En explorativ studie av språkgransk-ningsverktyg

Studie av olika arbetssätt för granskning av grammatik och stavning
Av

Annika Hansén-Eriksson
IPLab, NADA, KTH, Stockholm
annikah@nada.kth.se

Ola Knutsson
IPLab, NADA, KTH, Stockholm
knutsson@nada.kth.se

Innehåll

Problembeskrivning 3
Varför studera arbetssättet 3
Problem vid redigering och granskning 4
Några tidigare studier av språkgranskningsprogram 5
Gränssnittsfaktorer att studera 7
Några alternativa gränssnitt 7
Språkliga frågeställningar 10
Försökets utförande 10
Försök med MSWord 11
Försök med Granska 13
Försökspersoner 14
Gränssnittsaspekter - resultat 15
Försökspersonernas arbetsätt 15
Allmänna synpunkter och iakttagelser 15
Färgmarkeringar i MSWord 15
Användning av popupmeny 16
Användning av dialogen 17
Användning av Granska 18
Outnyttjade funktioner i MSWord 18
Användaren låter sig styras av programmet? 19
Språkliga aspekter - resultat 19
Datainsamling för den språkligt relaterade delen 19
Språkligt relaterade data insamlade med MSWord 19
Analys och diskussion av språkligt relaterade data – MSWord 22
Lång mening 22
Falska alarm 22
Felaktig diagnos 23
Ersättningsförslagen 23
Sammanställning av försöken med Granska 23
Analys och diskussion av insamlade data med Granska 25
Lång mening 25
Falska alarm 25
Felaktig diagnos 26
Ersättningsförslagen 27
Skillnader mellan programmen under försöken 27
Slutkommentarer för den språkliga delen 31
Diskussion om försöksuppläggningen 31
Litteratur 32

Problembeskrivning

Denna studie utfördes som ett moment i kursen Datorstöd för skrivprocessen. Försökspersonerna var alla deltagare i kursen och utförde försöket som ett laborativt moment i kursen. Vi som utförde studien arbetar med att utveckla ett nytt integrerat verktyg för att granska grammatiken och stavningen i en text skriven på svenska. Ola arbetar med de lingvistiska aspekterna och Annika med utformningen och programmeringen av gränssnittet. De språkliga aspekterna av Granska finns beskrivna på www [G 2000].

Under tiden som vi arbetat med projektet har den första kommersiella produkten för svensk grammatikgranskning lanserats på marknaden i form av MSWord 2000. Grammatikkontrollen i MSWord är utvecklad av det finländska språkteknologiföretaget Lingsoft AB [L 2000]. Dess gränssnitt tillåter att man utför granskningen på ett antal olika sätt och vi upptäckte att åsikterna inom vår arbetsgrupp gick i sär när det gällde det bästa sättet att arbeta med programmet. Vi beslöt därför att göra en liten studie av hur några olika användare arbetar med MSWord och att samtidigt göra en liten studie av vårt eget program Granska. De algoritmer och regler som beskriver ett grammatiskt fel i MSWord och i Granska är uppbyggda på olika sätt och vi ville undersöka skillnaderna.

Vi ville studera både de språkliga aspekterna och gränssnittsaspekterna av språkkontrollen samtidigt. Försöket gör inte anspråk på att vara en heltäckande studie men bör kunna fungera som en utgångspunkt för fortsatta studier.

Varför studera arbetssättet

Vi håller på att utveckla ett eget gränssnitt för grammatikgranskning. Arbetet utgår från det gränssnitt som Stefan Larsson utvecklade som sitt examensarbete 1998 [SL98]. En av de viktigaste anledningarna till att utveckla ett eget gränssnitt för grammatikkontroll är att det ger ett verktyg för att studera hur människor arbetar när de skriver på dator. Man får möjlighet att skapa en försöksmiljö som kan anpassas till det man vill studera. Man har också möjlighet att minska skillnaderna i bakgrundskunskaper hos försökspersonerna i motsats till när man använder ett verktyg som försökspersonerna har mer eller mindre stor erfarenhet sedan tidigare. Å andra sidan ger ju ett nytt gränssnitt ett inlärningsproblem vid försök.

Även om det funnits verktyg för granskning av engelsk text i MSWord under flera år så är svensk grammatikkontroll ny. MSWord har utvecklats under många år och när nya funktioner har lagts till så har man vanligtvis fortsatt att lägga till till redan befintliga delar. Resultatet har blivit ett program som är så omfattande att ingen vanlig användare av programmet kan använda alla funktioner.

Granskningen av grammatik och stavning kan göras på i huvudsak tre olika sätt med stöd av programmet. Man kan välja att bara använda understrykningar i texten och rätta de markerade (misstänkta) felen själv. Ett annat alternativ är att kombinera felmarkeringarna med den popupmeny som man kan aktivera för det felmarkerade området och rätta felen med hjälp av de alternativ som ges i poupmenyn. Det tredje och äldsta alternativet är att använda den dialogruta som innehåller mer omfattande felinformation.

Man kan också välja mellan ett antal olika granskningsvillkor i MSWord. Villkoren gäller vilka feltyper granskningen ska reagera på. Dessa väljs i en egen dialogruta och man kan inte förutsätta att den vanlige användaren av programmet vet hur man hanterar dessa valmöjligheter. Den version av programmet som nu finns tillgänglig har också den

oväntade egenheten att vilka fel som programmet upptäcker inte bara beror av de granskningsvillkor man kan ställa in utan också av vilken granskningsmetod man väljer. Detta är naturligtvis inte avsikten utan får ses som ett fel i programmet.

Problem vid redigering och granskning

Tidigare forskning har visat att när man skriver på dator tenderar man att arbeta mera lokalt i texten än när man skriver för hand. Eklundh [KSE 92] konstaterar att ändringarna visserligen blir fler men att de tenderar att vara rättelser av små lokala fel snarare än bearbetning av strukturen hos texten. Detta sammanhänger med användarens orientering i texten. Att läsa från en skärm innebär att man har en mindre del av texten synlig åt gången än när man har texten framför sig på skrivbordet i pappersformat. Dessutom är det rent fysiologiskt mera ansträngande att läsa från en skärm än från ett papper.

När man ska granska språkfelen i en text kan det lokala perspektivet snarare vara en fördel för direkta språkfel. Däremot är problemet detsamma när det gäller grammatisk inkonsekvens i en text t.ex. inkonsistent användning av numerus och tempus i ett stycke eller i texten som helhet. I dagsläget har dock inte varken de kommersiella verktygen eller de verktyg som är under utveckling ännu några anspråk på att kunna stöda språkriktigheten annat än på en lokal nivå.

Eftersom man i dag endast kan granska språkriktigheten automatiskt på meningsnivå är problemet snarare att presentera resultatet på ett sådant sätt att man minskar den kognintiva belastningen så mycket som möjligt. Ett problem är då att ju mer information man ska visa om ett fel desto mer av skärmytan kommer att upptas av informationen. Informationen bör presenteras nära det textavsnitt som är felaktigt utan att man för den skull skymmer den kringliggande texten och skribenten förlorar kontextöverblicken. Man måste överväga hur man placerar informationen och hur mycket information som presenteras. Presenterar man informationen nära felet stör man överblicken och presenterar man informationen lägre bort blir det ansträngande för skribenten att växla fokus mellan information och felmarkering.

Eftersom behovet av information dessutom kan skilja sig både mellan olika skribenter och beroende på feltypen är det önskvärt att skribenten själv kan styra vilken information som visas. Detta får å andra sidan inte vara så komplicerat att skribenten avskräcks från att styra informationsmängden.

Ett naturligt språk är inte lätt att beskriva i formella termer. Därför har alla verktyg för språkriktighet sina brister. Stavningskontrollen felmarkerar ord som är riktiga och är ofta inte särskilt bra på att generera förslag på de ord som är felstavade. Däremot brukar inte stavningskontrollen så ofta missa att ett ord är felstavat. När den missar brukar det bero på att man råkat skriva in ett annat riktigt ord än det avsedda. För grammatikfel är det betydligt svårare att formulera en heltäckande regelbeskrivning. Det betyder att även den bästa kontroll kommer att missa en del fel och ge falska alarm ganska ofta. Förslagen kan också lätt bli felaktiga.

Förväntningarna på ett granskningsprogram har effekter på hur man använder det och hur utfallet av användningen blir. Många användare av ordbehandlingsprogram har konstaterat att stavningskontrollen ger väldigt många falska alarm och ganska dåliga ändringsförslag. Speciellt de äldre versionerna av svensk stavningskontroll slutade många användare att använda efter några försök eftersom de falska alarmen var för många och förslagen som genererades kändes konstiga. Det finns tyvärr risker även med en bra kon-

troll eftersom man då kan luras att lita på programmet för mycket. Med införandet av grammatikkontroll blir problemet ännu mer komplicerat eftersom granskningen i sig är svårare och det också kan vara svårare för skribenten att bedöma programmets förslag. Ju mindre säker man är på språket desto värre är det om programmet är osäkert. MSWord verkar generera endast en förklaring och oftast endast ett förslag per grammatikfel vilket kan fresta användaren till en övertro på förslagen. Kongruensfel kan dock ibland resultera i två eller flera förslag.

Grammatikfel kan behöva förklaras även för den mycket språksäkra skribenten. De flesta användare av ordbehandlare har dock inte någon djupgående kunskap om grammatisk terminologi och det är därför en svår avvägning hur informationen ska presenteras.

Några tidigare studier av språkgranskningsprogram

Språkgranskningsprogram är idag kanske en av de mest använda tillämpningarna av språkteknologi som finns på marknaden. Trots att de flesta datoranvändare kommer i kontakt med dessa program finns det väldigt få studier som försöker analysera hur interaktionen mellan programmet och skribenten fungerar. Men det finns riktlinjer för hur denna typ av studier skall bedrivas. EU-projektet TEMAA [T 99] koncentrerade sig på utvärderingsmetoder för språkteknologiska tillämpningar, med fallstudier för stavningskontroll, grammatikkontroll och informationssökning. När det gäller utvärdering av användbarhet utgår TEMAA från kunskap från MDI-området och försöker applicera denna på stavnings- och grammatikkontroll. För att utvärdera användbarheten hos en stavningskontroll (och grammatikkontroll) utgår TEMAA ifrån följande högnivåvariabler:

- hur enkelt kan man köra stavningskontrollen (inkrementellt, sekventiellt, batch eller i ett annat program) och hur enkelt är det köra stavningskontrollen i varje typ av körning. (Invocation of spelling checker)
- selektion av text att kontrollera, kan man välja att kontrollera ett ord, en mening, ett stycke eller hela texten? (Selection of text for checking)
- dialogen med användaren, presenteras ersättningsförslagen i någon ordning, hur många operationer måste användaren utföra för att välja ett annat förslag än det högst rankade? Kan man lägga till egna ersättningsförslag? Kan man lägga till godkända ord till en stopplista? (Error dialogue)
- Möjlighet att förändra texten utifrån diagnos och ersättningsförslag. Åtgärdar programmet ändringen eller måste man själv editera texten (får man veta att programmet har påverkat texten?, vår kommentar). Detta är ju särskilt relevant om man har många regler som endast påpekar att något verkar fel i texten och att inget ersättningsförslag ges. Under denna kategori räknar man också in möjligheter att spara olika textförändringar utförda av programmet eller av användarna (viktigt om det är en grupp som skriver) (Text alteration)

För att utvärdera systemets enskilda delars användbarhet använder TEMAA följande begrepp:

- användbarhetspotential (experienced user performance), det vill säga hur användbart systemet är när användaren väl lärt sig det (man vill undvika novis/expert-problematiken med detta begrepp)

- gränssnittets förutsägbarhet (guessability), är det enkelt att lista ut hur man bär sig åt?
- går det att lära sig systemet på ett enkelt och effektivt sätt (learnability.)

