Institutionen för lingvistik
Stockholms universitet
HT -99

KÖPSTÖD FÖR SPRÅKGRANSKNINGSPROGRAM

Johan Borg & Helene Eriksson Belaieff

I denna uppsats konstrueras ett verktyg avsett för konsumenter. Detta verktyg ska hjälpa presumtiva kunder att testa diverse språkgranskningsprogram som finns på marknaden idag och i framtiden. Dessutom genomförs en grundlig studie av *Granska*, ett experimentellt program för datorstödd språkgranskning på KTH i Stockholm och en mindre studie av *Grammatifix,* ett kommersiellt språkgranskningsprogram från finska Lingsoft.

Påbyggnadskurs i allmän språkvetenskap med inriktning på datorlingvistik
Uppsats 10 poäng
Handledare: Rickard Domeij

INNEHÅLLSFÖRTECKNING

APPENDIX

1 SYFTE

Uppsatsens syfte är att skapa ett underlag, ett köpstöd, som ska kunna hjälpa köpare av språkgranskningsprogram o.d. att jämföra hur de olika systemen på marknaden behandlar olika typer av fel, vilka för- respektive nackdelar det ena systemet har jämfört med det andra och vilket eller vilka av dessa system som lämpar sig bäst för olika tänkbara användare av språkgranskningsprogram, t.ex. diverse yrkeskategorier och språkstudenter. Vår förhoppning är att, när uppsatsen har slutförts, kunna erbjuda verktyget som en tjänst på webben, där konsumenter kan testa diverse språkgranskningsprogram inte bara idag, utan även i framtiden, med tanke på språkgranskningsprogrammens kontinuerliga utveckling av nya versioner och upplagor. Vi har, i samråd med vår handledare, beslutat att begränsa vår uppsats till endast grammatikkontroll med anledning av att stavnings- och avstavningskontroll finns redan relativt väl utvecklat och dokumenterat och det torde vara mer intressant att utvärdera grammatikkoll. Uppsatsen kommer ej heller att behandla stilistiska fel[1]. I samspråk med vår handledare beslutade vi oss för att koncentrera oss på kongruensfel[2] och särskrivningar[3] och sedan komplettera med diverse andra olika typer av grammatiska fel.

För att detta ska gå att genomföra har vi konstruerat en testbänk, som inte är kopplad till någon specifik fabrikant och som förhoppningsvis ska kunna användas som vägledning i nära anslutning till tidpunkten för ett eventuellt inköp, bestående av exempelmeningar med olika typer av grammatiska fel.

Grammatiska fel finns på olika nivåer: inom ordet, inom frasen, inom satsen, inom meningen och mellan meningar. (Hultman, Westman, 1977 s. 230) Med testbänken som hjälpmedel, eller rättare sagt ett verktyg, skulle alltså konsumenterna kunna få precisa svar på hur det ena programmet reagerade på t.ex. bestämdhetskongruens och om det andra gav felmeddelanden vid felaktiga sammansättningar enligt Svenska Språknämndens *Svenska Skrivregler* § 97.

Vi har kört testbänken i KTH:s *Granska* och finska Lingsofts *Grammatifix* för att kunna dra en slutsats huruvida testbänken är tillförlitlig eller inte och om den eventuellt behöver utökas eller minskas och inte för att få en överblick om vilket system som passar våra behov allra bäst.

2 BAKGRUND

2.1 Historik

Redan under 1960-talet började man utforma datoriserade program för främst automatisk stavningskontroll. Senare började man också att arbeta fram grammatikkoll, s.k. språkgranskningsprogram. (Kukich, 1992). Sådan mjukvara för grammatisk analys analyserar dokument skapade i ordbehandlingsprogram, lokaliserar problem med grammatik, ordbruk och läsbarhet m.m. (Kohut, Gorman, 1995)

---

[1] Stilistiska fel innebär brott mot läsarens estetiska krav på texten. Vanligtvis rör det sig om upprepningar eller blandning av ord med olika stilvärde. (Hultman, Westman, 1977 s. 230)
[2] Kongruens betyder "överensstämmelse", d.v.s. ett ord rättar sin form efter ett annat (Florby & Fält, 1996 s. 52)
[3] Med särskrivning menas att man separerar delar av ett ord så att det blir flera ord, företrädesvis två.

2.2 Tidigare forskning

De program som idag finns på marknaden har i allmänhet stora brister. Anledningen till detta är den enorma komplexitet som komputationell språklig analys innebär. Programmen består av detektion och korrektion. Detektionen uppvisar vanligtvis ett betydligt bättre resultat än korrektionen, eftersom det är mycket lättare att matcha ord gentemot lexikon än att ”rätta” grammatiska fel. Tidigare rapporter tyder på att programmen vanligtvis inte finner mer än 50% av de fel som en text faktiskt innehåller. Även om ett fel detekteras kan felet vara av en annan typ än det som felmeddelandet anger. Dessutom händer det att programmet inte accepterar sådant som faktiskt är grammatiskt. (Kohut, Gorman, 1995)

Trots dessa begränsningar anser många att programmen är till stor hjälp och att dessa är kapabla att tillfredsställa deras behov. En del användare anställer korrekturläsare som komplement. Mjukvaran gör det enklare för dessa företag att formulera affärsbrev, rapporter, offerter m.m. Datorer och datoriserade nätverk installeras snabbare än anställda kan tränas för att nyttja dem. (Kohut, Gorman, 1995)

En av de första, Hanson, att utvärdera sådana system föreslog att systemen skulle presentera information noggrant och med tillägg som gör det, för användaren, lätt att förstå. De skulle också vara felfria, förståeliga, täcka många ämnen inom grammatiken, använda en nivå lämplig och passande för användaren, vara kapabla att stimulera och bibehålla användarens intresse och framställa tämligen tråkiga företeelser på ett roligare sätt. (Kohut, Gorman, 1995)

Även om programmens primära uppgift är att föreslå ändringar och i viss mån detektera icke-ord, så påstår vissa forskare att denna typ av mjukvara även har en pedagogisk funktion och resulterar i bl.a. ökad förmåga att skriva och förbättrad entusiasm inför skrivande. På så sätt får lärare mer tid över till mer avancerade ting såsom innehåll och disposition. (Kohut, Gorman, 1995)

Några av de mest kända programmen på den svenska marknaden är Norstedts *Skribent*, Postens *Prima Vista* och den finska producenten av tjänster för databehandling, Lingsofts, *Grammatifix*.

Det förekom också ett skandinaviskt projekt som hette *SCARRIE* där man utvecklade ett språkgranskningsprogram för de nordiska språken. Man använde autentiska tidningsartiklar från Sverige, Danmark och Norge. Tanken med *SCARRIE*:s program var att det skulle vara kompatibelt med så många olika processorer och operativsystem som möjligt, man skulle kunna köra det både på PC och på Macintosh, programmet skulle ha fyra olika fönster; ett för text, ett för förslagna alternativ, ett för förslag som passade bättre in i den givna kontexten och ett för förklaringar. Det skulle även innehålla adekvat lexikal täckning baserad på tidnings-texter och böcker, kontroll av interpunktions-[4] och stavfel, tillägg av ord och fraser, igen-känning av nya sammansättningar, nya egennamn och syntaktisk analys för att måhända identifiera fel som t.ex. homofoner.[5] (*SCARRIE*, 1999)

[4] Interpunktionsfel är alla fel som berör användningen av “skrivtecken” över ordnivå. (Hultman, Westman, 1977 s. 230)

[5] Med homofon menas ord som uttalas lika men stavas olika (Florby & Fält, 1996 s. 52)

2.3 Granska

Granska utvecklas på KTH:s Nada (Numerisk Analys och Datalogi) och är ett experimentellt program för datorstödd språkgranskning. *Granska* kan liknas vid ett ordbehandlingsprogram med menynamnen Arkiv, Redigera, Format, Verktyg och Hjälp och där innefattas även ordbehandlingsfunktioner såsom formatera och spara text samt ett hjälpsystem med möjlighet till sökning på ordklasser.

Genom att välja granskningskommandot i menyn aktiveras granskningsfunktionen. Sådana fel som programmet detekterar är felaktigt skrivna tecken, stilavvikelser och grammatikfel.

Programmet analyserar varje ord i texten vad gäller ordklass och böjningsform. Detta gör det möjligt att göra en lingvistisk sökning, t.ex. söka på specifika felaktigheter såsom inkongruens eller särskrivning i en given text.

