativa analyser under rubrikerna ”Stilistiska kommentarer” i analyskapiteln. *Relation* kallar jag berättandet som sker från ett utifrånperspektiv och inte är format efter en specifik rollfigurs tal eller tankar (Thelander 2015:117). Det som står efter ett streck och föreställer repliker kallar jag *direkt anföring* (Thelander 2015:23). *Indirekt anföring* syftar till när repliker vävs in i den löpande brödtexten och inte markeras med talstreck, citattecken eller dylikt (Thelander 2015:126). När läsaren närmar sig en person och får ta del av expressiva uttryck, dialekt, syntax och liknande kallar jag det *fri indirekt anföring*, efter Thelander som diskuterar olika möjliga termval (2015:128).

3.1 Statistik

Mycket av den här uppsatsen handlar om statistik. Hur många? Var är det flest? Vad är ovanligt? Som Leech & Short påpekar: olika läsare kan känna på sig olika, och utan belägg hävda att något är typiskt för ett visst verk (2007:35). Trots detta är statistisk stilistik att ta läsarens perspektiv, istället för författarens (Leech & Short 2007:34), då den studerar vad som faktiskt syns av språket, oavsett författarens uppsåt eller interna lexikon (Leech & Short 2007:31).

Leech & Short skriver följande, något humoristiska, kommentar om valet mellan frekvensmätning och närläsning:

‘the message carried by the frequency distributions and transitional probabilities of its linguistic features, especially as they differ from those of the same features in the language as a whole’ [citat av Bernard Bloch]. Such definitions appeal to the empiricist, who would like to reduce a subjectively perceived phenomenon to something objective, but they tend to alarm the student of literature. (Leech & Short 2007:35)

Det följs av en diskussion om att frekvensmätning inte enbart är mätning av några godtyckliga språkliga fenomen för att sedan inte göra något mer.

Statistikens främsta fördel är att den kan visa på skillnader och likheter. Som Lisa Holm (2015) med enkla siffror visar i sin artikel ”Rytm i romanprosa” tenderar vissa skönlitterära verk att ha mer gemensamt än andra. Sådant som är lätt att undersöka: ”hur varierar lexikonets storlek?” statistiskt faller inte nödvändigtvis under frågor om författarens medvetna stilistiska val – något som aldrig går att veta då en text inte bevisar en tankes förekomst – eller om språkets totala lexikon² eller struktur. Den språkvetenskap den här uppsatsen ägnar sig åt handlar alltså om språkbruk, det observerbara språket, enligt hypotesen att det är i språkbrukares vanor det går att hitta svar om hur språket är (Leech & Short 2007:12).

En uppenbar förtjänst med statistik är att det möjliggör jämförelser mellan större datamängder än vid manuell analys. Då inget projekt har oändligt med resurser måste några aspekter prioriteras och andra försakats. Prioriteten här är att jämföra ett flertal (här tolv) verk med varandra för att kunna få ett resultat som visar på hur språket i lättläst-litteratur ter sig. Att göra fallstudier i ett eller två verk skulle göra generaliserbarheten minimal. Generaliserbarhet prioriteras just för att kunna belägga eller motbevisa några av de påståenden om lättläst som går att hitta. Det i sin tur prioriteras eftersom lättläst i egenskap av politiskt sanktionerat fält bör kunna försvara sig, motbevisa påståenden, men också förändras utifrån välgrundade resultat.

3.1.1 Att mäta svårighetsgrad

Det verkar finnas en lockelse i att hitta ett sätt att mäta texters svårighetsgrad, så att skribenten och organisationen kan vara trygga i att ha skapat en bra text. Att få siffror på saker går också fort, och små justeringar kräver inte nödvändigtvis stora innehållsliga förändringar, till skillnad från Språkrådets mer kvalitativa Klarspråkstest.

Det klassiska verktyget för statistiska mått som relaterar till läsbarhet är Läsbarhetsindex, ofta förkortat LIX som ofta omnämns på språkutbildningar och då och då hos organisationer som försöker förbättra sina texter. Till exempel arbetar Nypon förlag, det största kommersiella förlaget för lättläst litteratur, med LIX, och använder förkortningen som ett verb: ”Våra böcker är i många fall lixade” (Nypon u.å.). Däremot framgår det inte om deras böcker måste uppnå ett visst LIX-värde, eller om de bara kontrolleras med en LIX-beräkning.

LIX-värdet räknas ut genom att medeltalet ord per mening adderas med andelen (i procent) ord över 6 bokstäver (Lix.se u.å.). Måttet blir en procentsats, där vissa omfång är typiska för barnböcker, andra för akademisk text och så vidare.

LIX ger alltså en aning om hur långa meningar och hur långa ord en text har, men det är svårt att utifrån det säga något om den faktiska svårighetsgraden i texten. Därmed anses LIX ofta, i min erfarenhet från främst språkkonsultutbildningen, vara ett väldigt förenklat verktyg vars resultat inte ska vara avgörande för hur en text bör bearbetas.

² Frågor om ett språks totala lexikon ställer lingvister ibland (och kanske journalister desto oftare) stora korpusar, se till exempel Andrés Kornais artikel ”How many words are there” (2002) för diskussion. Hans slutsats är att ett språks lexikon alltid, oavsett korpusens storlek är oändligt stor.

Jag räknar precis som LIX ut ord per mening, och visst skiljer det sig mellan verken. Däremot analyserar jag inte det som att det ena verket nödvändigtvis är svårare än det andra.

LIX är inte det enda verktyget. Produkten ”Validera text” använder bland annat LIX, men studerar också ordvariation och hittar förkortningar och annat som den sen uppmanar skribenten att ändra. Där kan LIX-resultatet styra skribenten att förändra ord- och meningslängd för att förändra LIX-värdet, till en nivå som programmet markerar med en grön bock istället för med en varningstriangel.

För engelska finns verktyget Coh-Metrix, som inte bedömer svårighetsgrad i sig, utan närliggande kohesion och läsbarhet, här baserat på över 250 språk- och ”cohesion features” (Crossley et. al 2007:16). På verktygets egen webbplats framgår dock inte vilka dessa språkdrag är. Verktyget visar att autentiska texter och texter bearbetade för andraspråkstalare (återigen, på engelska) skiljer sig åt: L2-texterna har färre ovanliga sambandsmarkörer, men också färre kausalitetsmarkörer. Texterna är lika abstrakta och dubbeltydiga (Crossley et. al. 2007:27). Studien av Crossley et. al. (2007) använder jag som inspiration i hur statistik på språkdrag kan analyseras på ett annat sätt än att kalla det ena med nödvändighet svårare än något annat.

3.2 Termer: variation & likhet

Jag använder termerna *variation* och *likhet* i alla tre frågeställningar, som bara skiljer sig åt med hänseende på studieobjekt: lexikon och syntax, rytm och interpunktion, färg och bild. Därmed får termerna stor roll i uppsatsen, och jag ska här reda ut närmare hur de används.

Språk varierar alltid. Olika ord används hela tiden, och dessutom varierar lexikon och uttal mellan personer, platser, tider, vilket ofta studeras inom sociolingvistik (Norrby & Håkansson 2010:12). Grammatik förändras, och en bok kanske använder långa meningar med semikolon medan en annan är korthuggen med huvudsatser utan underordnade satser. Olika böcker använder garanterat språk på olika sätt, och dessutom ser de olika ut. Självklart har alla tolv verk likheter inom sig. Självklart har de alla också variation.

Det är enkelt att föreställa sig att olika böcker kan innehålla olika grad av variation. En med ett mindre lexikon, inga bilder, samma färg och typsnitt har i min analys mindre variation än en med stort lexikon, bilder av varierande sort gällande motiv och framställningssätt (fotografi, teckning), och typsnitt som växlar vid olika funktioner. Jag studerar dels verkens interna variation, och också alla verks sammantagna variation.

Med *variation* menar jag två saker. Används olika semiotiska resurser i hög grad är variationen stor. Används samma semiotiska resurser för olika funktioner är variationen också stor.

Ordet *variation* rymmer på många sätt också *likhet* men jag väljer att fråga efter båda, för att kunna fånga bredden mellan de yttersta gränserna, men också klustret i mitten. Mycket stilistik handlar just om det ovanliga (Leech & Short 2007:39). Nackdelen med det är att det skapar en bild av att det som inte ingår i det ovanliga med självklarhet är vanligt, normalt, gängse och omarkerat, trots att där kan hända saker av yttersta stilistiska vikt. Därför frågar jag efter likheterna.

Ofta leder resonemang som dessa till att en skala får agera modell för resonemanget, där extremerna redan på förhand tycks uteslagna, vilket jag tycker tyder på att skalan inte helt fångar det som ska studeras. Dessutom leder skalor till att skillnader ska kunna synas utplacerade på dem, men samtidigt går det inte att dra gränser mellan dem. Inte heller nu vare sig kan eller vill jag skapa sådana flyktiga gränser. Genom att studera både likhet och variation hoppas jag kunna fånga vad i lättläst som verkar skilja sig verken emellan, och vad som verkar hålla sig inom ungefär samma område. De exakta gränserna för vad som är vanligt och vad som sticker ut får siffrorna visa, och är inget jag bestämmer på förhand.

Ofta när lättläst studeras bedöms de efter hur enkla eller begripliga de är, som till exempel Camilla Forsberg magisteruppsats "Myndigheten, texten och läsaren" (2012). Det kan ske dels genom textanalys och dels genom användartester. Jag studerar istället variation och likhet. Förstås finns en aspekt av att en text med hög variation kan innebära att texten är svårare, på det viset att den kräver att läsaren kan tolka fler semiotiska resurser. Men det rymms också något som kan tyckas motsägelsefullt: om en text har få resurser kan betydelskillnader istället försvinna och försvåra läsarens tolkningar. Dessutom kan dubbeltydighet skapas, när det är oklart vilken funktion en resurs har. Utifrån detta säger textens variation och dess likhet något om vad den kräver av läsaren. För att fånga det som varierar men döljs under samma yta utför jag en stilistisk läsning av verken, under rubrikerna "Stilistiska kommentarer" i mina tre analyskapitel.

4 Material

Här presenterar jag de tolv verk jag undersöker. De är från förlagets sida indelade i tre nivåer som innebär vissa uttalade skillnader, så först går jag igenom dem för att sedan gå igenom verken lite närmare.

Jag kallar böckerna för *verk* då det är ett ord brett nog att inrymma de olika typer av genrer som ryms innanför deras pärmar. Åtta av verken är romaner, där en berättelse om människor och händelser som inte inträffat (men nog skulle kunna ha inträffat) skildras.

Fyra verk (Abr1, Lag2, Eins3, Liv3) är av biografisk karaktär. Deras medverkan i materialet förklaras av att de kan sägas ingå i genren *litterär sakprosa*, efter Johan Tønnessons definition. Verken är förlagspublicerade av namngivna författare, och alla språkliga medel står till förfogande för att berätta berättelsen, så länge den ”direkte tilknytningskontrakten til virkligheten er opprettholdt” (2012:34). Kungliga Biblioteket definierar litterär sakprosa snarlikt: ”prosa som använder litterär stil och teknik för att på ett berättande sätt presentera ett fackämne på ett fängslande och levande sätt”.

För att benämna alla verken kallar jag genren för *litterär prosa*, då jag vill skilja dem från (icke-litterär) sakprosa.

4.1 De tre nivåerna

Lättläst har fått kritik för att vara en one size fits none-lösning, vilket Hanna Sofia Rehnberg tar upp i sin artikel i Språktidningen 2012:

Ett grundläggande problem är att en och samma text ska passa för alla med lässvårigheter. Det är sällan praktiskt genomförbart att skriva specifika texter anpassade för var och en av de olika undergrupperna, trots att de egentligen har olika behov. (2012)

Fyra år senare, till den här uppsatsen har LL-förlaget utgivning av lättläst litteratur på tre olika nivåer. Vad jag än försöker konstatera om det politiska läget kring verken 2016 kommer det troligen förändras i rask takt. Men verken kommer att bestå. Därför undersöker jag dem, då de oavsett politisk diskussion produceras och läses.

Nivåerna benämns av förlaget dels med adjektiv: ”lättast”, ”lättare” och ”lätt” och dels med siffror: nivå 1, 2 och 3. Nivå 1 är alltså de verk som anses lättast, nivå 2 är en mellankategori med ”lättare” verk, och ”lätt” är nivå 3 där de svåraste verken i utgivningen finns. Jag kommer att hänvisa till dem som ”nivå 1”, ”nivå 2” och ”nivå 3” för att förhoppningsvis undvika (min egen) förvirring kring att *lätt* betyder *svårast*.

Vad som är typiskt för nivåerna beskriver förlaget själva på sin webbplats. Jag har sammanställt beskrivningarna i nivåer i tabellen nedan, citaten i helhet återfinns i Bilaga 1: Nivåernas innehåll. Sammanställningen är indelad i tre kategorier för att återkoppla till mina frågeställningar, alltså lexikon och syntax, interpunktion och rytm, och färg och illustrationer.

Tabell 1: Skillnader mellan de tre nivåerna enligt LL-förlaget (u.å).

	Nivå 1	Nivå 2	Nivå 3
Lexikon och syntax	Enkla meningar. Utselas i vardaglig miljö. Få personer. Enkla ord. Korta meningar. Mycket lite text.	Enkel handling. Vanliga ord och uttryck. Korta meningar.	Ovanliga ord. Bisatser. Bildspråk. Handling kan hoppa i tid. Många personer.
Rytm och interpunktion	Korta rader. Kapitelböcker.	Väl disponerad handling.	Kapitelbok.
Färg och illustrationer	Bilder bär berättelsen. Alltid många illustrationer	Oftast illustrerade.	Fackböcker är alltid illustrerade.

Noterbart är att det tas upp olika parametrar för de olika nivåerna. För nivå 1 och 3 nämns antalet personer, medan nivå 2 inte verkar ha rekommendationer för detta. Däremot går det förstås att anta att antalet personer skulle ligga mellan nivå 1:s ”få” och nivå 3:s ”många”. För nivå 3 nämns inte hur ofta icke-fackböcker är illustrerade, men det går att anta att det är mer sällan än nivå 1 och 2. Huruvida verken på nivå 2 är kapitelböcker framgår inte heller.

Nästa fråga gällande nivåerna är vad som ryms i de olika beskrivningarna. Hur avgörs om en handling är enkel? Språkvetare vet att det inte är enkelt att avgöra om ett ord är enkelt, så exakt vilka ord som inryms i ”enkla” ord framgår inte. Däremot går det att jämföra med ”ovanliga ord” för nivå 3 och anta att ”enkelt” oftast betyder ”vanligt”. Intressant är att nivåerna på detta vis, trots att de nämner olika parametrar, fungerar i relation till varandra. Detta leder till att jag kan ställa upp följande hypotes:

Hypotes 1: Det finns skillnader mellan nivåerna på de tre områden jag ställt upp.

Vissa av parametrarna påminner mycket om parametrarna för att räkna ut LIX, och det finns säkert också här en tanke om att ett lätt innehåll hänger samman med specifika språkliga former.

4.2 De tolv verken

Materialet består av tolv verk jag fått tillgång till via LL-förlaget.

Som utgångspunkt i materialinsamlingen hade jag att verken skulle vara nytgivna. Det var viktigt för att lättläst hela tiden förändras, och jag ville att mina resultat skulle kunna generaliseras till det nutida arbetet med lättläst. Vidare ville jag ha verk av olika författare så att en författares individuella stil inte skulle råka misstas för gängse lättläst stil.

Verken ska inte vara facklitteratur eller brukstexter (alltså inte myndighetsmaterial, bankhemsidor, tidningar). Verken jag fick tillgång till är till stor del skönlitterära. Vissa av titlarna är snarare biografiska (*Anna på Brännö*, *Ny i laget* och *Mitt liv med Asperger*) och med fackkunskap likt en lärobok i vissa delar (*Albert Einstein*).

Tabell 2: Titlar och mina förkortningar

Nivå 1 (lättast)	Nivå 2 (lättare)	Nivå 3 (lätt)
Anna på Brännö, <i>Abr1</i>	Bengtssons hemlighet, <i>Behe2</i>	Albert Einstein, <i>Eins3</i>
Katten som inte ville dö, <i>Katt1</i>	Kickboxaren 2 – Fightern, <i>Kick2</i>	Bedragaren, <i>Bedr3</i>
Ny i butiken, <i>But1</i>	Ny i laget, <i>Lag2</i>	Mitt liv med Asperger, <i>Liv3</i>
I samma båt (1), <i>Båt1</i>	I samma båt (2), <i>Båt2</i>	Utvandrarna, <i>Utv3</i>

Här beskriver jag kort verken i tur och ordning. Varning för spoilers.

4.2.1 Nivå 1-verken



Anna på Brännö från 2016 med text av Agneta Slonawski och fotografier av Bibbie Friman. Läsaren får följa Anna som bor på Brännö. Vi får ta del av hennes hem och arbete. Anna berättar om vänner och kollegor, och några av dem pratar om Anna. Vi får höra hur det är att ha boendestödjare, och om en gång när Anna bråkade med en boendestödjare. Ett kapitel handlar om Annas födelsedagsfirande och ett annat om midsommarfirande.

Katten som inte ville dö från 2015 av Tony Samuelsson. Läsaren träffar familjen Andersson som upptäcker att vad de tror är deras katt Tiger ligger död längs med vägen. De blir ledsna och begraver katten i trädgården. Plötsligt kommer Tiger hem, vilket gör familjen glad och lättad. Pappa Peter går upp mitt i natten och gräver upp graven på misstanke att Tiger återuppstått, men kroppen ligger kvar. Katten i graven visar sig vara grannens.

Ny i butiken från 2014 av Sarah Andersson och Lars Lindqvist med bilder av Helena Davidsson Neppelberg. Verket är en del av en serie som heter *Sjöbergs affär* där läsaren får följa olika personer med koppling till affären i olika verk. *Ny i butiken* handlar om Helenas, som redan jobbar i affären, son Kalle som vantrivs i skolan och istället får göra en praktikperiod i affären. Han gillar inte arbetet i affären heller men det är bättre än att gå i skolan. När han träffar Karolin som jobbar extra och är duktig i skolan blir han motiverad att rita mer och satsa på det, och också därmed anstränga sig mer i skolan. Verket avslutas med att alla är på fest hos chefen.

I samma båt från 2013 av Sara Wadell med illustrationer av Peter Bergting. Verket finns i dubbel uppsättning i mitt material, en version på nivå 1 och en på nivå 2. Läsaren får följa Hassan som inte har något jobb och mår dåligt. När Hassan tidigare bodde i Iran byggde han båtar, och han är mycket stolt över det men har inte kunnat fortsätta med det i Sverige. Han börjar arbeta ideellt för att hjälpa hemlösa. Där bråkar han med Bosse som dricker mycket och anklagar Hassan för att som invandrare kosta samhället pengar. Bosse försvinner plötsligt, men Hassan träffar honom igen på en camping på sommaren. Bosse har börjat jobba med att bygga båtar, och han och Hassan kommer nu överens. De bestämmer sig för att börja bygga båtar tillsammans.

4.2.2 Nivå 2-verken



Bengtssons hemlighet från 2015 av Camilla Wallqvister. Verket handlar om Bengtsson och Johanna som bor på samma landsväg. Bengtsson är gammal och tycker att Johanna är nyinflyttad trots att hon bott där i tre år. Johanna har nyss skaffat får som rymmer till Bengtssons trädgård, vilket gör honom arg. Han är så arg att han prejar Johanna, i tron att det är några sommarturister, när hon letar efter fåren, ner i diket och hotar att slå sönder bilen och ha ihjäl henne. Han tar reda på att det är Johannas får och sprutar konstgödsel över hela hennes hem. Johanna är efter detta mycket sur på Bengtsson och önskar att han vore död. Bengtsson tar en simtur men får kramp och tror sig drunkna, men Johanna hör ropen på hjälp och kommer till undsättning. Efter viss tvekan när hon ser vem det är räddar hon honom. Bengtsson lämnar älgkött på hennes trappa som tack, vilket hon trots den tidiga timmen ser och bjuder på kaffe vilket han tackar ja till.

Kickboxaren 2: Fightern från 2015 av Niklas Krog och Maria Björn. *Kickboxaren 2* är en fristående fortsättning på *Kickboxaren* som handlar om Eremias. Irina tränar kickboxning och blir tillsagd att hon kommer förlora om hon låter ilskan ta över. Hennes styvpappa Ronny dricker mycket alkohol och är elak mot Irina, mamman och systemen. För att få pengar till träningen slår Irina ner en tjej och stjäl hennes väska, som sedan Salim på samma skola köper av henne. Ronny misshandlar mamman och Irina ingriper så slagsmålet är mellan henne och honom istället. Eremias kommer tillbaka till träningen efter att ha varit avstängd på grund av slagsmål, vilket Irina blir rädd ska hända henne. De blir tillsagda att hålla i en träning tillsammans. De planerar den hemma hos Eremias där det är mycket trevligare än hos Irina. Tjejen Irina misshandlade är på samma köpcentrum som Irina och Eremias, och flyr skrämt när hon ser Irina. Eremias ställer frågor och gissar att Irina slagit henne och att det inte är bra hemma. Irina utmanas av tränaren som ska avgöra om hon är redo för match. Hon håller sig lugn och blir godkänd. När Irina kommer hem gömmer sig lillasystemen i hallen efter att föräldrarna bråkat, och Irina tar direkt med sig systemen därifrån och möter upp Eremias. Verket slutar med att Eremias mamma kontaktar instanser som kan hjälpa Irina och systemen, och att Eremias övertalar Irina att försöka göra rätt mot tjejen hon misshandlade.

Ny i laget från 2015 av Mats Wänblad och Pekka Pääkkö. Läsaren får följa Malin som är nyinflyttad och ensam i Stockholm. Hon har tidigare spelat fotboll och provspelar för ett lag för att kanske börja igen. Hon trivs dock inte helt, men funderar i några dagar. Hon bestämmer sig för att börja spela i laget och blir glatt överraskad över att alla i laget är mer välkomnande efter hennes beslut. Vi får också följa med på en match.

I samma båt från 2010 av Sara Wadell. Samma handling som för nivå 1.

4.2.3 Nivå 3-verken



Albert Einstein: Det modiga geniet från 2015 med text av Bengt Fredrikson och bild av Erica Jacobson. Läsaren får Einsteins liv berättat för sig. Först är Einstein barn och olika problem med skolan beskrivs. Han studerade matematik utanför skolan. Föräldrarna flyttar till Italien och lämnar kvar honom ett år i Tyskland för att slutföra sin skolgång, vilken han avbryter och åker efter dem. På tekniska högskolan hamnar han återigen i bråk med lärarna, men studerar på egen hand och tillsammans med flickvännen Mileva. Efter examen försöker han söka tjänster på universitet men får inga, och jobbar istället på patentverket där han trivs. Han läser fysikers upptäckter och försöker lösa problemen de hittar i de gamla förklaringsmodellerna. Han publicerar artiklar vars innehåll refereras i gula rutor för läsaren. Efter artiklarna får han arbeta som lärare i fysik, och flyttar till Prag vilket inte blir bra för familjen. Så fortsätter Einsteins liv berättas, där vetenskapliga upptäckter, familjeliv och internationella oroligheter blandas. Verket slutar med att han dör.

Bedragaren från 2016 av Sara Lövestam. Återberättad av Sara Lövestam utifrån originalet *Önska kostar ingenting*. Det är del två i serien om den papperslösa detektiven Kouplan. Han behöver få in pengar till hormonmedicin. Han får i uppdrag att hitta borgarrådet Jenny Svärds flickvän Amanda som försvann med 200 000 kr – pengar det visar sig att Jenny fått som muta. Han upptäcker att Amanda är i Göteborg under ett annat namn, och åker dit. Han hittar henne. De visar sig båda gömma sig för polisen, arbeta svart och behöva pengar till medicin. Kouplan och Amanda blir vänner och hon tar honom till en doktor. Doktorn ringer senare Kouplan och säger att Amanda ska utvisas, men Kouplan hittar dem tillsammans på centralstationen och förstår att också han blivit bedragen. Kouplan åker tillbaka till Stockholm och planerar med Jenny att avslöja Amanda en gång för alla. Han spårar henne till Jönköping där hon har ny identitet. På sjukhuset får de veta hennes riktiga namn, Nathalie, och åker till hennes mamma i Sala. Där är också Nathalie, och Jenny börjar bråka med henne. I upplösningen kommer Jenny och Nathalie under hot om polis överens om hur pengarna ska delas, Jenny och Kouplan skiljs åt.

Mitt liv med Asperger från 2014 av Jerker Jansson med foto av Ulf Lundin. Jerker berättar om hur hans Asperger påverkar honom, och hur det påverkar hur han blir bemött av andra. Han berättar hur det var att utredas för Asperger, om uppväxt och olika situationer. I slutet tipsar han om vilken hjälp som finns att få.

Utvandrarna från 2015, återberättad av Cecilia Davidsson utifrån Vilhelm Mobergs original. *Utvandrarna* är den klassiska berättelsen om Karl-Oskar och Kristina som försöker skapa sig ett bra liv i Småland, men jorden är full av sten och deras dotter dör av svält. Tillsammans med några andra bestämmer de sig för att utvandra till Amerika. Alla har olika anledningar, men dåliga skördar, dåliga chanser till rimligt arbete och religiöst förtryck är vanligt. Resan över Atlanten är besvärlig, och gravida Kristina blir sjuk i skörbjugg. Verket slutar med att de ska gå i land i Amerika.

4.3 Korpuslingvistik

Eftersom jag använder mig av ett material som är för stort för att manuellt gå igenom, och därmed använder datorer för att analysera det, blir fältet korpuslingvistik relevant att ta upp. Korpuslingvistik handlar om hur stora språkmängder kan sammanställas och analyseras för att säga något om språk. Detta för att kunna diskutera vad mitt material egentligen säger om språket inuti och utanför verken.

Jag går nu igenom några grundläggande begrepp inom korpuslingvistik och kommenterar mitt material utifrån dem: maskinläsbarhet, autenticitet, representativitet, urval (engelska: *sample*), balans, specialiserad korpus.

Maskinläsbarhet handlar om att materialet ska gå att hantera med datorer, helt enkelt för att det är det enda sättet att hantera stora mängder data. Det räcker inte med att ha materialet, vi måste kunna göra något med det också. Ofta ingår det att nedteckna vad det är för texttyp, ursprungsår och dylikt, liksom ibland annotering av språkliga kategorier som ordklass. (McEnery, Xiao & Tono 2006:5)

Mitt material är maskinläsbart, med vissa brister som jag går igenom under rubriken ”programvara”.

Autenticitet handlar om att materialet ska vara skapat för ett annat syfte än att studeras som språkmaterial. Detta för att lingvisten inte ska bevisa sina teorier med material skapat efter samma utgångspunkter (McEnery, Xiao & Tono 2006:5). Mitt material är skapat för att läsas av människor som inte är språkvetare.

Representativitet handlar om att det material som ingår i korpusen ska kunna gå att generalisera till texttypen eller språkvarieteteten som helhet. En korpus med tidningstexter från förra året är inte representativ om vi vill studera ungdomars bruk av diskurspartiklar. (McEnery, Xiao & Tono 2006:13) Mitt material är representativt för modern lättläst litteratur då det är det det består av.