En del av de punkter som TEMAA kommer med känns lite svåra att använda sig av när MSWords gränssnitt är så pass etablerat för språkgranskning. Det är t.ex. självklart att stavning/grammatik finns i verktygsmenyn och i standard-verktyglisten (invocation of application). Det är också ganska självklart att man kan granska ett ord, en mening eller hela texten. Men några av de andra punkterna angränsar till den undersökning som vi genomfört och då framförallt dialogen (error dialogue), hur man kör själva applikationen (invocation of spelling checker) och hur man kan förändra texten (text alteriation). Vårt perspektiv har dock varit annorlunda, vi har velat undersöka hur användarna reagerar på programmen, vi har inte själva undersökt programmens användbarhet. Vad det gäller de intressanta begreppen *experienced user performance*, *guessability* och *learnability*, har vi inte haft möjlighet att undersöka dessa närmare. De nämns i texten ovan eftersom vi tycker att de är viktiga att nämna i kontexten av denna typ av studier.

Man skulle kunna se automatisk språkgranskning som en dialog mellan datorn och användaren om vad som är språkligt rätt eller fel. Domeij [RD 98] iakttar bland annat hur användarnas granskningsstrategier påverkar hur skribenterna förändrar texten efter programmets instruktioner. Om användarens granskningsstrategi stämmer överens med programmets är skribenten i högre grad benägen att följa programmets råd. Störst skillnader mellan olika användare uppstår när det handlar om stilistiska påpekanden från programmet. Granskningsstrategin borde ha ett mindre inflytande på grammatiska fel än på stilistiska, eftersom det finns fler åsikter kring stil än grammatik, speciellt nu när det finns ett bra referensverk för grammatik nämligen Svenska Akademiens Grammatik [UT, SH, EA 99], motsvarande saknas för stilistik (Svenska Akademiens grammatik tar dock upp en del av språkvårdarnas åsikter) . I våra försök kommer felen mestadels att bestå av grammatikfel, men det finns också fel som tangerar det stilistiska och därmed borde försökspersonerna följa programmets råd på ett mer likriktat sätt.

En annan viktig fråga som Domeij m.fl. undersöker i sin studie från 1996 [RD, OK, SL, KSE, ÅR 96] är feltypernas koppling till ett referensverk. Vill användarna veta om feltypen finns befäst i litteraturen? Domeijs studie visar att många användare tycker att det är bra med ett referensverk on-line integrerat med granskningsprogrammet. På grund av resursbrist har vi inte kunnat studera referensverkens relevans för grammatikfel i någon större utsträckning.

Språkets flertydighet och felens oförutsägbarhet tillsammans med den möjlighet till analys som dagens granskningsprogram har tillgång till medför att denna dialog på intet sätt kan förutsägas utan måste studeras på ett realistiskt sätt. Ordbildningen i svenskan är tillräckligt komplex för att utgöra en faktor som inte går att automatiskt hantera fullt ut, med andra ord ett ersättningsförslag till ett i mångas ögon enkelt stavfel kan ibland inte regelmässigt konstrueras utan programmet får göra en kvalificerad gissning (eller inget förslag alls). Hur dessa gissningar tas emot av användarna är mycket intressant att studera. Än mer intressant blir det ju mer språkligt komplext det detekterade felet är. Om ett ersättningsförslag saknas, medför det att användaren inte åtgärdar felet överhuvudtaget? Eller om ersättningsförslaget är dåligt, kan användaren själv avgöra det?

Gränssnittsfaktorer att studera

Vi ville studera ett antal faktorer som vi ansåg ha betydelse för den fortsatta programutvecklingen:

- Hur orienterar man sig i texten medan man granskar?
- Hur påverkar de olika alternativen för placering av granskningsinformationen kontextuppfattningen?
- Vilken information bör presenteras om felet och hur används den av användaren? Hur mycket information behövs? Om man försöker förklara det fel som detekteras, hur ska då förklaringen formuleras?
- Hittar man rätt alternativ bland förslagen om flera förslag presenteras? Finns det risk att välja fel alternativ? MSWord verkar leverera endast ett förslag för ett grammatikfel, medan Granska presenterar flera möjliga tolkningar av ett fel, hur påverkar det språkriktigheten i den färdiga texten?
- Hur hanterar man långa felmarkeringar? MSWord har en inställning där man kan välja att markera en mening som felaktig om den är för lång. Standardinställningen är 30 ord vilket kan göra att en ganska vanlig text får väldigt många grönmarkeringar för grammatikfel. Till saken hör att om ett markerat område innehåller flera fel så visas endast information om det största felområdet. Detta innebär i praktiken att kortare men viktigare fel inne i en lång markering kan missas om skribenten bara se noteringen om en lång mening. Problemet är långtifrån trivialt varken rent tekniskt eller teoretiskt eftersom en lång felmarkering även annars kan innehålla flera fel.
- Hur påverkar språksäkerhet och inrutade vanor användningen av gränssnittet?
- Vilken information skall presenteras? Många nivåer är möjliga: felområde (markering i texten), förslag, en kort förklaring, ett modellexempel för hur man bör skriva, en kort formulering av den grammatiska regeln, generella skrivregler (Svenska Språknämndens) och lexikal information om enskilda ord.

Några alternativa gränssnitt

Vår studie utgår från de gränssnittsalternativ som finns i MSWord och i Granska.

Dialogrutan är den äldsta varianten på felpresentation i MSWord. I äldre versioner av MSWord så kunde man inte växla mellan att arbeta i textfönstret och i dialogrutan utan hänvisades till att bearbeta ett fel i taget stegvis framåt i texten utgående från den textposition där man startade granskningen. Då innehöll den också bara ett litet inmatningsfält där felmarkeringen presenterades utan intilliggande text. I sitt nuvarande utförande är den mer avancerad, man kan växla mellan att arbeta i texten och att arbeta i dialogrutan. Inmatningsfältet i dialogrutan omfattar den närmaste kontexten kring felet och felet markeras i texten. Ett exempel på hur dialogen kan se ut när programmet hittat ett grammatikfel visas i figur 1.

Figur 1 Dialogen i MSWord

Dialogrutan kan användas när som helst under skrivandet och kräver inte att man granskar texten kontinuerligt medan man skriver. Popupmenyn och färgmarkeringarna däremot kräver att man har aktiverat den kontinuerliga granskningen.

Arbetshypotesen för försöket var att dialogrutan har vissa fördelar:

- Den är flyttbar och man kan därför undvika att den skymmer kontexten
- Ångra-knappen i dialogen är lättare att använda (speciellt för en ovan användare) än motsvarande knapp i verktygsfältet
- Mindre risk för feltryckning än i popupmenyn
- Den presenterar ett fel i taget vilket kan hjälpa koncentrationen
- Stegvis förflyttning framåt i texten vilket minskar risken att hoppa över ett fel

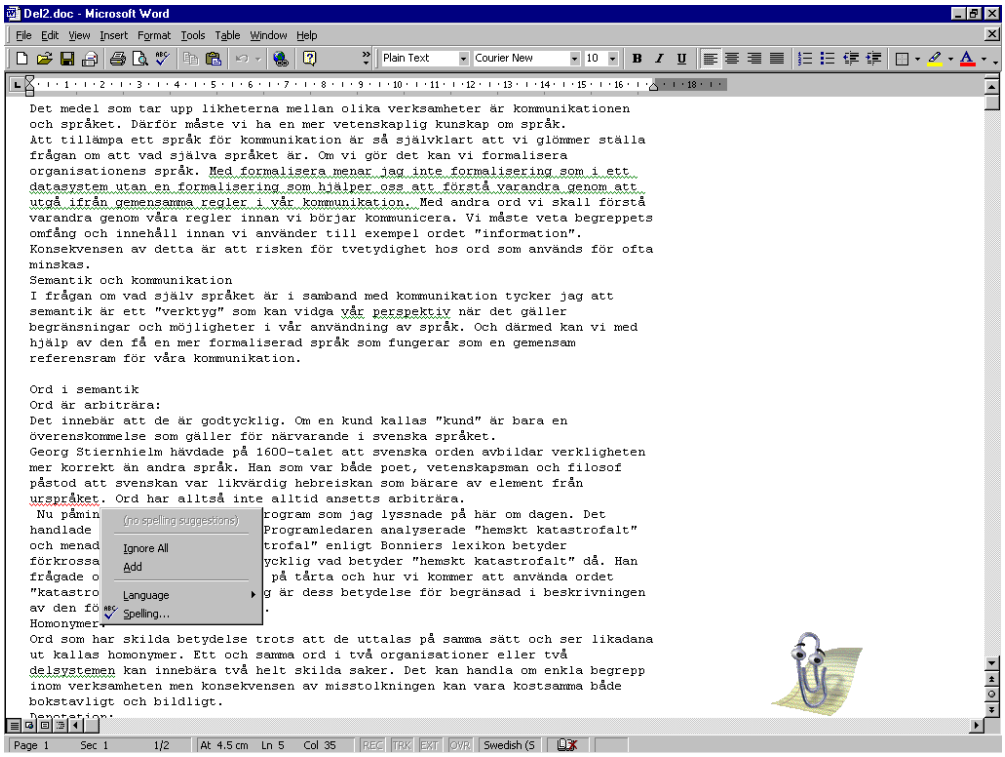
Dessutom hittar dialogen fler fel (bugg i Word). Denna skillnad är ju dock inte beroende av gränssnittsdesignen.

I hypotesen ingår också att dialogrutan har vissa nackdelar jämfört med färgmarkeringar i kombination med popupmenyn:

- Långsammare granskning
- Svårare att hoppa över falska alarm
- Tar stor plats
- För mycket information presenteras

Färgmarkeringar kan användas ensamma eller tillsammans med popupmenyn. Röda markeringar betyder stavfel och gröna markeringar betyder grammatikfel. Markeringarna visas när man använder kontinuerlig granskning av texten medan man skriver.

Popupmenyn kan endast användas i kombination med kontinuerlig granskning medan man skriver. Den fungerar så att man då ett fel är markerat med färgkod kan klicka på musens högerknapp (på PC) när markören befinner sig inom det markerade området. I figur 2 visas hur det kan se ut när man har hittat ett stavfel och popupmenyn inte innehåller något rättelseförslag.



Figur 2 Popupmenyn, stavfel utan förslag

Vi hade en arbetshypotes om fördelar och nackdelar även för detta sätt att granska.

Fördelar:

- Snabbt och bekvämt
- Man behöver inte kontrollera information om enkla fel närmare utan kan rätta direkt, vilket är speciellt bekvämt när man bara råkat åstadkomma ett stavfel genom att trycka på fel tangent
- Lättare ignorera falska alarm

Nackdelar:

- Popupmenyn aktiveras från högerknappen på musen på en PC och på annat sätt på en Mac
- Popupmenyn är inte flyttbar
- Risk att klicka på fel förslag eftersom det bara krävs en enkelklickning för att ändra
- Man får upp en annan meny utanför felområdet
- Saknar funktionen ignorera, har bara ignorera alla
- Placerar sig över delar av kontexten

Till nackdelarna hör också att funktionen missar vissa fel (bug i MSWord) se ovan. Eventuellt hittas vissa fel med den här funktionen som inte hittas med dialogen. Under försöket verkade dock dialogen hitta fler fel.