Granska är uppbyggt av en mängd olika regler som beskriver vanliga svenska skrivfel. *Granska* detekterar felaktigheter genom att undersöka om någon regel matchar någon ordsekvens i texten. För att detta ska kunna möjliggöras förses varje ord med en ordklass-tillhörighet eller *tagg* och regler matchas mot ordsekvenser. Ett välkänt problem hos lingvister är dock det att ett ord kan vara flertydigt eller *ambiguöst*. Detta löses genom att varje ord med olika tolkningar förses med alla tänkbara taggar. Ibland räcker det att endast en tolkning matchar, men i många fall måste alla tolkningar matchas för varje regel. Nackdelen med detta system är att då endast en tolkning måste matchas är sannolikheten stor att många fel aldrig upptäcks och då alla tolkningar måste matchas kan det lätt ge upphov till falska alarm. Ett sätt att komma runt detta problem är att man utgår ifrån de mest sannolika tolkningarna baserade på statistiska data.

Vad som vidare försvårar taggningen är att svenskan ideligen förändras genom att nya ord hela tiden konstrueras genom sammansättningar, och ord adopteras från andra språk. Detta innebär att *Granskas* lexikon kontinuerligt måste uppdateras. Återigen använder man statistiska data för att tagga ett ord som saknas i lexikon. Om ett ord slutar på *–ig* är sannolikheten relativt hög att det är fråga om ett adjektiv och således gissar taggaren att ordet har denna ordklasstillhörighet. (Nada/KTH, 1999)

3 DATA

3.1 Testbänkens uppbyggnad

Det var ett ansträngande och tidskrävande arbete att konstruera testbänken. Ett av problemen var att bestämma vad den skulle innehålla för typer av fel. Skulle den endast innehålla konstruktioner med de givna felen eller var det att föredra felen inbäddade i meningar? I många fall var det svårt att sätta en etikett på en feltyp då den kunde innebära att flera ordklasser och syntaktiska funktioner var inblandade. Vari felaktigheten består kan bero på vilket ord man utgår ifrån är det riktiga. I följande exempel kan man fundera över hur det var tänkt att meningen skulle vara.

Ex: 2a fel) *De lovar att *gamla telefonerna* blir som nya med den senaste teknikens kontantkort.

Är det så att det determinativa pronomenet[6] fattas eller skulle ”telefonerna” egentligen stå i obestämd form? Vilket är det rätta? I själva verket är det endast skribenten som vet det, om det inte klart framgår av kontexten vilket som är det rätta. I testbänken har det varit nödvändigt att helt enkelt bestämma sig för ett enda alternativ som det riktiga. I vissa fall kanske två. Vid närmare eftertanke beslöt vi oss för att ange varje feltyp var för sig under respektive rubrik och illustrera respektive typ i exempelmeningar. Varje feltyp med undantag av upprepning av stycke är alltså inbäddad i en exempelmening. För att konstruera meningar tog vi hjälp av de exempel som angavs hos *Granska* och Lingsoft och skapade ett motsvarande exempel där samma typ av fel placerades i en mening och markerades med en asterisk (*). Samma mening återges även i grammatisk version.

Ex: egennamn i genitiv + substantiv i bestämd form (rubricering av feltyp)
Klövbergs grottan (*Granskas* exempel)
Klarabergs gatan (Vårt exempel)
*Det var trafikstockning längs Klarabergs gatan i båda riktningarna.
Det var trafikstockning längs Klarabergsgatan i båda riktningarna.

3.2 Motivering av val av testbänkens feltyper

Vilka feltyper skulle vi då ha med i testbänken? Det relevanta var ju faktiskt sådana fel som ofta görs när man skriver på tangentbord och inte för hand. Eftersom man har sådan frihet i ordbehandlingsprogram och kan ändra i princip hursomhelst, är det vanligt att man ersätter ord (ofta substantiv) som man använt tidigare i texten med en lämplig synonym[7]. Vad användare av ordbehandlingsprogram tenderar att *inte* göra är att byta ut det determinativa pronomenet som bestämmer substantivet. Således blir konstruktionen ogrammatisk och inkongruens uppstår. En stor del av testbänken består därför av inkongruens i en mängd olika skepnader.

Ex: Ett stort hus.
*Ett stort villa.

På *Granskas* och Lingsofts hemsidor[8] återfinns uppräkningar av de grammatiska feltyper som respektive program “klarar av” och vilka de har förhoppningar om att en gång i framtiden kunna detektera. Om vi då använder dessa feltyper i korpusen[9] kan vi utreda vilka fel som detekteras idag och vilka som förhoppningsvis detekteras senare. En hel del av *Granskas* regler är baserade på Svenska Språknämndens *Svenska Skrivregler*.

---

[6] Determinativt pronomen syftar ofta framåt till en relativ bisats eller en att-sats (Florby & Fält, 1996 s. 15)
[7] En synonym är ett ord med nästan lika betydelse (Florby & Fält, 1996 s. 56)
[8] http://www.nada.kth.se/theory/projects/granska/popular.html
http://www.lingsoft.fi/grammatifix/features/grammatik.html
[8]En korpus är en text som används för lingvistisk forskning.

I *Granskas* uppräkning får särskrivningar ett stort utrymme och då föll det sig naturligt för oss att ha med sådana fel i testbänken.

Felaktiga verbfraskonstruktioner har vi också inkluderat i testbänken. Detta med anledning av det nuförtiden alltmer utbredda talspråkets utelämnande av t.ex. hjälpverbet "ha" i indirekt tal och finit verb i huvudsats. Har det möjligen påverkat språkgranskningsprogrammen i den riktningen att man formar dessa efter tendensen att många skriver på samma sätt som de talar?

Något som vi misstänkte redan innan vi testkörde bänken var att, åtminstone *Granska*, inte skulle detektera syftningsfel där en determinerare bestämmer en nominalfras från en tidigare sats i texten. Anledningen till att vi använde denna något långsökta feltyp var främst för att kunna fastslå vår egen teori om att språkgranskningsprogram inte har något "sinne" för kontext. Vi kan heller inte tänka oss hur man i språkgranskningsprogram har kunnat lösa det, inom datorlingvistiken, eviga problemet om världskunskap. (Se vidare Resultat och Diskussion.)

Ex: Ser du boken? De ligger på bordet.
Ser du häftena? Det ligger på bordet. (Andersson, 1993 s. 106)

Vissa feltyper är väldigt specifika och faller inte automatiskt under någon grupp av grammatiska fel, t.ex. kontaminationer[10], medan andra är varianter av en viss sorts fel och kan sorteras in i undergrupper, t.ex. felaktigt sammansatta särskrivningar. Vi har tagit med både och. Syftet med detta har varit att försöka täcka in så många feltyper som möjligt och att specificera dessa i undergrupper. Undergrupperna tillämpades för att tydliggöra skillnader inom vissa feltyper.

Vi har även använt svenska grammatikor som komplement i vårt arbete med att hitta och definiera vissa feltyper i testbänken.

Eftersom det relativt ofta förekommer att språkgranskningsprogram även markerar korrekta meningar som om de vore felaktiga har vi tagit med testmeningar utan grammatiska fel. I anslutning till varje felaktig exempelmening finns dess korrekta motsvarighet. Som rubriker och underrubriker i testbänken har vi använt följande:

- **Bestämdhetskongruens hos substantiv**
  1. Substantiv i genitiv + substantiv i bestämd form (*politikernas sparsamheten)
  2. Ingen artikel + adjektiv + substantiv i bestämd form (*ovana förhandlarna)
  3. Indefinit pronomen + substantiv i bestämd form (*många väktarna)
  4. Bestämd determinerare + adjektiv + obestämt substantiv (*den gamla dam)
- **Bestämdhetskongruens hos adjektiv**
  5. Obestämd artikel + adjektiv i bestämd form + substantiv i obestämd form (*en orörda chokladpudding)
  6. Possessivt pronomen "sin" + adjektivet "egen" + substantiv i obestämd form (*sin egna fantasi)
- **Numeruskongruens**
  7. Felaktigt pronomen (*Var är häftena? -Det…)

[10] Med kontamination menas, inom språkvetenskapen, sammanblandning av två ord eller uttryck. (Norstedts, 1992)

8. Felaktig determinerare + adjektiv + substantiv (*de största världshavet)
9. Felaktigt adjektiv + substantiv (*grön tomater)