Urval handlar om att en korpus omöjligt kan innefatta allt språk, precis som en enkätundersökning inte innefattar hela befolkningen den studerar. Ett urval kan anses vara bra om de slutsatser som går att dra ifrån det också visar sig gälla för texter som ska representeras av det och ligger utanför urvalet (McEnery, Xiao & Tono 2006:19).

Balans handlar om att uppnå representativitet. För att uppnå representativitet bör korpusen, om det är en allmän korpus som syftar till att visa på en brett definierad språkvarietetet, vara balanserad. För att en korpus ska vara balanserad ska urvalet så att säga inte väga över för någon del. Om modern svenska ska studeras bör inte bara skönlitteratur, eller för den delen bloggtexter, vara med eller uppta majoriteten av innehållet i korpusen. Det går inte att mäta eller bedöma korpusbalans, utan forskare beskriver urvalet de gjort och varför och låter sedan läsarna själva bedöma om det verkar rimligt (McEnery, Xiao & Tono 2006:17).

När det gäller en **specialiserad korpus** syftar den till att möjliggöra studier av mindre språkliga varieteter och kan därmed vara mindre. Det material jag sammanställt för den här uppsatsen fungerar som en specialiserad korpus, då den ska representera lättläst, modern litterär prosa från MTM. Balans bedöms annorlunda för specialiserade korpusar. En specialiserad korpus anses komplett när den är mättad (*closure* och *saturation* används på engelska). Den är mättad när varje nytt tillägg om säg tusen löpord bidrar med lika många nya typer (ungefär: ord, mer om detta senare) som föregående tusen löpord (McEnery, Xiao & Tono 2006:16). Därmed inte sagt att en mättad korpus skulle vara stabil eller visa på absoluta sanningar om de högfrekventa orden. Att dubblera mängden data i en korpus visar sig lätt ändra ordningen på de 25 mest frekventa orden (Kornai 2002:66). Eftersom jag inte använder mitt material för att visa på vilka språkliga enheter som används sammantaget i alla blir inte diskussionen om mättnad relevant.

4.4 Metod för att hantera materialet

Här skriver jag om en del tekniska detaljer, som nog främst är intressant för den som vill replikera mitt arbete. För den som vill hantera mer än några sidor av språk finns det sällan ett givet alternativ, tyvärr.

Jag valde bort att räkna ord, tecken och dylikt i Microsoft Word och andra ordbehandlare då jag sedan tidigare upplever dem som långsamma och något otillförlitliga. Jag tycker det är viktigt att alla verk får samma behandling, vilket jag inte ansåg att ordbehandlare kunde garantera. Dessutom vill jag använda mig av reguljära uttryck (se nedan) för att kunna identifiera vissa fenomen i verken.

Jag hade tillgång till verken i pdf-format, där texten gick att läsa maskinellt (och alltså inte som bild). Däremot gick den inte att kopiera in i ett textdokument med vanliga kopiera–klistra in-funktioner. Därför genomgick alla verken en extrahering för att få fram råtexten via programmet Pdf to text (googla ”pdftotext”) som är en del av Poppler-sviten.

Med hjälp av programmeraren Linn Engström (tack) blev det ett terminalprogram i Python som alla verken körs igenom. Python har en del funktioner som lätt räknar tecken vilket underlättade arbetet. Programmet körs på ett verk i taget.

Steg ett är ett kontrollsteg, där texten som programmet har identifierat ska kollas. Då kollar jag manuellt att början och slutet stämmer med hur original-pdf:en ser ut, och att skiljetecken efterföljs av mellanslag. Dessutom läser jag några stycken för att kontrollera att fel inte förekommer.

Steg två innebär att programmet visar alla de ord som enligt det reguljära uttrycket för egennamn identifieras (strängar med versal begynnelsebokstav som inte står i början av mening, mer om detta under Lexikon och syntax – Analysenheter – Egennamn). Där går jag manuellt igenom listan och väljer bort de ord som inte är egennamn men identifierats som det ändå, oftast för att de står som första ord efter en rubrik. Nu finns en korrekt kategori av egennamn.

Steg tre är själva räknandet, där programmet matchar alla reguljära uttryck och andra formler mot texten. Dessa siffror skrivs ut och jag ställer upp dem i tabeller med absoluta siffror i Excel.

De relativa siffror jag presenterar i resultatkapitlen är beräknade i Microsoft Excel, där också diagram och de flesta tabeller skapats (för att formateras till ett mer läsbart format i Word där uppsatsen är skriven).

4.5 Reguljära uttryck

Jag använder mig av en del reguljära uttryck för att studera materialet. Respektive uttryck redogörs vid dess analysenhet, och de finns samlade i Bilaga 2. Reguljära uttryck, på engelska *regular expressions*, ofta förkortat *regex* eller *regexp*, används för att identifiera vissa strängar i ett material.

Exempel: jag vill identifiera alla gånger bokstavssträngen ”Malin” används i en text. Då kan jag i många ordbehandlare och pdf-program söka på det. Säg att jag istället vill söka på alla ord med fem tecken där den första är versal. Då kan jag inte gissa mig till alla tänkbara ord, utan måste låta det vara öppet. Då skriver jag ”[A-ZÅÄÖ][a-zåäö]{4}” i ett verktyg som kan hantera reguljära uttryck. Det ska läsas så här: först en versal som är <a> till <z> eller <å>, <ä> eller <ö>. Därefter motsvarande gemener, 4 stycken.

Bokstäverna <å>, <ä> och <ö> har lagts till som unika tecken eftersom reguljära uttrycket ”a-ö” aldrig funkar, då programmen tenderar att analysera <ö> som <o>, och <å> och <ä> som <a>.

5 Lexikon och syntax

Det här är det första av tre kapitel som försöker besvara frågeställningarna. Det här kapitlet försöker besvara Frågeställning 1: Vilken variation och likhet finns det i lättlästa litterära prosaverks **lexikon** och **syntax**?

Lexikon och syntax säger något om textens variation. Innehåller den många olika ord kan den anses kräva ett större ordförråd av läsaren. Innehåller den många olika typer av syntaktiska former innebär det att texten är varierad. Det gör ofta en text enklare att hänga med i – enformighet engagerar knappast.

Kapitlet är uppdelat i tre delar, som kan ses som metod, resultat och avslutning. Först en del där jag går igenom de analysenheter jag använder mig av. I denna genomgång hänvisar jag till tidigare forskning som jag formulerar hypoteser utifrån. Därefter följer resultaten. Resultatdelen i sin tur är uppdelad i mindre delar. Först presenterar jag kvantitativa resultat för varje analysenhet. Därefter adderar jag varje verks värden för att kunna se hur det totalt sett tycks placera sig gentemot de andra när det gäller lexikon- och syntax-variation. Detta sker under rubriken Lexikon- och syntaxvärde (LSV). Därefter följer en stilistisk kommentar för varje verk, där instabiliteter som framgick i LSV-beräkningen kommenteras. Kapitlet avslutas med en sammanfattning.

5.1 Analysenheter och tillvägagångssätt

För den kvantitativa delen använder jag mig av en rad analysenheter. Genomgången här kan läsas som en förklaring till diagrammens nyckelord i resultatavsnittet.

De analyskategorier jag valt ut är: standardiserad typ/löpord, grafiska meningars längd, ordlängd, antalet långa ord, antalet löpord, antalet sidor, antalet pronomen, antalet egennamn och antalet konjunktioner.

Jag studerar inte radlängd automatiskt, trots att så kallat frasanpassat radfall är typiskt för de lättlästa verk jag studerar. Detta är en brist då det hade varit intressant, speciellt hur det korrelerar med annat. Det omöjliggjordes dock av att de verktyg jag använde mig av inte kunde läsa av pdf-filerna med verken i sig på ett sätt som möjliggjorde detta. Radbrytningarna var helt enkelt kodade på olika sätt inom samma verk och mellan verken att det inte gick att reda ut.

5.1.1 Standardiserad typ/löpord (STT)

Vad räknas?

Standardiserad typ per löpord (standardized type/token, STT) är ett mått på lexikonets storlek, alltså hur många olika ord som ryms inom en textmängd. Att mäta lexikonets storlek är inte enkelt att göra för hand ens i en liten textmängd. Det blir ännu svårare att göra i större mängder, men eftersom samma metod har använts på alla texterna torde felmarginalen också vara densamma så att resultaten ändå är jämförbara. Att mäta antalet löpord beskriver jag under rubriken Löpord hur det går till, och antalet typer räknas som följer:

Som typer (engelska *type*) räknas varje instans av en ny teckensträng. Samma sträng (ungefär: ord) räknas alltså bara en gång oavsett hur ofta den förekommer. I meningen ”Flickan i den rosa jackan satt

kvar och gungade, fram och tillbaka, fram och tillbaka.” (Båt2) finns det 11 olika strängar: *flickan* (1), *i* (1), *den* (1), *rosa* (1), *jackan* (1), *satt* (1), *kvar* (1), *och* (3), *gungade* (1), *fram* (2), *tillbaka* (2).

Böjningar och avledningar och dylika versioner av samma lexem kommer att räknas som olika typer. Utöver det brister analysen genom att helt förbise homografer, som alltså räknas som samma typ trots olika betydelser.

Hur har jag räknat?

Antalet typer har räknats ut automatiskt genom att alla unika ord togs ur gruppen löpord. Alla dubletter (och tripletter och så vidare) raderades alltså tills bara en förekomst av varje fanns kvar.

Eftersom verken i mitt material skiljer sig rejält åt i längd vore det orättvist mot de korta verken att jämföra dem med de långa, där fler ord onekligen får plats. Därför är verken uppstyckade i lika stora delar, om tusen löpord, och antalet typer räknade däri. För Abr1 återfanns följande mängd typer i de fyra tusen löporden (i storleksordning): 408, 413, 418 och 429. Jag placerade ut alla i just storleksordning för att kunna avgöra huruvida något verk helt plötsligt avvek, till exempel att orden tre till fyratusen skulle ha betydligt fler olika typer än resten av verken. Inget av verken hade sådana avstickare, och hela listan går att se i Bilaga 4. Om ett verk har 4500 löpord räknas inte de 500 sista, eftersom de inte går jämnt ut med tusen.

För att få fram ett standardiserat värde adderas värdena ihop och divideras sedan med antalet värden. För Abr1 innebär det: $408 + 413 + 418 + 429/4$.

Vad kan detta visa på? Vad används det för?

STT används för att jämföra lexikonets storlek mellan verken. Då MTM säger att nivå 1 ska ha ”enkla” ord, nivå 2 kan antas ha ”vanliga” och nivå 3 ”ovanliga” (se Tabell 1) är min hypotes som följer:

Hypotes 2: Nivå 3 har högst STT, följt av nivå 2, och nivå 1 har lägst STT.

5.1.2 Grafiska meningar

Vad räknas?

Som mening räknas det som finns mellan ordinledande Versal och meningsavslutande interpunktion.!?... Här finns viss felmarginal vid rubriker som ofta har inledande versal men inte avslutande interpunktion. Då har de räknats ihop med efterföljande mening. Till exempel här, där de första tre orden bildar rubrik, för att efterföljas av brödtext: ”Ronny kommer hem ³ Irina låg i sin säng när ytterdörren öppnades.” (Kick2). Rubrikerna är i alla verken tämligen korta och få och skulle därmed inte kunna förskjuta statistiken nämnvärt.

Hur har jag räknat?

Antalet meningar har räknats ut med hjälp av följande reguljära uttryck: ”[A-ZÅÄÖ][^\.!?]*[^\.!?]”.

Vad kan detta visa på? Vad används det för?

Antalet meningar använder jag för att räkna ut meningslängden. Det har gjorts genom att dividera antalet löpord med antalet grafiska meningar.

³ Snedstreck mellan mellanslag symboliserar radbrytning i den citerade texten.

Nivå 1 och 2 beskrivs ha korta meningar, vilket innebär att nivå 3 i så fall borde skilja sig från dem genom att ha längre meningar.

Hypotes 3: Nivå 1 och 2 har snarlik meningslängd, medan nivå 3 skiljer sig från dem och har längre meningar.

Holm fann att genomsnittliga meningslängden i samtida romaner varierade mellan 6,5 ord per mening till 16,7 (Holm 2015:224). Verken i mitt material borde ligga lägre, i alla fall lägre än de högsta, eftersom meningslängden rekommenderas vara kort för lättlästa verk.

Hypotes 4: Den genomsnittliga meningslängden är inte mer än 7 ord per mening.

5.1.3 Antal bokstäver

Vad räknas?

Varje bokstav (A–Ö) räknas.

Hur har jag räknat?

Antal bokstäver räknas ut automatiskt med den inbyggda Python-funktionen LEN som ställer upp alla bokstäver på en rad och sedan beräknar radens längd.

Vad kan detta visa på? Vad används det för?

Antalet tecken används för att räkna ut ordlängd.

5.1.4 Antal löpord och genomsnittlig ordlängd

Vad räknas?

Som löpord räknas varje instans av teckensträngar som separeras av mellanrum. I meningen ”Flickan i den rosa jackan satt kvar och gungade, fram och tillbaka, fram och tillbaka.” (Båt2) finns 15 löpord.

Hur har jag räknat?

Antalet löpord har räknats ut automatiskt med hjälp av reguljära uttryck. Uttrycket som använts är: `”[A-Za-zÅÄÖåäö0-9]+”`

Vad kan detta visa på? Vad används det för?

Antalet löpord i sig är inte så talande i absoluta tal, men i relativa tal jämfört med antalet typer och sidor säger det något om hur lång en sida och ett verk är. Det säger något om ordmängden som möter läsaren, och därmed om hur stor ordmängd skribenten antar att läsaren inte avskräcks av.

Antalet löpord används för att beräkna hur mycket text det är på sidorna genom att dividera antalet löpord med antalet sidor. Det kan visa på hur luftig texten är. Eftersom nivå 1 ska ha ”mycket lite text” ställer jag upp följande hypotes:

Hypotes 5: Nivå 1 har färre ord per sida än nivå 2 och 3.

Löpord och teckenberäkningen har använts för att räkna ut ordlängden. Antalet tecken har dividerats med antalet löpord, vilket ger ett genomsnitt.

Ordlängd använder jag för att kunna kommentera lexikonet. Om det skiljer väldigt mycket mellan verken kan det tyda på att lexikonet skiljer sig från varandra i verken.

Hypotes 6: Nivå 3 har längre ord än nivå 1 och 2.

Dessutom använder jag antalet löpord för att jämföra andra mängder gentemot.

5.1.5 Antal sidor

Vad räknas?

Som sidor räknas de sidor av verken som kan sägas tillhöra innehållet. Framsida, försättsblad, författares tack (Eins3), ordlistor (Utv3), tomma sidor vid pärmarna och dylikt räknas inte. Om det på sidan innan texten i första kapitlet finns en bild har den sidan räknats med, eftersom kapitelinledningar då ser ut så konsekvent i verken. Sidnumren i verken skiljer sig från mitt sidantal, då de ofta kan börja på till exempel 5 på första sidan med text. Mina sidhänvisningar till verken syftar till den siffra som står nederst i det tryckta verket.

Hur har jag räknat?

Jag har räknat genom att kolla på sidnumren i slutet av verket. I till exempel *Bedragaren* är sista utskrivna sidnumret längst ner på sidan 147, men sidan efter innehåller också text innan verket tar slut – alltså på sida 148. Första sidnumret som står utskrivet är 5, som också är första sidan som ingår i texten. Det är ett tväsidigt förord om huvudrollsinnehavaren Kouplan. De 4 sidor som föregår förordet räknas alltså in i de 148 som verket anser är del av texten, men jag vill inte räkna med dem. Därför räknar jag 148 minus 4, och antalet sidor landar på 144.

Vad kan detta visa på? Vad används det för?

Jag har räknat på mitt sätt för att kunna säga något om textens täthet i själva verken. Är det mycket text per sida, eller är den luftig? Förlaget uttrycker att de vill att sidorna ska se luftiga ut, enligt principen att en text som är tätt packad på en sida avskräcker en ovan och omotiverad läsare med sin blotta volym. Därför räknar jag ut vissa enheter per sida.

5.1.6 Pronomen

Vad räknas?

Pronomen är en ordklass som rymmer tämligen många ord, knappt tvåhundra uppskattar Svenska Akademiens *Språklära* det till (Hultman 2003:90). De flesta har dock inte räknats med. Jag har räknat på de pronomen som rör personer på ett direkt sätt, och kan sägas ersätta egennamn. Detta för att jag är intresserad av att studera hur mycket av en text som tycks handla direkt om personer och deras handlingar. De pronomen jag räknat är: *jag, sin, ni, min, mitt, han, hans, honom, hon, henne, hennes, vår, vårt, vårans, vi, er, ert, eran, de, deras, denne, dennes, denna, din, ditt, du, dens, dess*. Notera att *den* inte räknas, då träffarna för artikeln *den* bedömdes ta överhanden jämfört med pronomenen.

Hur har jag räknat?

Jag har använt följande reguljära uttryck för att identifiera matchande strängar:

```
”?<=&s)([jJ]ag|[sS]in|[nN]i|[mM](in|itt)|[hH](ans|ennes|enne|onom|an|on)|[vV](år|årt|åran|i)|[eE](r|rt|ran)|[dD](e|eras|enne|ennes|enna|in|itt|u|ens|ess))”(?=&s)”
```

Vad kan detta visa på? Vad används det för?

Alla pronomen uppskattas kunna uppta 10–25 procent av en text, med få pronomen i ”koncentrerat skriftspråk” och många i talspråksbetonat skriftspråk (Hultman 2003:90). Uppdelningen mellan

skriftspråk och skriftspråk med talspråkskaraktär är sällan en enkel uppdelning, men jag tror att mängden pronomen kan visa på hur mycket av texten som handlar om personer och deras handlingar. En text fokuserad kring människor lär även ha högre andel egennamn, trots att egennamn och pronomen sägs stå i komplementär relation till varandra (Hultman 2003:90). Det leder till följande antagande:

Hypotes 7: De verk med hög andel pronomen är också de som har egennamn bland de vanligaste orden och vars innehåll rör personer.

Jag presenterar antalet pronomen i relativ form, per tusen löpord. Detta genom att dividera antalet pronomen med antalet löpord, och multiplicera med tusen.

5.1.7 Långa ord

Vad räknas?

Jag räknar strängar med 6 eller fler bokstäver som långa. Det är en bokstav mindre än vanligtvis, där 7 sägs vara gränsen (Hellspong & Ledin 1997:71). Detta gör jag för att i beräkningen på så vis också kunna fånga hur många korta ord det finns. Egennamn räknas bort för att jag anser namn bara vara namn och inte variera gällande extensionens storlek.

Till exempel finns det i följande mening ”Jag är 51 år **gammal**, och jag **jobbar** med / att **skriva böcker**, spela musik och göra radio.” (Liv3 s. 7) fyra långa ord, i min fetning.

Hur har jag räknat?

Alla strängar med sex eller fler bokstäver identifieras. Från dem subtraheras alla instanser av egennamn som blivit definierade tidigare.

Vad kan detta visa på? Vad används det för?

Det har alltid varit och kommer alltid vara en godtycklig gräns var ett ord blir långt. Form och innehåll följs i språk ofta åt, och vanliga ord – tillhörande de stängda ordklasserna – blir ofta korta medan innehållsrika, specifika, ovanliga ord blir längre. Dessutom innebär svenska språket stora möjligheter till sammansättningar, och orden kan bli betydligt längre än sex tecken. Jag förväntar mig inte att någonsin hitta en text utan ord med sex eller fler tecken. Däremot förväntar jag mig att en hög andel långa ord kommer samverka med att texten i fler fall uppvisar tecken på variation. Därmed är mitt antagande som följer:

Hypotes 8: Ett verk med hög andel långa ord har också större lexikon och längre meningar.

Då nivå 3 enligt förlaget ska ”ovanliga ord” jämfört med nivå 1 och 2 antar jag följande:

Hypotes 9: Nivå 3 har högre andel långa ord än nivå 2, som i sin tur har högre andel än nivå 1.

5.1.8 Egennamn

Vad räknas?

Som egennamn räknas ord som föregås av gemener och inleds med versal för att följas av en eller flera gemener. En lista på vilka strängar som hittas genererades utifrån detta. Där finns också de ord som

följs av rubriker, eftersom rubrikerna som jag skrev ovan ofta inte avslutas med punkt. De ord som inte är namn, till exempel "När" raderades från listan med egennamn.

Hur har jag räknat?

Egennamn är notoriskt svåra att få fram via reguljära uttryck, men svenskan har fördelar framför engelskan (där de flesta svårigheter diskuteras) genom att vi mer sällan använder versaler för icke-egennamn av typen titlar, yrken, nationalitet, förkortningar. Därför vågade jag mig ändå på ett reguljärt uttryck baserat på principen att egennamn är ord som inuti meningar inleds med versal och följs av gemener. Reguljära uttrycket ser ut så här: "(?<=[a-zåäö]\s)[A-ZÅÄÖ][a-zåäö]+". Antal egennamn är varje instans av egennamn, så "Hassan" räknas många gånger i Båt1 & Båt2.

Vad kan detta visa på? Vad används det för?

Identifieringen av egennamn används på flera sätt. För det första exkluderades egennamnen från listan på "långa ord" då namn som "Sjöberg" (But1) inte kan anses ha mer innehåll eller komplexitet än "Katja" (Kick2).

För det andra räknas totala antalet träffar på egennamn fram, för att kunna jämföras med mängden personpronomen. Hur relaterar de till varandra? Innebär många namn få pronomen, eller tvärtom, eller följs de åt jämnt? Innebär hög andel egennamn att verket har tät koppling till verkligheten, där personer och platser är namngivna? Egennamn jämförs också med antal sidor, för att visa på hur personfokuserade texterna är. Nivå 1 ska ha få personer och nivå 3 många (och nivå 2 troligen därmed någonstans mellan), men det är oklart hur antalet namn fördelar sig i texten. En text med få personer kan ju handla mycket om olika personer utan att för den delen bestå av egennamn i hög utsträckning.

5.1.9 Konjunktioner

Vad räknas?

Jag räknar varje konjunktion i texterna. En konjunktion innebär en träff i statistiken. Konjunktionerna har jag sammanställt ur Svenska Akademiens *Språklära* (2003:68). Konjunktioner relaterar olika led till varandra. Jag har dock inte räknat med alla konjunktioner som finns. De särskiljande konjunktionerna *antingen*, *varken*, *vare sig* räknas inte med, då *eller* redan räknas och ingår i samma led, och därmed redan identifieras som sammanlänkande funktion. *Utan* räknas inte eftersom det sammanfaller med både adverb och preposition, likaså sammanfaller *för* och *så* med andra ordklasser.

De ord som räknas är: *och*, *samt*, *såväl*, *respektive*, *eller*, *men*, *fast*, *ty*. För *respektive* bedömer jag det osannolikt att substantivet skulle förekomma i så pass hög utsträckning att de skulle förvränga statistiken, likaså för adjektivet *fast*.

Hur har jag räknat?

Följande reguljära uttryck har använts för att identifiera konjunktioner:

```
"(?<=\s)[oO]ch|[sS]amt|åväl|[rR]espektive|[eE]ller|[mM]en|[fF]ast|[tT]y(?:=\s)"
```

Vad kan detta visa på? Vad används det för?

Det kan verka konstigt att vilja räkna på en ordklass som alltid förekommer på det här viset. Det stilistiska intresset ligger i att hög andel konjunktioner ofta innebär att texten är närmare talspråk än skriftspråk (Lagerholm 2008:118). Utifrån det går det att formulera följande antagande:

Hypotes 10: Verk med hög andel konjunktioner har också mycket direkt anföring.

Intressant vore om antagandet inte stämmer. Är det i så fall för att det anförda inte försöker ge sken av talspråk? Är det i så fall tecken på att texten inte har så stort stilistisk variation?

Det är lätt att föreställa sig att en text med hög andel samordning kommer ha längre meningar, vilket inte behöver innebära att texten skulle vara svårare eller enklare. Om meningslängden istället inte tycks korrelera med hög andel konjunktioner går det att gissa att meningarna inte i höge utsträckning innehåller samordnade led.

Hypotes 11: Verk med långa meningar har högre andel konjunktioner än de med korta.

5.2 Resultat

Här presenterar jag de resultat som gäller lexikon och syntax. Först presenteras relativa siffror, som värden för standardiserad typ/löpord och ordlängd. De presenteras i diagram, för att visa på hur verken förhåller sig till varandra för de olika analyskategorierna.

Därefter används förflyttningstabellerna för att skapa en sammanslagen bild av hur verken relaterar till varandra, under rubriken Lexikon- och syntax-värde. Där ställs de upp för att visa vilka verk som tycks ha mycket gemensamt och bildar en form av lättläst-medelväg, och vilka som tenderar att hamna utanför.

Dessa iakttagelser om vad som varierar och vad som tenderar att vara likadant leder in till den stilistiska analysen där jag studerar hur verken använder sig av olika semiotiska resurser.

5.2.1 Kvantitativa resultat

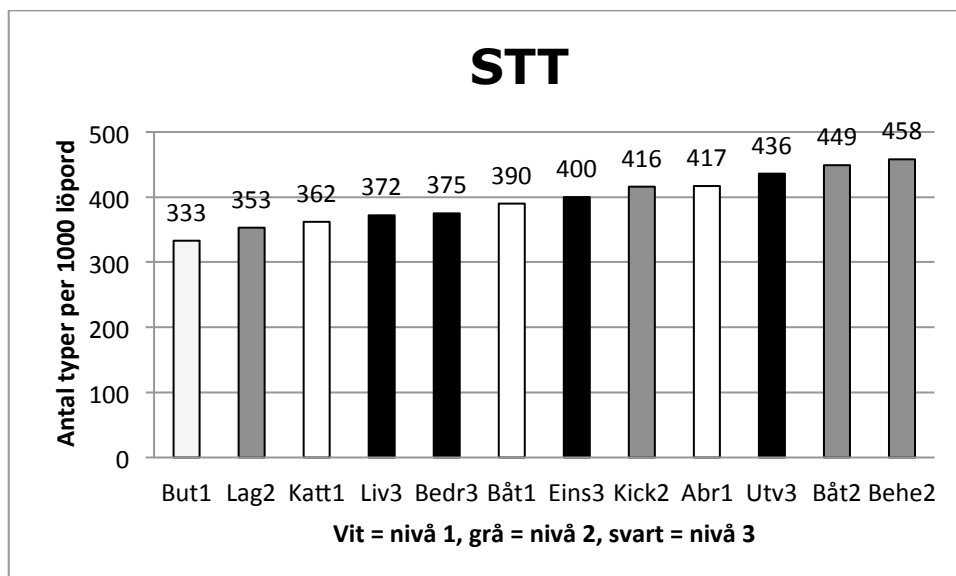
Här följer ett antal diagram. Verken i dem är placerade i storleksordning, där det verk som har högst värde för analysenheten ifråga är placerad längst till höger. Staplarna är färgade efter nivå, där vit stapel hör till nivå 1-verken, grå stapel till nivå 2-verken och svart stapel till nivå 3-verken. Om det finns tydliga mönster, att analysenhetens resultat följer verkens nivåer kommer det vara svart till höger, vitt till vänster och grått i mitten, medan om sådan struktur inte finns kommer färgerna att vara blandade, eller kanske till och med inverterade.

