

Inkongruens och särskrivna sammansättningar

– en beskrivning av två typer av fel och möjligheten att granska dem automatiskt

Rickard Domeij, Ola Knutsson och Lena Öhrman

Sammanfattning

Inkongruens och särskrivna sammansättningar är två feltyper som på senare år blivit allt mer frekvent förekommande i såväl icke-professionella som professionella texter skrivna av skribenter med svenska som modersmål. I artikeln belyser vi feltyperna utifrån empirisk grund och visar på intressanta regelbundenheter i realisationen av fel. Vi diskuterar orsakerna bakom dessa regelbundenheter och till felens uppkomst.

Vi visar också hur man kan detektera och korrigera en hel del förekomster av dessa feltyper i språkgranskningsprogrammet Granska, som håller på att utvecklas i ett projekt på Nada, KTH i Stockholm. Från svensk synvinkel är just dessa fel särskilt intressanta att detektera eftersom de inte förekommer på samma sätt i engelskan och därmed inte finns med i den uppsättning regler som härstammar från engelsk grammatikkontroll. Vi belyser slutligen problem som uppstår i detektionen av fel och i interaktionen mellan människa och dator vid användningen av grammatikkontroll.

Inledning

IPLab (Interaktions- och presentationslaboratoriet) på Nada, KTH har under drygt ett decennium bland annat varit ett centrum för forskare från olika discipliner som intresserar sig för forskning kring utvecklingen och användningen av datorverktyg för skrivande. Ett intressant exempel på sådana datorverktyg är grammatikkontroll med hjälp av vilken en skribent mer effektivt kan korrekturläsa sina texter. Det finns emellertid en hel del tekniska, språkvetenskapliga och användarrelaterade problem förenade med utvecklingen av en sådan tillämpning (se t.ex. Domeij 1997).

Ett första problem är begränsningar i tekniken och i den bakomliggande språkliga beskrivningen som gör att bara vissa typer av ytmässiga, väldefinierade

rade fel kan hittas med någorlunda precision. För att undersöka om ett fel är lämpligt att detektera med dator krävs bland annat kunskap om realisationer av förekommande fel och dessutom tester på precision och täckning hos de regler som definierar felen. Utöver detta behövs också kunskap om vilka effekter datorstödet får på språk och skrivförmåga.

I följande avsnitt fokuserar vi på två feltyper: nominalfrasinkongruens och felaktiga särskrivningar. Vi beskriver deras realisationer, hur de kan definieras med regler och användas i en grammatikkontroll. Som exempel visar vi hur programmet Granska som håller på att utvecklas i ett projekt på Nada, KTH i Stockholm (Domeij, Eklundh, Knutsson, Larsson & Rex 1998) hanterar inkongruens och felaktiga särskrivningar. Dessa fel är särskilt intressanta ur svensk synvinkel eftersom de inte förekommer på samma sätt i engelskan och därmed inte finns med i den uppsättning regler som härstammar från engelsk grammatikkontroll. De är också intressanta då de utgör stora problem för skribenter som inte har svenska som modersmål.

Inkongruens i nominalfraser

I en empirisk analys av språket i framför allt dagstidningar och vetenskapliga rapporter har vi observerat att kongruensfel i nominalfraser blivit relativt vanligt förekommande också i infödda skribenters texter. Detta är en ny iakttagelse och troligen också ett nytt fenomen som kommit med den ökade datoranvändningen. Tidigare har kongruensfel bara uppmärksammats i samband med andraspråksinläring där felen orsakas av bristande kunskaper i svensk grammatik, men för infödda skribenter handlar det snarast om slarv orsakat av de ökade redigeringsmöjligheter som datorn ger. En ändring i texten från *hus* till *villa* är till exempel lätt att göra med datorn, samtidigt som det är lätt att glömma påverkan på omgivande ord som kongruerar. En fras som *ett stort hus* kan därför lätt resultera i *ett stort villa*, och så är kongruensfelet ett faktum. Eftersom det dessutom ofta slarvas med den slutliga korrekturläsningen vid elektronisk textframställning (se Severinson Eklundh 1995) upptäcks ofta inte felet.