I den version av Granska som använts i försöket får man när man utfört granskningen dels färgmarkeringar i texten så att varje fel markeras med röd text oberoende av feltyp.

En figur som visar gränssnittet finns i kapitlet försökets utförande. Utöver färgmarkeringen i textfönstret får man, i fönstrets överkant, en yta med granskningsinformation där det finns två listboxar en som presenterar förslagen och en som berättar vilken regel man brutit mot. Ytan med granskningsinformation kan visas eller gömmas efter behov. Under försöket fanns ingen kontinuerlig granskning i Granska.

Hypoteser för Granska:

- Informationen skymmer inte texten och stör därför mindre än i MSWords dialogruta
- En förklaring till varje grammatikförslag gör det lättare för användaren att värdera rättelseförslagen
- Lättare att orientera sig i texten än i MSWord eftersom man kan välja att gå bakåt till föregående fel

Språkliga frågeställningar

Det finns givetvis en rad språkliga frågor som söker svar när det gäller språkgranskningsprogram, såväl språkteknologiska som användarrelaterade. I våra försök har vi koncentrerat oss på de språkliga aspekter som direkt berör användaren, men det är ju omöjligt att inte komma in på en del språkteknologiska spörsmål också. Gränssnittsutformningen och den språkliga analysen i den här typen av program hänger ju starkt samman och de olika delarna är beroende av varandra. Om programmet saknar innehåll så hjälper det inte hur väl gränssnittet är utformat, programmet blir ändå oanvändbart för användaren och det omvända gäller också, om den språkliga analysen är fantastisk men om det saknas ett gränssnitt som kan presentera den språkliga informationen på ett vettigt sätt så blir programmet också oanvändbart för användarna. Vi har varit tvungna att koncentrera oss på några få språkliga frågeställningar för att begränsa försöken. Först av allt ville vi ha svar på några övergripande frågor av allmän karaktär när det gäller användarstudier av de språkliga delarna i ett språkgranskningsprogram:

- Vad verkar möjligt och intressant att studera vidare i ett mer omfattande försök?
- Vad fungerar bra i försöksupplägget inför fortsatta försök?

Vi ville också försöka få svar på följande mer specialiserande frågor:

- **Lång mening:** Skulle en regel som enbart räknar orden i meningen och säger att meningen är ganska lång påverka försökspersonerna?
- **Falska alarm:** Hur kommer försökspersonerna att reagera på falska alarm?
- **Felaktig diagnos:** Om det detekterade felet får fler olika diagnoser, kan användaren avgöra vilken diagnos som är den rätta?
- **Ersättningsförslag:** Om programmet presenterar flera olika ersättningsförslag kan då användarna avgöra vilket/vilka som är rätt? Eller klickar de bara på det översta förslaget i förslagslistan?

Försökets utförande

Vårt försök utformades som ett tänka-högt-försök. Vår metodik baserar sig på den metodik som Flower och Hayes [HJR, FLS 80] använde för att studera skrivprocessen och på en artikel av Monk m.fl. [MA 93] som beskriver en metodik för att studera programs användbarhet under utvecklingen av nya program. Båda bygger på att man låter försöks-

personen utföra en uppgift samtidigt som man ber försökspersonen att tänka högt under försöket och försöka beskriva vad han eller hon gör samtidigt som han eller hon utför uppgiften. Skillnaden är att i Monks metod tillåts försökspersonen att ställa frågor till försöksledaren för att t.ex. ta reda på vad som händer om man gör på ett visst sätt. Försöket ska dokumenteras med bandningar av det som sägs och gärna också videofilmning.

Vi ansåg att videofilmning av våra försök dels skulle ta väldigt mycket tid och dessutom kan videofilmning ha en störande inverkan på försöket. Vårt försök dokumenterades i stället genom att försökspersonens kommentarer bandades. Samtidigt försökte vi också föra två olika typer av protokoll, ett som omfattade hur varje försöksperson arbetade med programmet och ett som omfattade vilka språkliga val som gjordes. Därutöver sparades alla ändringar i texten som försökspersonerna gjorde med MSWord. Efter försöket gjordes en intervju med försökspersonerna. Intervjun omfattade ca 45 frågor och en liten diskussion kring försöket. Se bilaga 1. Frågorna hade utformats som alternativvalsfrågor men fria svar tilläts och frågorna användes också som ett sätt att skapa diskussion kring programvara och arbetssätt.

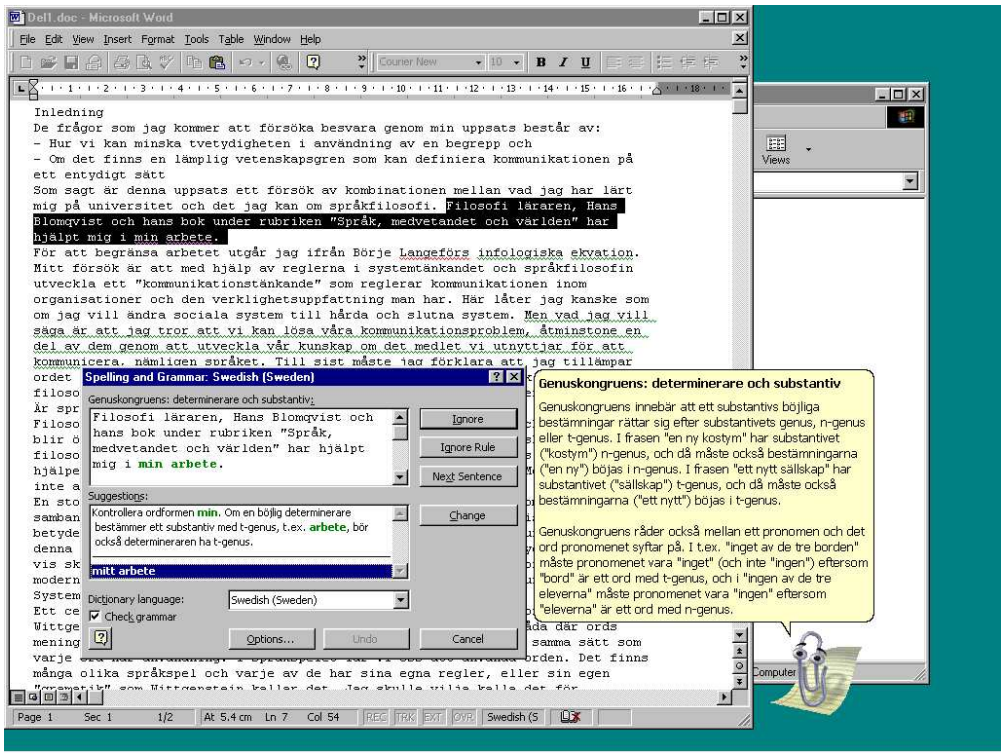
Uppgiften bestod av att granska en några sidor lång svensk text med MSWord och med en prototyp av vårt eget program, Granska. Texten var i MSWord-fallet uppdelad i tre delar och försökspersonerna instruerades om vilket sätt att arbeta som var tillåtet för vilken del. I Granska-fallet kördes samma text ouppdelad.

Den text som försökspersonerna skulle granska var en lingvistisk text. Texten innehöll många facktermer och var därför ganska svår att granska. Försökstexten preparerades med ytterligare fel, främst för att få till fler grammatiska fel. Kapitlet ”språkliga aspekter” kommer att beskriva hur försökspersonerna språkligt reagerade på den respons de fick på de fel som förekom i försökstexten.

Försök med MSWord

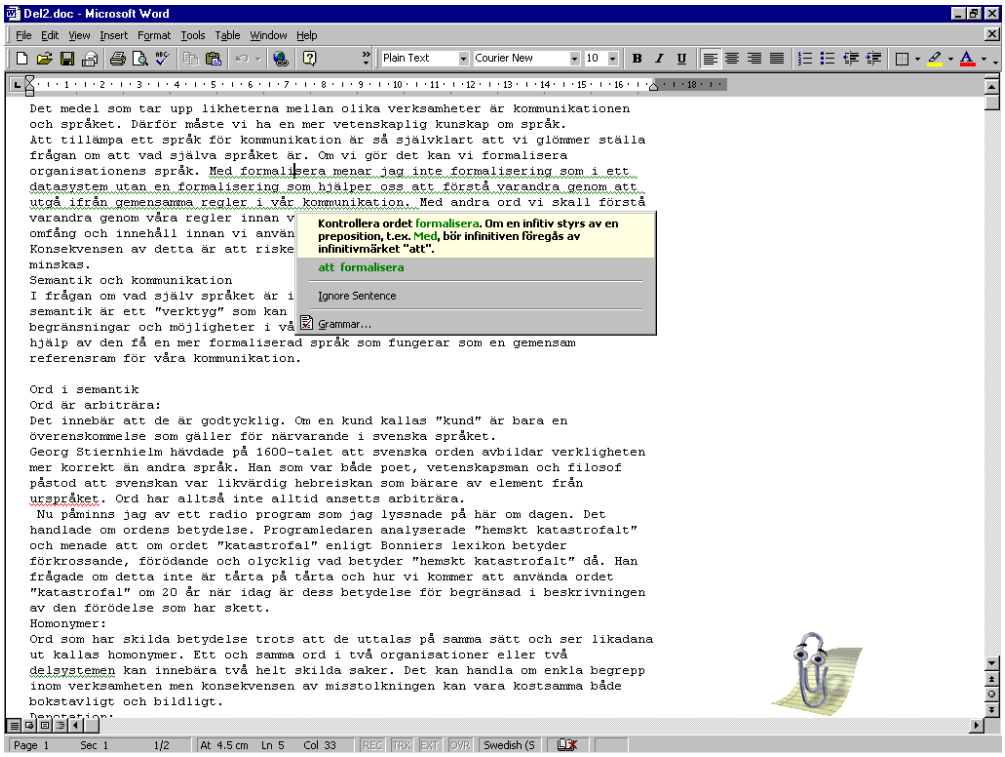
Eftersom det finns så många olika sätt att granska en text med MSWord så delades texten i tre delar, i tre olika filer. Alla försökspersonerna fick granska samma del av texten på samma sätt.

I den första delen av försöket stängde vi av den kontinuerliga granskningen och uppmanade försökspersonerna att starta granskningen antingen från menyn eller genom att trycka på granskningsknappen i verktygsfältet.



Figur 3 MSWord granskning med dialog

När försökspersonerna skulle granska den andra delen av texten aktiverade vi den kontinuerliga granskningen igen. Vi instruerade försökspersonerna att endast använda markeringarna och popupmenyn.