- **Genuskongruens**
  10. Felaktig determinerare + substantiv (*en problem)
  11. Felaktig determinerare + felaktigt adjektiv + substantiv (*ett tjockt bok)
  12. Felaktigt adjektiv + substantiv (*söt vin)
  13. Felaktigt pronomen + substantiv (*någon exemplar)
- **Predikativkongruens**
  14. Substantiv + kopula + felaktigt adjektiv/particip (*tillstånd var stabil)
- **Verbkongruens**
  15. Tempuskongruens (*åt…skall)
  16. Personkongruens (reflexiva verb) (*mannen…uppförde honom)
- **Verbkonstruktioner**
  17. Supinum utan ”ha” (*skulle stannat)
  18. Dubbelt supinum (*försökt löst)
  19. Dubbelt passivum (*bestämts att satsas)
  20. Infinitiv efter preposition (*till ge)
  21. Infinitiv utan ”att” (*hade för avsikt resa)
  22. Infinitiv med ”att” (*Sluta att viska…!)
  23. Inget verb (*Inga ljud. Ingenting)
  24. Inget finit verb (*redan åkt)
- **Maskulinform hos adjektiv**
  25. Maskulinform hos adjektiv + feminint substantiv (*den arbetslöse lärarinnan)
- **Ordföljd i bisats**
  26. Placering av adverb i bisats (*…att den blir förmodligen)
  27. Placering av negerat led i bisats (*…att det var inte)
  28. Ledföljd i indirekt frågesats (*…varför sköts planteringarna)
- **Övrigt**
  29. Dubbel negation (*inte varken)
  30. Prepositionsbruk vid tvådelad konjunktion (*på kvällen och morgonen)
  31. Pronomenets form efter preposition (*till han)
  32. Konstruktionen ”möjligast” + adjektiv (*möjligast troliga)
  33. Kontamination (*så snarast som möjligt)
  34. Felaktig preposition (*i enlighet till)
  35. Tautologi[11] (*just nu i detta ögonblick)
- **Särskrivningar**
  36. Ord som alltid skall sammanskrivas enligt *Svenska Skrivregler* (*var annan)
  37. Egennamn i genitiv + substantiv i bestämd form (*Klarabergs gatan)
  38. Substantiv i genitiv + substantiv i bestämd form (*bolags stämman)
  39. Substantiv i bestämd form + substantiv i obestämd form (*hus taken)
  40. Substantiv i obestämd form + substantiv i obestämd form (*bransch kännedom)
  41. Substantiv i obestämd form & genitiv + substantiv i obestämd form (*matlagnings vin)
  42. Substantiv + particip (*fond sparande)
  43. Egennamn/akronym i nominativ + substantiv (*Skansen terrariet)
  44. Adjektiv + substantiv (*stor stugan)
  45. Ofullständigt förled + innehållsord (*fåra kläder)
  46. Innehållsord + osjälvständigt efterled (*moder skapet)

[11] En tautologi är en upprepning av något redan sagt eller nämnt, fast med andra ord.

- **Upprepningar**
  47. Upprepningar av ett ord
  48. Upprepning av fras
  49. Upprepning av ett stycke
- **Ofullständiga satser**
  50. Ingen nominalfras
  51. Ofullständig bisats

3.3 Testmaterialets struktur

Med något undantag representeras alla sorters fel av två olika exempelmeningar, a) och b). Felaktigheten, ordet eller orden, är inte identiska i de två meningarna. Avsikten är att undersöka om programmet hittar felen i båda meningarna eller inte hittar något av dem. Vi finner det också intressant att undersöka huruvida programmen uppvisar inkonsekvens eller inte. För att få en tillförlitlig statistik inom respektive sorts fel skulle man naturligtvis ha många fler testmeningar, men det skulle innebära att testbänken skulle bli otymplig och svåranvänd för sitt syfte. Två testmeningar kan möjligen ge missvisande information om slumpen styr, men förhoppningsvis ger det ändå en bättre fingervisning av programmets förmåga att upptäcka en viss typ av fel än om endast ett fel funnits med.

Testbänken, d.v.s. det testmaterial som är avsett att köras i språkgranskningsprogrammet, består av en dokumentfil som rensats på rubriker, anvisningar, källhänvisningar och all annan text än just numreringarna av testmeningarna. Detta för att språkgranskningsprogrammet endast ska behandla just testmeningarna och inte övrig text, då analys av denna förvirrar testaren och dessutom är skrymmande och gör materialet svårare att överblicka.

Vid sidan av testbänksfilen finns också en fil, användarfilen, där samma exempelmeningar återfinns med sina respektive rubriker och underrubriker. Där finns också en möjlighet för testaren att fylla i resultatet för respektive mening i ett litet formulär som återfinns bifogat under varje rubrik. Se vidare Appendix 3.

Förutom själva urvalet av testmaterial har det funnits ett annat problem, nämligen hur den som använder denna test ska kunna få en överblick över det resultat han eller hon har fått. Om man utgår ifrån att testmaterialet kan hämtas hem från webben så krävs det att testaren kan fylla i resultatet för respektive feltyp och att denna information behandlas automatiskt och resulterar i någon sorts summering i form av tabell e.d. Ett alternativ är att användarfilen skrivs ut på papper innan testet genomförs och att resultaten skrivs in manuellt vartefter. Resultaten skulle sedan räknas ut och skrivas in i formulär manuellt, dels för överordnade grupper, dels för hela testbänken.

Den främsta anledningen till att vi valde olika alternativ med tillhörande rutor var att konsumenten själv skulle kunna använda vår testbänk för att utföra liknande experiment i framtiden. Detta för att kunna bestämma vilket program som passar dennes behov allra bäst. För en köpare är det ju inte särskilt intressant huruvida *Granska* klarar av att detektera felaktiga sammansättningar enligt *Svenska Skrivregler* § 97 och om *Grammatifix* ger felmeddelanden vid person- och verbkongruens i december -99. All form av programvara har en påtaglig tendens att ändras, och framför allt förbättras, bara under några månader. Idag är dessutom inte användandet av språkgranskningsprogram så utbrett att så många tester skulle göras. Vad som är poängen med testbänken är att konsumenter i framtiden ska kunna använda den för att utvärdera programmens kapacitet vid olika tillfällen och tidpunkter. Med största sannolikhet är inkongruens ett vanligt fel vid användande av ordbehandlingsprogram även om några år och då kan testbänken komma till användning. Köparen kan således använda verktyget på webben för att testa program och kryssa i respektive ruta för respektive test och utifrån detta utläsa ett resultat om vilket program som förefaller mest lämpligt. Vi vill dock inte påstå att testbänken är komplett. Vi hoppas däremot att vi har konstruerat en prototypisk testbänk som lämnar dörren öppen för senare utvidgning och förbättring.

4 METOD

4.1 Förberedelse för testkörning

När vi började arbetet med att beskriva grammatiska fel i olika feltyper förstod vi att det skulle bli svårt att genomföra vår tidigare ambition att ställa samman autentiska tidningsartiklar, rapporter o.d. till en korpus med fel i. Att hitta ”rätt sorts fel” i lagom mängd är osannolikt. Det skulle även vara viktigt att hitta ett slags balans mellan felen så att inte en viss typ av grammatiska fel var högfrekventa och andra typer var lågfrekventa. Detta skulle i så fall leda till svårigheter med att göra någon statistisk jämförelse mellan programmens förmåga att detektera och behandla olika typer av grammatiska fel. Dessutom skulle dessa artiklar förmodligen vara korrekturlästa. Att ”plantera in” fel skulle kanske gå men det skulle förutsätta omsorgsfull korrekturläsning och redigering för att den övriga texten inte skulle innehålla en massa oförutsedda fel. En testbänk däremot skulle vi kunna utforma helt och hållet efter egna behov och önskemål och därigenom skulle det bli lättare att testa materialet och systematisera detta, vilket vore omöjligt med en korpus innehållande artiklar skrivna av professionella journalister vars kunskaper i att utrycka sig i skrift torde vara alltför bra för att en mängd olika fel skulle påträffas. Därmed valdes detta sätt att genomföra undersökningen på. Med detta inte sagt att en korpus bestående av artiklar vore bristfällig till punkt och pricka. En annan lösning vore att komplettera testbänken med en korrekturläst text där man på så sätt skulle kunna styrka teorin att det förekommer fall där programmen anser att grammatiska satser faktiskt är felaktiga och av den anledningen har vi även bifogat en korpus av detta slag till testbänken.

Testbänkens sammansättning består av två exempel om 51 olika feltyper. Överst har vi specificerat feltypens karaktär och numrerat varje typ. Varje exempelmening utgörs av den felaktiga satsen med asterisk (*) och sedan den, som tidigare nämnts, korrekta satsen och i några fall även en komplex mening som om möjligt försvårar för språkgransknings-programmet. Den korrekta meningen är dock bara intressant i egenskap av grammatisk mening, d.v.s. avsikten var såväl att studera eventuella felmeddelanden vid korrekta satser, det som vi tidigare i uppsatsen kallade ”falskt alarm”, som att ange den korrekta motsvarigheten till den felaktiga satsen eller del av sats. Det felaktiga har vi även skrivit med kursiv text. Detta för att förtydliga vilket/vilka ord eller vilken fras i satsen som är fel. Det har också varit en stor fördel för vår egen del i och med att vi, med enkla medel, har kunnat rubricera respektive feltyp och på så sätt lyft fram testbänkens struktur.

Ex: **1. Substantiv i genitiv + substantiv i bestämd form**

1a fel) *- *Politikernas sparsamheten* har väl bidragit till den gynnsamma utvecklingen, sade han med ett skratt.

1a rätt) - Politikernas sparsamhet har väl bidragit till den gynnsamma utvecklingen, sade han med ett skratt.