En annan viktig observation vi gjort på basis av insamlade data är att de påträffade kongruensfelen visar upp starka regelbundenheter (se exempel i tabell 1). Det är till exempel mycket ovanligt att flera attribut i en substantivfras är inkongruenta både med varandra och med huvudordet, som i frasen *det stort villa*. När flera attribut är inkongruenta med huvudordet verkar de i regel behålla sin inbördes kongruens, till exempel *ett stort villa* eller *det stora villa*. Av 162 kongruensfel i enkla nominalfraser ser vi inte något exempel på motsatsen.

I materialet ser man dessutom sällan att ett attribut avviker från huvudordet i mer än ett särdrag. Detta gäller för hela 96 % av förekomsterna. En förklaring till detta kan vara att det sällan finns anledning att redigera ett ord i mer än ett särdrag. När ett huvudord till exempel byts ut mot en synonym sker föränd-

ringen endast i genus. Det är mera ovanligt att man gör flera ändringar av olika slag inom samma nominalfras, även om det förstås förekommer ibland.

Tabell 1.

Realiserade förekomster av inkongruenta nominalfaser med huvudord i utrum.

Obest. singular	#	Best. singular	#	Obest. plural	#	Best. plural	#
en bil		<i>en bilen</i>	5	<i>en bilar</i>		<i>en bilarna*</i>	1
en gul bil		<i>en gul bilen</i>	4	<i>en gul bilar</i>	2	<i>en gul bilarna*</i>	
<i>en gula bil</i>	3	<i>en gula bilen</i>		<i>en gula bilar</i>	2	<i>en gula bilarna*</i>	
<i>en gult bil</i>	7	<i>en gult bilen*</i>		<i>en gult bilar*</i>		<i>en gult bilarna*</i>	
<i>ett bil</i>	4	<i>ett bilen*</i>	1	<i>ett bilar*</i>		<i>ett bilarna*</i>	
<i>ett gul bil</i>	4	<i>ett gul bilen*</i>		<i>ett gul bilar*</i>		<i>ett gul bilarna*</i>	
<i>ett gula bil</i>		<i>ett gula bilen*</i>		<i>ett gula bilar*</i>		<i>ett gula bilarna*</i>	
<i>ett gult bil</i>	1	<i>ett gult bilen*</i>		<i>ett gult bilar*</i>		<i>ett gult bilarna*</i>	
den bil		den bilen		<i>den bilar</i>		<i>den bilarna</i>	1
<i>den gul bil</i>	2	<i>den gul bilen</i>	4	<i>den gul bilar*</i>		<i>den gul bilarna*</i>	
den gula bil		den gula bilen		<i>den gula bilar</i>		<i>den gula bilarna</i>	1
<i>den gult bil</i>		<i>den gult bilen</i>		<i>den gult bilar*</i>		<i>den gult bilarna*</i>	
<i>det bil</i>		<i>det bilen</i>		<i>det bilar*</i>		<i>det bilarna*</i>	1
<i>det gul bil*</i>		<i>det gul bilen*</i>		<i>det gul bilar*</i>		<i>det gul bilarna*</i>	
<i>det gula bil</i>		<i>det gula bilen</i>	9	<i>det gula bilar*</i>		<i>det gula bilarna*</i>	3
<i>det gult bil*</i>		<i>det gult bilen*</i>		<i>det gult bilar*</i>		<i>det gult bilarna*</i>	
<i>de bil</i>		<i>de bilen</i>		de bilar		de bilarna	
<i>de gul bil</i>		<i>de gul bilen*</i>		<i>de gul bilar</i>		<i>de gul bilarna</i>	
<i>de gula bil</i>	1	<i>de gula bilen</i>	6	de gula bilar		de gula bilarna	
<i>de gult bil*</i>		<i>de gult bilen*</i>		<i>de gult bilar</i>		<i>de gult bilarna</i>	
gul bil		<i>gul bilen</i>	2	<i>gul bilar</i>	3	<i>gul bilarna</i>	
gula bil		gula bilen		gula bilar		gula bilarna	
<i>gult bil</i>	4	<i>gult bilen*</i>		<i>gult bilar</i>		<i>gult bilarna</i>	

De observationer vi gjort kan sammanfattas i följande två generaliseringar:

1. Ett attribut avviker oftast bara i ett särdrag från huvudordet (särdragsavvikelsen= 1).
2. Om flera attribut avviker från huvudordet kongruerar de oftast med varandra.