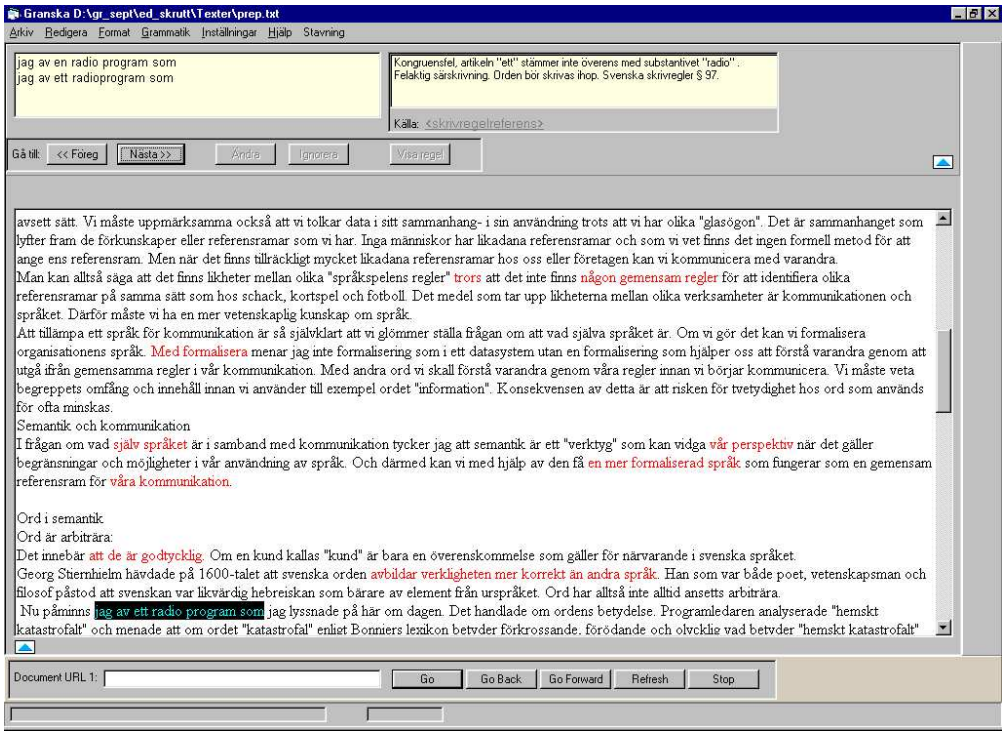


Figur 2 MSWord granskning med popupmeny

För den sista delen av texten fick försökspersonerna använda valfritt sätt att granska texten.

Försök med Granska

Hela texten granskades efter försöket med MSWord med en prototyp av Granska. Under försöket fungerade Granska så att granskningen startades från ett alternativ i menyn. Alla fel i granska markerades genom att texten i felområdet blev röd. De förslag Granska genererade för varje fel visades i en listbox ovanför textfönstret. Varje förslag motsvarades också av en förklaring till felet. Förklaringarna presenterades i en annan listbox till höger om förslagen. Eftersom förslaget beror av hur Granska tolkar felet så kan förklaringen vara olika för förslagen. De två listboxarna var därför kopplade så att om man valde ett förslagsalternativ genom att klicka i den ena listboxen så markerades motsvarande rad även i listboxen för förklaringar (se figur 5). Vi ville genom detta se vad försöks-personerna baserar sina ändringar på, förslag eller förklaring.



Figur 3 Granskas gränssnitt under försöket

Försökspersoner

Före det egentliga försöket gjorde vi ett förförsök med en person. Förförsöket skilde sig från det egentliga försöket genom att det inte bandades och att intervjun omfattade något färre frågor. Dessutom fördes inget omfattande protokoll över gjorda rättelser i texten under förförsöket. I det egentliga försöket deltog 5 försökspersoner. Eftersom betingelserna var så lika så har förförsöket tagits med bland resultaten när det gäller användaraspekterna men inte när det gäller de språkliga aspekterna.

Alla försökspersonerna är vana skribenter, vana datoranvändare och vana vid att använda MSWord. Av försökspersonerna använde 4 Macintosh för det mesta och en PC och en Mac och PC ungefär lika mycket. Alla PC-användare använder Windows som operativsystem. Alla hade dock vana vid båda miljöerna och även Unix-miljön.

Samtliga hade någon gång använt grammatikkontroll för engelska i MSWord. Eftersom ingen hade använt MSWord 2000 tidigare så har ingen tidigare testat svensk grammatikkontroll tidigare. De betecknar sig som vana användare av MSWord även om ingen ansåg sig kunna hålla en avancerad kurs i användningen. De frågor vi ställde om vilka funktioner man använt bekräftar detta.

Alla skriver texter på engelska och svenska. Av försökspersonerna har 5 svenska som sitt dagliga språk även om en av dem inte har det som sitt första språk. En av försökspersonerna håller dock på att lära sig svenska som främmande språk. Det verkar ganska klart även med detta lilla underlag att problematiken kring grammatikkontroll för den som håller på att lära sig språket och dem som kan språket skiljer sig och att de förklaringar som kan ges av programmet får en helt annan och väsentligare betydelse för den som kan tänkas använda verktyget för språkinlärning.

Gränssnittsaspekter - resultat

Försökspersonernas arbetsätt

Vi frågade om man brukar kontrollera sin text medan man skriver, när man skrivit hela texten eller när man tar en paus i skrivandet eller vid något annat tillfälle. Här varierade svaren ganska mycket. De som uppgav sig kontrollera texten efter att hela texten eller en del var klar verkade också använda dialogalternativet mera än de som föredrar att kontrollera texten kontinuerligt. En försöksperson uppgav sig alltid kontrollera texten när den var färdigskriven och samma person uppgav sig också alltid använda dialogen medan den försöksperson som uppgav sig alltid kontrollera texten under pågående skrivande uppgav sig alltid använda färgmarkeringarna och popupmenyn men aldrig dialogen. Dessa påståenden stämmer också med hur försökspersonerna faktiskt arbetade med verktygen när de själva fick välja. De flesta försökspersonerna verkade använda blandade alternativ men alla har sitt eget sätt att arbeta.

Det verkar också finnas en viss koppling till hur säker man känner sig på svenska språket och hur man arbetar. Vissa både ansåg sig vara säkra på språket vilket framgick av påståenden som att ”jag vet ju hur det ska vara” och det framgick också under försöket. De försökspersoner som är mycket säkra på språket verkar gärna ändra direkt i texten ofta utan att kontrollera vilka förslag som ges medan inte oväntat den försöksperson som håller på att lära sig svenska studerar alternativen och förklaringarna betydligt mer ingående än de andra och också spontant letar efter mera information som t.ex. synonymordlista.

Slutsatsen verkar vara att det är bra att erbjuda flera alternativ eftersom det verkar som om behovet finns. Det verkar finns lika många sätt att arbeta som det finns användare. De flesta använder en kombination av arbetssätt.

Allmänna synpunkter och iakttagelser

Hjälpfunktionen i MSWord irriterar de flesta användare. Det lilla ”gemet” som symboliserar en medhjälpare verkar irritera de flesta och vissa försökspersoner ville ha hjälp med att stänga av funktionen. När man trycker på hjälpknappen i dialogen får man när man granskar grammatikfel upp en pratbubbla som presenterar en förklaring. Pratbubblan kan dessutom dyka upp oväntat och de flesta verkar finna den störande. Man tycker dels att den är i vägen och dels att informationen som ges är svårtolkad.

Hur man använder granskningsverktyget påverkar resultatet i MSWord vilket vi märkte under försöket. Det är tydligen ett fel i programvaran som är känt sedan tidigare. Vi försökte dock inte kontrollera hur många fel som missades under granskningen utan endast hur man arbetade med de fel man faktiskt såg så det påverkade inte försöket i någon större utsträckning.

Färgmarkeringar i MSWord

I MSWord använder man röd understrykning för stavfel och grön understrykning för grammatikfel vid kontinuerlig granskning. Ingen av våra försökspersoner lider av färgblindhet men just att skilja på rött och grönt lär annars vara den vanligaste forman av färgblindhet så färgvalet kanske kan vara något att fundera på i vårt eget program. I trafiken betyder ju också grönt gå medan rött signalerar stopp.

Färgmarkeringarna upplevdes som irriterande av hälften av försökspersonerna. Det verkar dock inte finnas någon koppling mellan att man tycker markeringarna är störande och vilket verktyg man föredrar eftersom bland dem som fann markeringarna störande återfanns både en av dem som föredrog dialogen och en som nästan enbart använde markeringarna.

Några försökspersoner uppgav sig ignorera markeringarna till de själva ville börja rätta i texten. Det verkar inte som om markeringarna i sig är ett så stort problem, däremot är markeringarna av långa meningar ganska tydligt ett problem.

Korta felmarkeringar verkar annars fungera väldigt bra som markör av fel. Ofta behöver inte granskaren någon ytterligare information utan rättar felet direkt i texten utan att använda vare sig poupmenyn eller dialogen. Av intervjuerna att döma utnyttjar man det i än högre grad när man skriver själv och bara råkar göra tangentbordsmissar. En försöksperson anmärkte dock det kan hindra ”flödet” under skrivandet om man direkt uppmärksammar de fel som uppstår.

Under försöket lade vi märke till att flera försökspersoner (dock inte alla) även läste intilliggande text kring en felmarkering. Detta gällde speciellt då ett längre område var markerat. Man kan möjligen se det som att en lång felmarkering i viss mån utökar kontextperspektivet dock inte längre än till det stycke där markeringen ingår.

Vi noterade att flera försökspersoner gjorde stora ansträngningar att få bort alla felmarkeringar i texten. När man ställdes inför problemet med att ett avsnitt markerats som lång mening gjorde man försök att få bort markeringen genom att formulera om meningen, genom att försöka bryta meningen med punkter och genom att lägga in parenteser. Man missade även verkliga fel i sådana markeringar och ägnade oproportionerligt stor tid åt dem. Även i de fall när man konstaterade att det kanske inte fanns något annat fel i meningen försökte man ofta få bort markeringen. Ingen försöksperson reagerade med att direkt trycka på ignorera när man ställdes inför problemet med lång mening. Vissa försökspersoner missade även stavfel som ingick i en längre markering trots att dessa markerats med rött i en lång grön markering. Man försökte också att formulera om i övrigt korrekta meningar för att få bort markeringen.

På ett ställe i texten förekom en markering för lång mening som sträckte sig över två meningar som programmet uppfattade som en sammanhängande mening. En försöksperson noterade att det fanns två mellanslag mellan meningarna i stället för ett vilket i så fall tyder på en mindre önskvärd egenskap hos programvaran.

Slutsatsen blir att även om det kan vara bra att programmet anmärker på långa meningar i den meningen att man försöker kontrollera avsnittet noggrannare så verkar långa markeringar leda till fler problem än de löser. En möjlig reservation mot resultaten kan vara att försökspersonerna under försöket kan ha varit mer angelägna att få bort alla felmarkeringar än de normalt skulle vara.

Användning av popupmeny

Två försökspersoner uppgav sig aldrig använda popupmenyn. Av dem angav den ena som skäl placeringen i texten och försök att flytta den medan den andra försökspersonen inte angav något skäl. De övriga fyra försökspersonerna uppgav att de brukade använda popupmenyn. Av dem ansåg tre att placeringen är bra medan en av dem sade sig ha försökt flytta den ibland. Under försöket hände det en gång att någon försöksperson försökte

flytta popupmenyn och att en försöksperson försökte stänga den och då oavsiktligt råkade trycka på ignorera alla.

Under försöket hände det ett antal gånger att någon försöksperson försökte få upp popupmenyn utan att befinna sig inom markerat område och fick då upp en annan meny än den önskade.