Positionen med lägst värde, längst till vänster, hänvisar jag till som ”position 1”, näst längst till vänster som ”position 2” och så vidare. Utöver det använder jag mig av *fack* för att beskriva de fyra vänstra positionerna (fack 1), de fyra mittenpositionerna (fack 2) och de fyra högra positionerna (fack 3).

Standardiserad typ/löpord (STT)

Här presenterar jag värdena för STT.

Figur 1: Lexikon och syntax, STT



Lägst STT har But1 med ett värde på 333 typer per tusen löpord i genomsnitt, och högst har Behe2 med 458. Noterbart är att Liv3 har flyttat till fack 1 från nivå 3, med sitt värde på 379, och att Abr1 har flyttat till fack 3 från nivå 1 med ett värde på 417. Fördelningen mellan lägsta och högsta är tämligen jämn.

Det går nu att avgöra om antagandet stämmer:

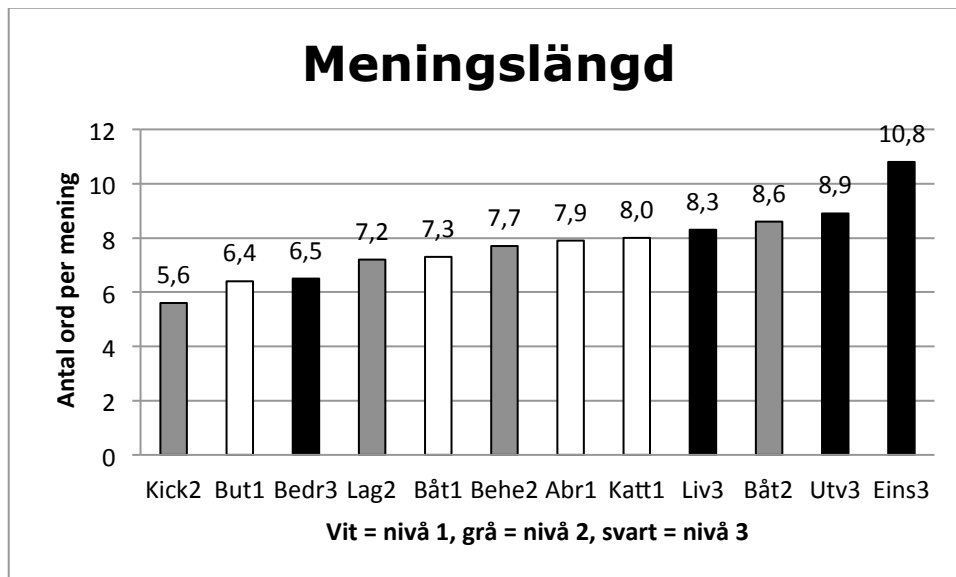
*Hypotes 2: Nivå 3 har högst STT, följt av nivå 2, och nivå 1 har lägst STT. **Stöds inte.***

Jag bedömer att hypotesen inte stöds, då det är två nivå 2-verk som har högst STT. Dessutom är det stor spridning mellan verk från samma nivåer, där alla nivåerna är representerade i varje fack.

Meningslängd

Här presenterar jag värdena för meningslängd.

Figur 2: Lexikon och syntax, meningslängd



Kortast meningar har Kick2, med ett genomsnitt på 5,6 ord per mening, och längst meningar har Eins3 med nästan dubbelt så långa: 10,8 ord per mening. Bedr3 är det verk som förflyttat sig över två fack, och hamnat i fack 1. Fördelningen är också här jämn, och inga stora glapp finns att hitta mellan verken. Störst glapp är det mellan Utv3 och Eins3 med 1,9.

Stöds hypoteserna?

*Hypotes 3: Nivå 1 och 2 har snarlik meningslängd, medan nivå 3 skiljer sig från dem och har längre meningar. **Stöds mestadels.***

Hypotesen stöds i stor utsträckning eftersom tre av verken med längst meningar är nivå 3-verk, medan verken från nivå 1 och 2 blandas i fack 1 och 2. Däremot placerar sig inte nivå 1-verken sig i det lägsta facket, men tenderar att ha lägre värden än nivå 1 och 2.

*Hypotes 4: Den genomsnittliga meningslängden är inte mer än 7 ord per mening. **Stöds inte.***

Hypotesen stöds inte eftersom bara tre verk har kortare meningar än sju ord. Däremot ligger många verk inom spannet 6–9, och inget verk kommer upp i så höga siffror som hos Holm (2015).

Ordlängd

Ordlängden visar ett litet mått av variation mellan verken, och varierar mellan 4,13 för Abr1 till 4,50 för Behe2. Jag bedömer variationen vara så liten att en uppställning i storleksordning bara skulle betona en skillnad som knappt förekommer.

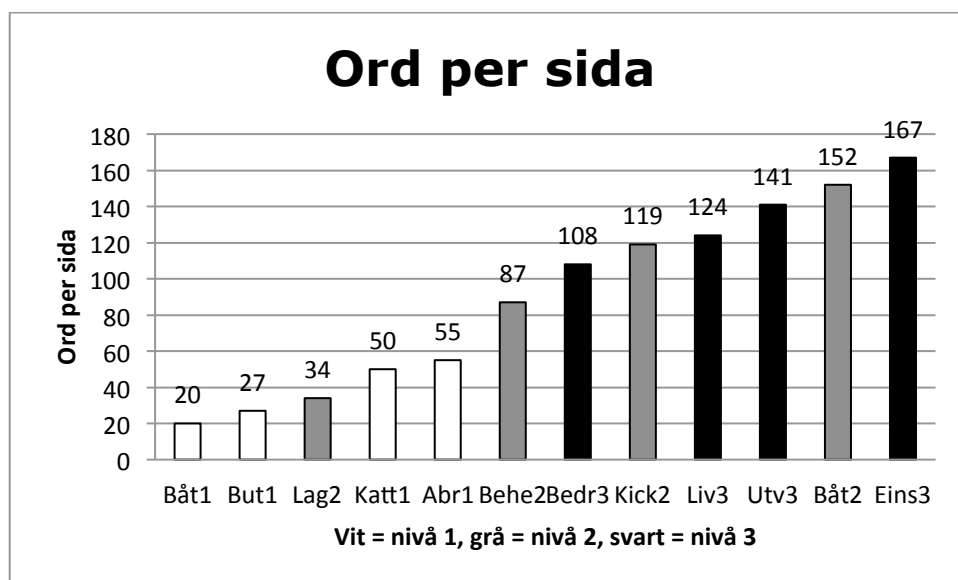
*Hypotes 6: Nivå 3 har längre ord än nivå 1 och 2. **Stöds inte.***

Hypotesen stöds inte eftersom sådana skillnader inte går att finna.

Ord per sida

Här presenteras genomsnittliga antalet ord per sida.

Figur 3: Lexikon och syntax, ord/sida



Båt1 har färst ord per sida med 20 ord i genomsnitt. Flest ord per sida har Eins3 med 167 ord, alltså drygt åtta gånger fler.

Här är fördelningen ganska ojämn. Det går att skönja ett hopp mellan Lag2 och Katt1 på 16 ord, och mellan Abr1 och Behe2 på 32 ord. På andra sidan Behe2 skiljer det 21 ord innan nästa verk, Bedr3, kommer. Behe2 ligger alltså ganska ensamt i mitten. Bland verken med flest ord finns det också skillnader, men de är av så pass liten andel jämfört med totala mängder ord att de nog inte framstår som lika stora skillnader för läsaren som lika stora hopp mellan de ordfattigare verken gör.

Det går nu att besvara hypotesen:

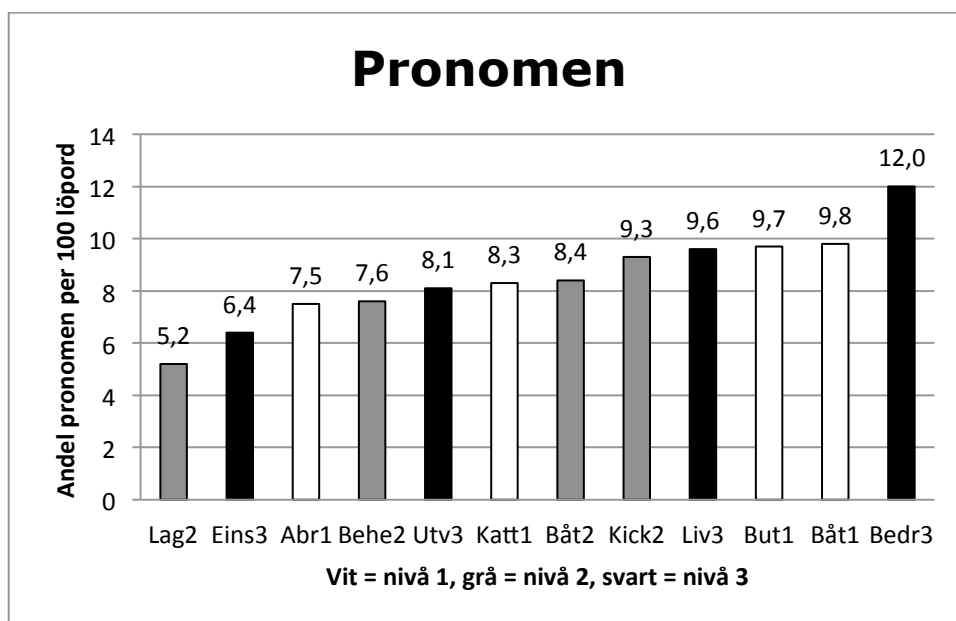
*Hypotes 5: Nivå 1 har färre ord per sida än nivå 2 och 3. **Stöds.***

Hypotesen stöds, då nivå 1-verken (inklusive Abr1 som placerar sig nära de andra) har lägre antal ord per sida än resterande, med undantag för Lag2.

Pronomen

Här presenteras andel pronomen av totala antalet löpord.

Figur 4: Lexikon och syntax, pronomen



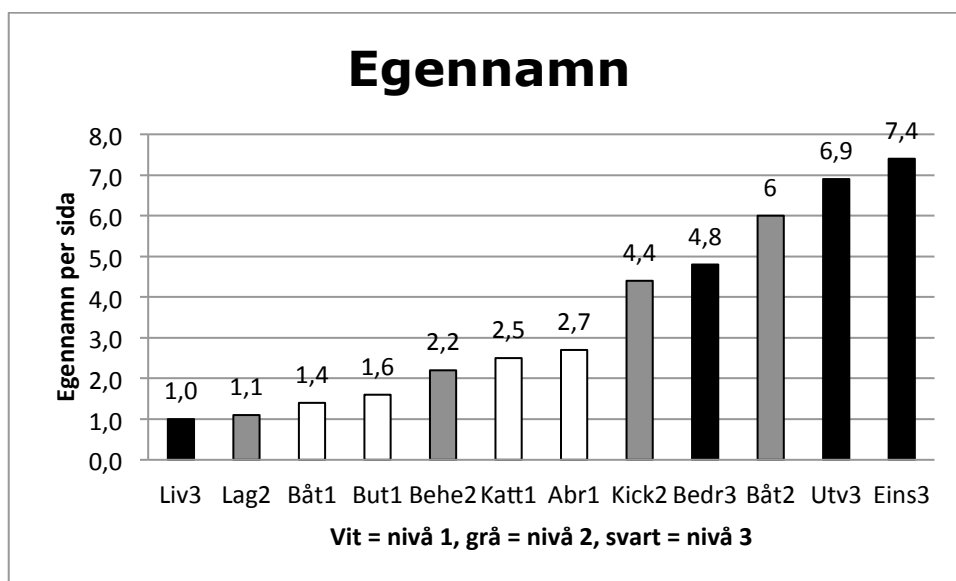
Lägst andel pronomen har Lag2, med 5,20 procent, och högst har Bedr3 med 12,03 – alltså mer än dubbelt så mycket. Det har skett stora förflyttningar och trenden verkar vara att verken på de låga nivåerna har många pronomen och verken på de högre nivåerna har få pronomen. Jag vill dock betona att de flesta verk klustrar sig mot mitten.

De flesta verk tycks ha pronomen mellan 7 och 10 procent, med Lag2 – i viss mån Eins3 – och Bedr3 som avvikelser.

Egennamn

Här presenteras antal egennamn per sida.

Figur 5: Lexikon och syntax, egennamn



Liv3 har färst egennamn per sida, med ett värde på 1,0, och Eins3 har flest egennamn per sida med sina 7,4. Liv3 har sjunkit två fack från nivå 3 till fack 1.

Det kan tyckas godtyckligt att räkna antalet egennamn per sida, allra helst eftersom få antal ord per sida borde sammanfalla med få antal egennamn. Liv3 är beviset på att det inte är så enkelt.

Långa ord

Här presenteras andel långa ord per tusen löpord.

Lexikonet i alla verk är snarlikt när det gäller andel långa ord. Alla verken har runt 200 långa ord per 1000 löpord. Variationen sträcker sig mellan 170 (Lag2) och 260 (Behe2). Jag bedömer att det råder stor likhet mellan verkens lexikon i detta avseende.

När texten blir längre ryms också fler långa ord, och läsarens förmodade intryck av en text kan därmed bli relevant. En läsare kommer möta många fler långa ord i vissa verk, eftersom de innehåller mer text totalt och per sida.

Hypotesen går nu att besvara:

*Hypotes 9: Nivå 3 har högre andel långa ord än nivå 2, som i sin tur har högre andel än nivå 1. **Stöds inte.***

Hypotesen stöds inte av materialet då sådana skillnader inte går att urskilja.

Det innebär att också Hypotes 8 går att besvara:

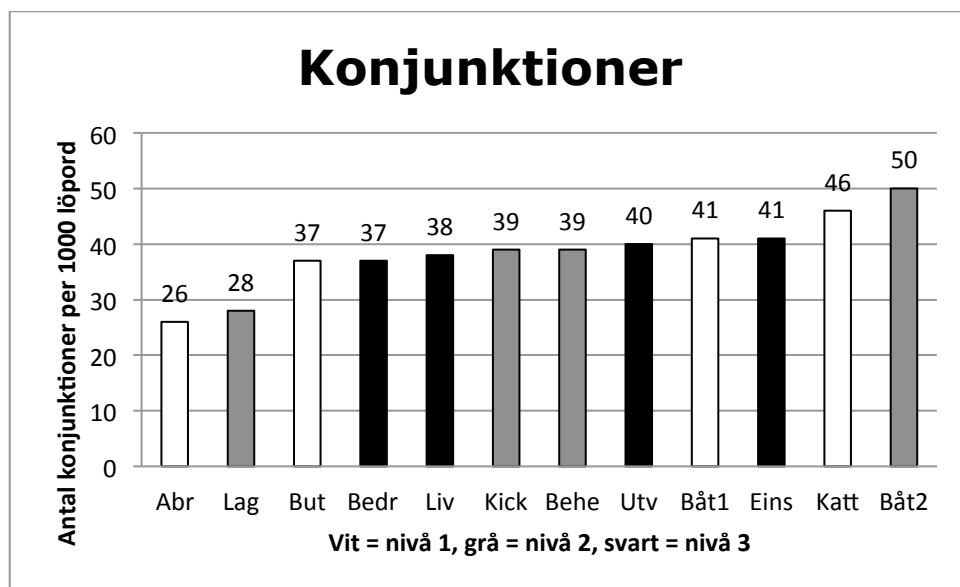
*Hypotes 8: Ett verk med hög andel långa ord har också större lexikon och längre meningar. **Stöds inte.***

Hypotesen stöds inte eftersom andelen långa ord är konstant.

Konjunktioner

Här presenteras antal konjunktioner per tusen löpord.

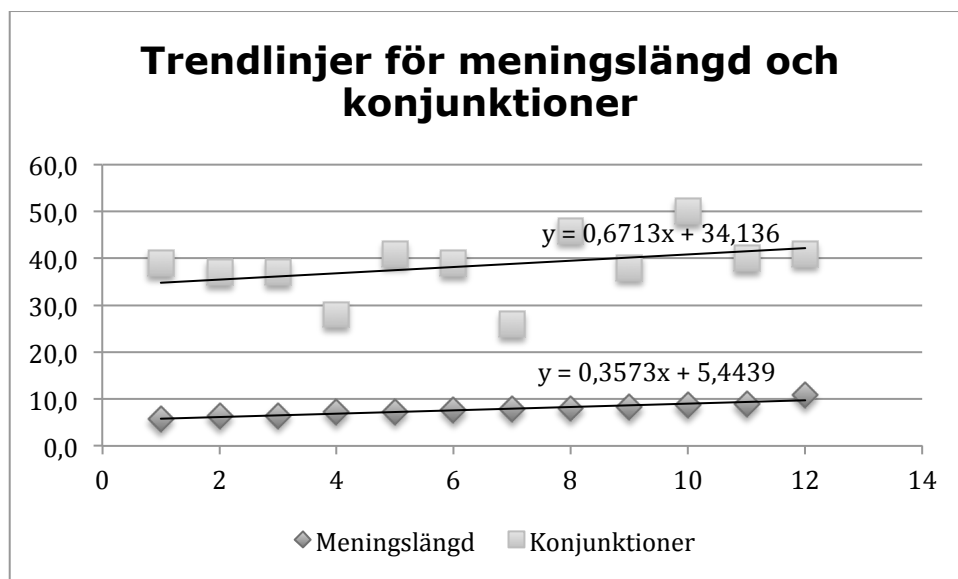
Figur 6: Lexikon och syntax, konjunktioner



Lägst antal konjunktioner har Abr1 med 26 per tusen löpord, och högst antal har Båt2 med 50. Ett stort mittenkluster kan anas, där de flesta verk har 37–41 konjunktioner per tusen ord.

För att besvara Hypotes 11 jämför jag meningslängd med konjunktioner. Det gör jag genom att rita upp trendlinjer för båda.

Figur 7: Lexikon och syntax, trendlinjer för meningslängd och konjunktioner



Romberna representerar meningslängden och ligger i storleksordning. Kvadraterna är konjunktionerna, som har placerats ovanför respektive verk. Det går att utläsa att när meningslängden ökar ökar också antalet konjunktioner. Verket med få konjunktioner på position 7 är Abr1.

Hypotes 11: Verk med långa meningar har högre andel konjunktioner än de med korta.

Stöds.

5.2.2 Lexikon- & syntaxvärde (LSV)

Här räknar jag ut ett sammanlagt lexikon- och syntaxvärde. Värdet speglar hur verken placerar sig i relation till varandra. Uträkningen sker i flera steg: först skapas en tabell som har samma innehåll som de ovanstående diagrammen, men med bara verken kvar i storleksordning. Därefter tilldelas verken utifrån den tabellen en rank för de olika analyskategorierna. Till sist adderas rankerna ihop till det jag kallar lexikon- och syntaxvärde, LSV.

Nu slår jag samman ovanstående tabeller, för att studera närmare vilka verk som placerar sig var.

Tabell 3: LSV-position

Fack	1				2				3			
	1	2	3	4	5	6	7	8	9	10	11	12
STT	But1	Lag2	Katt1	Liv3	Bedr3	Båt1	Eins3	Kick2	Abr1	Utv3	Båt2	Behe2
Meningslängd	Kick2	NiyB	Bedr3	Lag2	Båt1	Behe2	Abr1	Katt1	Liv3	Båt2	Utv3	Eins3
Ordmängd	Båt1	But1	Lag2	Katt1	Abr1	Behe2	Bedr3	Kick2	Liv3	Utv3	Båt2	Eins3
Pronomen	Lag2	Eins3	Abr1	Behe2	Utv3	Katt1	Båt2	Kick2	Liv3	But1	Båt1	Bedr3
Egennamn	Liv3	Lag2	Båt1	But1	Behe2	Katt1	Abr1	Kick2	Bedr3	Båt2	Utv3	Eins3
Konjunktioner	Abr1	Lag2	But1	Bedr3	Liv3	Kick2	Behe2	Utv3	Båt1	Eins3	Katt1	Båt2

Varje verk tilldelas en rank för vardera analysenhet. För STT får But1 1, Lag2 2, Behe2 12, och så vidare. Detsamma gäller för resterande analysenheter.

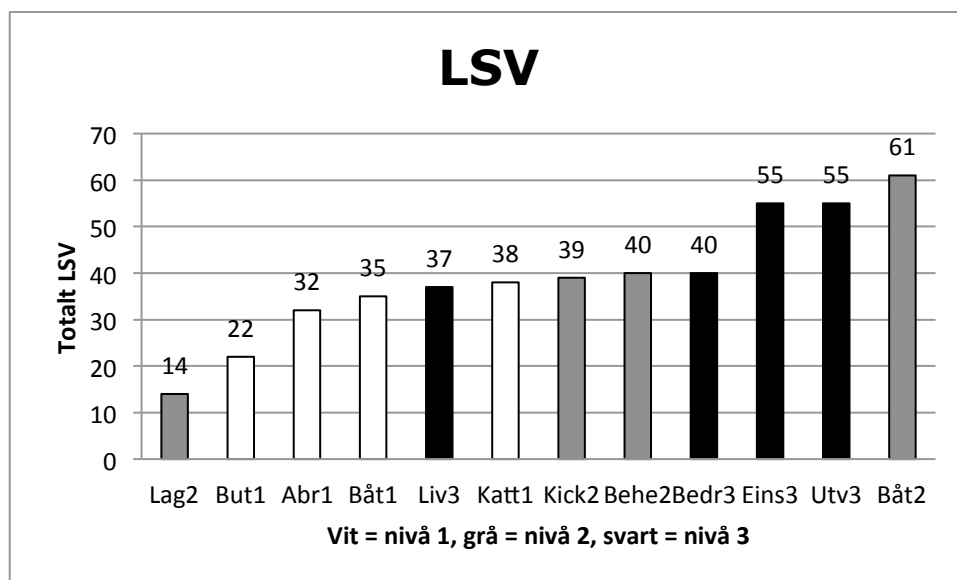
Nedan visas en uppställning av verken, deras positionsvärde för varje analyskategori och det totala LSV:t. Tabellen innehåller samma information (till och med mindre) som Tabell 3.

Tabell 4: LSV-rank

Verk	Abr1	Katt1	But1	Båt1	Behe2	Kick2	Lag2	Båt2	Eins3	Bedr3	Liv3	Utv3
STT	9	3	1	6	12	8	2	11	7	5	4	10
Meningslängd	7	8	2	5	6	1	4	10	12	3	9	11
Ordmängd	5	4	2	1	6	8	3	11	12	7	9	10
Pronomen	3	6	10	11	4	8	1	7	2	12	9	5
Egennamn	7	6	4	3	5	8	2	10	12	9	1	11
Konjunktioner	1	11	3	9	7	6	2	12	10	4	5	8
LSV	32	38	22	35	40	39	14	61	55	40	37	55

Utifrån den LSV-beräkningen hamnar verken i följande ordning:

Figur 8: LSV



Det blir tydligt att verken sinsemellan uppvisar stor variation, men också att vissa verk tycks vara mer lika varandra än andra. Tre kluster går att urskilja. Ett kluster där Lag2 och But1 placeras sig, och där verken tycks uppvisa liten variation gällande lexikon och syntax. Ett mittenkluster med LSV från 32 till 40 där de flesta verk placeras sig. Till sist ett högsta kluster där Utv3, Eins3 och Båt2 placeras sig, med värden från 55 till 61.

Generellt tycks nivå 1-verken ha låga LSV:n, medan nivå 2 och 3 blandas, med viss tendens för nivå 3 att placera sig högre. Båt2 på högsta position och Lag2 på lägsta position är båda nivå 2-verk och avviker från tendensen.

Med det sagt är alla verken i mitten inte lika varandra. Båt1 har kanske mer gemensamt med But1 än med Bedr3 och Liv3. Anledningen till att det ändå är relevant att placera dem tillsammans är att de tycks vara del av en större kategori där de flesta lättlästa verk hamnar, och där ytterligheterna hamnar utanför. Den vanliga lättlästa stilen finns troligtvis att finna i mittenklusterverken.

5.2.3 Stilistiska kommentarer

Här utgår jag från de instabiliteter som gick att uttolka från Tabell 3 och Tabell 4. Jag frågar nu de enskilda verken vad det är som händer där de uppvisar instabilitet. Verken presenteras i nivå-ordning.

Abr1: Nivå 1-verket i fack 2

Abr1 placerar sig nästan aldrig i fack 1, trots att det är ett nivå 1-verk. Istället beter det sig som ett stabilt nivå 2-verk när det gäller lexikon och syntax, förutom gällande antal konjunktioner och pronomen. Då är det intressant att notera att det vanligaste ordet i verket är ett pronomen: *hon*.

De vanligaste orden i Abr1 är: *hon, Anna, på, är, det, i, en, och, att, de, med, jag, har, till, som, om, när*.⁴

Att Abr1 placerar sig så lågt för pronomen har nog att göra med att verket främst handlar om en person, som omnämns som *Anna* och *hon* (och många av de andra personerna är också kvinnor och omnämns med *hon*). Att *jag* också placerar sig högt på listan kan förklaras med att det är mycket direkt anföring där *Anna* pratar om sitt liv och sig själv. En annan förklaring till att Abr1 har höga värden är att det handlar om många olika saker: födelsedagsfirande, arbete, att åka båt. Dessutom blandas repliker med berättande, vilket skapar stor variation i lexikonet och syntaxen. För att säga olika saker på olika sätt krävs olika resurser.

En förklaring till att Abr1 placerar sig högre än resterande nivå 1-verk är att Abr1 helt enkelt är längre, och att detta trots mina relativa siffror gör att mer variation har rymts innanför verkets pärmar.

Katt1: Få typer men långa meningar

Katt1 handlar bara om ett skeende och några få utvalda människor, vilket tycks förklara varför lexikonet är litet. Meningarna som möter läsaren på sidorna är inte iögonfallande långa, men meningslängderna varierar hela tiden vilket skapar utrymme för de relativt långa meningarna i verket:

1. Då såg Peter en mörk skugga närma sig Bea. Det var en katt. (Katt1 s. 24)

Korta meningar tycks användas för en del direkt anföring och när innehållet ska betonas, likt exemplet ovan. Därmed blir också den omarkerade meningen något längre.

De vanligaste orden i Katt1 är: *och, han, det, hon, Tiger, var, en, på, i, Peter, hade, Bea, att, som, inte, de, är, katt, katten, den*.

De namn och substantiv som bredvid småorden finns visar på verkets fokus på några få personer och deras katt (vid namn Tiger).

But1: Många pronomen

But1 handlar om många olika personer som alla gör och säger saker. Mycket av handlingen kretsar kring personer, vilket kan förklara att mängden pronomen är hög. Det tycks som om antagandet att många pronomen innebär få egennamn stämmer för But1. Här är ett citat som kan visa på hur egennamn och pronomen används i kombination:

⁴ Den uppmärksamma läsaren noterar att jag listar olika många ord när jag listar verkens vanligaste ord. Jag extraherade de tjugo vanligaste orden från verken, men då inkluderas i fallet Lag2 en del dolda tecken som verkar ha fungerat som instruktioner för något i tryckprocessen (eller faktiskt, rent skräp, det kan jag inte avgöra). Det som inte är ord av bokstäver har jag alltså tagit bort från listan på de vanligaste orden. Detta "skräp" tycks inte ha påverkat identifieringen av interpunktionstecken i nästa kapitel, vilket jag kontrollerat.