1b fel) **Pistens kungen* hade börjat fjälla på näsan.
1b rätt) Pistens kung hade börjat fjälla på näsan.

Trots vår selektering vad gäller feltyper har det ändå varit svårt att få med så mycket som möjligt och att inte råka utelämna något viktigt område. Dessutom ska det helst råda något slags jämvikt mellan de olika typerna, för att det ska gå att utvärdera resultatet av testkörningen på ett bra sätt, d.v.s. det ska vara relativt lätt att kunna se att språkgransknings-programmet klarar en viss sorts fel bra och en annan sorts fel mindre bra.

4.2 Testkörning

4.2.1 Testkörning av Granska

Granskas grammatikkoll finns i en demo-version som en tjänst på webben[12] och därför var det enkelt att testköra i detta program. Testbänken hade vi sparat i en fil i txt-format och det var bara att ange sökvägen och filnamnet. Sedan kom svaret ut på en separat sida och vi kunde då utläsa resultatet. *Granska* markerar misstänkta områden i texten med *rött*, presenterar ersättningsförslag med grönt, ger en kortfattad kommentar som skrivs efter varje rättningsförslag i *blått*. Ett F efter förslaget betyder att inga ändringar har gjorts, vilket indikerar ett möjligt falskt alarm. Ett E efter förslaget betyder att rättelsen kan ge ett annat fel. Om en mening innehåller misstänkta fel presenteras samma fel flera gånger. Så här kan ett rättningsförslag se ut:

Ex: 3b fel) *Varje utställaren fick 30 biljetter att dela ut.
önskad böjningsform saknas Om Varje syftar på utställaren är det kongruensfel
Varje utställare fick Om Varje syftar på utställaren är det kongruensfel
utställaren fick Om Varje syftar på utställaren är det kongruensfel
önskad böjningsform saknas Om Varje syftar på utställaren är det kongruensfel

3b rätt) Varje utställare fick 30 biljetter att dela ut.

Som tidigare nämnts händer det också ibland att språkgranskningsprogram detekterar och ger felmeddelanden till sådan text som *inte* är felaktig. Det gäller framför allt väl korrekturlästa tidningsartiklar. Därför ville vi använda ett utdrag ur den s.k. SUC-korpusen[13]. Utdraget normaliserade vi i UNIX-miljö och sparade på skrivbordet så att det skulle kunna användas i Windows-miljö, varpå det sparades som ett dokument i txt-format. Även här angavs sökvägen

[12] http://www.nada.kth.se/theory/projects/granska/demo.html
[13] SUC (Stockholm - Umeå Corpus) är en textsamling som består av ca 1 miljon ord modern skriven svenska. (Stockholms universitet, 1999)

och filnamnet till demo-versionens webbplats och det testkördes på exakt samma sätt som testbänken. Utdraget ur SUC-filerna är dock inte del i testbänken. En stor skillnad mellan utdraget ur SUC-filerna och testbänken var att i testbänken hade vi staplat upp exempel-meningar, medan utdraget ur SUC-filerna innehöll löpande text. Se utdrag ur SUC-fil och testbänken i Appendix 1.

4.2.2 Teskörning av Grammatifix

Till skillnad från *Granska* finns inte *Grammatifix* i en demo-version som en tjänst på webben. För att få access till *Grammatifix* installerade vi detta program i befintligt ordbehandlingsprogram. Dokumentet med testbänken och SUC-filen öppnade vi i ordbehandlingsprogrammet. För att sedan testköra dokumentet markerade vi allt och angav via Vektygs-menyn kommandot ”Grammatifix”. Detta för att vi skulle köra hela dokumentet. Man kan, om man så önskar, välja att markera den del av texten som man anser intressant att kontrollera. På skärmen dök då en dialogruta upp med det första i texten som programmet anmärkte på. (Om det visar sig att texten inte innehåller något som *Grammatifix* anmärker på så ger programmet endast en statistisk översikt över det markerade avsnittet.) *Grammatfix’* dialogruta innehåller tre fönster med tillhörande rubriker samt ett antal funktionsknappar.

Huvudrubriken, där feltypen anges, t.ex. bestämdhetsform hos substantiv står ovanför det största fönstret som är placerat högst upp i dialogrutan. I meningsfönstret återfinns den mening som programmet kommenterar. Den del av meningen eller det ordet kommentaren avser markeras med *blått*. I förklaringsfönstret, som placeras under meningsfönstret, redovisar programmet kommentaren om varför anmärkning har skett, t.ex. *Kontrollera ordformen sparsamheten. Om ett substantiv styrs av en genitiv, t.ex. Politikernas, bör det stå i obestämd form.* I förslagsfönstret allra längst ned i dialogrutan anges ibland även ett ersättningsförslag.

4.3 Behandling och redovisning av resultat

Med *Granskas* respektive *Grammatifix’* resultat på papper kunde vi börja gå igenom varje exempel var för sig och på så sätt se

- i vilka fall alarm[14] gavs och rättningsförslaget var tillfredsställande,
- i vilka fall alarm gavs men rättningsförslaget inte var tillfredsställande,
- i vilka fall alarm gavs utan rättningsförslag ,
- i vilka fall alarm inte gavs trots felaktighet,
- i vilka fall alarm inte gavs i korrekt text
- i vilket/vilka fall det fanns skäl för andra anmärkningar och egna kommentarer, inte beträffande detektion av satsen eller angivet rättningsförslag, utan sådant som är svårt att dela in i kategorier och lämnas till användarens subjektiva tolkning.

Givet dessa utfall summerade vi resultatet i en tabell där vi för varje exempel angav det eller de alternativ som passade in. För varje exempel kryssade vi i den ruta som överensstämde med testets resultat. Tack vare detta kunde vi också göra uträkningar om hur många fel som

[14] ”Alarm” innebär att felaktighet upptäcktes

detekterades, dra slutsatser om vilka typer av fel som inte korrigerades och om programmet vid tillfälle uppvisade inkonsekvens, etc. Se vidare Resultat.

5 RESULTAT OCH DISKUSSION

5.1 Vilket är resultatet?

Till att börja med måste vi poängtera att det egentliga resultatet för detta arbete består av själva testbänken. Att sammanställa hur testbänken behandlades av *Granskas* och *Grammatifix* språkgranskningsprogram är av intresse i första hand för att se om testbänken är ett fungerande verktyg för att testa den här sortens program. För att undersöka detta är det viktigt att se hur programmen arbetar idag, hur de behandlar texten i stort och hur de markerar och kommenterar fel. Det bör noteras att *Granskas* grammatikgranskare är en preliminär version i ett pågående projekt och att man inte påstår sig kunna hantera så väldigt många typer av grammatikfel än så länge. Gränssnitt och rättningsförslag är inte heller färdigarbetade.

Att köra testbänken i *Granska* över internet visade sig vara mycket enkelt. Hemsidan var tydlig och anvisningarna för att testa text var enkla. Vi valde när vi körde vår testbänksfil att få alla meningarna presenterade. Det går även att få endast de misstänkt felaktiga meningarna presenterade, vilket i vårt fall kanske hade gått lika bra. Vi valde dessutom att få rättnings-förslagen men inte ordklasstaggarna eller de matchande reglernas namn för att inte få en alltför omfattande textmängd tillbaka. Vi tror att en presumtiv kund inte har någon glädje av taggar eller regler. För att själva kunna förstå varför programmet hanterat feltyperna som det gjort genomförde vi ändå en andra körning av testbänken där resultatet kompletterades med taggar och regler. Därigenom fick vi förklaringar till vissa felaktiga rättningsförslag.

När det gäller *Grammatifix* så är det ett kommersiellt program vars senaste version vi fick testköra precis innan det släppts på marknaden. Dock hade vi inte tillgång till programmet förrän i ett sent skede av vårt arbete varför vi begränsat analysen av testresultaten. Se vidare avsnitt 5.9.

5.2 Resultatets aktualitet i *Granska* och *Grammatifix*

Då *Granska* är ett projekt under utveckling pågår ständigt arbete med dess uppbyggnad. Detta medför att förutsättningarna för en testkörning inte bara ändras från månad till månad utan kanske från ena dagen till den andra. Vi fick i efterhand höra att om man vetat exakt när vi skulle testköra *Granska* så hade en del regler möjligen tagits bort först. Detta medförde att vi fick vissa felmeddelanden som inte var relevanta, men detta har knappast haft någon större betydelse för undersökningen i stort, eftersom det inte var vår målsättning att utvärdera *Granskas* regelverk.

Det resultat som ges för *Grammatifix* kan sägas gälla så länge inte en ny version av programmet finns tillgänglig.