Man använde popupmenyn på olika sätt. Några använde den som väntat för att välja det ersättningsförslag man ville använda. Andra använde den för att se förslagen men ändrade ändå genom att själva skriva in ändringen i texten. I bland använde man den för att enbart titta på förklaringen till ett grammatikfel och några använde den för att starta dialogrutan. Förklaringen till ett grammatikfel verkar lättare att hitta i popupmenyn än i dialogrutan där den gömmer sig bland en massa annan information.

Det verkade inte som om man i allmänhet hade något större problem med att välja rätt alternativ när det fanns flera att välja mellan. Den befarade risken för att man kunde välja fel alternativ verkade inte så stor.

Vi ställde frågan om man vill att förslaget i en popupmeny ska väljas genom att man enkelklickar på förslaget eller genom att man dubbelklickar. Lika många ville ha enkelklickning som dubbelklickning. När vi frågade om placeringen av popupmenyn så ville de flesta att menyn ska finnas nära felet dock gärna en liten bit ifrån så att den inte skymmer kontexten. De flesta var positiva till att man skulle kunna flytta popupmenyn.

Slutsatsen blir att popupmeny är ett bra alternativ men att man kanske kan ändra utformningen litet, speciellt göra den flyttbar.

Användning av dialogen

En enda försöksperson uppgav sig aldrig använda dialogrutan, två mera sällan. De övriga uppgav sig använda den alltid. Av dem som sade sig använda den mera sällan använde den ena den i osäkra fall och den andra försökspersonen använde den för stavning.

Man använder dialogen på olika sätt i bland enkelklickar man på förslaget och trycker ändra, ibland dubbelklickar man. Ibland skriver man in sin ändring i dialogens inmatningsfält, ibland växlar man till textfönstret och ändrar där. De flesta försökspersonerna verkar växla beteende.

En del tycker att dialogrutan placerar sig bra i förhållande till texten andra tycker den är i vägen. Alla utom en (den som inte brukar använda dialogen) anger att de flyttar den vid behov.

Man verkar anse dialogrutan mer användbar för stavfel än för grammatikfel. Detta kan kanske förklaras genom att grammatikfelen presenteras med bara ett förslag och en förklaring i MSWord. Den korta förklaring som presenteras är dessutom så väl gömd att vi under försöket kunde iaktta att många inte uppmärksammade den.

Det verkar också som om man använder dialogen för att inte missa några fel. Det kan dels bero på att man i en längre text kan missa färgmarkeringarna speciellt små sådana, dels på att man kan stega sig igenom alla fel från början till slut i en längre text. Dessutom verkar det ju som om MSWord i nuläget hittar fler fel med dialogrutan än annars.

Slutsatsen blir att dialogrutan fyller en funktion men att den upplevs som stor och grötig.

Användning av Granska

Eftersom Granska endast förelåg i prototyp under försöket är det svårt att göra någon rättvis jämförelse. Orienteringen i texten verkar enklare än i MSWord trots att man växlar uppmärksamhet mellan textfönster och informationsfältet. Alla verkade tycka att kopplingen mellan aktuellt fel och motsvarande område i texten, som svartmarkerats fungerar bra. Ingen av försökspersonerna ville markera aktuellt fel i texten på något annat sätt. Man tycker att det verkar bra att kunna gömma informationsdelen av fönstret.

Kopplingen mellan förslag och förklaring verkar tydlig och förstods av alla. I försöket ingick att man kunde välja det alternativ man föredrog antingen genom att klicka på förslaget eller genom att klicka på motsvarande regel. Med regel avses här en mycket kort kommentar eller diagnos som indikerade vilken språkregel man brutit mot t.ex. kongruensfel. De flesta använde någon gång möjligheten att välja utgående från regel snarare än förslag även om förslagen klart var det man tittade mest på. Den försöksperson som håller på att lära sig svenska var mycket mer intresserad av reglerna an de övriga. Av det kan man troligen dra den slutsatsen att det är intressant att låta användaren själv ta ställning till vilken regel man brutit mot och att det dessutom är extra intressant när man arbetar med ett främmande språk. I det fallet är det inte så ovanligt att man känner till regeln även om man inte vet den riktiga formuleringen.

Granska hade ingen regel för att markera långa meningar. Detta verkar ha underlättat orienteringen i texten.

Försökspersonerna var i allmänhet mera kritiska till förslagen i Granska. Detta kan bero på många faktorer. Man är i allmänhet mera kritisk till ett program som föreligger i prototyp än ett kommersiellt program som man dessutom är van vid genom daglig användning. Det kan också vara så att om programmet ställer användaren inför flera förslag så värderar man förslagen annorlunda än om programmet bara presenterar ett förslag. Alla vet väl att om en politiker låter tillräckligt säker så ökar sannolikheten att bli han eller hon blir trodd, det samma gäller förmodligen också för programvara. Slutsatsen blir att det är bra att presentera flera alternativ eftersom en användare bör vara kritisk till alla automatiska språkhjälpmedel. Försökspersonerna var positiva till flera förslag.

På frågan om vilken information man vill få om ett fel svarade alla att de ville ha rättelseförslag. Alla ville också ha tillgång till en kort förklaring även om en del tyckte att det var onödigt i vissa fall och att det skulle vara valbart. Mera omfattande förklaringar var två försökspersoner mera tveksamma till och till att se omfattande skrivregler var tre tveksamma till. Alla tyckte dock att det kan vara bra om det finns tillgängligt men bara visas på begäran.

Både för Granska och för MSWord anmärkte man på att man inte alltid förstod förklaringarna till de grammatiska felen. De flesta tyckte sig inte kunna den grammatiska terminologin tillräckligt för att tillgodogöra sig informationen som ges. Flera ansåg att en förklaring bör belysas med exempel.

Outnyttjade funktioner i MSWord

Det visade sig att även vana användare av MSWord missar vissa möjligheter i granskningen. Det var bara två försökspersoner som sade sig använda möjligheten att ändra inställningar i verktygsmenyn. Möjligheten att ändra inställningarna för granskningen verkar ganska okänd trots att försökspersonerna är vana att använda engelsk grammatikkontroll. Ingen verkade känna till att man t.ex. kan välja längd på det som betraktas som

en lång mening. Möjligheten att ”ignorera regel” i dialogrutan var det heller ingen som kom på tanken att använda för långa meningar. Slutsatsen av detta blir att man tydligare bör visa hur man kan ändra granskningsinställningar än vad MSWord gör.

Användaren låter sig styras av programmet?

I vårt försök instruerade vi inte försökspersonerna att försöka upptäcka fel som missats av programmet. Detta borde vi ha gjort eftersom det är svårt att ur resultaten säga varför man inte upptäckte de fel som programmet missade. En del fel upptäcktes men andra missades. Vi uteslöt manuell granskning eftersom det skulle ha tagit för mycket tid. Därför kan vi nu inte ta ställning till huruvida man litar för mycket på programmet eller inte. Efter försöket fick försökspersonerna frågan om det trodde att det var någon skillnad i vilka fel MSWord och Granska missade men ingen vågade ta ställning till frågan.

I Granska-fallet ingick ett antal fel där programmet gav felaktiga förslag. De uppenbara fallen framkallade några leenden men ett mera svårtolkat fall där inget alternativ var riktigt och alternativen dessutom var snarlika vållade en del huvudbry och en del felaktiga val. Det är svårt att jämföra med MSWord eftersom det i allmänhet inte är stavfelen som är svåra att ta ställning till och man bara får ett alternativ för grammatikfallen.

Språkliga aspekter - resultat

Datainsamling för den språkligt relaterade delen

Den huvudsakliga datainsamlingen för den språkliga delen bestod av direkt observation med tillhörande protokoll. Protokoll fördes över vilka fel som försökspersonerna accepterade som fel, vilket ersättningsförslag som försökspersonen valde bland ersättningsförslagen samt om försökspersonen kom med egna ersättningsförslag. Nedan följer två tabeller där samtliga försökspersoners reaktioner på programmen presenteras. Datainsamlingen kretsar kring det som TEMAA kallar *error dialogue*, dock något utvidgat till att gälla hur användaren reagerar på olika händelser i feldialogen.

Språkligt relaterade data insamlade med MSWord

Något som försvårade insamlingen av data var om försökspersonen utförde en ändring så påverkade det vilka övriga fel som upptäcktes av MSWord. Nya fel markerades också som en konsekvens av en annan ändring, detta försvårade möjligheten att jämföra de olika försökspersonernas reaktioner på programmet. Det protokoll som användes under försöken med MSWord var baserat på en körning av programmet utan att några ändringar i texten gjordes. Detta medförde att protokollet inte alltid stämde med det som hände i försökssituationen. Ett annat problem var att granskningen skiljer sig åt om man använder den sekventiella granskningsdialogen eller arbetar med den kontinuerliga granskningen. Man får olika resultat beroende på vald granskningsmetod. Men för det mesta fungerade metoden med direkt observation och protokoll. Att försöka spåra vad som egentligen hände vid dataförlusterna (markerade med ? i tabellen, inom parentes anges val av ersättningsförslag eller egna ersättningsförslag) har vi inte gett oss på i denna kursuppgift. Om vi vid ett senare tillfälle vill förbättra databearbetningen har vi tillgång både till bandinspelningar av försökspersonerna när de tänker högt och också de datafiler som innehåller de ändringar som försökspersonerna har utfört. Försökspersoner anges som 1, 2, 3, 4 och

5 i tabellerna nedan. Följande rubriker används i tabellerna (rubrikerna gäller även för protokollet från försöken med Granska nedan):

Feltyp Under rubriken Feltyp beskrivs det detekterade områdets verkliga feltyp, bedömt av Ola. Om det rör sig om ett falskt alarm finns det med här också. Om det är svårt att utröna vad det egentligen är för fel så har vi kallat dessa detektioner för *Oklart fall*.

Text Under rubriken **Text** finns den text som har detekterats av programmen.

Diagnos Under rubriken **Diagnos** finns programmets diagnos av det detekterade området.

Förslag Under rubriken **Förslag** finns det/de ersättningsförslag som programmen presenterar för användaren.

Fp ändrar Under rubriken **Fp ändrar** anges vilka försökspersoner som har utfört en ändring p.g.a. att programmet har markerat ett detekterat område i texten.

Fp ignorerar Under rubriken **Fp ignorerar** anges vilka försökspersoner som inte har utfört en ändring p.g.a. att programmet har markerat ett detekterat område i texten.

Vi har valt att särskilja de falska alarmen från de alarm som upptäcker ett fel men ger en falsk diagnos. Det finns en skillnad mellan dessa två typer av alarm, nämligen att den språksäkra användaren kan hantera såväl falska alarm som felaktiga diagnoser. De falska alarmen lämnas medan alarm med felaktig diagnos åtgärdas på egen hand. Det räcker alltså med att felet markeras för den språksäkra användaren för att slutresultatet skall bli bra. Hur detta fungerar i dessa försök presenteras i avsnitten om *Felaktig diagnos*.