2. **Kalle** älskar att rita. / **Han** ritat flera timmar varje dag. / **Han** är duktig på det. (But1 s. 18) [Min fetning].

I övrigt är But1 ett kort verk med lite text på varje sida. Varje skeende ägnas få ord, vilket gör att det inte tycks finnas utrymme för variation.

De vanligaste orden i But1 är: *Kalle, på, och, han, är, i, hon, det, att, du, säger, jag, inte, Karolin, Helena, med, till, en, skolan, vill.*

Båt1: Korta sidor med många pronomen

Båt1 är en bok med lite text på sidorna, och med många bilder som tar upp en hel sida. Intrycket är därmed att boken är luftig. Berättelsen innehåller dock många delar och scener, och växlar mellan repliker, berättande i tredje person och att läsaren får inblick i huvudpersonens tankar:

3. Han är orolig. Den sista kvällen de träffades / var Bosse sjuk. / Tänk om det har hänt honom något? (Båt1 s. 31).

Båt1 är ett instabilt verk, med många pronomen. Antagligen – liksom för flera andra nivå 1-verk – är det många pronomen eftersom berättelsen i mångt och mycket är väldigt bunden till personer och deras handlingar.

De vanligaste orden i Båt1 är: *Hassan, är, han, och, en, Bosse, på, det, att, jag, har, de, i, du, som, men, till, båtar, ett, Noor.*

Egennamn, pronomen, prepositioner och konjunktioner står som vanligt högt. Båtar, centrala i handlingen, placerar sig också högt.

Behe2: Det stabila nivå 2-verket med störst ordförråd av alla

I Behe2 växlar handlingen mellan två olika personer. När läsaren följer Bengtsson ges det ofta inblick i hans tankar, i tredje person:

4. Ett lite brummande ljud. / Bengtsson blir nästan rädd. / Håller han på att bli tokig? / Eller kanske dement? / Han är ju inte precis någon ungdom längre. (Behe2 s. 27).

Likaså när läsaren följer Johanna:

5. Johanna är så rädd att hon nästan kräks. / Håller han på att skjuta ihjäl alla hennes får? (Behe2 s. 38)

Skiften i personer och perspektiv bidrar säkert till den höga typmängden.

De vanligaste orden i Behe2 är: *han, och, det, på, i, att, hon, Bengtsson, är, en, som, har, Johanna, till, med, sig, den, inte, de, upp.*

Att bara formord, förutom två namn, är de vanligaste orden tyder på att resterande ord uppvisar stor variation (både när det gäller vilka ord som används och i vilka former de används).

Kick2: Det stabila nivå 2-verket med kortast meningar av alla

Kick2 har många grafiska meningar som bara består av ett eller två ord. Ofta är det en upprepning av något från meningen innan, för att betona det:

6. – Jag kommer aldrig att veva mer, sa hon snabbt. / Aldrig. (Kick2 s. 9).

Dessutom finns det många meningar som inte är grammatiskt fullständiga, vilket förstås bidrar till att meningarna kan bli kortare:

7. Men Irina tänkte på framtiden. På sin kommande match. (Kick2 s. 10).

Texten är på så vis korthuggen, med många pauser. Rytmen kanske fungerar som en allegori för kickboxningen verket handlar om.

De vanligaste orden i Kick2 är: *hon, och, Irina, det, på, i, att, inte, en, var, han, som, med, sig, jag, men, hade, Eremias, sa, till.*

Lag2: Stabilt i fack 1

Lag2 har lägst LSV av alla verken, och placerar sig genomgående i fack 1. Att verket är ett nivå 2-verk är alltså inget som kan ha med lexikonet eller syntaxen att göra. Läsaren följer Malin som provtränar för ett nytt lag, funderar på hur det var och sen bestämmer sig för att gå med i laget. Det är alltså inte så mycket handling som ryms, och inte heller finns några perspektivbyten. Det bidrar säkert till den låga variationen.

De vanligaste orden i Lag2 är: *Malin, det, hon, är, att, i, och, de, om, har, en, på, laget, inte.*

Att innehållsordet *laget* ingår i de vanligaste orden tyder på att verket använder några få ord många gånger.

Båt2: Stabilt i fack 3, med få pronomen

Båt2 är det verk som har högst LSV, och alltså borde innehålla störst variation när det gäller lexikon och syntax, men har färre pronomen än andra verk.

De vanligaste orden i Båt2 är: *och, han, det, Hassan, att, på, en, i, jag, med, inte, du, var, är, som, de, för, till, om, sig.*

Bristen på pronomen syns inte i listan på de vanligaste orden, där *han, jag, du* och *sig* ingår. Det tyder däremot på att resterande ord, de som inte är pronomen är många, men också olika. De vanliga orden ackompanjeras alltså av ett stort lexikon som används i meningar som tillåts vara långa:

8. ”På soffbordet framför dem låg färgglada broschyrer / om camping i Bohuslän. Mitt på golvet / stod ett splitternytt fyrmannatält uppslaget.” (Båt2 s. 51)

Som synes är många av de längre orden sammansättningar: *soffbordet, färgglada, splitternytt, fyrmannatält.* Det kan förklara både ordlängden och ordrikedomen.

Eins3: Högsta verket i alla kategorier förutom i STT & pronomen

Eins3 är ett verk som skiljer sig från de andra på flera sätt. Det är en biografi (liksom Lag2, Liv3 och Abr1) men som rör en känd person och hans upptäckter. Fokus är på det storartade och inte på det vardagliga, som i Liv3 och Abr1, även om familjeliv skildras. Dessutom förklaras flera matematiska upptäckter, och verket rör stora internationella skeenden. Det är gott om egennamn och ont om pronomen, vilket stöder tesen att de står i komplementär distribution.

Att STT är lågt verkar spontant förvånande. Dessutom växlar verket tempus mellan nutid och dåtid vilket borde skapa ett högt STT-värde, även om samma verb används så skulle det uppträda i olika

former. Inte heller är det någon tusen-mängd som sticker ut och drar ner värdet för resten – Eins3 är inte det verk med störst lexikonspridning mellan de olika mängderna på tusen ord.

De vanligaste orden i Eins3 är: *att, han, och, det, i, Albert, är, på, en, som, inte, de, om, har, till, med, för, men, den, ett.*

Inte heller denna lista visar på något som kan förklara det låga STT-värdet. Kanske används många infinita verb, vilket har puttat upp *att* till ledande position som gör att samma former återanvänds. Jag tror att svaret på frågan snarare ligger i något som blir tydligare när jag visar resultaten för rytm och interpunktion: det är liten variation när det gäller kommatering: få kommatecken, utropstecken, frågetecken och talstreck används, vilket tyder på att berättelsen trots allt inte är speciellt språkligt varierad.

Bedr3: Instabilt fack 2-verk

Bedr3 hamnar oftare i fack 2 än i fack 3, trots att det är ett nivå 3-verk. Den totala LSV-placeringen sätter Bedr3 i mittenklustret, dock angränsande till det högsta. Sammanlagt verkar det röra sig inom mitten. Kanske kan det förklaras med att verket är en bearbetning, där originalet på vissa sätt gjort det instabilt (men då borde kanske samma process ha skett för Utv3). Mycket av verket består av dialoger, vilket i det här fallet tycks begränsa språket till saker som folk ska verka kunna säga på riktigt. Då sjunker meningslängden och andelen pronomen går upp. Det tycks kunna förklara Bedr3:s LSV-position.

Inget förvånande sker bland de vanligaste orden. Det är formord och egennamn på rollfigurer: *han, är, hon, och, på, en, Kouplan, det, jag, att, i, inte, Amanda, har, du, till, Jenny, säger, som, de.*

Liv3: Instabilt med många pronomen men få egennamn

Liv3 är ett verk som kretsar kring en person som berättar om sitt liv. Pronomenet *jag* är det vanligaste ordet, och få andra personer eller namngivna platser eller märken nämns. Ordet *Asberger* nämns flera gånger och skrivs med versalt A, men det är fortfarande färre namn i jämförelse med övriga verk.

De vanligaste orden i Liv3 är: *jag, att, det, är, på, och, inte, de, en, i, med, som, för, men, om, har, kan, mig, människor, vara.*

Noterbart är att *människor* är ett vanligt ord. Så här kan det användas:

9. ”Jag försöker verkligen höra vad människor säger, / men allt blir bara oljud.” (Liv3 s. 10)
10. ”Människor med Aspergers syndrom / kan till exempel ha svårt att förstå kroppsspråk,” (Liv3 s. 12).

Människor tycks vara en anonymiserad mängd personer, lite vilka som helst, som Jerker har träffat eller hört talas om, snarare än människa som art jämfört med katter eller hundar.

Meningarna är långa men lexikonet litet. Den talspråkliga tonen förstärks därmed, och läsaren ska därmed antagligen känna sig inbjuden att höra på Jerkers erfarenheter.

Utv3: Stabilt verk med få pronomen

Utv3, ett verk som i original är känt (i alla fall under min grundskolegång) för att vara långt och tungt har här blivit kortare och lättare. Det placerar sig stabilt i fack 3, och har få pronomen, vilket med övriga verk i minnet inte är förvånande.

Lexikonet varierar dessutom på dialekt i replikerna:

11. ”– Dä ä omöjligt. Ingen kan vandra te Amerika, sa Robert. / – Inte på nået vis? försökte Arvid.”
(Utv3 s. 52).

De vanligaste orden i Utv3 är: *och, han, i, var, att, på, det, de, inte, en, hon, som, hade, till, Karl-Oskar, för, sig, om, med.*

5.3 Sammanfattning av lexikon och syntax

Det finns verk som uppvisar mer intern variation, och verk som uppvisar mer intern likhet. Variationen inom verk tycks ofta bero på att verken handlar om många olika saker, som kräver ett större lexikon. Till detta hör ofta perspektivbyten.

De verk som bara har ett egennamn bland de vanligaste orden är: Utv3, Lag2, Eins3, Abr1, och Båt2. Liv3 har inget egennamn, vilket samverkar med att Jerker berättar i första person. Visst sammanfall mellan få pronomen och få egennamn bland de vanligaste orden tycks finnas.

Alla verken ”rör personer”, men på olika sätt. I biografierna Lag2, Eins3, Abr1 (liten andel pronomen för dessa tre) och Liv3 följer läsaren bara en person, även om andra förekommer vid sidan av. Också i Katt1, But1, Kick2, Båt1 och Båt2 utgår berättelsen från en person. I Utv3 och Behe2 skiftar perspektivet regelbundet. Bedr3, som har störst andel pronomen, handlar främst om Kouplan, men i vissa stycken får läsaren följa Jenny.

Mönstret är mycket svagt, men om det finns tycks verk med få pronomen, få egennamn bland de vanligaste orden sammanfalla med biografierna. Det är dock inte det hypotesen förutsäger. Därmed bedömer jag att den inte stöds av materialet.

*Hypotes 7: De verk med hög andel pronomen är också de som har egennamn bland de vanligaste orden och vars innehåll rör personer. **Stöds inte.***

De verk som placerar sig i mellanklustret är ofta stabila, men inte nödvändigtvis. Däremot är de verk som placerar sig i det högsta och lägsta klustret i hög utsträckning stabila. Det verkar alltså som om verk har antingen en stabil stil med låg variation, en stabil stil med hög variation eller en instabil stil med vissa drag från de andra klustren.

Typiskt för det låga klustret (Lag2 och But1) är litet ordförråd, korta verk, korta meningar, få konjunktioner och egennamn. Lag2 har få pronomen, men But1 har många, så utrymme för variation gällande pronomen tycks finnas. Båda verken handlar om en begränsad mängd människor. I dessa verk finns stor lexikal och syntaktisk likhet.

Typiskt för det höga klustret (Utv3, Eins3, Båt2) är stort ordförråd, långa verk, långa meningar, många konjunktioner och egennamn och få pronomen. I dessa verk ryms mer variation. Utv3 och Eins3 har båda utöver personnamn också platsnamn vilket placerar innehållet nära den fysiska världen, även om skeendena i dem skedde för länge sedan.

Inom lättlästa verk finns alltså variation för dessa analysenheter, och de varierar mellan högt och lågt. Det som är konstant är genomsnittliga ordlängden, som hålls runt 4,13 och 4,50 bokstäver per ord.

6 Rytmm och interpunktion

Det här kapitlet söker svara på Frågeställning 2: Vilken variation och likhet finns det i lättlästa litterära prosaverks **rytm** och **interpunktion**?

Här presenterar jag analyser kring rytmm och interpunktion. Rytmmen och interpunktionen rör delvis olika företeelser, men går också ihop mycket, så jag presenterar dem tillsammans. Rytmmen kan sägas vara det läsaren känner av; en bok kan vara snabb, stillsam, hetsig och dylikt. Det finns troligtvis något i texten som ger upphov till att böcker på detta vis skiljer sig åt.

Interpunktion är ständigt närvarande i skrivet språk, men har inte en självklar plats när det gäller textanalys. Kanske för att interpunktionen är för omarkerad, för ointressant, och utan direkt motsvarighet i talat språk som anses vara språkvetenskapens främsta studieobjekt, föreslår Dahl som förklaring till detta (2015:11). I interpunktionen ryms dock mycket. Känslor uttrycks! Åtskillnad mellan olika textdelar skapas med punkter, kommatecken, styckeindelning. Samband visas; kolon, semikolon och tankstreck används ofta, och så en och annan pil. Interpunktionens tecken har dock långt ifrån en enkel betydelse, utan varierar liksom alla språkliga enheter med sitt sammanhang.

Det som gör rytmm och interpunktion intressant för de tolv lättlästa verken i den här studien är att interpunktionen uppmanas hållas på ett tydligt och entydigt vis i kursmaterial jag fått från MTM om lättläst språk. Samtidigt finns det lite att hitta i instruktionerna för lättläst som rör rytmm och interpunktion, utöver att texterna och deras beståndsdelar – stycken, rader, meningar – ska vara korta. Interpunktionen kan användas för att göra texten varierad, där utropstecken, frågetecken, kolon och ellipstecken i icke-lättläst skönlitteratur kan kopplas till en mer elaborerad stil där böckernas huvudsyfte inte tycks vara att förmedla en spännande handling (Holm 2015:324).

Utifrån detta är det oklart hur de lättlästa verken skulle använda interpunktion och rytmiska element: är variationen stor för att särskilja innehåll från varandra, eller är variationen liten för att låta läsaren tolka så få språkliga enheter som möjligt? Jag undersöker två preciserade hypoteser:

Hypotes 12: De verk som har höga LSV har också höga RIV.

Hypotes 13: Nivå 3-verken har högst RIV, följt av nivå 2-verken, och lägst RIV har nivå 1-verken.

Kapitlet är indelat likt föregående, i tre avsnitt: ett metodavsnitt, ett resultatavsnitt och ett sammanfattande avsnitt. Först en del där jag går igenom de olika analysenheter jag använder mig av. I denna genomgång hänvisar jag till tidigare forskning som jag formulerar hypoteser utifrån. Därefter följer resultaten. Resultatdelen i sin tur är uppdelad i mindre delar. Först presenterar jag kvantitativa resultat för varje analysenhet. Därefter adderar jag varje verks värde för att kunna se hur det totalt sett tycks placera sig gentemot de andra när det gäller rytmisk och interpunktionell variation. Detta sker under rubriken Rytmm- och interpunktionensvärde (RIV). Därefter följer en stilistisk kommentar för varje verk, där instabiliteter som framgick i RIV-beräkningen kommenteras. Kapitlet avslutas med en sammanfattning.

6.1 Analysenheter och tillvägagångssätt

Här presenterar jag de analysenheter jag använder mig av.

Jag har använt samma reguljära uttryck för att identifiera flera av tecknen jag går igenom i det här kapitlet. Utropstecken, frågetecken, ellipstecken (samt tre punkter), samt vanlig punkt, har identifierats med följande reguljära uttryck: "[?!]\.\.\.|...\."

Parenteser, dubbla och enkla citattecken, tankstreck, bindestreck och asterisker har identifierats med följande reguljära uttryck: "[()\"'\\-*]"

Kolon, semikolon och kommatecken har identifierats med följande reguljära uttryck: "(:;|,)"

6.1.1 Utropstecken och frågetecken

Här presenterar jag både utropstecken och frågetecken, då mycket av diskussionen om dem gäller för båda. De presenteras individuellt i resultatet.

Vad räknas?

De vanliga tecknen "!" och "?" undersöks.

Hur har jag räknat?

Jag har identifierat utropstecken och frågetecken med hjälp av reguljära uttryck.

Vad kan detta visa på? Vad används det för?

Utropstecken och frågetecken kan tyda på att stilen är elaborerad, där skribenten har använt interpunktion och rytm för att uttrycka känslor, perspektivbyten och dylikt. Många utropstecken och frågetecken kan tyda på att berättelsen ligger nära en persons inre, snarare än ett utifrånperspektiv (Dahl 2015:113). Det kan användas både i direkt, indirekt och fri indirekt anföring. Därför ställer jag upp följande hypotes:

Hypotes 14: Utropstecken och frågetecken används för att förmedla expressivitet.

Det är ett antagande som Dahl diskuterar som något forskare inte ska ta för givet (2015:47). Det är lätt att säga att ett tecken används på ett visst sätt och sen gå vidare, vilket jag inte vill göra. I och med att faktiskt undersöka om så är fallet undviker jag förhoppningsvis den fällan. Dessutom följer frågan huruvida verk utan utropstecken är utan expressivitet, vilket Holms resultat tyder på (2015:324). Självklart handlar verken hon studerar också om känsllosa händelser, men stilen används där på ett återhållsamt sätt.

6.1.2 Kolon

Vad räknas?

Kolontecknet ":" har undersökts.

Hur har jag räknat?

Tecknet har identifierats med reguljära uttryck.

Vad kan detta visa på? Vad används det för?

Kolon undersöker jag för att Holm fann att det användes ofta i vissa verk, och sällan i andra. Där det används ofta används det systematiskt inom verket, men med olika funktioner emellan. Ett verk använder det likt ett dramamanus, där namn följs av kolon som följs av direkt anföring. Där sammanfaller högt antal kolon med många utropstecken, frågetecken och ellipser, vilket inte är fallet för det andra verket. Det andra verket använder kolon som ersättning för formord, självklara finita verb och bisatsinledare (Holm 2015:228–229). Att kolon kan användas som ett sätt att skapa rytm i

verk med stil som i övrigt också uppvisar skillnader mot den normalprosan Holm fann gör att kolon blir intressant att undersöka i lättlästa verk. Tillåts kolon användas på det sättet? Eller används kolon bara där det måste, vid en och annan anföring och vid genitiv?

Hypotes 15: Det används få kolon, och när de används är det vid förkortningar i genitiv, eller för att presentera en bild, citat eller liknande, och inte löpande i brödtexten.

6.1.3 Kommatecken

Vad räknas?

Det vanliga kommatecknet ”,” undersöks.

Hur har jag räknat?

Kommatecken har identifierats med reguljära uttryck.

Vad kan detta visa på? Vad används det för?

Jag studerar kommatecken då de har en central roll för att skapa rytm i texter. De visar på under- och överordning, de uttrycker fart (Dahl 2015:153), de kan signalera tydlighet och därmed också i sin frånvaro skapa ett rörligt intryck (Dahl 2015:154). Det verkar alltså som om kommatecken är mycket mer ett stilistiskt val än ett syntaktiskt måste för författaren att använda för vissa konstruktioner. Å andra sidan ryms inte kommatecken i korta meningar. Därför frågar jag verken om följande hypotes stämmer, som skulle innebära att kommatecken används som ett stilistiskt valfritt medel:

Hypotes 16: Högt antal kommatecken sammanfaller inte med långa meningar.

6.1.4 Tre punkter/ellips

Vad räknas?

Jag har räknat både ellipstecknet ”...” och när tre vanliga punkter skrivs efter varandra, eftersom jag inte visste hur förlaget valt att skriva tecknet. Inte heller spelar det någon roll på vilket sätt de skrivs i min undersökning, däremot vill jag förstås hitta alla.

Hur har jag räknat?

Ellips och tre punkter har identifierats med reguljära uttryck.

Vad kan detta visa på? Vad används det för?

Ellipser visar på något uteslutet, något som läsaren ska förstå utan att det står utskrivet, och uttrycker samtidigt expressivitet (Dahl 2015:173). Ryms tecken med sådant otydligt innehåll i lättlästa texter? Hur används de i så fall?

6.1.5 Tankstreck/talstreck

Vad räknas?

Det långa tecknet ”—” räknas. Det som används för tankstreck, och beroende på förlag också för direkt anföring.

Hur har jag räknat?

Tankstreck/talstreck har identifierats med reguljära uttryck. Urskillning mellan om tecknet används i början av rad och följs av versal (som vid repliker) eller inuti rad mellan gemener har inte gjorts.

Vad kan detta visa på? Vad används det för?

Främst visar antal tankstreck antagligen på mängden dialog. Det är intressant med direkt anföring då det kan ställas mot andra perspektivskiften. Är direkt anföring den enda anföringstypen i verket?

Hypotes 17: Verk med många tankstreck/talstreck har också många utropstecken och frågetecken.

6.1.6 Bindestreck

Vad räknas?

Det korta bindestrecket ”-” undersöks.

Hur har jag räknat?

Bindestreck har identifierats med reguljära uttryck.

Vad kan detta visa på? Vad används det för?

Bindestreck kan visa på om sammansättningar separeras med bindestreck, och en granskning av dem kan i så fall visa på om det tycks ha gjorts av tydlighetsskäl. Bindestreck kommer också kunna förekomma vid dubbelnamn, vilket jag inte gör någon stilistisk analys av: platser och rollfigurer heter vad de heter.

6.1.7 Versalord

Vad räknas?

Ord i versaler är kort och gott teckensträngar där minst två bokstäver efter varandra står i versaler. Enstaka versalerade *i*, *å* och dylikt syns alltså inte i statistiken. En sträng i versaler innebär en träff och en siffra i statistiken.

Hur har jag räknat?

Versalord har identifierats genom följande reguljära uttryck: `”\s[A-ZÅÄÖ]{2,}\s”`.

Vad kan detta visa på? Vad används det för?

Ord i versaler analyseras ofta som framskjutna, då de står ut från resten av texten (Dahl 2015:41). Framskjutenheten fungerar ikoniskt på så vis att versalernas framskjutenhet i texten symboliserar framskjutenhet i den scen de skildrar, med typexemplet att någon skriker. Versaler kan också användas för att signalera nytt kapitel, där de dessutom skjuts fram med hjälp av typstorlek och radbrytningar. Versaler används alltså ofta för att skapa variation i texter, och det är därför jag studerar om de över huvud taget förekommer. Eftersom lättlästa texter sägs vilja använda text på ett tydligt sätt, där något betydelsebärande (som versaler) inte innehar alltför många betydelsenysanser är min hypotes som följer:

Hypotes 18: Verken innehåller knappt versaler alls, och om de gör det är det inte i den vanliga brödtexten utan vid rubriker och illustrationer.

6.1.8 Asterisker

Vad räknas?

Tecknet ”*” räknas.

Hur har jag räknat?

Asterisker har identifierats med reguljära uttryck.

Vad kan detta visa på? Vad används det för?

Främst gissar jag att asterisker används – om de används – för att separera stycken från varandra. Om så är fallet är det intressant att studera vilket avstånd mellan olika stycken signalerar, speciellt i jämförelse med styckeindelning utan asterisker och jämfört med kapitelindelning. Om asterisker dessutom visar sig användas på något annat sätt vore det självklart också intressant.

6.1.9 Parenteser, citattecken och semikolon

Jag sökte på liknande vis också efter parenteser, citattecken (dubbla och enkla) och semikolon i verken, men hittade inga. Det tyder på att viss återhållsamhet med interpunktionstecken kan urskiljas i materialet.

6.2 Resultat

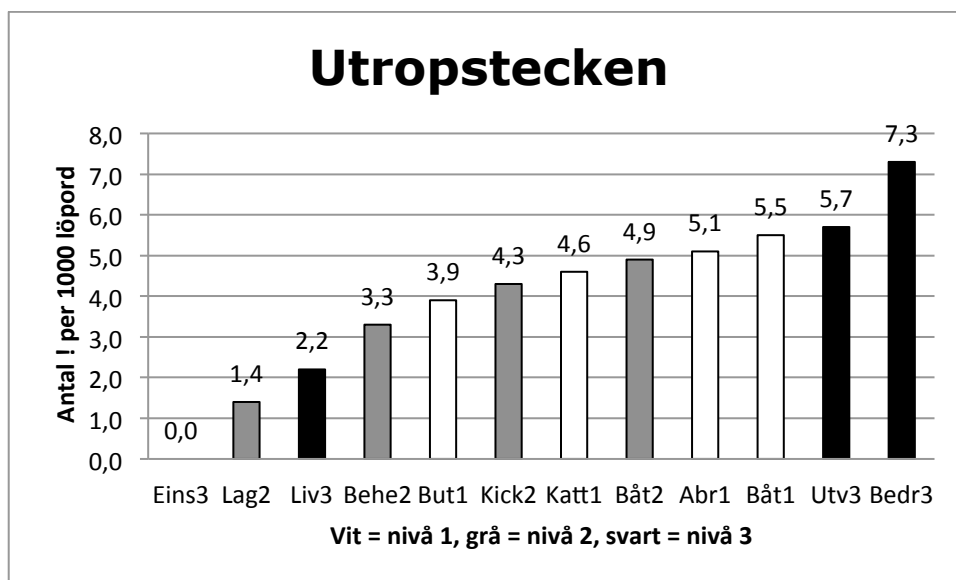
Här presenterar jag resultaten för rytm och interpunktion. Alla relativa värden i kapitlet är per tusen löpord. De absoluta värdena och alla relativa finns uppställda i tabeller i Bilaga 5.

Här nedan följer diagram som fungerar på samma sätt som i förra kapitlet. Verken är nu uppställda i storleksordning, med verket med färst förekomster står till vänster och verket med flest till höger. Verken är färgkodade, där vit fyllning är verk på nivå 1, grå fyllning verk på nivå 2, och svart fyllning verk på nivå 3. På det viset går det förhoppningsvis att snabbt se om nivåerna placerar sig i ordning eller är blandade i hög utsträckning.

Utropstecken

Här presenteras antal utropstecken per tusen löpord.