5.3 Hur svaren placerades in i kategorierna

Svaren från testkörningen delades in i de fem kategorier som nämnts tidigare. I många fall kunde svaret på en exempelmening sättas in under mer än en kategori, t.ex. kunde fler rättningsförslag anges för en viss sorts fel och av dessa var kanske ett bra och ett annat felaktigt i sammanhanget. Vi valde att låta det vara möjligt för ett fel att ha svar i flera kategorier samtidigt. Detta visade sig nödvändigt av flera skäl. Dels för att få med det faktum att både bra och dåliga rättningsförslag förekom till samma fel, dels för att vissa meningar innehöll fel som aldrig upptäcktes samtidigt som programmen gav falskt alarm i en annan del av meningen. Av rättningsförslag som angivits hände det också att något av dem var en korrekt rättning i sig men att det var en långsökt lösning. Det var många gånger svårt att sätta gränser för vad som skulle vara en tillfredsställande rättning och vad som inte kunde anses vara det. Se Appendix 2.

För att inte få alltför många svarskategorier hade vi valt att låta a- och b-versionen av meningen samsas om en rad i svarsformuläret. Det visade sig i efterhand att detta kunde medföra förvirring om svaret förutsattes hamna i kolumnen med bocken, nämligen där programmet givit falskt alarm. De fyra första kolumnerna handlar bara om de planterade felen, medan en markering i femte kolumnen kunde vara ett falskt alarm antingen i en helt korrekt b-mening eller i korrekt text i en a-mening som således innehöll fel någon annanstans. Vi hade visserligen lagt till en kolumn för egna noteringar, men frågan är om det räcker.

5.4 *Granskas* felmeddelanden

5.4.1 Alarm med rättningsförslag

Det optimala är att programmet alarmerar fel och ger bra rättningsförslag. Det är dock inte alltid lätt att följa logiken i Granskas presentation av dessa förslag. Det är t ex intressant att jämföra hur programmet kommenterar 24a fel) och 24a rätt). Notera att programmet behandlar ett ställe i meningen i taget. Om det är fler ställen i samma mening som kommenteras så upprepas först exempelmeningen.

24a fel) *Det hänt att jag funderat på att sluta röka.
Finit verb verkar saknas i satsen
*24a fel) *Det har hänt att jag funderat på att sluta röka. E supinum utan verbet ha*
*24a fel) *Det hade hänt att jag funderat på att sluta röka. E supinum utan verbet ha*

24a fel) *Det hänt att jag funderat på att sluta röka.
**har funderat Supinum utan verbet ha*
**hade funderat Supinum utan verbet ha*

24a fel) *Det hänt att jag funderat på att sluta röka.
sluta att röka E om röka står i infinitiv bör det föregås av att

24a rätt) *Det har hänt att jag funderat på att sluta röka.
sluta att röka Om röka står i infinitiv bör det föregås av att

Den första kommentaren om att finit verb verkar saknas ges endast vid första tillfället. Detta är inget att anmärka på egentligen. Det intressanta är att sedan anges hela meningen med inledande meningsnumrering och programmet föreslår ett finit verb i antingen presens eller preteritum. De finita verben sätts in direkt vid det första verbet i meningen. Vid nästa felmarkering däremot anges först den ursprungliga meningen igen, men sedan varken numrering eller hela meningen utan bara nästa verbkonstruktion med ett föreslaget finit hjälpverb. Detta upprepas sedan ytterligare en gång i det tredje rättningsförslaget *sluta att röka*. Om detta är ett system så är det gott så. I den korrekta meningen som följer efter den felaktiga har *Granska* dock inte upprepat hela meningen. Det har man för övrigt inte gjort på något annat ställe i vår testbänk. Inte heller har något finit verb föreslagits till verbet *funderat.* Om den första tanken är att programmet kommer ihåg en tidigare rättning så kan man snabbt glömma detta eftersom *Granska* återigen anmärker på konstruktionen *sluta röka*. Det framgår dessutom från övriga testbänken att något sådant minne inte finns.

5.4.2 Alarm utan rättningsförslag

Under ovanstående rubrik samlades svaren från *Granska* där programmet upptäckt felet vi åsyftat i meningen och givit en kommentar men inte givit något förslag på hur rättningen skulle se ut. Det blev inte många sådana svar, endast fem stycken. När man jämför kommentarerna till dessa meningar, så är de av helt olika karaktär, vilket i sin tur beror på felens art. En kommentar till 19a fel) och 19b fel) är *Dubbelpassivum anses osnyggt av många.* När det gäller 23a fel) och 23b fel), är kommentaren *Verb verkar saknas i satsen.* Att ge ett rättningsförslag för att undvika dubbelpassivum borde gå att utveckla, men om ett verb saknas kan programmet knappast föreslå ett.

En kommentar som syftar på 22b fel) är *Verbfrasfel kanske*, och verbfrasen som åsyftas är *började att avta.* Om programmet tycker sig ha hittat ett verb som inte bör följas av "*att* "så borde det vara enkelt att göra ett rättningsförslag till detta. Å andra sidan kan det förstås finnas en poäng i att inte komma med specifika rättningsförslag om programmet inte är helt säkert på felets art.

Två meningarna som vi haft med som korrekta exempelmeningar, men som ändå blev markerade var följande:

33b rätt) Som tur var anlände tåget mer försenat än jag.
det är ovanligt med ett verb efter verbet var

33b rätt) Som tur var anlände tåget mer försenat än jag.
önskad böjningsform saknas Komparativformen är kanske att föredra

35a rätt) *Ett brott begås i detta ögonblick.
Finit verb verkar saknas i satsen

Den första kommentaren är fullt berättigad även om det i just det här fallet inte innebar något fel. I den andra kommentaren föreslår *Granska* att man ska fundera på om komparativformen är bättre. Det är den naturligtvis inte. Det tycks som om detta svar bygger på regler som inte är helt färdiga.

Den tredje kommentaren som anger att det kanske saknas ett finit verb är förvirrande. *Begås* är ju ett finit verb, men kanske finns det inte i *Granskas* lexikon?

5.4.3 Falska alarm

Falska alarm är de markeringar Granska gjort på grammatiskt korrekt text oavsett om den är i en mening som benämnts "rätt" eller "fel". *Granska* lyckades mycket bra med att lämna grammatiskt riktiga meningar orörda. Endast tre av de meningar som benämnts *rätt* blev markerade som oriktiga av programmet Att svarsfrekvensen totalt blev så hög, 43 st, är missvisande. Fr.o.m. exempel 10 markerades alla b-alternativ "fel" som en felaktig särskrivning, sammanlagt 40 st.

Ex. 10b fel) *Ur det synvinkeln är förslaget förkastligt.
10bfel) Felaktig särskrivning Svenska Skrivregler §97

Det blev totalt 40 st felaktigheter som egentligen är ett och samma fel, och som dessutom inte är relevanta i sammanhanget p.g.a. att de inte ingår i själva meningarna. Programmet vet inte skillnaden på numreringen och själva texten, i synnerhet inte som vår numrering inte endast består av siffror. Alltså betraktas numreringen som ett ord och förses med en tagg. Taggarna anger att 10b är ett substantiv när det följs av *fel*, men ett adjektiv när det följs av *rätt.*

10b<nn.utr.sin.ind.nom< fel<nn.neu.sin.ind.nom<)<pad> Ur<pp>

10b<jj.pos.utr.sin.ind.nom> rätt<nn.utr.sin.ind.nom>)<pad> Ur<pp>

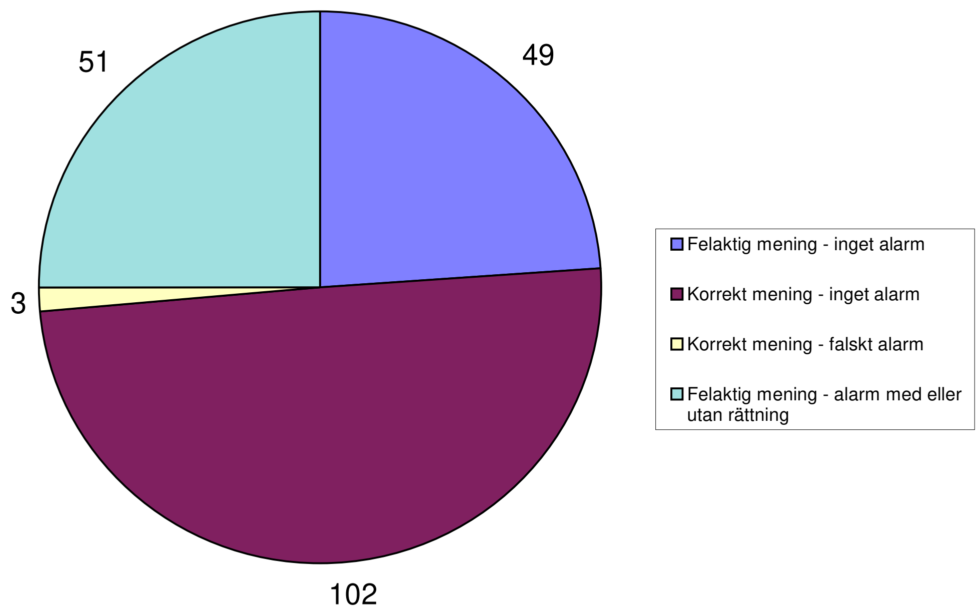
Ändå är både *fel* och *rätt* taggade som substantiv, det ena i neutrum och det andra i utrum.