Tabell 1 visar en delmängd realiserade förekomster av inkongruenta nominalfaser med huvudord i utrum. Konstruktionsmöjligheter som avviker från generaliseringarna har markerats med asterisk. Korrekta konstruktionsmöjligheter är okursiverade. Observera att det stora flertalet konstruktioner följer mönstret. Ett undantag är *det gula bilarna* som har tre förekomster. Detta kan möjligen förklaras av att *de* och *det* lätt förväxlas eftersom de kan uttalas på samma sätt.

De två generaliseringarna ovan gör att vi kraftigt kan reducera antalet teoretiska felmöjligheter genom att bortse från de mest osannolika och på så sätt reducera antalet regler i granskningsprogrammet. Färre regler för osannolika konstruktioner gör att falska alarm kan undvikas i högre grad, vilket kan vara att föredra om antalet missade fel därmed inte blir för stort. Generaliseringarna kan också användas för att rangordna förslag till rättelser i grammatikkontrollen så att de förslag som ligger på närmast särdragsavstånd från felet läggs först i för-

slagslistan. Vår hypotes är att en sådan rangordning av förslagen oftast överensstämmer med skribentens intentioner.

Särskrivna sammansättningar

Särskrivna sammansättningar är något man ser allt oftare. Engelskan tycks påverka svenska språket på många områden, kanske framför allt i IT-sammanhang. Ord som är direktöversatta från engelska verkar ofta leda till felaktiga särskrivningar. Sådana särskrivningar dyker ofta upp i datorsammanhang och i tekniska sammanhang där författaren/översättaren har sett dessa ord på engelska och sedan direktöversatt. Till exempel *home page* blir *hem sida*. I svenskan, till skillnad från engelskan, utgör felaktig särskrivning av sammansättningar ett språkproblem.

Engelskan är nog inte det enda som påverkar till fler särskrivningar. Sättet på vilket text framställs i reklam, till exempel utan avstavningar, kan vara en annan orsak. Även de som skriver restaurangmenyer verkar ha en viss förkärlek för särskrivna sammansättningar. Ibland kan man se dussinet felaktiga särskrivningar på en och samma matsedel. Särskrivningar blir allt vanligare också i löpande text. Man ser dem på Internet, i uppsatser, i e-postmeddelanden och så vidare.

För att analysera särskrivningar och möjligheten att detektera dem med hjälp av språkgranskningsprogrammet Granska, gjorde vi en egen undersökning där vi samlade ihop en mängd felaktiga särskrivningar (Öhrman 1998). Den resulterande felsamlingen består av 389 särskrivningar med kontexter hämtade från Internet. Vi använde oss av sökord som till exempel *dator program* i sökverktyget AltaVista och fick på så vis fram ett antal dokument som innehöll denna särskrivning. Då vi gick igenom varje dokument fann vi nästan i samtliga fall fler meningar i samma dokument som innehöll alla möjliga sorters särskrivningar. Vi upptäckte att de flesta författare med felaktiga särskrivningar varit mer eller mindre konsekventa med att skriva isär sammansättningar.

Vi delade in särskrivningarna i kategorier efter ordklass typer. Med ordklass menas ordklassen för de ord (för- och efterled) som särskrivningen består av. Därefter beräknade vi frekvens för de olika typerna. På så vis fick vi fram en typologi, helt och hållet baserad på vår insamling. Vi har även tagit med sammansättningar bestående av tre led som listas separat. Tabell 2 visar frekvensen för de vanligaste särskrivningstyperna.

Som framgår av tabell 2 är särskrivningar av typen *substantiv + substantiv* i särklass vanligast. De representerar 70 % av de felaktiga särskrivningarna i vår felsamling. Man kan också se att åtminstone två substantiv även förekommer i de flesta särskrivningar med tre led. Substantiv ingår i något av leden i sammanlagt 18 av de 31 olika typerna och i 364 av särskrivningarna. Med andra ord ingår substantiv i 94 % av särskrivningarna. Det näst vanligast förekommande är

adjektiv som finns i 15 % av fallen. Det är med andra ord stor skillnad mellan frekvensen för särskrivningar som innehåller substantiv och övriga.