Figur 9: Rytm och interpunktion, utropstecken



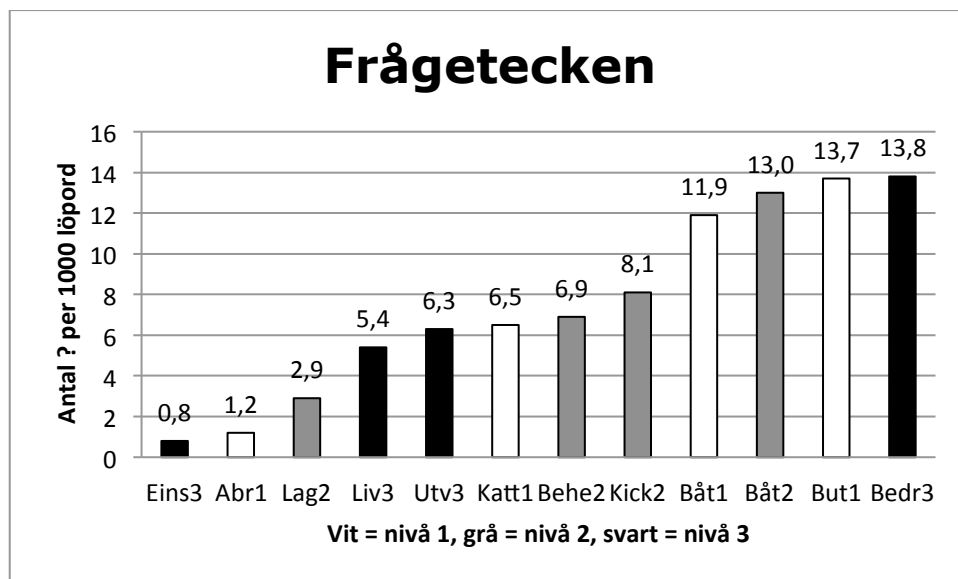
Eins3 är ensamt om att inte ha några utropstecken i diagrammet, men i själva verket finns det ett utropstecken i verket. Lag2 är det verk med utropstecken som har färst, 1,4 per tusen löpord. Bedr3 har flest: 7,3 per tusen löpord.

Fördelningen är jämn och det finns inga stora hopp mellan verken som ligger jämte varandra i tabellen. Två nivå 3-verk ligger i toppen, men följs av två nivå 1-verk vilket tyder på att viss systematik kan finnas, men att två nivå 3-verk placerar sig väldigt lågt tyder på att det är svårt att avgöra om så är fallet. Om systematik finns beror den kanske inte på nivå-indelning.

Frågetecken

Här presenteras antal frågetecken per tusen löpord.

Figur 10: Rytms och interpunktion, frågetecken



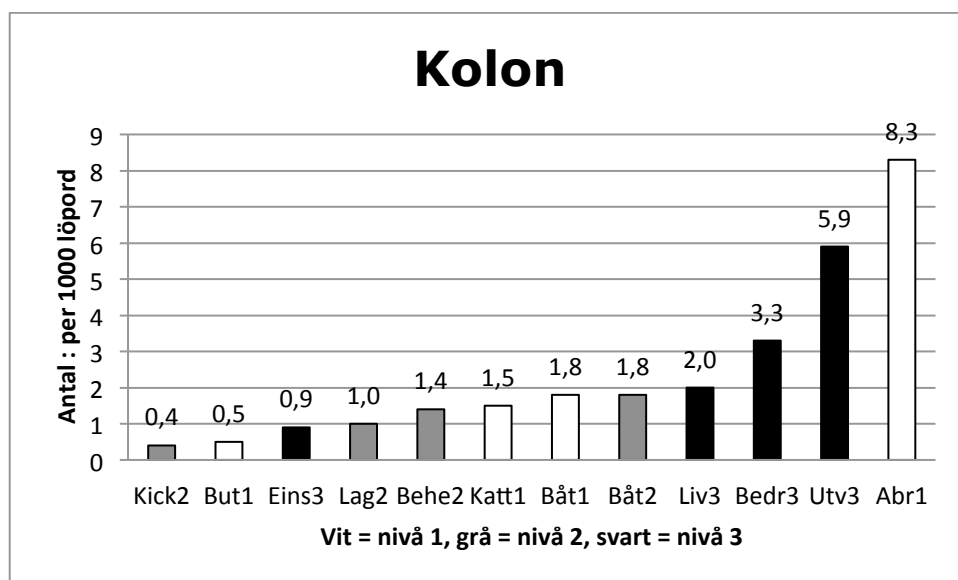
Eins3 har färst frågetecken, 0,8, och Bedr3 har flest på 13,8, dock tätt följd av But1 och Båt2 på 13,7 respektive 13,0. Ett mittenspann kan anas runt 6 frågetecken per tusen löpord. Nivåerna är väldigt utspridda.

Det är ganska stora skillnader, från knappt 1 till närmare 14. De i fack 3 som dessutom har många utropstecken är: Bedr3 (i toppen för båda) och Båt1, och i viss mån Båt2, som tycks uppvisa stor expressivitet.

Kolon

Här presenteras antal kolon per tusen löpord.

Figur 11: Rytms och interpunktion, kolon



Kick2 har färst kolon per tusen löpord med 0,4, medan Abr1 har flest på 8,3. Det ska dock sägas att Kick2 har totalt sju stycken medan But1 bara har ett enda, vilket förstas påverkar läsarens uppfattning av hur många kolon det är i texten (förutsatt att läsaren lägger märke till det).

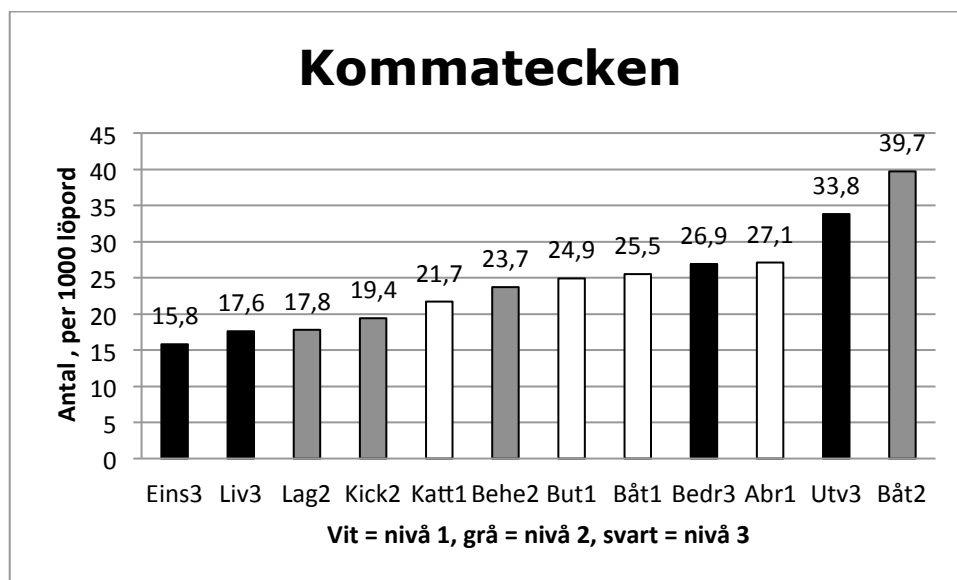
Fördelningen är ojämn. De flesta verk har två eller färre kolon per tusen löpord, förutom Bedr3 som har 3,3, Utv3 som har 5,9 och Abr1 som har 8,3. Tre av fyra nivå 3-verk placerar sig i det högre facket, men flest kolon har Abr1 från nivå 1.

Det förvånade mig att kolon användes i alla verken, då två av de skönlitterära verken i Holms studie inte innehåller kolon alls (2015:223).

Kommatecken

Här presenteras antal kommatecken per tusen löpord.

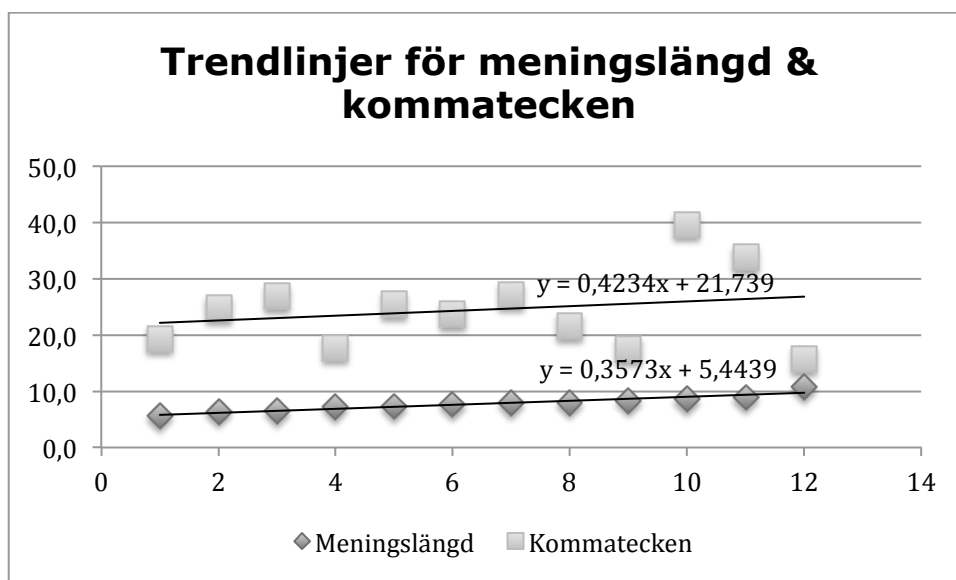
Figur 12: Rytms och interpunktion, kommatecken



Eins3 har färst kommatecken, 15,8, och Båt2 har flest på 39,7, alltså över dubbelt så många.

För att kunna avgöra om kommatecken och meningslängd hör samman måste antalet kommatecken jämföras med meningslängden. Har verk med långa meningar fler kommatecken? För att ta reda på detta skapar jag ett trenddiagram:

Figur 13: Rytm och interpunktion, trendlinjer för meningslängd och kommatecken



Översta linjen är kommatecken. Nedersta linjen är meningslängden, med värden och ordning tagna från Figur 2. Kommatecknen har placerats ut efter respektive verk, som tyvärr inte går att namnge i Excel-diagram av den här typen. Läst från vänster finns verket med kortast meningar (Kick2 på drygt 5 ord per mening) som ovanför har ett kommateckensvärde på 19,4.

Det som går att utläsa från diagrammet är att det verkar finnas viss korrelation då båda trendlinjerna sakta stiger. Meningslängden ökar jämnt, men kommatecknen gör det inte. De tre högra verken visar på stor spridning för kommatecknen, men dessförinnan ser trenden tydligare, om än något spretig ut. Det verkar utifrån detta som att kommateckenmängden för verk med korta meningar är ganska förutsägbar, medan verk med längre meningar (Eins3, Utv3 och Båt2) har utrymme att använda antingen få eller många kommatecken. Kommateckenbruket uppvisar alltså stor variation för verk med längre meningar, och förutsägbarhet för verk med kortare meningar. Detta tycks i viss mån sammanfalla med verkens nivå, då de fyra högra verken tillhör nivå 3 och 2.

*Hypotes 16: Högt antal kommatecken sammanfaller inte med långa meningar. **Stöds.***

Tre punkter/ellips

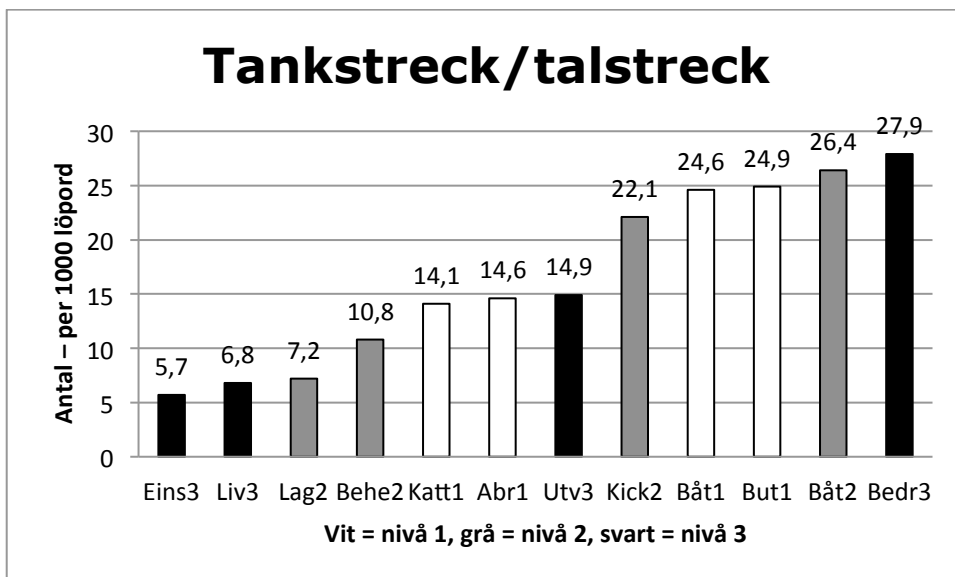
När det gäller tre punkter är variationen väldigt liten, så jag redovisar inte den i ett diagram då det bara skulle betona skillnader som i själva verket är väldigt små.

Liv3 har lägst värde på 0,2, och Utv3 högst på 1,7. Förutom de som inte har några, förstås.

Tankstreck/talstreck

Här presenteras antalet tankstreck/talstreck per tusen löpord. Samma tecken används för båda funktionerna och särskiljs alltså inte ifrån varandra.

Figur 14: Rytm och interpunktion, tankstreck/talstreck

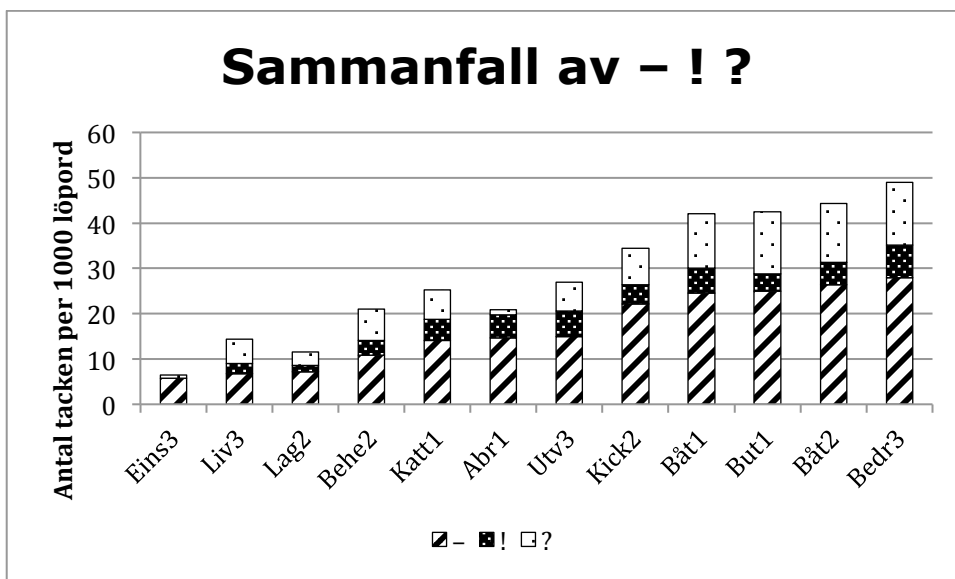


Först tankstreck/talstreck har Eins3, med 5,7 per tusen löpord, och flest har Bedr3 med 27,9.

Det är lätt att utifrån den här tabellen dra slutsatsen att antalet tankstreck/talstreck inte sammanfaller med nivåindelningen. Troligen rör det sig om direkt anföring, och då beror det på handlingen snarare än nivån vilka verk som har många. Verken från nivå 2 och 3 är utspridda och tycks kunna ha både många och få tankstreck. Nivå 1-verken placerar sig mer systematisk i mitten, och tycks alltså alltid innehålla dialog i ungefär samma mängd.

Med tankstreck/talstreck presenterat går det nu att avgöra om höga placeringar för (förmodad) direkt anföring också innebär höga placeringar för utropstecken och frågetecken.

Figur 15: Rytm och interpunktion, sammanfall av tankstreck, utropstecken och frågetecken

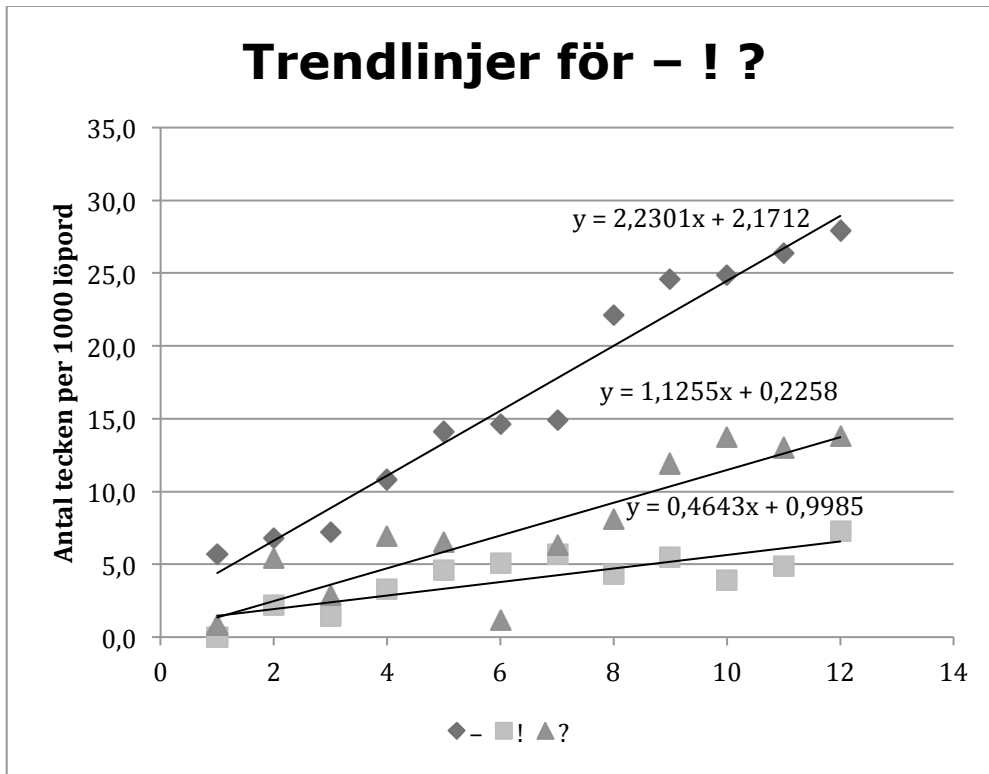


För åtta av tolv verk verkar hög andel direkt anföring innebära hög andel utropstecken och frågetecken, och motsvarande för låga värden. Det gäller Eins3, Liv3, Lag2, Katt1, Kick2, Båt1, Båt2 och Bedr3.

Fyra av verken faller bort: But1, Utv3, Abr1 och Behe2. But1 har många tankstreck och frågetecken men få utropstecken. Utv3 har medelmånga tankstreck, många utropstecken men ganska få frågetecken. Abr1 har medelmycket direkt anföring, många utropstecken men få frågetecken. Behe2 har ganska lite direkt anföring och utropstecken men ganska många frågetecken, och ligger på gränsen till att jag skulle vilja kalla det stabilt.

De verk där sammanfall finns ser ut att ha en mängd direkt anföring, en mindre mängd utropstecken och en mängd frågetecken som kanske är dubbelt så stor som antalet utropstecken. Men de verk där sambandet inte tycks finnas drar förstås ner den trenden. Jag skapade trendlinjer för alla verken:

Figur 16: Rytms och interpunktion, trendlinjer för tankstreck, utropstecken och frågetecken



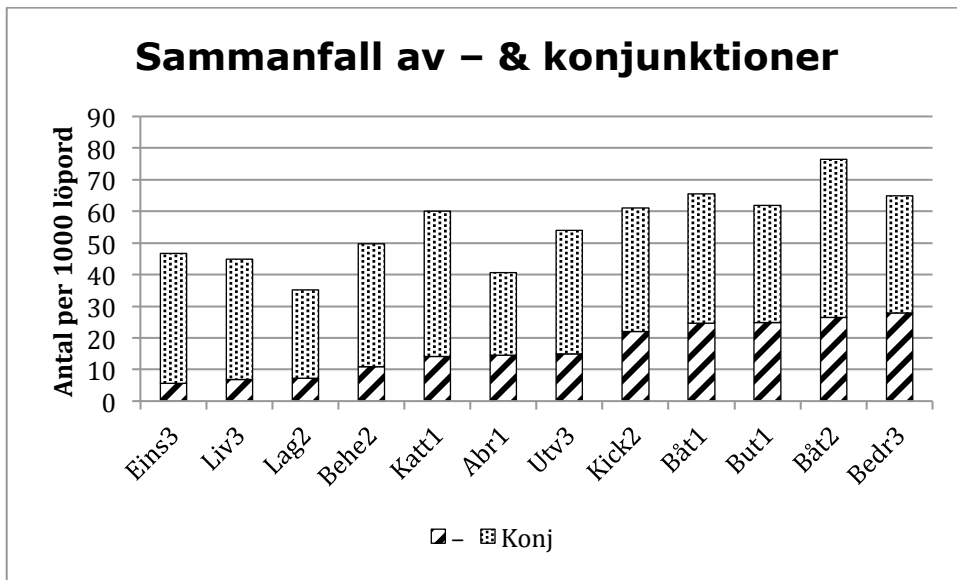
Eftersom trendlinjerna tycks öka allihop verkar visst samband finnas: om något ökar, ökar också de andra. För varje frågetecken i texten förutspår trendlinjerna att det kommer finnas 1,98 tankstreck. För varje utropstecken kommer texten enligt mönstret ha 4,80 tankstreck.

Utifrån detta besvarar jag hypotesen:

*Hypotes 17: Verk med många tankstreck/talstreck har också många utropstecken och frågetecken. **Stöds.***

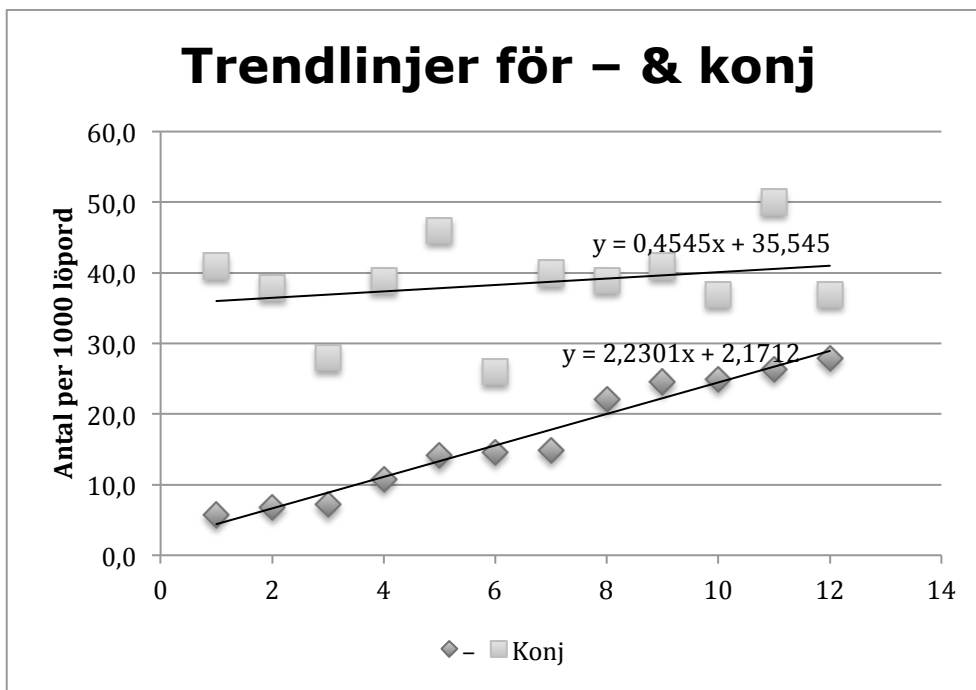
Också Hypotes 10 handlar om direkt anförings samverkan med andra språkliga enheter, i det här fallet konjunktioner. Motsvarande uppställning som för utropstecken och frågetecken:

Figur 17: Rytme och interpunktion, sammanfall av tankstreck och konjunktioner



Här är återigen tankstrecken nederst i diagrammet, här med konjunktioner ovanför. Inget tydligt mönster utkristalliserar sig i diagrammet. För att kontrollera detta ritas jag återigen trendlinjer:

Figur 18: Rytme och interpunktion, trendlinjer för tankstreck och konjunktioner



Här syns det att medan replikerna ökar håller sig konjunktionerna på en stabil nivå, med en svag ökning. Utifrån detta drar jag slutsatsen att om texter med mycket direkt anföring är mer talspråksnära än de med lite direkt anföring, är det av andra orsaker än genom hög användning av konjunktioner.

Hypotes 10: Verk med hög andel konjunktioner har också mycket direkt anföring. **Stöds inte.**

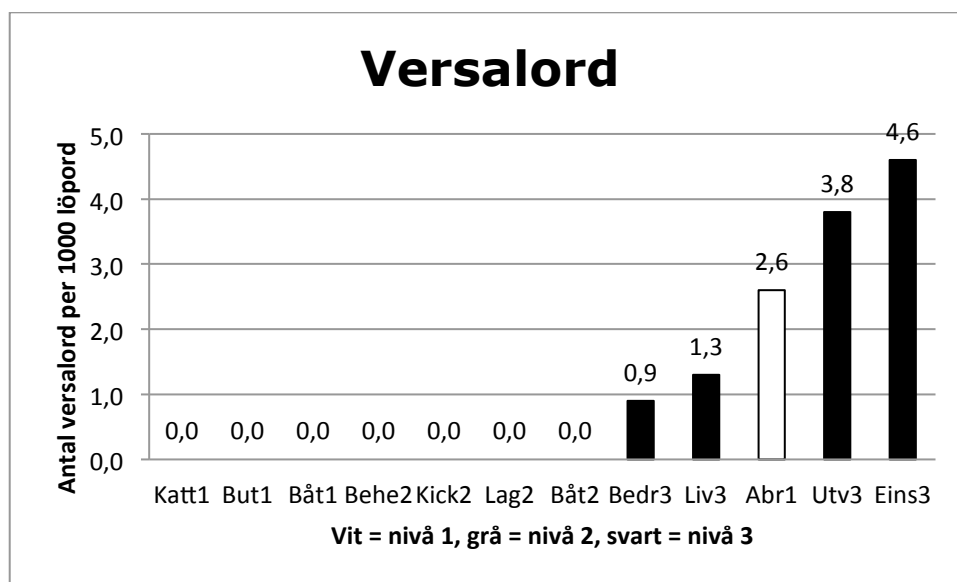
Bindestreck

Variationen inom bindestreck är liten, förutom att Abr1 sticker ut med 20,4. Utv3 är det verk med näst flest, på 5,4 och resterande ligger lågt, med bara Liv3 ovanför 1,0 på 1,6. För både Utv3 och Abr1 handlar det främst om dubbelnamn.

Versalord

Här presenteras antalet ord som står i versaler, per tusen löpord.

Figur 19: Rytm och interpunktion, versalord



Det är många verk som är helt utan versaler, tre av fyra nivå 1-verk och alla nivå 2-verk: Katt1, But1, Båt1, Behe2, Kick2, Lag2 och Båt2. Av de som har ord i versaler har Bedr3 fåst med 0,9, och Eins3 flest med 4,6. Abr1 är det verk som förflyttat sig mest, och hamnar bland nivå 3-verken trots att det är ett nivå 1-verk.