Att ha en fungerande numrering i testbänksfilen var en nödvändighet för att hantera materialet, både för att härleda meningarna till användarfilen med rubriker och för att kunna orientera sig i programmets svar, där upprepningar av en och samma mening förekommer beroende på antal fel som upptäcks. Dessutom är det svårt att skilja den korrekta testmeningen från den felaktiga vid en snabb överblick om man inte tydligt kan skilja på rätt-meningar och fel-meningar. Således har vi behållit numreringen och redovisar svaren även om de på ett sätt kan anses irrelevanta.

5.5 Resultatet av testkörningen i *Granska*

Av de totalt 205 exempelmeningar som ingår i testbänken finns 100 st felaktigheter, d.v.s. 100 st meningar innehåller vardera ett fel och 105 st meningar innehåller inget fel. Om programmet hade markerat endast de planterade felen och dessutom endast givit ett riktigt och inget felaktigt rättningsförslag skulle alltså resultatet ha medfört 100 st felmarkeringar. I figur2 kan man se i hur många av exempelmeningarna *Granska* larmat - även falska alarm. Den faktiska siffran visar dock att *Granska* givit 91markeringar för felaktigheter. Detta betyder inte att nästan alla felaktigheter upptäckts. I själva verket är nästan hälften, 49 st, av felen oupptäckta. Programmet upptäckte 51 av de 100 planterade felen. I 39 fall av dessa 51 gjorde *Granska* en riktig grammatisk analys av felet och gav bra rättningsförslag. I sex fall gavs förutom ett riktigt förslag även oacceptabla rättningsförslag. Tre falska alarm gavs i meningar som benämnts "rätt". Vad som ej framgår av figur 2 är att 40 falska alarm gavs i meningar med inplanterade fel, se vidare avsnitt 5.4.3 Falska alarm.

Det förekom också falska alarm på stavfel i texten på några ställen, men då det faller utanför vår undersökning har vi inte räknat med dessa.

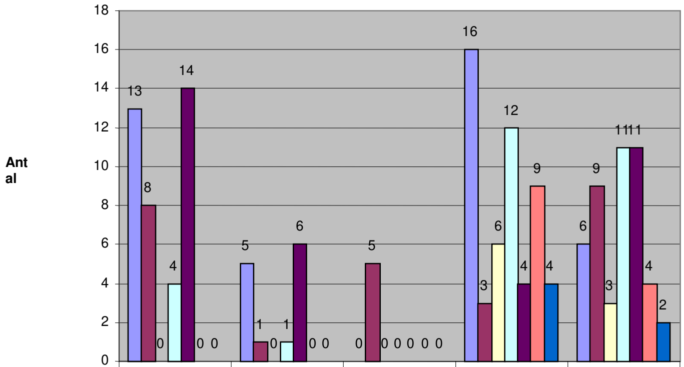


Figur 1

Resultat för testkörningen av de 205 exemplen i Granska .

5.6 Resultatet som statistikunderlag

Såsom testbänken är konstruerad är inte de olika delarna av testmaterialet avvägda så att det går att göra jämförande statistik över resultaten. Från början var avsikten att detta skulle låta sig göras, men efter hand framstod det som orimligt att vikta de olika delarna. Större områden som kongruenser och särskrivningar är svåra att jämföra med specifika grammatiska företeelser som kontaminationer, tautologier eller upprepningar av ord. Ändå har vi tittat på vissa områden för sig, då vi tror att det kan vara av intresse för en presumtiv köpare av ett språkkontrollprogram. En överblick över de olika områden som behandlats ges i Figur 2 nedan.



	Alarm, rättning tillfredsställande	Alarm, rättning ej tillfredsställande	Alarm, inget rättningsförslag	Inget alarm trots felaktighet	Falskt alarm i korrekt text
Kongruenser (mening 1-16)	13	5	0	16	6
Verbkonstruktioner (mening 17-24)	8	1	5	3	9
Ordföljd I bisats (mening 25-27)	0	0	0	6	3
Övrigt (mening 28-35)	4	1	0	12	11
Särskrivningar (mening 36-46)	14	6	0	4	11
Upprepningar (mening 47-49)	0	0	0	9	4
Ofullständiga satser (mening 50-51)	0	0	0	4	2

Figur 2
Granska: Resultaten fördelade på respektive grammatiska rubriceringar

5.7 Diskussion kring respektive feltyp

Som tidigare nämnts är de olika feltyperna mycket olika till sin karaktär. Vi har inte funnit några exakta siffror på hur vanliga de olika feltyperna är i förhållande till varandra. En del får större utrymme i vår undersökning än andra. Nedan kommenterar vi *Granskas* behandling av respektive feltyp.

5.7.1 Kongruensfel

Exempel 1-16 handlar om kongruensfel och där hade vi trott att programmet skulle hitta fler fel än det faktiskt gjorde, eftersom kongruenser hör till de fel som *Granska* anser sig klara, åtminstone till viss del. Visserligen finns det ett stort antal olika kongruenser men i våra exempel är de flesta sådana att orden står i nära anslutning till varandra, varför det skulle vara lättare att hitta felen än om de funnits t.ex. inskjutna satser emellan.

Det var i påfallande många fall som exempelmening a och b inte gav likadana resultat. För att jämföra de olika exempelmeningarnas a- och b-resultat har vi först räknat bort exempel 6 som inte består av något a- och b-alternativ. Av de övriga 15 meningarna har fyra upptäckts och rättats tillfredsställande likadant för a- och b-alternativet. Fem meningar har innehållit fel som varken hittats i a- eller b-alternativet. Sex meningar har innehållit fel som endast upptäckts i det ena fallet.

Vi förvånades t ex över att *många väktarna* inte uppmärksammades medan *varje utställaren* rättades. Likaså anmärktes på *bara grön tomater* men inte på *flera sen nätter*, där *sen* taggats som adverb istället för adjektiv, likaså på *en problem* men inte på *det synvinkeln.* Andra exempel som förvånar är att *någon exemplar* inte avslöjas som fel men däremot *vilket färg som helst.*

Ett annat exempel är att *en överskådlig diagram* hittades av programmet men rättningsförslaget blev inte det förväntade *ett överskådligt diagram*, utan istället *några överskådliga diagram.* Detta framstår som en onödigt långsökt rättning, om än korrekt grammatisk i sig med tanke på att kontexten inte utesluter vare sig det ena eller det andra.

Exempel 7 består egentligen av två meningar: *Ser du boken? De är på bordet*. Eftersom vi visste att *Granska* analyserar varje mening för sig var det väntat att inget fel skulle hittas. Ändå valde vi att behålla exemplet i den formen. Det kan vara intressant för en eventuell konsument att bli uppmärksammad på detta. Utan att veta exakt hur långt utvecklade olika språkgranskningsprogram är kan man anta att den här typen av kontextkänslighet kan komma att lösas i programmen. Speciellt i enklare fall som detta där pronomenet kommer alldeles efter korrelatet[15].

5.7.2 Verbkonstruktioner

Bland de 16 felaktiga verbkonstruktionerna upptäcktes åtta som rättades på ett bra sätt. Det var bl.a. fel med anknytning till supinumformen där *Granska* upptäckte att "ha" saknades före supinum, t ex *kunde ha slutat, skulle ha stannat.* Dubbelsupinum upptäcktes också men bara i ett av två fall. *Hade velat köpt* markerades men inte *hade försökt löst.* När det gäller dubbelt passivum upptäckte *Granska* detta men gav inget rättningsförslag utan endast kommentaren att det inte anses snyggt**.**

Att infinitiv inte föregåtts av "att" har *Granska* hittat och rättat vad gäller mening 21b fel) där programmet vill sätta in ett "att" mellan *kommer* och *innebära. Ha för avsikt resa* undgår denna anmärkning. I dagligt språkbruk stöter man ofta på exempel som dessa och eftersom människor i allmänhet inte entydigt anser att de är felaktiga kan det möjligen vara tveksamt om det ska betraktas som absoluta fel. Avsaknad av infinitiv med "att" efter en preposition har markerats och rättats i de båda exemplen i testbänken. Här är det tydligare att det handlar om en grammatisk felaktighet som bör korrigeras.

Exempelmeningarna 23a fel) och 23b fel) saknar verb och det är den kommentar som *Granska* ger. Detta innebär att fel upptäckts men inte rättats och det är det enda rimliga då det inte finns något som talar för vilket verb som skulle vara lämpligast.