Tabell 1 Granskning med MSWord

Feltyp	**Text**	**Diagnos**	**Förslag**	**FP ändrar**	**FP ignorerar**
Falskt alarm	infologiska ekvation	Kongruensfel	infologisk ekvation | infologiska ekvationer	5	1,2,3,4
Falskt alarm	sema	Stavfel	Selma |semla | sena | sera | sega		1,2,3,4,5
Falskt alarm	I=i	stavfel (dock grönt också)	Ii	3(sätter in kolon)	1,2,4,5
Falskt alarm	formalisera	Infinitiv efter preposition	att formalisera	1,2,4(egen ändring ”ordet formalisera”),5	3
Falskt alarm	urspråket	Stavfel	-	5(ett s.k. urspråk)	1?(2),3,4
Falskt alarm	”bokhet”	Stavfel	-		1,2,3,4,5
Falskt alarm	är	Versal vid meningsbörjan	Är	1(tar bort citattecken), 4(ändrar själv), 5(Detta är…)	2?,3?
Falskt alarm	Stipulativ	Stavfel	-		1?,2,3,4,5
Falskt alarm	stipulativ	Stavfel	-		1,2,3,4,5
Falskt alarm	klassifikatorisk	Stavfel	-	2 (delar upp ordet), 5(endast kommentar till skribenten)	1,3,4

Falskt alarm	."innebär	Versal vid menings-början	Innebär	?ingen notering	?
Falskt alarm	."har	Versal vid menings-början	Har	?ingen notering	?
Kongruens-fel	vår perspektiv	Kongruensfel	vårt perspektiv	1(ändrar själv), 2,3,4,5(våra perspektiv)	
Kongruens-fel	två delsystemen	Kongruensfel	två delsystem	1(ändrar själv), 2,3,4,5	
Kongruens-fel	ett bra situa-tionsanalys	Kongruensfel	en bra situa-tionsanalys	1,2,3,4,5	
Kongruens-fel:	min arbete	Kongruensfel	mitt arbete	1,2?,3,4?,5	
Kongruens-fel	någon gemen-sam regler	Kongruensfel	någon gemen-sam regel	1(egen ändring), 2(några …regler), 3,4,5(något gemensamt regelverk)	
Oklart fall	Begreppet bok …	Lång mening		1,2,3,	4(ignore efter att ha ändrat), 5(ignore efter att ha kom-menterat me-ningen på an-nat ställe i texten.
Oklart fall	Men vad jag vill säga	Lång mening	-	1(orsakas av en annan ändring),	2,4
Oklart fall	Med formali-sera menar jag	Lång mening	-	1(dock ej nöjd med meningen),5	?(2),3,4(ver-kade inte upp-stå)
Oklart fall	Ordets "bok"s konnotation …	Lång mening	-	1(ändrar själv), 2 (delar meningen), 3,4,5	
Oklart fall	Ju abstraktare …	Lång mening		3,5(lägger till kom-mentarer)	1,2,4
Oklart fall (fel diagnos –bör ha varit teckenfel, 'bok's ??)	"bok"s	Stavfel	boks	1(ändrar själv), 2 (bok),3,4, 5(stryker –s)	
Oklart fel	Den brittiske	Lång mening	-	1,2,4,5	3?
Falskt alarm	Langeförs	Stavfel	Langekörs		1,2,3,4,5
Stavfel	"gramatik"	Stavfel	grammatik	1,3,4,5	2(eftersom "…"
Stavfel	sturuktuera	Stavfel	strukturera	1,2,3,4,5	
Stavfel	trors	Stavfel	trots (6:e för-slaget), förslget ej med vid hö-germusknapp)	1(egen ändring),2, 3,4,5	
Stavfel	bedömmer	Stavfel	bedömer	1,2, 3,4,5	
Stavfel	överblikbar-heten	Stavfel	överblickbar-heten	1, 2,3,4,5?	

Stavfel	detaljnings-graden	Stavfel	-	1 (detaljgraden), 2(försöker rätta själv), 3(sätter "…"), 4(detaljeringsgraden), 5?	
Stavfel	systemveteren	Stavfel	?	?	?
Särskrivning	avgränsnings problematiken	Bestämdhets-fel	avgränsnings problematik	1(en rad ändringar), 3(sätter ihop orden), 4(eget: problematiken med avgränsning), 5	2(programmet buggar ur till-fälligt)
Teckenfel	information ,	Mellanrum före skilje-tecken	information,		1?, 2
Teckenfel	sammanhang-	Stavfel	sammanhang, sammanhangs	1(egen ändring), 3, 4(sätter in mellanslag), 5	2?,
Teckenfel	laddade-	Stavfel	laddade | lad-dades	1(tar bort bindestrecket) 2(bort med -), 3,4(sätter in mellanslag) 5	
Teckenfel	verktyg-	Stavfel	verktyg | verk-tygs	1(tar bort bindestrecket) 2(bort -) 3(sätter in mellanslag) 5	4?
Verb saknas	verb saknas	verb saknas	-	1(stoppar in komma-tecken) 2(lägger till sub-jekt+predikat), 5(endast kommentar)	4?

Analys och diskussion av språkligt relaterade data – MSWord

Lång mening

Skulle en regel som enbart räknar antalet ordet i meningen påverka försökspersonerna? Jo det visade sig att samtliga personer går in och börjar ändra i meningar som MSWord har signalerat som långa (fler än 30 ord). Vi har inte bedömt om meningarna blev bättre efter de försök som gjordes att ändra meningarna. Det intressanta är att användarna försöker förbättra något med diagnosen *Meningen är ganska lång.* Försöksperson 4 kommenterar att "Lång mening behöver ju inte innebära att meningen är grammatiskt felaktig". Försöksperson 4 är också försiktigare med åtgärder i dessa meningar. Men mot slutet av försökssessionen börjar också försöksperson 4 att gräva i de långa meningarna, vilket skulle indikera försökspersonens skepsis mot "feltypen" mattas när den har dykt upp tillräckligt ofta.

Falska alarm

De flesta falska alarmen härstammar från stavningskontrollen och förvånansvärt många av dessa alarm blir besvarade av användarna. En försöksperson med svenska som modersmål sväljer ett falskt kongruensfel *Langeförs infologiska ekvaktion* av bara farten.

Om diagnosen ser bra ut och ersättningsförslagen ser bra ut är ju detta inte så konstigt, speciellt inte i en försökssituation. Ett intressant falskt alarm är *Infinitiv efter preposition* i meningen *Med formalisera menar jag inte formalisering som i ett datasystem utan en formalisering som hjälper oss att förstå varandra genom att utgå ifrån gemensamma regler i vår kommunikation.* Tre försökspersoner köper det falska alarmet medan två personer genomskådar problemet med granskning av metaspråk. Försöksperson 4 sätter in ordet ”ordet” före *formalisera* för att förbättra texten, imponerande!

MSWord ger en felaktig diagnos på särskrivningen *avgränsnings problematiken* och tolkar det som ett bestämdhetsfel. 3 av försökspersonerna genomskådar den felaktiga diagnosen och skriver om, men endast en av dessa (försöksperson 4) inser att det är en felaktig särskrivning och sätter ihop orden.

Felaktig diagnos

De detekterade felen får inte flera olika diagnoser i MSWord, vilket kan tänkas medföra färre felaktiga diagnoser. MSWord försöker att detektera kongruensfel och när dessa sammanfaller med särskrivningar kommer man att i åtminstone strikt mening få felaktiga diagnoser. Detta sker dock endast en gång på försökstexten. Tre personer ändrar, endast en person gör den troligaste ändringen – sätter ihop orden. En person sväljer MSWords förslag.

Ersättningsförslagen

När det gäller ersättningsförslag till stavfel klarar försökspersonerna att välja rätt ersättningsförslag även om det bästa inte står överst i listan. (Möjligen skulle man kunna tänka sig att det tar lite längre tid att välja ett förslag en bit ner på rankningen). Förslagen till kongruensfel valdes ofta enhälligt av försökspersonerna, men förvånansvärt ofta valde försökspersonerna att själva sätta in förslag i texten, ofta samma förslag som programmet gav.

Sammanställning av försöken med Granska

Att samla in och sammanställa data från försöken med Granska beredde mycket färre problem än med MSWord. En stor fördel var att Granska uppträdde konsekvent i alla försökstillfällena.

Tabell 2 Granskning med Granska

Feltyp	Text	Diagnos	Förslag	FP ändrar	FP ignorerar
Falskt alarm	sema	Stavfel	sena | sega | sexa| semla		1,2,3,4,5
Falskt alarm	formalisera	Verbfrasfel	att formalisera	1,2,5	3,4
Falskt alarm	mer korrekt	Kongruensfel	mer korrekta	5	1,2,3,4
Falskt alarm	mindre positivt eller negativt laddade	Kongruensfel | Särskrivningsfel	mindre positiv eller negativ laddade- | negativ laddade- | laddade emotivt	2(alt.2), 3(alt.2)	1,4,5
Falskt alarm	Stipulativ	Stavfel	Stipelakaltiv | Stipeldativ | Stipelnativ		1,2,3,4,5
Falskt alarm	stipulativ	Stavfel	Stipelakaltiv | Stipelnativ		1,2,3,4,5

Falskt alarm	klassifikatorisk	Stavfel	klassifikatorsk | klassifikatonisk | klassifikatoxisk | klassfikadorisk		1,2,3,4,5
Falskt alarm	Performativa	Stavfel	Parnormativa | Pärnormativa		1,2,3,4,5
Falskt alarm	Langshow	Stavfel	Lagshow | Långshow | …		1,2,3,4,5
Särskrivning	Filosofi läraren	Felaktig särskrivning	Filosofiläraren	1,2,3,4,5	
Kongruensfel	en begrepp	Kongruensfel	ett begrepp	1,2,3,4,5	
Kongruensfel	själv språket	Kongruensfel	själva språket | själv språk	1(alt.1), 2(alt.1), 3(alt.1), 5(alt.1)	4
Kongruensfel	vår perspektiv	Kongruensfel	(vårt perspektiv, ej i PC-versionen)		1,2,3,4, 5(våra perspektiv)
Kongruensfel	en mer formaliserad språk	Kongruensfel	ett mer formaliserat språk	1,2,3,4,5	
Kongruensfel	våra kommunikation	Kongruensfel	våra kommunikationer	2,5	1,3, 4(vår kommunikation)
Kongruensfel	skilda betydelse	Kongruensfel	skild betydelse | skilda betydelser | skilda betydelserna	1(alt.2), 2(alt.2), 3(alt.2), 4(alt.2), 5(alt.2)	
Kongruensfel	båda betydelsen	Kongruensfel	bådade betydelsen| båda betydelser | båda betydelserna	1 (alt.3), 2(alt.3), 3(alt.3), 4(alt.3), 5(alt.3)	
Kongruensfel	två delsystemen	Upptäcks ej!			
Kongruensfel	ett bra situationsanalys	Kongruensfel	en bra situationsanalys	1, 2,3,4,5	
Kongruensfel	laddade kommunikation	Kongruensfel	laddad kommunikation | laddade kommunikationer | laddade kommunikationer	1(alt.1), 2(alt1), 3(alt.1), 4(alt.1), 5(alt.1)	
Kongruensfel:	min arbete	Kongruensfel	mitt arbete	1, 3,4,5	2
Kongruensfel	någon gemensam regler	Kongruensfel	någon gemensam regel, några gemensamma regler	1(alt.1), 2(alt.2) 3(alt.1), 4(alt.1) 5(alt.1)	
Meningsbyggnadsfel	och varje av de har sina egna regler	Felaktig pronomensform använd subjektsform	av dem har	1(av har), 2, 3 (var och en), 4,	5(vill ha bättre förslag)