6.2.2 Rytm- & interpunktionsvärde (RIV)

Det här avsnittet handlar om verkens relation till varandra. För att studera detta närmare slår jag ihop de tabeller jag presenterat ovan, för att skapa ett värde som sammanfattar verkens variation när det gäller rytm och interpunktion. Detta kallar jag rytm- och interpunktionsvärde, RIV.

Utplacerade i storleksordning fördelar sig verken som följer:

Tabell 5: RIV-position

Fack	1				2				3			
Position	1	2	3	4	5	6	7	8	9	10	11	12
!	(Eins3)	Lag2	Liv3	Behe2	But1	Kick2	Katt1	Båt2	Abr1	Båt1	Utv3	Bedr3
?	Eins3	Abr1	Lag2	Liv3	Utv3	Katt1	Behe2	Kick2	Båt1	Båt2	But1	Bedr3
:	Kick2	But1	Eins3	Lag2	Behe2	Katt1	Båt1	Båt2	Liv3	Bedr3	Utv3	Abr1
,	Eins3	Liv3	Lag2	Kick2	Katt1	Behe2	But1	Båt1	Bedr3	Abr1	Utv3	Båt2
–	Eins3	Liv3	Lag2	Behe2	Katt1	Abr1	Utv3	Kick2	Båt1	But1	Båt2	Bedr3
VERSAL	(Katt1)	(But1)	(Båt1)	(Behe2)	(Kick2)	(Lag2)	(Båt2)	Bedr3	Liv3	Abr1	Utv3	Eins3

Precis som för LSV innebär placeringen i tabellen en rakning och därmed en siffra. För frågetecken får Eins3 1 poäng, Abr1 2 poäng, Lag2 3 poäng och så vidare.

Att ett verk står inom parentes innebär att det inte har några förekomster av den språkliga enheten, och får alltså 0 (noll) poäng, oavsett placering i tabellen ovan.

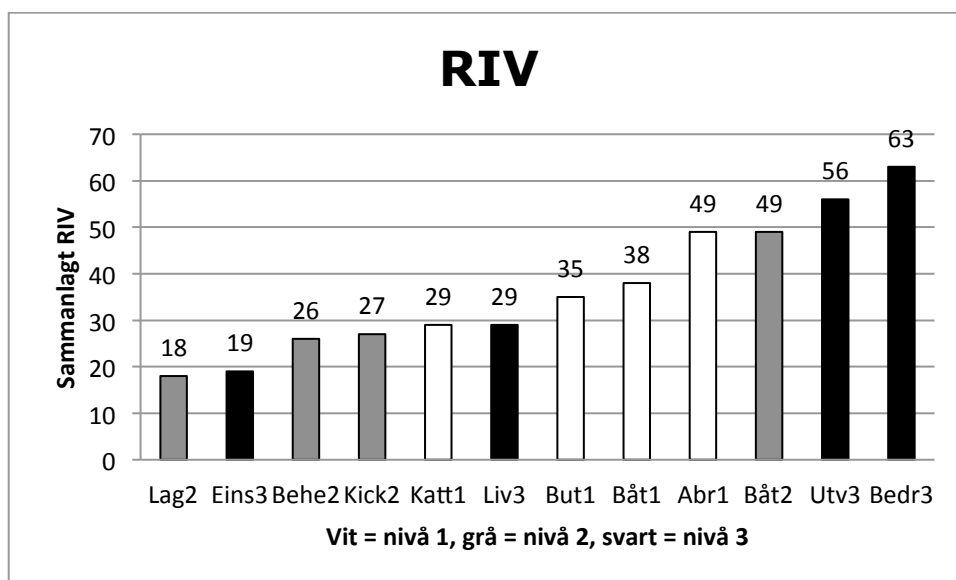
Här räknar jag ut RIV:

Tabell 6: RIV-uträkning

Nivå	1				2				3			
Verk	Abr1	Katt1	But1	Båt1	Behe2	Kick2	Lag2	Båt2	Eins3	Bedr3	Liv3	Utv3
!	9	7	5	10	4	6	2	8	0	12	3	11
?	2	6	11	9	7	8	3	10	1	12	4	5
:	12	6	2	2	5	1	4	8	3	10	9	11
,	10	5	7	8	6	4	3	12	1	9	2	11
–	6	5	10	9	4	8	3	11	1	12	2	7
VERSAL	10	0	0	0	0	0	0	0	12	8	9	11
RIV	49	29	35	38	26	27	15	49	18	63	29	56

Utplacerade i storleksordning fördelar sig verken så här:

Figur 20: RIV



Lag2 och Eins3 skiljer sig från övriga verk genom att ha lågt RIV. Ett mellankluster går att urskilja, där Behe2, Kick2, Katt1, Liv3, But1 och Båt1 hamnar, med RIV mellan 26 och 38. Därefter finns ett glapp till nästa verk, som är Abr1 och Båt2 på vardera 49 poäng. De delar det högsta klustret med Utv3 på 56 och Bedr3 på 63.

I toppen finns två nivå 3-verk, Bedr3 och Utv3. Nivå 1-verken placerar sig i mitten, med dragning uppåt snarare än nedåt. Nivå 2-verken upptar tre av fyra av de lägsta positionerna, men sticker ut med Båt2 med tredje högsta värdet.

Utifrån detta bedömer jag hypotesen:

*Hypotes 13: Nivå 3-verken har högst RIV, följt av nivå 2-verken, och lägst RIV har nivå 1-verken. **Stöds inte.***

Hypotesen stöds inte eftersom verken fördelar sig ojämnt. Det går att ana att nivå 1-verken är mer lika varandra än nivå 2- och nivå 3-verken som tycks kunna ha både högt och lågt RIV.

6.2.3 Stilistiska kommentarer

Här följer stilistiska kommentarer för varje verk. Verken presenteras i nivå-ordning.

Abr1: Gott om allt utom frågetecken

Precis som att Abr1 uppvisade mer variation än övriga nivå 1-verk för lexikon och syntax placerar det sig högst av nivå 1-verken för rytm och interpunktion. Det finns en hel del repliker i verket, men jämfört med de andra i materialet är det bara medelmånga.

Dessutom skiljer sig Abr1 från andra verk och uppvisar få frågetecken – totalt sex stycken – trots att det finns mycket direkt anföring. De frågetecken som finns används alla – förutom en gång – i direkt anföring. Den direkta anföringen är ofta ensam, och visar på hur en scen brukar gå till, snarare än som del i dialog:

12. Hon brukar gå rakt fram till nyinflyttade öbor och säga: / – Hej! Jag heter Anna. Vad heter du? (Abr1 s. 46)

Det frågetecknet som inte är i direkt anföring används i en sång. Det är på sätt och vis direkt anföring då också, men det är en stor grupp sångare som sjunger en text som är bestämd sen innan, och är alltså inte en person som ställer en fråga.

Abr1 innehåller jämförelsevis många kolon. De används på varierande vis. Som anföring (13), för att introducera uppradning (14), för att illustrera en text som Anna läser (15) eller skriver (16):

13. Anna ville ha rosa väggar. / Men då sa hennes mamma: Stopp! (Abr1 s. 12)
14. Det syns i ansiktet: / hennes mun ler, ögonen glittrar. (Abr1 s. 16)
15. Där står det vad hon ska göra: / Dammsuga på måndagar. / Handla på tisdagar. (Abr1 s. 34)
16. I dag skriver Anna: / Käraste vackra Karin, (Abr1 s. 59)

Kolon tycks alltså ha en ganska gängse funktion av att introducera listor och olika typer av citat.

Styckeindelningen sker med blankrad. Dessutom finns här och var rosa ensamma femuddiga stjärnor mellan styckena. De tycks signalera byte av scen, men scenbyte sker också utan stjärnorna. Verket har kapitelrubriker som skrivs i versaler.

Katt1: Stabilt fack 2-verk

Katt1 har mellanhöga värden i alla kategorier, men tycks vara utan versalord. Verket har dock kapitelrubriker i versaler, men de är skrivna med ett versaltypsnitt vilket gör att de inte läses som versaler av mitt program. Katt1 tycks på så vis uppvisa en behärskad stil, där interpunktionen och rytmen inte sticker ut någonstans.

Det kan tyckas märkligt att Katt1 inte har fler expressiva skiljetecken, då handlingen rör många personers känslor kring det upprörande att tro sin katt vara död. Det visar sig att utropstecknen används i kluster, för att där uttrycka mycket känslolägg på liten yta:

17. – Tiger! skrek hon, det är du! / Det är Tiger! (Katt1 s. 24)

Dessutom upprepas vissa meningar vilket skapar extra emfas:

18. Det var verkligen Tiger! (Katt1 s. 25)
19. Det var ju verkligen Tiger! (Katt1 s. 26)

Dialogen består ofta av minst två repliker, även om de ofta varvas med relation eller fri indirekt anföring där läsaren får ta del av Peters tankar.

But1: Få kolon, många utropstecken, frågetecken, kommatecken och tankstreck

But1 har högre RIV än LSV, och främst verkar det bero på att But1 har många frågetecken, och tankstreck. Består But1 således av frågor i dialog?

Verket inleds med att presentera familjen och Kalles skoltrötthet. Här finns ingen direkt anföring, utan de kommer först på sida 28, när Kalle är på möte med rektorn. Där är det mycket direkt anföring, och många är mycket riktigt frågor till Kalle där rektorn och mamman försöker fråga Kalle vad han vill i livet. ”De andra pratar men han lyssnar inte” (s. 27), vilket också förstärks av att samtalet är något stiltigt med frågor och korta svar:

20. – Vill du inte jobba? säger rektorn. / – Jo, det vill jag kanske, säger Kalle tyst. (But1 s. 29)

När Kalle senare träffar Karolin ställer också hon frågor. Också de är frågor om Kalles livsval:

21. Jag pluggar varje kväll. / Och jobbar varje helg. / Tror du att jag har roligt hela tiden? frågar hon.
(But1 s. 65)

På detta svarar Kalle att han brukar rita och tycker det är roligt, vilket följs av ifrågasättande från Karolin. Kalle blir arg och vill bevisa för Karolin att han är duktig på att rita.

Frågetecknen används i But1 för frågor, ofta av ifrågasättande karaktär, vilket bidrar till att personerna blir uttråkade eller arga. Interpunktionen samspelar därmed starkt med känslor. Eftersom But1 är ett kort verk med låga värden för lexikon och syntax kanske rytmen och interpunktionen kompenserar för det och skapar variation.

Båt1: Frågetecknen och utropstecken

Båt1 påminner om But1. Också här kan frågorna tolkas ifrågasättande (22), men vissa är också vänligare (23):

22. – När hade du själv ett jobb senast? svarar Hassan. (Båt1 s. 24)

23. – Vill du jobba för mig? / Ska vi bygga båtar tillsammans? (Båt1 s. 50)

Vid ett tillfälle är frågan något Hassan tänker och inte yttrar, och vi får inblick i hans oro, och en gång används frågetecknen för ett citat som inte markeras som citat:

24. Tänk om det har hänt honom [Bosse] något? (Båt1 s. 31)

25. En av de som jobbar natt i villan är sjuk. // Kan Hassan komma? (Båt1 s. 22)

Det är intressant att citatet är utan anföringstecken. Inget verk har citattecken, men det vore inte orimligt att föreställa sig en replik med streck. Detta används dock inte. Kanske räcker frågetecknet för att placera frågan hos en rollfigur? Förbudet mot citattecken verkar vara så starkt att det väljs bort även när det hade varit tydligt avgränsande mot resten av texten.

Frågor används på ett mer varierat sätt i Båt1 än i But1.

Båt1 har relativt många utropstecken, men i absoluta tal är de bara sex stycken. Utropstecknen uppvisar samma varierade funktioner, för repliker och för fri indirekt anföring:

26. Det är ett äventyr! svarar Hassan. (Båt1 s. 35)

27. Hassan vänder sig om. / Det är Bosse! (Båt1 s. 44)

Interpunktionen används för frågor i direkt anföring och fri indirekt anföring i Båt1.

Behe2: Stabilt verk med ganska lite av allt

Behe2 har ganska lite av alla värden, och blir därmed ett stabilt verk. Frågetecknen och utropstecken används både i direkt anföring och för att signalera både Bengtssons och Johannas tankar:

28. Blundar och försöker att tänka klart. / Vad måste hon göra först? (Behe2 s. 6)

Alla utom ett kolon presenterar direkt anföring, där olika verb används för talet: ”säger” (till exempel s. 16), ”gormar” (s. 46), ”gråtskriker” (s. 50), ”frågar” (s. 52), ”avslutar” (s. 53), ”vråla” (s. 72), och ”skriker” (s. 86). Verben signalerar det känslomässiga innehållet i verket, och kombineras antagligen med kolon för att förtydliga att det rör sig om direkt anföring. Mycket annan direkt anföring är utan kolon men har också mindre laddade verb, eller är helt utan verb. Det enda kolonet som inte introducerar direkt anföring anför texten på en skriven lapp.

Behe2 har alltså lågt RIV men medelhögt LSV med stort lexikon, och tycks använda lexikon istället för interpunktion för att variera texten.

Kick2: Snarlik But1 och Båt1 men med färre av allt

Kick2 har få kolon och versalord, och medelmånga utropstecken, frågetecken, direkt anföring, samt ganska få kommatecken. Att kommatecknen är få kan förklaras med att meningslängden är kort. De kolon som finns används för att anföra repliker.

Frågetecken används inte på ett ifrågasättande vis av samma karaktär som i But1:

29. – Sover Emma? (Kick2 s. 12)

30. – Men varför blev du avstängd? frågade Amid. (Kick2 s. 38)

Utropstecken används främst i direkt anföring av utropande karaktär, och ofta finns flera utrop på samma sida. I kombination med de korta meningarna får Kick2 en korthuggen rytm. På samma sida återfinns:

31. – Bra jobbat!

32. – Dags för plankan!

33. – Tre minuter från nu! (Kick2 s. 52)

Lag2: Lägst RIV och låga värden för allt

Lag2 har lägst RIV och LSV, och saknar höga värden överlag. Frågetecken används i direkt anföring och för fri indirekt anföring:

34. Hon vill berätta de roligaste historierna. Kan hon bli en sån person i det här laget? (Lag2 s. 21)

Lag2 tycks vara ett verk där texten har stor likhet. Inte heller finns variation mellan olika typer av innehåll. Enskild direkt anföring strösslås i texten utan att längre dialogpartier ges utrymme. Perspektivet är hela tiden Malins, och läsaren följer hennes tankar. Det finns inget i Lag2 hittills som kan förklara varför det är ett nivå 2-verk, då det hela tiden håller sig inom snäva språkliga gränser och därmed snarare hör hemma bland nivå 1-verken där färre semiotiska resurser ska användas.

Båt2: Höga värden och utan versaler

Båt2 placerar sig på tredje högsta plats för RIV, jämfört med överst för LSV. Det är en hög placering, allra helst med tanke på att verket är ett nivå 2-verk. Det har höga värden i alla rytm- och interpunktionsanalyskategorier (utom för versalord), och rör sig mellan position 8 och 12. Att verket saknar versalord verkar typiskt, då alla på nivå 3 men få andra har versalord. Verkets kapitel markeras med stor typgrad och kursiv stil.

Båt2 har långa partier med dialog, där repliker, relation och fri indirekt anföring (Hassans tankar) blandas. Det kan förklara det höga bruket av tankstreck/talstreck. Ofta används flera interpunktionstecken tillsammans, som i den här repliken:

35. – Och jag frågar dig: hur skulle någon, ens en idiot, vilja vara där gratis? Som slav? (Båt2 s. 13)

Här finns direkt anföring, kolon, kommatecken och två frågetecken. Tillsammans skapar de en rytm som ska efterlikna talspråk, med anföring till sig själv, under- och överordning, samt förtydligandet för effekt i slutet av repliken, ”Som slav?” som kan ses som en upprepning av innehållet i satsen innan.

Eins3: Låga värden för allt förutom versaler som det har flest av alla för

Eins3 kan sägas vara motsatsen till Båt2. Det är också ett långt verk, och det placerar sig i toppen för LSV. Medan Båt2 är i toppen för RIV också har Eins3 sjunkit till position 2. Det enda Eins3 har gott om är versalord, som det har flest av alla av! Eins3 tycks alltså vara ett paradoxalt verk.

Det finns rader som markeras med tankstreck/talstreck, men de innehåller inte yttranden utan är främst citat från brev. Att de markeras som tal gör att texten tycks närmare den vanliga romanens dialog, samtidigt som verbet *skriver* hela tiden påminner läsaren om att detta är en biografi som ska innehålla autentiska citat och korrekt fakta. Kolon används på motsvarande funktioner som anföring av text (som antas vara bevarad och reproducerad) och kommer ibland flera i närhet till varandra:

36. På blankettens rad för religion skriver han: Ingen. / På raden för yrke skriver han: / Ger privata lektioner i matematik. (Eins3 s. 33)

När det gäller frågetecken används de ofta för att visa på vilka fysiska och matematiska problem Einstein funderade på, och därmed kunna leda läsaren genom samma frågor:

37. I Aarau tänkte Albert att han åkte bredvid en ljusstråle / i samma fart som ljuset. / Vad skulle hända då? (Eins3 s. 55)

Mestadelen av verket består dock av grafiska meningar som avslutas med punkter.

När det gäller versalerna används de för rubriker. Rubrikerna till kapitlen har två nivåer, och båda skrivs med versaler. Ett exempel:

38. PROFESSOR EINSTEIN / ZÜRICH OCH PRAG 1909–1912 (Eins3 s. 65)

Översta raden är skriven i fetad stil, och nedre raden i normal. Övre är större än nedre, som i sin tur är större än brödtexten. Inuti kapitlen finns ytterligare en rubriknivå, som är fetad men inte i versaler. Versaler används däremot för rubrikerna i de gula faktarutor som då och då bryter av den vanliga berättande texten. Sammantaget finns alltså många, ganska långa rubriker som förklarar versalbruket.

Bedr3: Högst RIV med höga värden för allt

Bedr3 är det verk med högst RIV-värde, och placerade sig högt (delad position 8–9 med Kick2) för LSV, om än i mellanklustret. Bedr3 placerar sig högst för tre rytm- och interpunktionsanalysenheter, och högt i övriga.

Versaler används inte i rubrikerna, då de bara är numrerade och inte namngivna. Versalerna används alltså i löptexten, vilket skiljer verket från övriga. Versalerna är namn på ett återkommande företag i berättelsen, ”BEON Bygg”. Därmed signalerar versalerna ett namn som skiljer sig från övriga texten, men de signalerar inte engagemang, känslor eller avgränsning av olika textpartier som i övriga verk. Vid ett tillfälle används versalt ”INGENTING” (s. 137) där versalerna fungerar betonande.

Dialogerna är ofta snabba, med korta turer vilket samspelar med att alla rollfigurer har hemligheter och försöker att inte avslöja sig för de andra:

39. – Det säger jag inte. / – Varför inte? / – Du kan ju gå till polisen. / – Så det var mutor? / – Det säger jag inte. (Bedr3 s. 108)

Enkla asterisker används mellan stycken för att markera scenbyte.

I Bedr3 citeras ibland sms, som skrivs med stort radavstånd och fetad stil, men utan citattecken. Återigen tycks citattecken undvikas till varje pris.

Liv3: Instabilt med många kolon och versaler men låga värden för resten

Liv3 ligger i mittenklustret av RIV och har låga värden, förutom för kolon och versalord. Kolon används i vissa fall för bestämd form vid ordet ”DJ”, men anför annars en tanke:

40. Sen kom besvikelsen: / Varför fick jag inte reda på detta förrän nu? (Liv3 s. 21)

Kolon används också för direkt anföring och introducerande av upprädningsrubriker. Versalerna används i rubriker.

Utv3: Höga värden förutom för frågetecken och medelmånga tankstreck

Utv3 verkar vara ett verk med höga resultat, förutom för frågetecken och tankstreck där det placerar sig i fack 2. Ofta används många utropstecken i samma scen, och kombineras med korta repliker:

41. – Lägg ifrån dej yxan! / – Jag ska hugga ihjäl kärringen! (Utv3 s. 45)

Det gör att texten i perioder är mer intensiv än under mer berättande avsnitt. Dialog, känslor, utropstecken, och också ellipser används för att snabba upp handlingen.

Frågetecken används i direkt anföring, och när läsaren får ta del av Karl-Oskars tankar:

42. Han gick ut och tittade på marken i Korpamoen. Fanns det mer att odla upp? (Utv3 s. 21)

Versalerna används för rubriker, och Utv3 är det enda verket som har rubriker mitt på en sida, alltså med brödtext både ovan och under.

Sammantaget tycks Utv3 kombinera variation för lexikon- och syntax med rytm- och interpunktion på ett sätt som ger utrymme för olika typer av berättande och perspektiv.

6.3 Sammanfattning av rytm och interpunktion

Precis som för LSV bildades ett större mellankluster, med ett lågt och ett högt kluster vid sidorna. Båt2 och Utv3 är återigen i det högsta klustret, tillsammans med Bedr3. Lag2 är återigen i det lägsta klustret, nu tillsammans med Eins3 som alltså sjunkit från högsta LSV-klustret till lägsta RIV-klustret.

Typiskt för det låga klustret är få utropstecken, frågetecken, kommatecken (även om Eins3 har långa meningar) och tankstreck. Viss direkt anföring finns, men det upptar en liten del av verken. Eins3 har gott om versalord i rubriker, vilket tyder på att stor likhet inom rytm och interpunktion inte nödvändigtvis innebär få versalord.

I det höga klustret har verken genomgående många utropstecken, kolon, kommatecken och tankstreck. Det varierar dem emellan om de har många frågetecken och versalord.

Mittenklustret innehåller verk som är mer instabila. De tenderar att ha mycket av något och lite av något annat. Behe2 är exempel på ett stabilt verk i mittenklustret, som uppvisar medelmycket av allt. Det har ett stort lexikon som tycks kompensera för den ganska låga variationen för rytm och interpunktion.

Nu kan jag besvara den hypotes som inledde kapitlet.

*Hypotes 12: De verk som har höga LSV har också höga RIV. **Stöds.***

Hypotesen stöds då tre verk i fack 3 är samma för LSV och RIV. Det gäller Utv3, Båt2 och Bedr3. Eins3 har högt LSV men lågt RIV, vilket motsäger trenden. Jag vill ändå säga att hypotesen stämmer, och att verk med hög variation för lexikon och syntax också tenderar att ha hög variation för rytm och interpunktion. När det gäller Eins3 verkar hög STT-placering kompensera för lågt RIV.

Det finns visst sammanfall av utropstecken och frågetecken, där verk med mycket direkt anföring tenderar att ha mer utropstecken och frågetecken. Detta gäller dock inte alla verk, och några har många frågetecken och knappt några utropstecken. Frågetecken används i vissa verk för ifrågasättande och bidrar till att bygga upp konflikter i berättelsen (Båt1 och But1), medan andra frågor är av artigare karaktär. I Bedr3 används ofta rapp dialog med frågor och undvikande svar. Utifrån detta bedömer jag att utropstecken och frågetecken används för att förmedla expressivitet.

*Hypotes 14: Utropstecken och frågetecken används för att förmedla expressivitet. **Stöds.***

Intressant är att kommatecken tycks användas som valfritt stilistiskt grepp, men bara för verken med längre meningar. Dessförinnan ökar andelen kommatecken jämnt medan meningarna blir längre, till en viss gräns där kommatecknen i verk med långa meningar börjar fluktuera.

Kolon används för att signalera direkt eller fri indirekt anföring. Utifrån detta bedömer jag att hypotesen ifråga stöds.

*Hypotes 15: Det används få kolon, och när de används är det vid förkortningar i genitiv, eller för att presentera en bild, citat eller liknande, och inte löpande i brödtexten. **Stöds.***

Rytm skapas genom att liknande enheter används i kluster, med exempel från Kick2 och Katt där utrop förekommer i grupp och därmed förstärker varandra.

Versalord används främst för rubriker, förutom i Bedr3 där de främst används i ett företagsnamn. Därmed bedömer jag att hypotesen ifråga stöds av materialet.

*Hypotes 18: Verken innehåller knappt versaler alls, och om de gör det är det inte i den vanliga brödtexten utan vid rubriker och illustrationer. **Stöds.***

7 Färg och illustration

Det här avsnittet söker svara på Frågeställning 3: Vilken variation och likhet finns det i lättlästa litterära prosaverks användning av **färg och illustrationer**?

Eftersom förlaget specificerar att bilder används på olika sätt för de olika nivåerna studerar jag illustrationer och deras färger. Studiet av färg och illustration utgår från en multimodal modell, där olika drag i bilder analyseras som semiotiska resurser, precis som jag gjort för lexikon och syntax och för interpunktion och rytm.

Kapitlet är likt tidigare uppbyggt i tre avsnitt, som kan ses som metod, resultat och avslutning. Först en del där jag går igenom de olika analyskategorierna jag använder mig av. I denna genomgång hänvisar jag till tidigare forskning som jag formulerar hypoteser utifrån. Därefter följer resultaten av analyskategorierna. Resultatdelen i sin tur är uppdelad i mindre delar. Först presenterar jag kvantitativa resultat för varje analyskategori. Därefter adderar jag varje verks värde för att kunna se hur det totalt sett tycks placera sig gentemot de andra när det gäller bild- och färg-variation. Detta sker under rubriken Bild- och färg-värde (FIV). Därefter följer en stilistisk kommentar för varje verk. Kapitlet avslutas med en sammanfattning.

Jag använder ordet *illustration* som inkluderande för fotografier, tecknade bilder, målade bilder och dylikt.

Utifrån hur förlaget beskriver nivåerna ställer jag upp följande hypotes. Nivå 1 ska ”alltid” ha många illustrationer, att jämföra med nivå 2 som ”oftast” är illustrerade, alltså i lägre grad än nivå 1. För nivå 3 ska vad förlaget kallar fackböcker ”alltid” vara illustrerade, och hur det är med nivå 3-verk som inte är fackböcker framgår inte. Jag förväntar mig att dessa instruktioner följs, då det är relativt enkelt från ett förlags sida att bestämma att ett verk ska ha illustrationer.

Hypotes 19: Alla nivå 1-verken är illustrerade. Några men inte alla nivå 2-verken är illustrerade. Nivå 3-verken är inte illustrerade, förutom Eins3 som har fackboksdrag och därmed är illustrerad.

7.1 Tillvägagångssätt

Arbetet med färg och illustration har varit manuellt, och till skillnad från lexikon och syntax samt rytm och interpunktion alltså inte automatiserats.

Jag har sida för sida gått igenom ett verk i taget och för varje illustration noterat placering, distans och innehåll samt färg. Detta har noterats i tabeller som jag sedan arbetat utifrån.