5.7.3 Ordföljd i bisats

Felaktig placeringen av adverb i bisatser har inte alarmerats av *Granska*, vilket förvånade oss lite. Reglerna för ordföljden tycks inte vara kompletta.

[15] Ett korrelat är en referent som ett pronomen syftar tillbaka på (Andersson, E, 1993 s. 40)

5.7.4 Övrigt

Exempel 28 - 35 omfattar relativt specifika feltyper som egentligen inte kan behandlas i grupp. Att hitta en felaktigt maskulin böjning som skall kongruera med ordet kvinna gick bra. Däremot lämnades exemplet *den arbetslöse lärarinnan* utan notering. Varken lärarinna eller kvinna är taggat som feminimum, medan däremot *arbetslöse* respektive *vidsynte* är taggat som maskulinum.

Användning av dubbel negation, t ex *inte…varken eller*, eller *inga kunde inte berätta* borde ha markerats, kan man tycka. Däremot är det mer tveksamt om språkbruket i exempel 30a är felaktigt. Det är enligt vår åsikt tämligen vanligt både i skriftspråk och talspråk att utelämna preposition framför det andra av två substantiv där en konjunktion är placerat emellan.

Sammanblandningar av uttryck och tautologier (tårta på tårta-uttryck) var tydligen för svåra att upptäcka. I vissa fall kan här förekomma vacklande språkbruk där åtminstone tautologier används för att understryka något i texten.

5.7.5 Särskrivningar

Av de 22 st felaktiga särskrivningarna från 11 st feltyper har 14 st hittats och rättats tillfredsställande. Fyra upptäcktes inte alls, och sex särskrivningar tillskrevs rättningsförslag som var oacceptabla. 16 av särskrivningarna återfinns i åtta av feltypsexemplen och innebär alltså att programmet har behandlat a- och b-alternativen konsekvent, och av dessa är endast två rättade konsekvent men på ett oriktigt sätt. Hur konsekventa *Granskas* svar är för de olika grammatiska områdena i testbänken framgår i figur 3.

Särskrivningar som programmet hade svårt att hitta var de som är konstruerade med egennamn eller akronymer när dessa stod i genitiv. *Strindbergs museet* fick som rättningsförslag *Strindbergs museum*, vilket är en förståelig slutsats, om än felaktig i det här sammanhanget. I stället för *Klarabergs gatan* föreslogs *Klarabergs gata.* Utan tidigare nämnda omvärldskunskap kan det även i framtiden bli svårt att få sådana exempel att bli rätt.

5.7.6 Upprepningar

Granska alarmerade inte för några typer av upprepningar. Eftersom programmet arbetar med en mening i taget så förväntade vi oss inte att det skulle upptäcka upprepningar av meningar. Däremot hade vi trott att enstaka ord efter varandra skulle hittas.

Upprepning av fraser och även av hela stycken tror vi är ett så vanligt förekommande fel när man skriver och redigerar texter i ordbehandlingsprogram att vi anser det vara viktigt att ha med i testbänken även om det i dagsläget är en svårighet för programmen. Det är också viktigt för användare att veta vad programmen inte klarar.

5.7.7 Ofullständiga satser

Frånvaron av verb i satser upptäcktes av *Granska*. Däremot fick “meningar“ passera där nominalfras saknades. Meningar med en fullständig huvudsats men med endast en inledande konjunktion till en förmodad bisats passerade också utan anmärkning. Även dessa typer av fel tror vi är vanligt förekommande när det redigeras i texter i ordbehandlingsprogrammen.

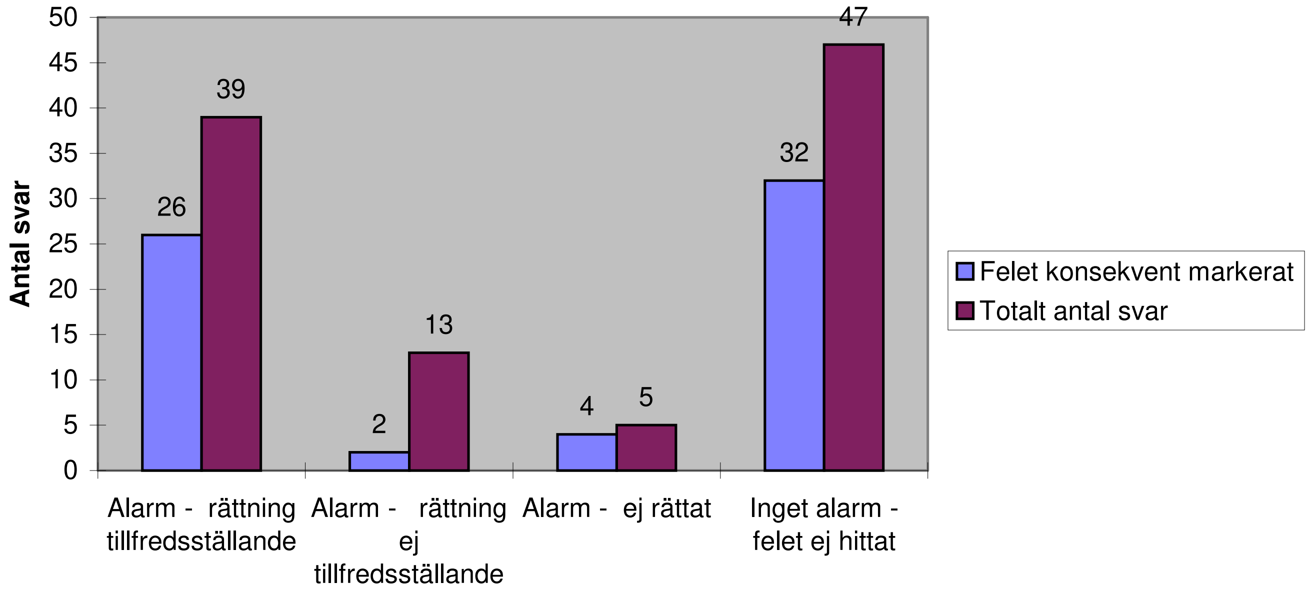
5.7.8 Jämförande testkörning av korrekturläst text

Testkörningen av en korrekturläst text hämtad ur SUC-korpusen medförde få alarm om felaktigheter från *Granska*. Texten innehöll drygt 2000 ord uppdelade i 100 meningar varav fem bedömdes innehålla felaktigheter. Av dessa fem är det endast en mening där *Granskas* larm är befogat. Meningen innehåller inget verb. I vissa typer av texter kan det vara författarens medvetna val och konstnärliga frihet att utelämna verbet i en sats, men det är naturligtvis viktigt att ett språkkontrollprogram alarmerar. Övriga alarm gällde kongruensfel, felaktig särskrivning och flera konjunktioner efter varandra. Vid de två markerade kongruensfelen var det predikativ som kongruerade med en hel sats eller fras och inte med ett närliggande ord, som utgjorde svårigheten för *Granska*. Ett annat alarm bestod i att *Granska* bedömt en punkt i *t.ex.* som en meningsavskiljare. Detta medförde att resterande del av meningen saknade ett finit verb.

Vi bedömer att *Granska* behandlat vår testbänk och SUC-filen någorlunda lika, d.v.s. att programmet alarmerar ungefär lika många fel i korrekt text i båda fallen. Att exakt bedöma detta är svårt, dels är det ett visst mått av godtycke i hur man uppfattar rättningsförslagen, dels är materialet inte så stort att någon statistik blir tillförlitlig.

5.8 Konsekvens i *Granskas* svar

Anledningen till att vi valt att i de flesta fall använda två exempelmeningar, a respektive b, är för att se om programmen skulle behandla feltyperna konsekvent i båda varianterna. Svarens fördelning kan ses i figur 3. Siffran ovanför respektive stapel anger antalet alarm som faller under de olika rubrikerna



Figur 3
Konsekvens i Granskas resultat för a- respektive b-exempel

Således får man i den vänstra av stapeln för de olika rubrikerna dela antalet med två för att få reda på hur många exempel som behandlats konsekvent. Stapeln längst till vänster anger 26 alarm, vilket innebär att både a- och b-exemplet markerats i 13 feltyper.