Tabell 2. Frekvenser för olika typer av särskrivningar

<i>Ordklasser</i>	<i>Antal</i>	<i>%</i>	<i>Exempel</i>
subst. + subst.	276	70%	polis stationen
subst. + adj.	31	8%	färg glada
namn + subst.	14	3,6%	Tunabergs veteranen
adj. + subst.	10	2,6%	snabb maten
adj. + adj.	8	2,1%	mörk hårig
verb + subst.	6	1,5%	vandrar kängor
subst. + verb	3	< 1%	tjuv jagade
subst. + particip	3	< 1%	maskin beroende
adj. + verb	3	< 1%	offentlig gjort
adverb + verb	3	< 1%	samman faller
adverb + adj.	2	< 1%	sönder brända
prep. + subst.	2	< 1%	inom hus
<i>Tre led:</i>			
subst. + subst. + subst.	6	1,5%	favorit data spel
räkneord + subst. + subst.	3	< 1%	två mans kajak
adj. + subst. + subst.	3	< 1%	god natt hälsningar

Det finns en nackdel med den metod vi valde för att få fram särskrivningarna: vi kan ha fått ett snedvridet resultat då vi framför allt sökt på särskrivningar av typen *substantiv + substantiv*. Det stora flertalet särskrivningar vi fick fram var dock inte de vi hade som sökord i AltaVista. Även om just de särskrivningar som vi från början använde som sökord är i minoritet i vår felsamling så finns det en risk för att de övriga särskrivningarna är av samma sort eftersom de som skriver kan tänkas ha en tendens att upprepa samma typ av fel.

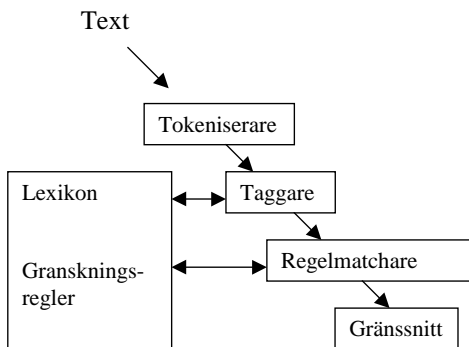
Särskrivna sammansättningar utgör ett växande språkproblem. I undersökningen kom vi fram till att särskrivningar innehållande två substantiv är i särklass vanligast. Därför är det naturligt koncentrera sig på att med hjälp av språkgranskningsprogrammet detektera och korrigera just dessa särskrivningar.

Granskasystemet

Både inkongruens och felaktiga särskrivningar kan i viss utsträckning hanteras i språkgranskningsprogrammet Granska. Granska är ett hybridssystem som kombinerar probabilistiska och regelbaserade metoder för att uppnå hög effektivitet och robusthet (Domeij, Knutsson, Carlberger & Kann 1999). Detta är en nöd-

vändig förutsättning för en grammatikkontroll som ska fungera i realtid i direkt interaktion med användaren (t.ex. Kukich 1992).

Hur systemet är uppbyggt i moduler visas i figur 2. Först görs en igenkänning av alla textens ord i tokeniseraren. I nästa steg tilldelas orden information om ordklasser och morfologiska särdrag med hjälp av en s.k. taggare som slår upp ord i lexikon och på statistisk grund räknar ut vilken av flera möjliga ordklassstilldelningar ett ord får i sin kontext. Den taggade texten skickas sedan vidare till regelmatcharen som genomsöker texten och matchar den mot de förväntade grammatiska avvikelser som definierats i granskningsreglerna. Granskningsreglerna genererar också felbeskrivningar och rättningsförslag som slutligen presenteras för användaren i ett grafiskt gränssnitt. Systemet innehåller också en rättstavningsmodul som kan hantera svenska sammansättningar (Kann, Domeij, Hollman & Tillenius 1998) och som används i granskningsreglerna för att kontrollera felaktiga särskrivningar.