7.2 Analyskategorier

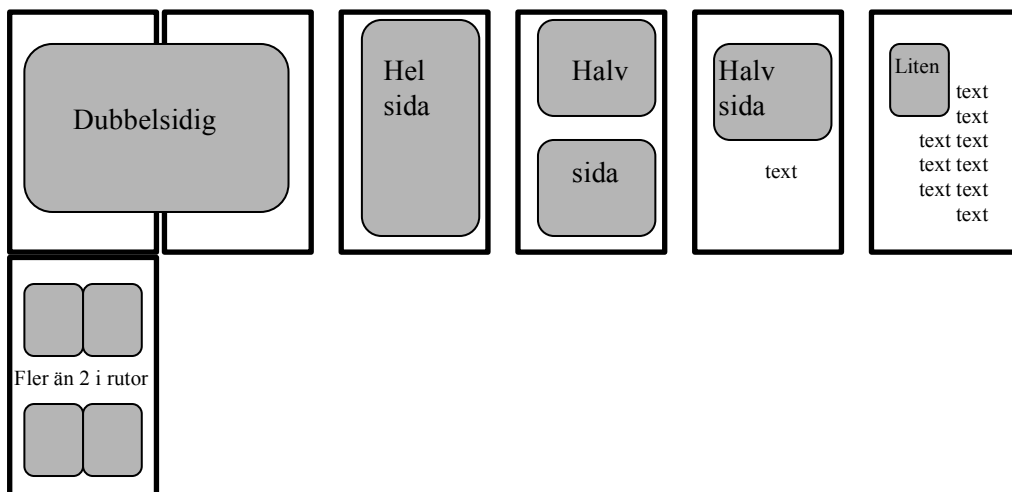
Här går jag igenom de analyskategorier jag använder mig av för att analysera färg och illustration. De är: illustrationens placering och antal, distans, och färg.

7.2.1 Placering och antal

Placering och antal hör ihop på så vis att *en* bild som tar upp *en hel sida* eller *två bilder* som tar upp *en sida* räknas som hel sida respektive två halva bilder.

Jag har gjort små illustrationer för att visa vad jag menar:

Figur 21: Bild och färg, placerings- och antalsmodell



Om ett verk har en markering för **dubbelsidiga** bilder betyder det att samma bild upptar båda sidorna på ett uppslag. Ingen text finns över eller under illustrationen. **Hel sida** innebär att verket har en illustration som är ensam på en sida, och inte ackompanjeras av text. **Halv sida** innebär att verket har en illustration som upptar en halv sida. Det kan antingen vara två sådana bilder, som då markeras som två i räkningen, eller en som ackompanjeras av text på resterande sidan. **Liten** innebär att en liten illustration finns, som upptar mindre än halva sidan och ackompanjeras av text. **Flera** innebär att det finns fler än två illustrationer på samma sida, men som är separerade med ramar, vitt tomt utrymme eller dylikt. Dessa kan vara med eller utan text.

Detta är de typer av placeringar som återfinns i materialet. Självklart kan en annan indelning vara aktuellt för ett annat material.

7.2.2 Distans och innehåll

Distans handlar om avståndet mellan det avbildade och läsaren. Porträttbilder där den avbildade möter läsarens blick kan illustrera en inbjudan till deltagande, men innebär samtidigt ett krav på möte och respekt. En helfigur med blicken bort tillåter i jämförelse att läsaren granskar personen, och signalerar knappast inbjudan.

Jag har inspirerats av den multimodala teorin som presenteras i Anders Björkvalls *Den visuella texten* (2009), men gör ingen särskiljning mellan vissa kategorier. Vanligtvis finns kategorierna *personlig distans*, *social distans* och *distanserade* bilder. Bilder med personlig distans finns i två undertyper, där *extrem närbild* bara visar delar av ansiktet, och där *nära distans* visar huvud och axlar. Bilder med social distans har tre undertyper, där *nära* innebär en beskärning i midjehöjd, *medel* innebär beskärning vid knäna och *långt* innebär att hela kroppen syns. Distanserade bilder innebär att hela kroppen syns, men upptar mindre än hälften av totala bilden (2009:44). Självklart är gränserna inte exakt dragna, för när slutar axeln och blir arm, var exakt sitter den avbildades midja?

Kategorierna är inte enkla att använda för att besvara mina frågeställningar, då vissa verk har bilder från många sammanhang och som är beskurna på många olika sätt, medan andra är mycket mer likartade genomgående. Det är det jag vill fånga. Då blir ett fokus på huruvida det är midja eller knän som syns inte användbart. Istället gör jag en grov abstraktion från dem, och använder mig av följande: *byst in*, *byst bort*, *helkropp*. *Byst* innebär att personens kropp syns ned till max höft-området, och *helkropp* tar vid när benen syns. I många fall är bilden distanserad, alltså att den avbildade står långt borta och upptar liten bildyta, men benen syns ändå inte för att något är i vägen. De är analyserade som *helkropp*, för att hela kroppen ryms på bildytan, även om den inte syns. Aktuellt är därmed

avståndet. Orden *in* och *bort* hänvisar till vart den avbildades blick är riktad. *In* betyder att personen tittar in i kameran och alltså möter läsaren blick vilket tyder på att högre engagemang förväntas mellan läsaren och den avbildade, och *bort* innebär att blicken är riktad bort från kameran och också bort från mötet med läsaren. Därmed hoppas jag kunna skildra vilken typ av distans och engagemang som skildras.

Dessutom använder jag mig av följande kategorier för att fånga annat än människor: djur och icke-animat. Djur är djur i gängse betydelse, och icke-animat är saker, vyer och dylikt utan människor eller djur. I många fall är gränsdragningen också här oklar. Är en bild på en hand som håller i en stekspade för att röra om i en stekpanna (Abr1) en bild på människa eller är den icke-animat? Jag har analyserat det som en bild på mat, alltså icke-animat. Människans närvaro är viktig, men det har i min analys av variation varit viktigt att fånga om det finns bilder på annat än människor. Antalet icke-animata illustrationer är dock litet, vilket gör att jag inte använt mig av en noggrannare uppdelning av innehållet i dem. Den multimodala analysen intresserar sig inte alltid för vad som avbildas, men jag anser att det är intressant för detta material, för att kunna fånga variation och likhet i hur illustrationer används. Då är motivet intressant: är det alltid av samma typ eller varierar det?

7.2.3 Färg

Jag analyserar färg som betydelsebärande.

Färgfotot har en särställning bland illustrationer, då det symboliserar en skildring av världen så som den är. Olika manipulationer av dem ökar modaliteten, och förflyttar dem från att skildra verkligheten till att illustrera en version av verkligheten. Det svartvita fotot kan därmed innebära att modaliteten har ökat: mellan läsaren och verkligheten har inte bara beskärning skett, utan färgen har försvunnit (Björkvall 2009:119). Utöver det har det svartvita fotot också associationer till gammeldags fotografier, där färg inte gick att fånga på bild.

Idag är svartvitt fortfarande billigare att trycka än färg, vilket kan påverka intrycket av verket som produkt som lyxigt eller alldagligt. Därmed kan mycket färg – ofta fungerar förstärkning av färg som en modalitetsmarkör – signalera att produkten är påkostad eller rolig.

För att kunna diskutera variation gällande modalitet snarare än att spekulera i hur dyr tryckningen var använder jag mig inte av *svartvitt* som term utan av *monokrom*. Då blir en sepiabild precis som en svartvit bild monokrom.

7.3 Resultat

Här presenterar jag resultaten för färg och illustration. Till skillnad från i de två tidigare kapitlen presenterar jag här tabeller med absoluta siffror. Ett kort bindestreck i en tabellcell betyder noll, och tecknet "Σ" betyder summa.

Först presenterar jag illustrationernas placering, därefter deras innehållstyp och när det är människor deras distans, och sedan bildernas färg. Dessa analyskategorier sammanställs sedan till ett färg- och illustrationsvärde (det är för att det ska få uttalbara förkortningen FIV de står i den ordningen, trots att färg innebär en mindre analys än illustration). Utifrån detta följer kvalitativa kommentarer för verkens illustrationer.

7.3.1 Illustrationernas placering och antal

Här presenterar jag vilka verk som har illustrationer, vilket antal de har och hur de är placerade.

Tabell 7: Färg och illustration, placering av illustrationer

Nivå	1				2				3				Σ
	Abr1	Katt1	But1	Båt1	Behe2	Kick2	Lag2	Båt2	Eins3	Bedr3	Liv3	Utv3	
Dubbel	7	-	-	1	-	-	3	-	-	-	-	6	17
Hel	25	4	15	17	-	-	31	-	13	-	8	-	112
Halv	17	2	-	-	-	-	-	-	3	-	-	-	23
Liten	1	-	-	-	-	-	-	-	2	-	-	-	2
Flera	3	-	1	1	-	-	-	-	-	-	-	-	5
Σ	53	6	16	19	-	-	34	-	18	-	8	6	-

Flera av verken är helt utan illustrationer: Behe2, Kick2, Båt2 och Bedr3. Alla verken i nivå 1 har illustrationer, med flest i Abr1 med 53, följt av Båt1 med 19, But1 med 16 och Katt1 med 6. Bara ett verk i nivå 2 har illustrationer, Lag2 som har 34 stycken. På nivå 3 har tre verk illustrationer, och flest har Eins3 med 17. Liv3 har 8 och Utv3 har 6.

Abr1 är det verk som utöver att ha flest illustrationer är det enda verk som har illustrationer på alla typerna av placeringar. Det verkar som om Abr1 är ett verk med stor variation. Liv3 och Utv3 har bara en typ av placering vardera, vilket istället tyder på stor likhet.

Klart vanligaste placeringen, 70 procent, är att illustrationen tar upp en hel sida, och alltså inte ackompanjeras av text. Att illustrationen tar upp två sidor bredvid varandra och en illustration som tar upp ungefär halva sidan är ungefär lika vanligt. Att fler än två bilder ryms på samma sida och att en bild är liten och samsas med huvudsakligen text på sidan förekommer sammantaget bara sju gånger.

7.3.2 Illustrationernas innehåll och distans

Här presenterar jag verkens innehåll och de avbildade personernas distans till läsaren. Totala antalet illustrationer för verken överensstämmer inte här med totala antalet i Tabell 7. Detta då jag för placering räknade en sida med många illustrationer som "flera" som en sådan förekomst, där alltså fyra och åtta illustrationer räknades som en förekomst vardera av "flera" illustrationer. Här har jag istället räknat alla enskilda illustrationer, eftersom de ibland skiljt sig åt gällande innehåll på samma sida. De verk detta gäller är markerade med en asterisk.

Tabell 8: Bild och färg, innehåll i bilder

Nivå	1				2				3				Σ
Verk	Abr1	Katt1	But1	Båt1	Behe2	Kick2	Lag2	Båt2	Eins3	Bedr3	Liv3	Utv3	-
Byst in	12	-	8	-	-	-	-	-	4	-	5	-	29
Byst bort	20	-	9	12	-	-	9	-	4	-	3	-	57
Helkropp	17	-	4	5	-	-	25	-	6	-	-	3	60
Djur	-	-	-	-	-	-	-	-	1	-	-	-	1
Icke-animat	13	6	-	4	-	-	-	-	3	-	-	3	29
Σ	62*	6	21*	21*	-	-	34	-	18	-	8	6	-

Återigen uppvisar Abr1 stor variation, där alla utom en innehållstyp finns rikligt representerad. Också Eins3 har stor variation mellan innehållen. Illustrationerna i But1, Lag2 och Liv3 är bara på människor. Katt1 har bara icke-animat innehåll och tycks därmed ha stor likhet.

7.3.3 Färg

Här presenterar jag färgfördelningen i verken.

Tabell 9: Färg och illustration, färg i bilder

	Abr1	Katt1	But1	Båt1	Behe2	Kick2	Lag2	Båt2	Eins3	Bedr3	Liv3	Utv3
Flerfärg	57	-	-	-	-	-	34	-	16	-	-	-
Monokrom	5	6	21	21	-	-	-	-	2	-	8	6

Abr1 och Eins3 sticker igen ut och tycks mer varierade än övriga då de har illustrationer både i färg och som är monokroma. Övriga verk har antingen eller.

7.3.4 Färg- och illustrationsvärde (FIV)

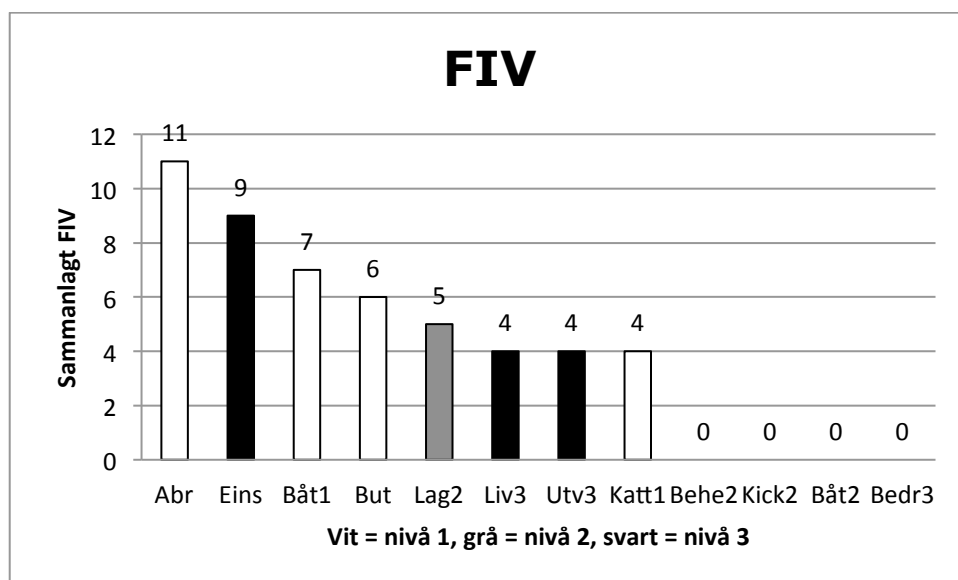
Jag räknar ut ett FIV, men inte på samma sätt som LSV och RIV. Poängen med dessa tre typer av värden är att visa på variation. Då är frågan här: hur varierar bruket av illustrationer i verken? För att kunna besvara det tilldelas varje verk poäng på ett nytt binärt sätt jämfört med tidigare. Varje gång det har bild i en kategori får det en 1:a, som sedan adderas ihop. Om ett verk uppvisar stor variation kommer det få många ettor, medan ett verk med liten (till exempel illustrationer i färg på helkropp genomgående) kommer få färre. Ett exempel: Lag2 får 1 poäng för följande kategorier: dubbelsida, helsida, byst bort, helkropp, färg, vilket blir 5 poäng. Jag markerar nedan 1:or i tabellen, och visar sammanlagda värdet, alltså FIV, nederst.

Tabell 10: FIV-position

Nivå	1				2				3			
Verk	Abr1	Katt1	But1	Båt1	Behe2	Kick2	Lag2	Båt2	Eins3	Bedr3	Liv3	Utv3
Dubbelsida	1	-	-	1	-	-	1	-	-	-	-	1
Hel sida	1	1	1	1	-	-	1	-	1	-	1	-
Halv sida	1	1	-	-	-	-	-	-	1	-	-	-
Liten m text	1	-	-	-	-	-	-	-	1	-	-	-
Flera	1	-	1	1	-	-	-	-	-	-	-	-
Byst: in	1	-	1	-	-	-	-	-	1	-	1	-
Byst: bort	1	-	1	1	-	-	1	-	1	-	1	-
Helkropp	1	-	1	1	-	-	1	-	1	-	-	1
Djur	-	-	-	-	-	-	-	-	-	-	-	-
Icke-animat	1	1	-	1	-	-	-	-	1	-	-	1
Färg	1	-	-	-	-	-	1	-	1	-	-	-
Monokromt	1	1	1	1	-	-	-	-	1	-	1	1
FIV	11	4	6	7	0	0	5	0	9	0	4	4

Precis som tidigare ställer jag upp värdena i storleksordning. Eftersom nivå 1 borde ha fler illustrationer och större variation (jämfört med tidigare då nivå 3 skulle ha mest variation) ställer jag verket med högst värde till vänster, där nivå 1-verken enligt förlaget bör placera sig. Så här ser FIV-diagrammet ut:

Figur 22: FIV



Högst FIV har Abr1, därefter Eins3. Det kan anas ett glapp mellan Eins3 på 9 poäng och Båt1 på 7. Eins3 är ett verk med mycket färg, där många illustrationer är kollage-artade, där de varierar i storlek, innehåll och placering. Båt1 är istället svartvit, där varje illustration är i samma avskalade stil med svarta tecknade streck på luftig bakgrund. Därför är jag villig att placera Eins3 i ett högre kluster än Båt. Det innebär att klustren fördelar sig som följer: FIV på 9 eller mer innebär det högsta klustret, där verken uppvisar stor variation, FIV på 7 eller mindre innebär att verken har illustrationer men att de är ganska likartade, FIV på 0 innebär att verken hamnar i det lägsta klustret och inte har några illustrationer alls. Precis som för LSV och RIV placerar sig flest verk i mellanklustret.

Hypotes 19: Alla nivå 1-verken är illustrerade. Några men inte alla nivå 2-verken är illustrerade. Nivå 3-verken är inte illustrerade, förutom Eins3 som har fackboksdrag och därmed är illustrerad. Stöds mestadels.

Jag bedömer att hypotesen mestadels stöds då alla nivå 1-verk mycket riktigt är illustrerade, även om Katt1 är det i låg utsträckning. Eins3 har mycket riktigt gott om illustrationer, men två andra nivå 3-verk har också det: Liv3 och Utv3. Det innebär att fler nivå 3-verk än nivå 2-verk har illustrationer, då bara Lag2 har illustrationer av nivå 2-verken. Därför stöds inte hypotesen fullt ut.

7.3.5 Stilistiska kommentarer

Här följer stilistiska kommentarer för de verk som har illustrationer. De står i nivå-ordning, med nivå 1-verken först. De verk som är utan illustrationer har uppenbarligen inte illustrationer som går att kommentera.

Abr1: Allt utom djur

Abr1 innehåller alla placeringar och innehållstyper – förutom djur – och har både flerfärgade och monokroma fotografier. Det är det verk med högst FIV.

De allra flesta bilderna är på Anna i olika sammanhang. Verket handlar om Annas liv och visar på många olika situationer, från både vardag och högtider, där illustrationerna kopplas till textens innehåll. Verket innehåller stor interpunktionsvariation och har högt STT-värde och långa meningar jämfört med andra nivå 1-verk. Den stora mängden bilder lättar i kontrast upp intrycket.

Dessutom innehåller Abr1 rosa stjärnor för att markera styckeindelning, och rubrikerna är skrivna med rosa text.

Katt1: Samma tecknade tassavtryck genom hela

Katt1 är det nivå 1-verk som har lägst FIV. Den som bläddrar i verket märker fort en stor likhet bland illustrationerna. Det är alltid tecknade svarta tassavtryck som slingrar sig över sidan. De förekommer vid kapitelbyten och vid inledning och avslutning av verket. Självklart hör verkets handling om den förmodat döda katten samman med katt-temat i illustrationerna, men illustrationerna förändras inte efter handlingen, utan är fristående.

But1: Tecknade människor

I But1 är alla illustrationerna stiliserat svarta och tecknade. Alla illustrationerna är på människor, och handlingen är också nära kopplad olika människor. Stor likhet finns också när det gäller placeringen. Alla upptar en hel sida, även om motivet inte upptar hela sidan fylls kanterna inte av text utan lämnas vita. Även den sida som innehåller flera bilder upptar en hel sida. Den sidan visar porträttbilder på rollfigurer.

Båt1: Tecknade människor

Båt1:s illustrationer är också svartvita tecknade med stora vita områden runt sig. De flesta illustrationer är en bild som är ensam på en sida. De fall som inte är det är jämte varandra. På vänstra sidan av uppslaget är det en stor bild på Bosse som spottar mot Hassan, vilket förstås är en grov handling. På högra sidan följer tre mindre bilder, som separeras med vita områden. På dem fortsätter bråket. Intressant är att bråkets detaljer som framgår på illustrationerna inte är utskrivna i texten, där står bara att ”Bosse knyter näven och vill slåss. / Hassan har aldrig slagit någon, / men nu är han nära.” (Båt1 s. 24).

Text och illustrationer är alltså tätt sammanlänkade, och kan bitvis ersätta varandra snarare än bara komplettera.

Lag2: Fotografier på avstånd

Lag2 innehåller fotografier på Malin i en rad situationer, främst på träningen och i hemmet. Det är många helkroppsbilder, vilket förklaras av att bilderna från träningen ofta är tagna på avstånd. Malin tittar aldrig in i kameran, vilket kan bidra till upplevelsen att läsaren följer efter henne, snarare än deltar tillsammans med henne.

Eins3: Kollageartade blandningar

Eins3 innehåller stor variation när det gäller bilder. I princip alla är av kollagetyp, där skrift, fotografi, tecknade illustrationer och färgade bakgrunder kombineras. De fotografier som finns från Einsteins liv är svartvita, men intrycket är ändå att mycket färg förekommer då fotografierna alltid kombineras med något i färg i kollaget, eller har tilldelats sepiatoning. Bilderna är ofta tätt kopplade till handlingen. Förutom kollagen har Eins3 gula bakgrunder bakom text som refererar hans upptäckter, vilket skiljer rutorna från övriga texten.

Eins3 blandar resurser med olika modalitet och sammantaget skapas intrycket att en människa tydligt varit inblandad för att manipulera fotografier och övriga illustrationer.

Liv3: Fotografier med varierade posering

Liv3 innehåller likartade bilder. De är alla fotografier på Jerker som står mot en grå bakgrund. De är beskurna ungefär vid höften, och alla är svartvita. Det som varierar är Jerkers hållning. Ibland tittar han in i kameran, ibland inte. På en bild har han händerna för öronen, på en annan är armarna hängande nedåt och händerna utanför beskärningen. Kopplingen till texten är svag.

Utv3: Tecknade berg och profiler

Utv3 innehåller vid kapitelbyten och verkets början och avslut dubbelsidiga illustrationer. På dem syns berg i bakgrunden, och gråsvarta människosiluetter står framför dem. De sidor med bara berg har jag markerat som icke-animat och de med siluetter som helkropp. Illustrationerna ska säkert ge intryck av att människor är på resande fot, men är i övrigt inte länkade till handlingen.

7.4 Sammanfattning av färg och illustration

Hypotes 19: Alla nivå 1-verken är illustrerade. Några men inte alla nivå 2-verken är illustrerade. Nivå 3-verken är inte illustrerade, förutom Eins3 som har fackboksdrag och därmed är illustrerad.", visade sig mestadels stödjas, och alla nivå 1-verken hade illustrationer.

Ett mittenkluster framkom av FIV-beräkningen, där verken uppvisar stor likhet inom varje verk. De är antingen monokroma eller flerfärgade, oftast samma typ av innehåll och samma typ av placering. Ofta förekommer de också i kapitelskiften och är inte lika tydligt kopplade till den handlingen i den omedelbara närheten, vilket Båt1 och But1 är undantag från.

Eins3 och Abr1, från nivå 3 och nivå 1, är de verk som uppvisar stor variation. Där används illustrationer med många typer av placeringar och innehåll. Där är de också tydligare kopplade till vad texten i anslutning handlar om.

En rad verk är helt utan illustrationer, och har FIV 0.

Enligt förlaget ska bilder ”bära berättelsen” i nivå 1-verken, vilket inte är fallet för Katt1. Bland nivå 2-verken är det bara ett som är illustrerat, Lag2, vilket också det bryter mot förlagets utsaga att de flesta verken på den nivån ska vara illustrerade. Där skulle fler än fyra nivå 2-verk behöva undersökas för att bekräfta om så är fallet. Men det är noterbart att nivå 3-verken i högre utsträckning än nivå 2-verken är illustrerade.

Tre av nivå 3-verken har illustrationer, trots att de enligt förlaget bara måste ha det om de är av fackbokstyp. Det kan förklara att Eins3 har så riktigt med varierade illustrationer.

8 Sammanställning av resultat

Här presenterar jag en sammanställning av resultaten, där jag slår ihop resultaten från de tre analyskapitlen. För sammanfattningar hänvisar jag till respektive kapitel avslutningar.

Vilka mönster går att utläsa? Innebär stor variation inom ett område stor variation också inom de andra? För att diskutera detta placerar jag ut verken i en tabell. Kluster 1 innehåller de verk som för LSV och RIV har låga värden, men höga för FIV. Detta för att verk med stor språklig likhet enligt förlaget bör vara nivå 1-verk, och alltså ha mycket illustrationer. På grund av detta kallar jag inte längre klustren för låg, mellan och hög, utan för 1, 2 och 3. De verk som är i fetstil är de som alltid placerar sig i samma kluster.

Tabell 11: Sammanställning LSV, RIV & FIV

	Kluster 1 (låg LSV & RIV, hög FIV)	Kluster 2	Kluster 3 (hög LSV & RIV, låg FIV)
LSV	Lag2	Båt1	Båt2
	But1	Katt1	Utv3
		Liv3	Eins3
		Behe2	
		Abr1	
		Kick2	
		Bedr3	
RIV	Lag2	Båt1	Båt2
	Eins3	Katt1	Utv3
		Liv3	Bedr3
		Behe2	Abr1
		Abr1	
		Kick2	
		But1	
FIV	Abr1	Båt1	Båt2
	Eins3	Katt1	Bedr3
		Liv3	Behe2
		But1	Kick2
		Lag2	
		Utv3	

Det finns tre verk som alltid hamnar i mittenklustret. Dessa är: Båt1, Katt1 och Liv3, som alltid placerar sig i kluster 2. De uppvisar alla viss variation inom de tre områdena jag studerat. Då återstår frågan hur lika de är varandra. Alla har tämligen litet lexikon med STT på 362–390, med längre snarare än kortare meningar på 7,3–8,3 ord. För Katt1 och Båt1 sammanfaller meningslängden med hög andel konjunktioner, vilket inte är fallet för Liv3. Alla har ganska många pronomen och desto färre egennamn. Liv3 har många långa ord, vilket de andra inte har. De använder rytm och interpunktion olika. Katt1 och Båt1 använder utropstecken och frågetecken på snarlikt sätt, och har en hel del direkt anföring. Liv3 har få utropstecken, frågetecken och anföringstecken, men desto fler kolon och versaler. De har alla ett fåtal illustrationer som används på ett genomgående likartad sätt.

Standardverket på lättläst svenska tycks alltså ha ett fåtal illustrationer, litet lexikon och många pronomen. Det finns återhållsamhet i hur rytm och interpunktion används, men utropstecken, frågetecken, kolon och versal kan användas. Anföringstecken (tankstreck) kan användas, men det beror på berättelsens innehåll.

Utöver det hamnar Båt2 alltid i det kluster som skulle innehålla variation typisk för nivå 3-verk, vilket är anmärkningsvärt då det är ett nivå 2-verk.

I övrigt rör sig de flesta verk mellan två angränsande kluster. Avvikelse från detta är Eins3 som förekommer i både kluster 1 och 3.

Det verkar som om verken, även om de är extrema i ett hänseende, närmar sig resterande verk i andra. Den standardiserade lättlästa stilen anas när verken läggs ihop, men inget verk representerar den i sin helhet. De verk som placerar sig spretigast, Eins3 och Abr1 tillåts kanske vara spretiga just eftersom de inte är nivå 2-verk. Nivå 2-verken tycks nämligen stabila, även om det innebär att Lag2 ofta placerar sig lågt och Båt2 högt.

Intressant är att de verk som är biografier av olika slag, Eins3, Liv3, Abr1 och Lag2, alla betar sig annorlunda jämfört med de andra verken i samma nivå. Här är det troligt att genren har påverkat.