Predikativfel	konsekvensen av misstolkningen kan vara kost-samma	Predfel	kostsam	1,3,4,5	2
Predikativfel	de är godtycklig	Predikativfel	de är godtyckliga	1,2,3,4,5	
Falskt alarm	Langeförs	Stavfel	Langekörs		1,2,3,4,5
Stavfel	gramatik	Stavfel	grammatik	1,2,3,4,5	
Stavfel	någon sort ordning	Felaktig sär-skrivning	någon sortordning	1(sorts ord-ning), 2,4	3(det är dock ett fel),5
Stavfel	sturuktuera	Stavfel	strukturera	1,2,4	3,5
Stavfel	trors	Stavfel	trots (6:e förslaget), förslget ej med vid högermusknapp)	1,2,3,4,5	
Stavfel	bedömmer	Stavfel	bedömer, berömmer	1,3,4	2,5
Stavfel	systemveteren	Upptäckas ej			
Stavfel	överblikbarheten	Stavfel	överläkbarheten| över-blickbarheten | överblekbarheten | övervikbarheten	1,2,3,5	4
Stavfel	detaljningsgraden	Stavfel	dentaljningsgraden | deltaljningsgraden | detaljnyingsgraden | dentaljningsgraden		1,2,3,4, 5(vill ändra direkt i texten)
Särskrivning	en cirkel defintion	Felaktig sär-skrivning	en cirkeldefinition	1,2,3,4, 5	
Särskrivning	avgränsnings pro-blematiken	Felaktig sär-skrivning	avgränsningsproble-matiken | avgränsnings pro-blematik	1,2,3,4, 5(vill egent-ligen skriva om)	
Särskrivning	ett radio program	Felaktig sär-skrivning | Kongruensfel	ett radioprogram | en radio program	1(alt.1), 3(alt.1), 4(alt.1), 5(alt.1)	2
Teckenfel	sammanhang- i	Felaktig sär-skrivning	sammanhang-i		1,?,?,?,5ej med
Teckenfel	ett verktyg- se-mana	Kongruens | Felaktig sär-skrivning	den verktyg- sema-tiken | ett verktyg-se-mantiken		1,2,3,4,5

Analys och diskussion av insamlade data med Granska

Lång mening

Granska var inte utrustat med en regel som kontrollerade längden på meningarna och frågeställning undersöktes därför inte med Granska.

Falska alarm

Samtliga försökspersoner ändrade texten vid något av de falska alarm som uppstod. En försöksperson tyckte dock att just den meningen med det falska alarmet var konstig överhuvudtaget. Detta kan vara ett problem med försöksupplägget, genom att försökspersonerna inte vara vana vid innehållet i texten. Vi gör en noggrannare genomgång av de

falska alarm som Granska genererade än vad vi gjorde för MSWords falska alarm, eftersom vi känner till hur Granska är uppbyggt i detalj.

De falska alarmen bestod av:

- 6 Stavfel
- 2 Kongruensfel
- 1 Verbfrasfel
- 1 Särskrivningsfel (samförekomst med ett kongruensfel)

Ingen försöksperson utförde någon åtgärd på falska stavfel, däremot svaldes tre falska alarm för grammatikfel av olika försökspersoner. En försöksperson accepterade inte något av dessa falska alarm överhuvudtaget. De tre falska alarmen på grammatisk nivå som lurade användarna var följande.

Falskt alarm 1
Med formalisera menar jag inte formalisering som i ett datasystem utan en formalisering som hjälper oss att förstå varandra genom att utgå ifrån gemensamma regler i vår kommunikation.

Granskas respons: Verbfrasfel, *att formalisera* skall det vara.

Teknisk orsak: Detta alarm kan egentligen tolkas som ett korrekt alarm. Programmet kan inte förstå att det rör sig om en metaanvändning av ordet *formalisera*.

Kommentar om användarnas reaktion: Det är samma personer som i fallet med MSWord som väljer att ignorera detta falska alarm.

Falskt alarm 2
Georg Stiernhielm hävdade på 1600-talet att svenska orden avbildar verkligheten mer korrekt än andra språk.

Granskas respons: Kongruensfel, *mer korrekta* skall det vara.

Teknisk orsak: För grov tolkning av konjunktionen i en sammansatt nominalfras. Den tillämpade regeln är nu förfinad och liknande falska alarm skall inte uppstå igen.

Kommentar om användarnas reaktion: Det är endast en försöksperson med svenska som modersmål som sväljer detta falska alarm.

Falskt alarm 3.
Ord kan vara mer eller mindre positivt eller negativt laddade- emotivt laddade.

Granskas respons: Kongruensfel och särskrivningsfel.

Teknisk orsak: bindestrecket tolkas inte som ett tankstreck. Förbättrad tokenisering i Granska medför numera att det falska alarmet inte uppstår längre.

Kommentar om användarnas reaktion: Försöksperson 2 och 3 sväljer samma diagnos och väljer samma förslag i detta fall (särskrivning). Försöksperson 3 reagerar med kommentaren ”Konstig mening”. Försöksperson 2 litar på referensen till Svenska skrivregler och klickar snabbt på åtgärdsknappen. Det kan vara problem med referenser till kända böcker när programmet inte är tillräckligt bra. Hur referensen presenteras bör vi noga tänka igenom framöver. Det bör framgå att detektionen, diagnosen och ersättningsförslaget är programmets egna och referensverket kan endast förklara felet närmare.

Felaktig diagnos

Kongruensfel och särskrivning sammanfaller ganska ofta, och ofta verkar kongruensfel vara en felaktig diagnos i strikt mening. Det är ju så att det ibland uppstår ett kongruens-

fel på grund av särskrivningen och om det då är en felaktig diagnos eller inte kan man diskutera. Trots allt verkar det som att användarna kan välja rätt genom att förslagen visar vilken diagnos som är rätt.

Ersättningsförslagen

När Granska presenterar flera olika ersättningsförslag verkar användarna kunna välja ett förslag som gör meningen grammatiskt korrekt. Försökspersonerna är överens i alla fall utom ett när det finns flera olika ersättningsförslag presenterade till kongruensfelen.

Skillnader mellan programmen under försöken

En jämförelse mellan de två programmen kan inte bli annat än orättvis. När det gäller de lingvistiska verktygen som används i respektive grammatikkontroll har Lingsoft program som har utvecklats sen tidigt 1980-tal. Granskas grundläggande analys bygger i stort på program utvecklade under 1998 och 1999. För en överblick se tabell 3 nedan. När det gäller själva språkgranskningsreglerna påbörjades i stort sett arbetet samtidigt, dock med de olika förutsättningarna som finns mellan högskolevärlden och industrin när det gäller utveckling av dataprogram. I det här fallet handlar det väl mycket om antalet timmar som har lagts ner på själva utvecklingen av språkgranskningsreglerna och kanske framförallt testning och utvärdering, där Granska har haft mycket små resurser till sitt förfogande. Just testning och utvärdering är mycket viktigt då man i denna process upptäcker och eliminerar de förhatliga falska alarmen. Just detta kan faktiskt få Granska att framstå som ett bättre program vid utvärdering på en förhållandevis kort text. Det beror på att elimineringen av falska alarm inte har varit så omfattande när det gäller Granska, vilket medför att täckningen av felen inte har påverkats av denna faktor. Få falska alarm innebär också en något sämre täckning, vilket man bör ha i bakhuvudet vid läsning av den fortsatta texten.

Tabell 3 **En översikt över de grundläggande lingvistiska modulerna i respektive program.**

Lingvistiskt program	Lingsoft	Granska
Lexikal analys	Swetwol (1991)	SUC + SAOL genererat lexikon 1998-2000
Lexikal generering	Swetwol (1991?, osäkert om man kunde generera ord från denna version)	Carlbergers morfologiska genereringsregelsystem (sommaren 1999-2000)
Taggdisambiguering	SWECG (1998) Constraint grammar-formalismen utvecklad sen tidigt 1980-tal. De svenska reglerna färdiga under 1998	Granskas taggare (1998-2000)
Regelspråk:	Constraint grammar (tidigt 1980-tal)	Granskas regelspråk (sommaren 1998-2000)
Språkgranskningsregler	Regelarbete påbörjades 1997	Ett mer omfattande regelarbete påbörjades under 1998. En del kongruensregler fanns skrivna i dels Prolog 1996 och dels i ett enkelt regelspråk 1996-1997.

Det finns också många andra punkter en jämförelse blir orättvis, en av de viktigaste är kanske valet av försökstext/testtext. Vi hävdar att det i stort sett är omöjligt att konstruera en försökstext lämpad för ett försök som skulle kunna ge en rättvis bild av de olika pro-

grammens förmåga. Om texten t.ex. innehåller tio kongruensfel, så kanske det ena programmet har tur så att de valda orden i konstruktion analyseras på ett förutsägbart sätt, felets placering i satsen stämmer överens med vad granskningsreglerna förutspår och programmet kan generera ett förslag som passar in i kontexten.

Om man skulle göra en jämförelse av programmen vore det rimligt med åtminstone ett hundratal autentiska felexempel för varje feltyp (grova feltyper). Ett sådant material skulle tas fram av en oberoende person utan insyn i de båda programmen. Ett sådant material skulle dock inte lämpa sig för en försökssituation där andra krav ställs på en testtext. Vi valde därför att välja en text med många autentiska fel i, som vi försökte balansera med avseende på de båda programmen, vilket innebar att vi lade till en del kongruensfel i några nominalfraser samt några fall av inkongruens i predikativ. Detta blev ju inte heller rättvist, eftersom vi själva låg bakom reglerna i Granska, men i våra försök koncentrerade vi oss på effekterna av utformningen av det grafiska gränssnittet samt hur de språkliga reaktionerna från programmen bemöts av användarna. En jämförelse av programmet var alltså inget som vi i större utsträckning tog med i designen av försöken. Vi hade med en intervjufråga till försökspersonerna om de hade uppmärksammat några skillnader men ingen kunde besvara frågan.

Trots det orättvisa i en jämförelse av programmen, tycker vi att det är intressant att visa hur programmen skiljer sig åt på just den testtext som vi konstruerat. I den fortsatta texten presenteras först de gemensamma fel och de gemensamma falska alarm som båda programmen genererar, därefter presenteras de fel och falska alarm endast upptäcks av ett av programmen.

Det som programmen är ense om är mestadels kongruensfel och stavfel, se tabell 4. Om man tänker sig att användaren hade tillgång till båda programmen, så borde dessa detektioner betecknas som säkra.

Tabell 4 **Fel som båda programmen detekterar korrekt**

Text	**Diagnos**
min arbete	Kongruensfel
ett bra situationsanalys	Kongruensfel
vår perspektiv	kongruensfel (Granska ger dock ett otillfredsställande ersättningsförslag)
någon gemensam regler	Kongruensfel
Gramatik	Stavfel
Sturuktuera	Stavfel
Trors	Stavfel
Bedömmer	Stavfel
Överblikbarheten	Stavfel
Detaljningsgraden	Stavfel
avgränsnings problematiken	särskrivning (Granska), bestämdhetsfel (MSWord) om man rättar felet genom att sätta ihop orden blir satsen helt grammatiskt korrekt, väljer man att skriva det andra substantivet i obestämd form kvarstår det felaktiga intrycket av satsen.
sammanhang-	teckenfel, MSWord och Granska säger enstämmigt att det är ett stavfel
verktyg-	teckenfel, MSWord och Granska säger enstämmigt att det är ett stavfel

Jämförelsen visar vidare att programmen gör några helt rimliga falska alarm, dessa presenteras i Tabell 5. De falska alarmen kan förklaras genom att försökstexten innehöll några facktermer som de båda stavningskontrollerna inte kände till. När det gäller det

falska alarmet ”infinitiv efter preposition” tvingas programmen brottas med en av många metaspråkliga satser i texten och det falska alarmet är inte mycket att göra åt. Det går inte att eliminera det falska alarmet utan att veta vad texten handlar om.