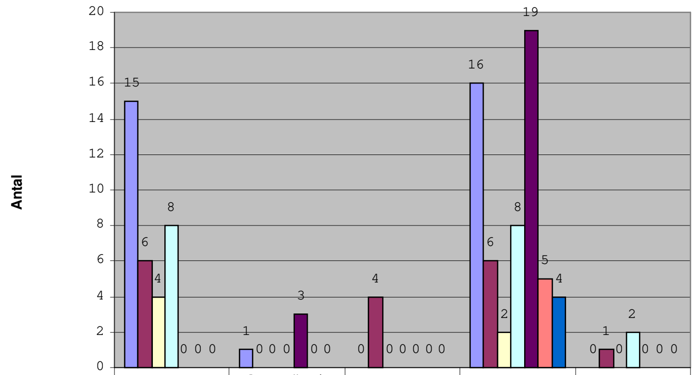
5.9 Resultatet av testkörningen i *Grammatifix*

Grammatifix upptäckte endast 40 av de 100 planterade felen i testbänken och således har 60 fel passerat utan att uppmärksammas. Av de 40 felen som alarmerades gavs 33 st en tillfreds-ställande rättning och ett av dessa 33 fick ett av felen även ett rättningsförslag som inte var acceptabelt. I fyra fall alarmerade programmet men gav endast oacceptabla rättningsförslag. I ytterligare fyra fall hittades felet men inget rättningsförslag gavs. Liksom för resultatet av *Granska* hade *Grammatifix* inte några rättningsförslag i exempel 23a fel) och 23b fel) där verbfras saknades. I exempel 24a fel) och 24b fel) där finit verb saknades gavs bara en kommentar till skillnad från *Granska* som på samma ställe gav förslag på verb.

Falskt alarm i korrekt text gavs endast i tre fall. Två av dessa alarm gjordes i exempel 33 på verbet “anlände“ som programmet trodde var ett felböjt adjektiv. Det tredje alarmet gällde exempel 23b fel) och ordet “tom“, som antogs vara en förkortning istället för ett adjektiv.

När det gäller de olika feltyperna kan de två programmens resultat jämföras. *Granskas* resultat finns i figur 2 och resultatet för *Grammatifix* i figur 4. Vi kan konstatera att *Grammatifix* klarade kongruensfel något bättre än *Granska*. Då ska man dock notera att drygt hälften av kongruensfelen inte upptäcktes alls. *Grammatifix* hade även bättre resultat i kategorierna “övrigt“ och “ordföljd i bisats“. Verbkonstruktioner klarade dock *Granska* något bättre. Särskrivningar var ett område som *Grammatifix* inte behandlade alls. Likaså var upprepningar och ofullständiga satser områden som inte gav önskvärt resultat. *Grammatifix* har dock inte utgett sig för att klara dessa feltyper.

Samma text ur SUC-korpusen som testades i *Granska* testkördes också i *Grammatifix*. Inga alarm gavs för felaktig grammatik.



	Alarm, rättning tillfredsställande	Alarm, rättning ej tillfredsställande	Alarm, inget rättningsförslag	Inget alarm trots felaktighet	Falskt alarm i korrekt text
Kongruenser (mening 1-16)	15	1	0	16	0
Verbkonstruktioner (mening 17-24)	6	0	4	6	1
Ordföljd I bisats (mening 25-27)	4	0	0	2	0
Övrigt (mening 28-35)	8	0	0	8	2
Särskrivningar (mening 36-46)	0	3	0	19	0
Upprepningar (mening 47-49)	0	0	0	5	0
Ofullständiga satser (mening 50-51)	0	0	0	4	0

Figur 4
Grammatifix: Resultaten fördelade på respektive grammatiska rubriceringar

5.10 Framtida förbättringar av testbänken

Vi har under arbetets gång blivit medvetna om att testbänkens utformning behöver förbättras. Det har dock inte varit möjligt för oss att inom tidsramen för denna uppsats själva genomföra dessa förbättringar.

För att göra testbänken till ett användbart verktyg skulle den dessutom behöva kompletteras med något slags program som åskådliggjorde resultaten i siffror eller figurer.

6 SAMMANFATTNING

Vi har haft för avsikt att skapa ett första utkast till en testbänk för grammatikkontroll, som ska kunna fungera som ett köpstöd för inköpare av språkgranskningsprogram. För detta ändamål har vi sammanställt 51 exempel på fel som vi bedömer kan förekomma i texter som produceras i ordbehandlingsprogram. Vi har även tagit med några exempel på feltyper som i princip inte berör grammatik, nämligen upprepningar. Exemplen består i de allra flesta fallen av en a- och en b-variant för att det eventuellt ska gå att urskilja ett mönster i hur konsekvent språkkontrollprogrammet behandlar olika typer av fel. Dessa två varianter är i sin tur uppdelade i ett korrekt och ett felaktigt utförande för att man dels ska kunna se vad som kan vara en riktig version av en felaktig mening, dels för att se om programmet ger falska alarm, d.v.s.

anmärker på korrekt text. Urvalet av de olika feltyperna är gjort dels med hjälp av två språk-kontrollprograms egna listade feltyper som satts in i exempelmeningar och dels med konstruerade exempel utifrån våra egna funderingar. Det huvudsakliga resultatet av denna uppsats är testbänken och för att ta reda på hur den fungerar i praktiken och få uppslag till hur den kan förbättras har vi gjort en testkörning i *Granska*, som är ett program under utveckling på Nada, KTH i Stockholm, och analyserat resultaten. Vi har även gjort en jämförande testkörning på en text från SUC-korpusen som är en korrekturläst och i stort sett korrekt text, och sedan bedömt resultatet från denna. Samma testkörning har gjorts i *Grammatifix,* ett språkgranskningprogram som finns på marknaden, från finska Lingsoft. Resultaten från denna testkörning har sammanställts och en mindre analys har gjorts. Se vidare Appendix 2.

För att testbänken ska vara ett attraktivt verktyg för att testa språkkontrollprogram krävs många förbättringar. Detta gäller både för innehåll och utformning, och framför allt måste själva tillämpningen av testbänken, t.ex. från en webbplats, fungera tillfredsställande så att den blir ett lättanvänt hjälpmedel. Det vore önskvärt att resultaten av en testkörning på något sätt kunde bearbetas automatiskt så att man snabbt fick en överblick över vad det testade programmet klarar respektive inte klarar.

När det gäller urvalet av feltyper skulle det naturligtvis vara av intresse att veta hur vanligt förekommande olika fel är i texter. Dessutom skulle möjligen den korrekta texten sätta programmen på hårdare prov, t.ex. genom att göra mer komplexa meningsstrukturer, inskjutna satser etc. Om man ändå ville behålla fördelningen med ungefär hälften av meningarna korrekta så skulle man antingen behöva frångå att den korrekta meningen var mycket snarlik den felaktiga, eller så skulle även meningen med ett planterat fel vara mer komplex.

Analyserna av de båda programmen visar att inget av dem detekterade ens hälften av felen. Gemensamt för båda programmen är att de inte klarade upprepningar eller ofullständiga satser med undantag av utelämnande av verb. Inget program kunde heller hantera kongruensfel över meningsgränser eftersom varje sats analyseras för sig. Kongruenser och verbkonstruktioner behandlades relativt likvärdigt. *Granska* klarade av felaktiga särskrivningar övervägande bra medan *Grammatifix* inte behandlade dem alls. Inte något av programmen kunde heller uppmärksamma feltyper såsom kontaminationer och tautologier. Resultaten visar att program för grammatikkontroll i dagsläget inte kan anses fungera tillfredsställande. Med så många fel oupptäckta kan en skribent inte känna sig försäkrad om att en text som kontrollerats i ett språkgranskningsprogram verkligen är korrekt. Vi anser att man som skribent kan få viss hjälp av ett språkkontrollprogram men att man än så länge främst måste lita på sina egna språkliga kunskaper.

7 REFERENSER

Andersson, E, 1993, *Grammatik från grunden*, Uppsala, Hallgren & Fallgren
Florby, G, Fält, E 1996, *Grammatiskt ABC*, Uppsala, Reklam & Katalogtryck
Westman, M, Hultman, T 1977, *Gymnasistsvenska*, Lund, Bröderna Ekstrands Tryckeri AB

Norstedts uppslagsbok, nionde upplagan, 1992, Norstedts förlag
Svenska språknämnden, 1991, *Svenska skrivregler*, Uppsala, Almqvist & Wiksell

Kohut, Gary F., Gorman, Kevin J., "The Effectiveness of Leading Grammar/Style Software Packages in Analyzing Business Students' Writing.", *Journal of Business and Technical Communication*, vol 9 nr 3, juli 1995, sid 341-361
Kukich, Karen, "Techniques for Automatically Correcting Words in Text", *ACM Computing Surveys*, vol 24, december 1992, sid. 377-439

Lingsoft, 2000,
"Grammatiska fenomen som *Grammatifix* kontrollerar",
URL: http://www.lingsoft.fi/grammatifix/features/grammatik.html
Nada/KTH, 1999,
"*Granska* – ett program för svensk språkgranskning",
URL: http://www.nada.kth.se/theory/projects/granska/popular.html
Nada/KTH, 1999,
"Testa vår grammatikgranskare *Granska*, version 1.0"
URL: http://www.nada.kth.se/theory/projects/granska/demo.html
SCARRIE, 1999
"SCARRIE – Project Summary",
URL: http://www.linglink.lu/le/projects/scarrie/summary.html
Stockholms universitet, 1999,
"Vad är SUC?",
URL: http://www.ling.su.se/Dali/projects/SUC/vadaersuc.htm