När jag ser diagrammen över LSV, RIV och FIV slår det mig att de påminner om normalfördelningskurvor, alltså att det är flest som placerar sig i mitten, och enstaka extremfall placerar sig på kanterna.

Tydligt är att verken inte placerar sig efter förlagets nivåer. I många fall finns vissa tendenser till en sådan rankning, men i många fall är verken helt blandade. Därmed besvarar jag hypotesen från uppsatsens tidiga kapitel:

*Hypotes 1: Det finns skillnader mellan nivåerna på de tre områden jag ställt upp. **Stöds inte.***

9 Diskussion

Här diskuterar jag tre frågor utifrån mina resultat. Kapitlet avslutas med en metodutvärdering med rekommendationer för den som vill replikera eller inspireras av metoden.

9.1 Hur hänger form och innehåll ihop?

Stilistiken antar ofta att där formen ändras förändras innehållet. Tydligast exempel på där form och innehåll hänger ihop skulle jag säga är Kick2, med sina korta meningar som förstärker både kickboxningen och Irinas ilska. Med längre meningar hade innehållet förändrats – kickboxningen och ilskan skulle vara kvar men inte simuleras för läsaren i rytmen. Många verk har på det här sättet något som förstärker stilen, och gör stilen unik även om den i andra fall dras till mellanklustret. Det gäller de korta utropen i Katt1, dialekten i Utv3, de många varierade illustrationerna i Abr1 och Eins3.

Svårare är det med de verk som är väldigt stabila. Stilen i dem ter sig lugnare, där inget sticker ut. Lag2 är inget hetsigt verk, och handlingen går lugnt framåt, vilket verkar avspeglas i den stabila stilen. Behe2 är också ett stabilt verk, med en avstickare för högt STT. Handlingen i den är desto intensivare, där många saker sker och vi får följa två personer genom en rad (negativa) känslor. Här verkar högt STT kompensera för låga värden på resterande expressiva tecken, då känslorna verkar avspeglas i ord. Samma innehåll, ilska, skildras alltså olika i Behe2 och Kick2, men känslan är med största säkerhet tydlig för läsaren oavsett teknik.

Det finns alltså stort utrymme för författare att använda olika resurser för att förmedla ett innehåll. Den samlade repertoaren i de tolv verken innehåller många olika tekniker.

9.2 Hur hårt går det att normera litterär prosa?

Lättläst litterär prosa normeras med höga ambitioner enligt principer som skiljer sig från de för andra texter. Genom att ha studerat lättlästa texter går det kanske säga något om vad som går att normera. Hur mycket går det att bestämma över texter?

Lättlästa texter regleras av vanliga svenska skrivregler och utöver det instruktioner för att skriva lättläst. Författare och redaktör arbetar båda med att skapa en text som ska vara korrekt och intressant för läsarna. Normering sker alltså hela tiden, och berättelsen ska kunna berättas inom de uppsatta språkramarna.

Verken skiljer sig som vi sett en hel del ifrån varandra. En del har dessutom många olika sorters partier inuti sig, medan andra uppvisar större intern likhet. Mellan extremvärdena för varje analysenhet ryms vad som är möjligt att göra i litterär prosa från MTM. Bland dessa möjligheter tycks en del vara troligare att förekomma än andra, då mittenklustret alltid visar sig bli störst. Det som inte ryms i de tre klustren är det som är omöjligt, eller i alla fall inte troligt. Vore det till exempel omöjligt att ha ett verk utan direkt anföring? Det verkar så, i och med att breven i Eins3 formuleras som talat språk, och inte som citat från brev. Det verkar omöjligt att ha ett verk utan kolon, eller ett verk med citattecken.

Jämfört med Holms (2015) resultat visade sig de lättlästa verken ha korta meningar, vilket tyder på att meningslängd är något som går att normera. Hur meningarna är uppbyggda verkar delvis svårare, då kommatecken tycks användas strukturerat i verken med kortast meningar och mer som stilistiskt valfritt grepp i verken med längre meningar. Alla verk saknar parenteser, citattecken och semikolon,

vilket visar att det går att normera bort vissa tecken från en text. Hur och i vilken mängd kolon, utropstecken och frågetecken ska användas verkar däremot tillåtas variera verken emellan.

Huruvida färg och illustrationer förekommer verkar också gå att normera. Nivå 1-verken har alla illustrationer, även om Katt1 bryter mönstret genom att inte använda illustrationer som del i berättandet. Det verkar enkelt att bestämma att ett verk ska ha illustrationer, och sen också genomföra det. Svårare är det för lexikon och syntax samt rytm och interpunktion.

Typiskt för mittenklustren är att verken har höga värden i en analysenhet och lägre för en annan, vilket tyder på att det finns en återhållsamhet men också möjligheter att använda enskilda stilistiska funktioner oftare i samma text.

Den lexikon- och syntaxvariation och -likhet som går att urskilja är antagligen resultatet av författare och redaktörers gemensamma tolkningar av instruktionerna. Jämfört med resultaten för kolon, att de ofta används vid direkt anföring, är det svårare att uppfatta exakt hur långa meningar ett verk har, hur stort lexikonet är eller hur konjunktioner (inte) påverkar meningslängden. Vad som går att urskilja i lexikon- och syntaxkapitlet är alltså hur instruktioner och föreställningar om hur språket på de olika nivåerna ska vara realiserats.

Att det rör sig om tolkningar beläggs ytterligare genom att hypoteserna i så hög utsträckning inte stöds av materialet. De har alla grundats i tidigare forskning, förlagets utsaga och dylikt. Totalt har jag använt mig av nitton hypoteser, varav bara åtta stöds av materialet. Två bedömer jag stöds mestadels, men nio stöds alltså inte. Det borde göra alla som vill uttala sig om lättläst varsamma, för mycket av det vi kan tro visar sig alltså sällan stämma.

Att det framträder ett normalfördelningsmönster anser jag tyder på att verken, oavsett normeringsambitioner, uppvisar en systematik som tycks stå utanför den medvetna nivåindelningen. Lättläst litterär prosa beter sig alltså på ett systematiskt sätt som står utanför den enskilda författarens eller redaktörens nuvarande kontroll.

Mina resultat kan hjälpa den som försöker normera texter. Instruktionerna bör inte bara ställa upp vilka betydelsebärande enheter som får användas, utan också i vilken mån samma enhet kan användas med varierad funktion.

9.3 Vad innebär förlagets nivåer för läsaren?

Eftersom skillnader mellan lexikon och syntax, rytm och interpunktion samt färg och illustration inte går att urskilja för de olika nivåerna är det oklart vilket syfte nivåerna fyller. Språket är av stor vikt i att göra lättlästa texter lättlästa, men språket är aldrig ensamt. Eftersom jag inte kvalitativt studerat verken med frågor om vilka förkunskaper de förutsätter, om hur nya ord förklaras och dylikt, kan jag inte avgöra huruvida det är där nivåerna skiljer sig åt. Alltså i det som rör språkets sociala aspekter, hur texten förhåller sig till läsaren och vem läsaren förväntas vara. Jag utgår däremot från att stora sådana skillnader mellan nivåerna skulle ge avspiegling i det ytspråk jag studerat, allra helst eftersom vissa genreeffekter går att urskilja.

I mina initiala kontakter med MTM efterfrågade de en analys av skillnader mellan de tre nivåerna. Skillnader har inte framträtt, annat än för enstaka analysenheter. Om förlaget väljer att fortsätta använda de tre nivåerna med deras nuvarande beskrivningar bör hårdare normering av verken ske så att de uppfyller beskrivningarna. Då blir det tydligare vad en läsare kan förvänta sig, och en läsare som öppnar en nivå 2-bok kommer inte mötas av text som placerar sig högre och har mer variation än många nivå 3-verk, så som Behe2 gör idag. Det andra alternativet är att förändra nivåsystemet. Om nivåerna explicit förhåller sig till varandra i högre utsträckning än idag bör författare, redaktör och läsare ha enklare att förstå och uppfylla nivåernas krav.

Att bevara systemet som det ser ut idag, med verk som inte placerar sig nivåenligt, rekommenderar jag inte. Den stora poängen med nivåerna är som jag ser det att läsaren kan förutsäga vilken typ av text ett verk innehåller. Om gränserna mellan nivåerna på detta vis suddas ut omöjliggörs detta.

9.4 Metodutvärdering

Jag påminner först kort om metoden innan jag utvärderar den nedan. De tolv lättlästa verken kördes genom samma program som genererade absoluta värden för ett antal analysenheter inom lexikon och syntax samt inom rytm och interpunktion. De absoluta värdena omvandlades till relativa värden, ofta en språklig enhet jämfört med tusen löpord, ibland jämfört med sidor. Färger och illustrationer har beräknats manuellt, sida för sida av mig. Värdena sammanställs för vardera av de tre frågeställningarna, till LSV, RIV och FIV. Utifrån detta går det att urskilja tre kluster för vardera värde: ett med låga värden, ett större kluster i mitten med mellanhöga värden, och ett kluster med höga värden. Utifrån detta analyserar jag verken kvalitativt där de uppvisar instabiliteter.

Reliabilitet uppnås genom att beräkningarna är automatiserade och att samma procedur genomförs på alla verken, vilket begränsar risken för fel till ett verks favör eller nackdel. När det gäller den kvalitativa analysen bygger den på kvantitativa iakttagelser jag redogör för, och jag redogör också varför jag analyserar verken som jag gör. Den som vill upprepa studien kan använda sig av samma reguljära uttryck som jag använt mig av. En klar förbättring vore om det gått att analysera radlängd, men det visade sig som jag varit inne på omöjligt. Den som inspireras och vill replikera uppmuntrar jag kontrollera huruvida det är möjligt att göra i det aktuella materialet.

Jag ämnar studera variation och likhet. Många värden jag tog fram identifierar främst likhet, då det ofta är medelvärden. Andra typer av beräkningar hade kunnat visa på kvantitativ variation på ett mer direkt sätt, genom att identifiera extremvärdena för varje analysenhet ("Hur lång är den kortaste och längsta meningen?"). Den kvantitativa delen studerar alltså inte nödvändigtvis variation, vilket kan göra frågeställningarna svåra att besvara och dra ner studiens validitet. Däremot fungerar den kvalitativa stilistiska analysen, men också jämförelserna verken emellan för att identifiera variation och likhet. Därmed behövs metodens alla steg, och bara den kvantitativa kan inte nödvändigtvis säga något om variation. Om jag skulle studerat extremvärdena skulle jag säkerligen fått givande resultat, men eftersom jag ändå hittat stora områden som liknar och skiljer sig från varandra bedömer jag att en sådan beräkning endast skulle vara kompletterande och inte omkullkasta de slutsatser jag kunnat dra.

Nackdelar med statistik av den typen jag kan få fram ur materialet är att den blir intetsägande utan jämförelser. Säg att ett verk har 20 procent längre genomsnittsmening än alla andra, vad innebär det? Och kan det visa på att lättläst som avgränsat område skiljer sig från andra texter? Vad skulle det i så fall innebära? När det gäller jämförelser mellan verken ifråga tror jag att resultaten kan visa att vissa verk är mer lika varandra än andra, vilket säger något om vad lättläst stil innebär gällande variation. Om variationen är liten verken emellan säger det något om lättläst stil. Om den är stor ter sig lättläst stil kanske inte avgränsad andra stilar. Jämförelser med andra stilar hämtar jag ur tidigare forskning där jag också hämtar mina analyskategorier. Jag arbetar efter antagandet att lättläst stil är en del av svenskspråkigt textmaterial, något jag inte tror är kontroversiellt, och därmed blir jämförelser med "resten av svenska språket" – omöjligt att samla in – inte heller aktuellt.

Ytterligare nackdel är att statistiken kan fångsla forskaren till den grad att innehållet siffrorna är tagna ur glöms bort, vilket Östman skriver har varit vanligt förekommande (2015:7). För att inte fastna i den fällan rör jag mig i ett U: från initiala intryck av materialet, till statistiken som är längre bort från materialet, för att utifrån den närma mig materialet på nya sätt för den kvalitativa analysen. Det bedömer jag som obligatoriskt för den som inspireras.

10 Avslutning

De huvudsakliga slutsatserna det går att dra av den här studien är enligt mig:

Ett. De tolv verken bildar alltid ett stort mittenkluster som påminner om mitten av en normalfördelningskurva, där de flesta verk placerar sig. Den lättlästa stilen tycks därmed eftersträva en viss stil, antagligen på grund av normeringen.

Två. Det är inte enkelt att normera texter. Det är enkelt att bestämma om texter får ha vissa tecken (till exempel att vanliga kolon gärna får användas men aldrig semikolon eller parentes) eller om de ska vara illustrerade. Det är svårt att bedöma hur språkliga resurser används när de väl förekommer. Ju mer normeringen tillåter desto större tycks också variationen bli, med typexemplet att långa meningar har större variation i kommateckenanvändning än korta meningar.

Tre. Vi som diskuterar lättlästa texter bör vara försiktiga med tvärsäkra uttalanden om hur lättlästa texter ser ut, eftersom många av hypoteserna inte stöds av materialet. Vad vi tror oss veta stämmer sällan.

Primärlitteratur

- Andersson, Sarah och Lindqvist, Lars. 2014. *Sjöbergs affär: Ny i butiken*. Stockholm: LL-förlaget.
- Fredrikson, Bengt. 2015. *Albert Einstein: Det modiga geniet*. Stockholm: LL-förlaget.
- Jansson, Jerker. 2014. *Mitt liv med Asperger*. Stockholm: LL-förlaget.
- Krog, Niklas och Björn, Maria. 2015. *Kickboxaren 2: Fightern*. Stockholm: LL-förlaget.
- Lövestam, Sara. 2016. *Bedragaren*. Stockholm: LL-förlaget.
- Moberg, Vilhelm. 2015. *Utvandrarna*. Stockholm: LL-förlaget.
- Samuelsson, Tony. 2015. *Katten som inte ville dö*. Stockholm: LL-förlaget.
- Slonawski, Agneta. 2016. *Anna på Brännö*. Stockholm: LL-förlaget.
- Wadell, Sara. 2013. *I samma båt*. Nivå 1. Stockholm: LL-förlaget.
- Wadell, Sara. 2010. *I samma båt*. Nivå 2. Stockholm: LL-förlaget.
- Wallqvister, Camilla. 2015. *Bengtssons hemlighet*. Stockholm: LL-förlaget.
- Wänblad, Mats. 2015. *Ny i laget*. Stockholm: LL-förlaget.

Litteratur

- Björkqvall, Anders. 2009. *Den visuella texten: Multimodal analys i praktiken*. Stockholm: Hallgren och Fallgren.
- Bornhöft, Malin. 2014. *Legitima språkråd*. Kandidatuppsats, Stockholms universitet.
- Bornhöft, Malin. 2015. *Demokratiskt klarspråk*. Magisteruppsats, Stockholms universitet.
- Crossley, Scott A, Max M Louwerse, Philip M Mccarthy, och Danielle S Mcnamara. 2007. A Linguistic Analysis of Simplified and Authentic Texts. *The Modern Language Journal*, 91 (1). Blackwell Publishing Inc. 15–30.
- Dahl, Alva. 2015. *I skriftens gränstrakter: Interpunktionens funktioner i tre samtida svenska romaner*. Diss., Uppsala universitet.
- Forsberg, Camilla. 2012. *Myndigheten, texten och läsaren: Myndighetsinformation i lättläst version*. Magisteruppsats, Malmö högskola.
- Hellspong, Lennart och Ledin, Per. 1997. *Vägar genom texten: Handbok i brukstextanalys*. Lund: Studentlitteratur.
- Holm, Lisa. 2015. Rytmi i romanprosa: en studie av rytmiska signalement i tio samtida svenska romaner. I Östman, Carin (red.). *Det skönlitterära språket: tolv texter om stil*. Stockholm: Morfem, 215–235.
- Hultman, Tor G. 2003. *Svenska Akademiens språklära*. Stockholm: Svenska Akademien.
- Kungliga biblioteket. u.å. Litterär sakprosa. <https://id.kb.se/term/saogf/Litter%C3%A4r%20sakprosa> (Hämtad 2016-05-16).
- Kornai, András. 2002. How Many Words Are There? *Glottometrics* 4: 61–86.
- Lagerholm, Per. 2008. *Stilistik*. Lund: Studentlitteratur.
- Leech, Geoffrey och Short, Mick. 2007. *Style in fiction*. 2. uppl. Harlow: Longman.
- Lix.se. u.å. *LIX räknare*. (Hämtad 2016-05-16).
- LL-förlaget. u.å. Vad är lättlästa böcker? <http://lattelast.se/ll-forlaget/vad-ar-lattlasta-bocker> (Hämtad 2016-03-07)
- McEnery, Tony, Xiao, Richard och Tono, Yukio. 2006. *Corpus-based Language Studies: an Advanced Resource Book*. London: Routledge.
- Myndigheten för tillgängliga medier. u.å. *Vanliga frågor om lättläst: Vem behöver lättläst?* [http://lattelast.se/om-lattelast/vanliga-fragor-om-lattelast#Vem behöver](http://lattelast.se/om-lattelast/vanliga-fragor-om-lattelast#Vem%20behöver) (Hämtad 2016-03-14).

- Reichenberg, Monica. 2000. *Röst och kausalitet i lärobokstexter. En studie av elevers förståelse av olika textversioner*. Diss., Göteborgs universitet.
- Norrby, Catrin och Håkansson, Gisela. 2010. *Introduktion till sociolingvistik*. Stockholm: Norstedts.
- Nypon förlag. u.å. *Om Nypon*. <http://www.nyponforlag.se/om-nypon/> (Hämtad 2016-05-09).
- Rehnberg, Hanna Sofia. 2012. Inte klart vad som är lätt. *Språktidningen*. Augusti. <http://spraktidningen.se/artiklar/2012/08/inte-klart-vad-som-ar-latt> (Hämtad 2016-05-16).
- Språkrådet. 2014. *Vad är språkvård?* <http://www.sprakochfolkminnen.se/sprak/vad-ar-sprakvard.html> (Hämtad 2016-05-05).
- Språkrådet. 2015. *Klarspråkstestet*. <http://www.sprakochfolkminnen.se/sprak/klarsprak/att-skriva-klarsprak/klarsprakstestet.html> (Hämtad 2016-05-16).
- Svenska språknämnden. 2005. *Språkriktighetsboken*. Stockholm: Norstedts akademiska förlag.
- Teleman, Ulf. 2003. *Tradis och funkis: svensk språkvård och språkpolitik efter 1800*. Stockholm: Norstedts akademiska förlag.
- Tønnesson, Johan. 2012. *Hva er sakprosa*. 2. uppl. Oslo: Universitetsforlaget.
- Thelander, Mats. 2015. Empati i dialektspåret: en analys av framställningsformer i Stina Stoors novell Ojura. I Östman, Carin (red.). *Det skönlitterära språket: tolv texter om stil*. Stockholm: Morfem, 117–137.
- Wengelin, Åsa. 2015. Mot en evidensbaserad språkvård? *Sakprosa* 7 (2). <https://www.journals.uio.no/index.php/sakprosa/article/view/983> (Hämtad 2016-05-16).
- Östman, Carin. 2015. Inledning. I Östman, Carin (red.). *Det skönlitterära språket: tolv texter om stil*. Stockholm: Morfem, 7–13.

Bilagor

Här följer bilagorna.

Bilaga 1: Nivåernas innehåll

Bilagan består av MTM:s egna beskrivningar av vad de tre lättläst-nivåerna innebär, och är alltså fullständiga citat.

Lätt = 3

De skönlitterära böckerna är tjockare kapitelböcker med längre meningar och ovanligare ord. Bisatser och bildspråk förekommer. Handlingen kan hoppa i tid och rum och ha ett större persongalleri. Fackböckerna tar upp ämnen som kräver viss förförståelse men de är alltid illustrerade så att bilden stöder texten.

Lättare = 2

Handlingen är enkel och väl disponerad, ord och uttryck är välkända, meningarna korta. Böckerna är oftast illustrerade men pocketböcker och klassiker kan sakna illustrationer.

Lättast = 1

Här finns böckerna där bilder är det som bär berättelsen. Mycket lite text med enkla meningar och korta rader. Lätta kapitelböcker där handlingen utspelas i vardaglig miljö och har få personer. Ordvalet är enkelt, meningarna korta. Böckerna har alltid många illustrationer.

Källa:

<http://lattlast.se/ll-forlaget/vad-ar-lattlasta-bocker> (Hämtad 2016-03-07)

Bilaga 2: Reguljära uttryck

Här är de reguljära uttryck jag använt mig av sammanställda. Samma finns i uppsatsen vid respektive avsnitt.

Löpord:

```
”[A-Za-zÅÄÖåäö0-9]+”
```

Pronomen:

```
”(?<=\s)([jJ]ag|[sS]in|[nN]i|[mM](in|itt)|[hH](ans|ennes|enne|onom|an|on)|[vV](år|årt|åran|i)|[eE](r|rt|ran)|[dD](e|eras|enne|ennes|enna|in|itt|u|ens|ess))”
```

Egennamn:

```
”(?<=[a-zåäö]\s)[A-ZÅÄÖ][a-zåäö]+”
```

Konjunktioner:

```
”(?<=\s)[oO]ch|[sS]amt|åväl|[rR]espektive|[eE]ller|[mM]en|[fF]ast|[tT]y”
```

Grafiska meningar:

```
”[A-ZÅÄÖ][^\.\!]*[^\.\!]+”
```

Utropstecken, frågetecken, ellipstecken (samt tre punkter), samt vanlig punkt:

```
”[?!]\.\.\.\.\.”
```

Parenteser, dubbla och enkla citattecken, tankstreck, bindestreck och asterisker:

```
”[()"'- \- \*]”
```

Kolon, semikolon och kommatecken:

```
”(:|;|,)”
```

Versalord:

```
”\s[A-ZÅÄÖ]{2,}\s”
```


Bilaga 3: Lexikon och syntax, absoluta siffror

Här är de absoluta siffrorna som de relativa siffrorna i Kapitel 5 bygger på.

Tabell 12: Lexikon och syntax, absoluta siffror för nivå 1-verken

	Abr1	Katt	But1	Båt1
Sidor	92	52	75	56
Tecken	20873	11037	8962	4723
Löpord	5054	2624	2047	1097
Meningar	614	330	318	152
Tecken utan egennamn	19835	10289	7776	4170
Egennamn	247	128	121	77
Långa ord	1189	571	482	248
Versalord	13	0	0	0
Konjunktioner	132	120	75	45
Pronomen	378	217	199	107

Tabell 13: Lexikon och syntax, absoluta siffror för nivå 2- & 3-verken

	Behe2	Kick2	Lag2	Båt2	Eins3	Bedr3	Liv3	Utv3
Sidor	93	148	61	75	132	144	93	232
Tecken	36387	78096	8708	50190	98260	67498	49466	141496
Löpord	8090	17569	2076	11429	21985	15557	11561	32706
Meningar	1051	3131	249	1333	2029	2394	1396	3678
Tecken utan egennamn	33950	71590	6267	46233	89679	60892	48855	126935
Egennamn	204	657	68	451	973	694	94	1602
Långa ord	2098	4098	349	2733	5367	3436	2740	7071
Versalord	0	0	0	0	101	14	15	124
Konjunktioner	318	681	58	574	902	569	442	1303
Pronomen	617	1630	108	961	1415	1871	1107	2653

Bilaga 4: STT

Här är de siffror som STT (som är ett genomsnitt) räknades ut ifrån. Uppställda i storleksordning.

Tabell 14: STT i storleksordning

Abr1	Katt	But1	Båt1	Behe2	Kick2	Lag2	Båt2	Eins3	Bedr3	Liv3	Utv3
408	361	323	390	440	372	346	405	372	344	348	392
413	363	342		444	381	360	431	378	350	350	409
418				448	382		442	380	350	356	409
429				452	387		442	385	354	360	414
				454	392		443	387	355	369	420
				466	402		449	388	359	370	420
				474	409		449	389	365	375	420
				489	412		455	389	371	384	422
					414		458	389	373	388	423
					415		482	392	386	393	424
					419		485	397	390	403	425
					434			401	395		427
					434			407	399		429
					441			407	401		431
					442			416	429		432
					458			417			432
					475			418			433
								418			439
								420			441
								425			444
								430			446
											447
											449
											451
											452
											452
											452
											455
											462
											462
											467
											476

Bilaga 5: Rytm och interpunktion

Här följer två tabeller med värden för rytm och interpunktion. Den första är absoluta siffror, den andra är samma siffror i relativ form per tusen löpord.

Tabell 15: Rytm och interpunktion, absoluta siffror

Nivå	1				2				3			
Verk	Abr1	Katt1	But1	Båt1	Behe2	Kick2	Lag2	Båt2	Eins3	Bedr3	Liv3	Utv3
.	611	304	296	140	979	2929	213	1145	2010	2134	1337	3339
?	6	17	28	13	56	143	6	149	18	214	62	207
!	26	12	8	6	27	76	3	56	1	113	25	187
:	42	4	1	2	11	7	2	21	20	51	23	192
,	137	57	51	28	192	341	37	454	348	419	204	1106
;	-	-	-	-	-	-	-	-	-	-	-	-
.../...	3	-	3	-	13	15	-	15	-	8	2	57
–	74	37	51	27	87	389	15	302	125	434	79	486
"/	-	-	-	-	-	-	-	-	-	-	-	-
()	-	-	-	-	-	-	-	-	-	-	-	-
-	103	-	-	-	1	4	1	8	12	6	19	176
*	5	-	-	-	-	-	-	-	-	4	-	47
VERSAL	13	0	0	0	0	0	0	0	101	14	15	124

Här visas alla värden per löpord gånger 1000. Tabellen läsas alltså så här: I Abr1 finns det 5,1 utropstecken för varje tusen löpord, i Båt2 finns det 13 frågetecken per tusen löpord, och så vidare.

Tabell 16: Rytm och interpunktion, relativa siffror, per tusen löpord

	Abr1	Katt1	But1	Båt1	Behe2	Kick2	Lag2	Båt2	Eins3	Bedr3	Liv3	Utv3
!	5,1	4,6	3,9	5,5	3,3	4,3	1,4	4,9	-	7,3	2,2	5,7
?	1,2	6,5	13,7	11,9	6,9	8,1	2,9	13,0	0,8	13,8	5,4	6,3
:	8,3	1,5	0,5	1,8	1,4	0,4	1,0	1,8	0,9	3,3	2,0	5,9
,	27,1	21,7	24,9	25,5	23,7	19,4	17,8	39,7	15,8	26,9	17,6	33,8
...	0,6	-	1,5	-	1,6	0,9	-	1,3	-	0,5	0,2	1,7
–	14,6	14,1	24,9	24,6	10,8	22,1	7,2	26,4	5,7	27,9	6,8	14,9
-	20,4	-	-	-	0,1	0,2	0,5	0,7	0,5	0,4	1,6	5,4
VERSAL	2,6	-	-	-	-	-	-	-	4,6	0,9	1,3	3,8

Stockholms universitet
SE-106 91 Stockholm
Telefon: 08 – 16 20 00
www.su.se



**Stockholms
universitet**