Tabell 5 Gemensamma falska alarm

Text	**diagnos**
Sema	stavfel
Formalisera	infinitiv efter preposition
stipulativ	stavfel, 2 gånger
klassifikatorisk	stavfel
Langeförs	stavfel

MSWords språkkontroll gör 5 korrekta detektioner som Granska inte gör (se tabell 6 för detaljer). Det är svagt av Granska att inte upptäcka *två delsystemen*[1], det saknades helt enkelt en regel för denna typ av konstruktion vid försökens genomförande. När detta skrivs har vi en regel som upptäcker felet. Teckenfel (mellanrum för skiljetecken och bindestreck istället för tankstreck) hanterades inte av denna version av Granska. Det sammansatta ordet systemveteren kan delas upp i system-vete-ren och godkänds av Granska. Felet ”verb saknas” skall däremot detekteras av nuvarande version av Granska.

Tabell 6 Fel som endast MSWord detekterar korrekt

text	**Diagnos**
två delsystemen	Kongruensfel
systemveteren	Stavfel
information ,	mellanrum före skiljetecken
laddade-	Stavfel (fel diagnos, vara teckenfel)
Inte bara … i organisationer.	verb saknas

Granska lyckas detektera 13 fel som Word inte upptäcker (se tabell 7 för detaljer). Det kan tyckas lite förvånande att Granska hittar så pass många fler kongruensfel än MSWord, men det hänger delvis ihop med att arbetet med elimineringen av falska alarm i Granska inte har varit så omfattande som i Lingsofts program. Med en annan text och andra fel hade kanske MSWord hittat fler kongruensfel än Granska. När det gäller inkongruens i predikativ hittar Granska två fel medan MSWord inte markerar något. Det verkar nästan som att subjektet måste stå först i satsen för att MSWord skall hitta denna feltyp. Med andra placeringar av felkonstruktionerna i satsen hade MSWord säkert kunnat hitta fler fel. Särskrivningar detekteras av Granska men inte alls av MSWord, detta visste vi, men det är ändå lite förvånande att Lingsoft inte har gett sig på särskrivningar, då denna feltyp verkar vara mycket vanlig i svenskan.

Tabell 7 Fel som endast upptäcks av Granska

Text	**Diagnos**
en begrepp	Kongruensfel
själv språket	Kongruensfel
en mer formaliserad språk	Kongruensfel
våra kommunikationer	Kongruensfel

[1] När vi skriver denna text upptäcker inte Word *två delsystemen*, åtminstone inte med den kontinuerliga granskningen. Detta kan bero på att felkonstruktionen har en annan placering i denna text än vad den hade i försökstexten.

skilda betydelse	Kongruensfel
båda betydelsen	Kongruensfel
laddade kommunikation	Kongruensfel
filosofi läraren	Särskrivning
en cirkel defintion	Särskrivning
ett radio program	Särskrivning
varje av de	felaktig pronomenform (Granska ger felaktig diagnos och felaktigt förslag
konsekvensen av misstolkningen kan vara kostsamma	Predikativfel
de är godtycklig	Predikativfel

Fem falska alarm signaleras enbart av MSWord (se tabell 8) och fyra signaleras enbart av Granska (se tabell 9). Det är inte förvånande att ordet *bokhet* markeras av Word, då det är ett synnerligen konstigt ord, men i försökstexten är det korrekt, författaren har dessutom satt citationstecken omkring för att visa att det inte är något vanligt ord. Det bör tilläggas att Granska genererar en hel del konstiga ersättningsförslag, t.ex. för felstavade sammansättningar, Word väljer i många fall att inte föreslå något alls vid felstavade sammansättningar. Detta visar att även dagens stavningskontroller har en bit kvar till att vara verkligt användbara för användaren. Granska ger också en del felaktiga diagnoser vid särskrivningar, men eftersom försöken visar att användaren kan välja rätt mellan två förslag (där det ena är felaktigt), så presenterar vi inte de felaktiga diagnoserna bland de falska alarmen. Granska ger också en hel del felaktiga förslag t.ex. vid kongruensfel som t.ex. *bådade betydelsen*. Detta får ses som ett utslag av att de regler som genererar ersättningsförslag inte har testats under tillräckligt lång tid (se vidare i tabell 3 ovan).

Tabell 8 **Falska alarm som endast genereras av MSWord**

Text	**Diagnos**
Infologiska ekvation	Kongruensfel
I=i	stavfel, dock markerat med grönt också.
Urspråket	Stavfel
”bokhet”	stavfel, (konstigt ord)
Är	versal vid meningsbörjan

När det gäller användarnas reaktion på de olika programmen verkar de kunna svälja falska alarm från såväl Granska som MSWord. Försökspersonerna verkade inte tro mer på t.ex. MSWord än Granska, utan de förhöll sig ungefär på samma sätt till de två programmen vad det gäller det språkliga. Granska ger ofta flera olika tolkningar av feltypen, detta verkar inte besvära användarna utan snarare tvärtom, det verkar bli enklare att välja rätt alternativ även om ett tveksamt presenteras.

Tabell 9 Falska alarm som endast genereras av Granska

Text	**Diagnos**
mer korrekt	Kongruensfel
Mindre positivt eller negativt laddade	Kongruensfel, särskrivning
Performativa	Stavfel
Langshow	Stavfel

Något som är intressant är att programmen tillsammans hittar fler fel än enskilt, dock bidrar varje program med några egna falska alarm, så det kanske går jämt upp. Båda pro-

grammen ger en del falska alarm och en del falska diagnoser, vilket användarna får stå ut med även i framtiden.

Slutkommentarer för den språkliga delen

De falska alarmen är intressanta eftersom vi observerade att det var stor skillnad i bearbetningstid mellan t.ex. ett stavfel och ett besynnerligt kongruensfel, speciellt för den försöksperson med svenska som andraspråk.

De falska diagnoserna är också intressanta för det verkar som att språkligt skickliga skribenter nöjer sig med detektionen av ett fel, åtgärden klarar de snabbt av på egen hand. Dessa två problem skulle det vara intressant att gräva ner sig djupare i. Det finns en del kvar att hämta i denna undersökning, men framförallt vore det spännande att göra en studie som var mer inriktad på användarnas bedömning av falska alarm och felaktiga diagnoser och sätta det i en kontext av språklig kunskap.

Diskussion om försöksuppläggningen

Hur man ser på sin roll påverkar hur man använder verktygen. Detta framgick speciellt tydligt när en av försökspersonerna vid vissa feltyper, speciellt vid långa markeringar, inte försökte rätta i texten utan i stället skrev in kommentarer i texten riktade till författaren. Denne försöksperson skulle ha velat använda verktyget för kommentarer men detta tillät vi inte under försöket. Även andra skribenter fällde kommentarer av typen man vet ju inte hur skribenten har tänkt här. Man kan av detta dra den slutsatsen att man om man ska studera hur användare använder granskningsverktygen så blir resultaten förmodligen något annorlunda om man ombeds att skriva en egen text och granska den än om man får en färdig text som ska granskas.

Att rollen påverkar granskningen var vi medvetna om från början när vi planerade försöket. Vi ville dock studera hur en användare granskar en text som innehåller fel och vi ville att alla försökspersoner skulle granska samma text. Om man i stället hade valt att låta försökspersonerna själva producera texten hade man förmodligen fått ganska olika typer av fel och vi hade kanske fått texter som i huvudsak innehållit stavfel av typen tangentbordsfel eftersom vi hade vana skribenter som försökspersoner.

Ambitionsnivån påverkar säkert hur vi granskar en text. Normalt kan man utgå från att en text som ska publiceras för en större publik granskas noggrannare än en som ska läsas av en mera förlåtande krets. Det är troligt att även detta påverkar sättet att använda verktygen. Detta är svårt att simulera i ett försök. Vi valde att inte försöka påverka försökspersonernas ambitionsnivå. Möjligen kan vi se en koppling i resultaten till att de som lade ned mycket energi på att försöka lösa problemen med de långa meningarna hade en högre ambitionsnivå under försöket än de som gav upp och tryckte på ignorera. Man kanske borde definiera ambitionsnivån om man ska göra om försöket. Speciellt om man fritt låter försökspersonen välja arbetssätt.

De invanda rutinerna verkar påverka valet av arbetssätt ganska mycket. Vi hade också förutsatt att man följer sina vanor så mycket att man inte upptäcker de alternativ som finns tillgängliga. Vi försökte därför fråga om våra försökspersoner trodde sig ha blivit påverkade av försöket och trodde sig använda något arbetssätt mer efter försöket än tidigare. Ingen svarade jakande på den frågan. Eftersom alla var så pass skickliga användare av programmet så kände förmodligen alla till alternativen tillräckligt väl för att inte bli påverkade. Med en mindre van användargrupp hade resultatet kanske blivit annorlunda.

Vi försökte därför ta reda på hur våra försökspersoner normalt arbetar genom intervjufrågor.

När det gäller de språkliga aspekterna är studien som redan nämndes alldeles för liten för att kunna säga något om de olika programmens kvalitet. Det är trots det intressant att se att programmen ger stora skillnader i resultat beroende på hur granskningsalgoritmen ser ut.

Litteratur

[G 2000] Granska, URL: http://www.nada.kth.se/theory/granska/

[HJR, FLS 80] Hayes, J. R., & Flower, L. S. (1980). Identifying the organization of writing processes. In L. W. Gregg & E. R. Steinberg (Eds.), Cognitive Processes in Writing (pp. 3-30). Hillsdale, N.J: Lawrence Erlbaum.

[KSE 92] Severinson Eklundh K. (1992) Problems in Achieving a global Perspective of the Text in Computer-based Writing. . Instructional Science, 21, 73-84.

[L 2000] Grammatifx, URL: http://www.lingsoft.fi/swegc

[MA 93] Monk A., Wright P., Haber J., Dawenport L. (1993) Cooperative Evaluation: how to do it. In Improving your Human-Computer Interface. Prentice Hall

[RD 98] Domeij R Detecting and presenting errors for Swedish writers at work. Technical report TRITA-NA-P9629, NADA, 1996.

[RD, OK, SL, KSE, ÅR 96] Rickard Domeij, Ola Knutsson, Stefan Larsson, Kerstin Severinson-Eklundh and Åsa Rex.Granskaprojektet 1996-1997. Technical report, NADA, June 1998.

[SL 98] Larsson S. Interaktivitet och användbarhet vid datorstödd språkgranskning och redigering en integrerad skrivmiljö, Technical report TRITA-NA-E9833

[T 99] TEMAA Final Report Case studies. URL: http://www.cst.ku.dk/projects/temaa/D16/d1616exp-3.html

[UT, SH, EA 99] Teleman, U. Hellberg, S. & Andersson E. Svenska Akademiens grammatik. Band 1-4, Nordstedts Ordbok, Stockholm 1999.