Figur 1. Granskasystemets uppbyggnad.

Systemet finns implementerat på Unix och det finns också ett webbgränssnitt där man kan pröva programmet via en webbläsare (<http://www.nada.kth.se/theory/projects/granska/>). Vi arbetar dessutom på att ta fram ett gränssnitt för PC som kan användas interaktivt i skrivsituationen; bland annat är det tänkt att användas som ett forskningsverktyg vid användartester.

Granskningsregler för inkongruens och felaktiga särskrivningar

Granska har särskilda granskningsregler som genomsöker resultatet från taggningen på jakt efter grammatiska felaktigheter. En granskningsregel som kan detektera kongruensfelet *en hus* ser ut som i regel 1 (se äv. Knutsson 1996).

Regel 1:

```

kong22@inkongruens
{
X(wordcl= dt),
Y(wordcl= jj)*,
Z(wordcl= nn & (gender!= X.gender | num!= X.num | spec!= X.spec))
-->
mark(X Y Z)
corr(X.get_ form(gender:= Z.gender, num:= Z.num, spec:= Z.spec) Y Z)
info("Artikeln" X.text "stämmer inte överens med substantivet" Z.text)
action(granskning)
}

```

Regeln har två led som avskiljs med en pil. Det första ledet innehåller ett matchningsvillkor, det andra ledet specificerar den åtgärd som utlöses när villkoret uppfylls. I exemplet uppfylls villkoret när en determinerare följs av ett nomen (ev. föregånget av ett eller flera adjektiv) som skiljer sig i genus, numerus eller species från determineraren. I åtgärdsledet anges först att den felaktiga frasen i texten ska markeras. Sedan anropas en ordböjningsfunktion som med hjälp av lexikonet och ett antal regler genererar ett ersättningsförslag som kongruerar med substantivet. Detta presenteras som korrektionsförslag för användaren tillsammans med en diagnostisk text som förklarar felets art.

Lokalisering av särskrivna sammansättningar där det första ledet är ett sammansättningsled som inte kan stå själv, till exempel *skid åkning*, görs helt enkelt genom uppslagning i en förledsordlista. I övriga fall utgår strategin för att hitta felaktiga särskrivningar från att två nomen sällan står intill varandra. Om det dessutom går att sätta samman dessa två nomen så att de bildar en godkänd sammansättning som passar i kontexten finns ytterligare skäl att tro att det faktiskt rör sig om ett fel. Nedan följer ett exempel på en regel (regel 2) som hittar felaktiga särskrivningar av typen *Han såg de stora sten bumlingarna*:

Regel 2.

```

sär2@särskrivning
{
X1(wordcl= dt),
X2(wordcl= jj)*,
X3(wordcl= nn & (gender!= X1.gender | num!= X1.num | spec!= X1.spec),
X4(wordcl= nn & gender= X1.gender & num= X1.num & spec= X1.spec &
correctly_ spelled(concat(X3.text, text)))
-->
mark(X3 X4)
corr(X1 X10 X2 concat(X3.text, X4.text))
info("Särskrivningsfel, sätt ihop" X3 X4 "till " concat(X3.text, X4.text))
action(granskning)
}

```

I särskrivningsregeln ovan beskrivs ett särskrivningsfel där determineraren X1 är inkongruent med det första nomenet X3 men kongruerar med det andra nomenet X4 som i exemplet *de stora sten bumlingarna*. Inkongruensen mellan determinerare och det efterföljande nomenet tillsammans med kongruensen mellan

determineraren och det andra nomenet ger indicier på att ett särskrivningsfel har detekterats. För att ytterligare testa att det rör sig om ett särskrivningsfel sätts de två nomina samman och kontrolleras i stavningsfunktionen Stava. Om Stava godkänner det sammansatta ordet så felmarkeras särskrivningen och det sammansatta ordet ges som rättningsförslag.

Ibland kan två olika regler matcha samma område i texten, som i exemplet *ett cykel ställ*. För att inte systemet ska tolka detta fel som ett kongruensfel behöver granskningsreglerna ordnas så att särskrivningsreglerna får företräde framför kongruensreglerna i fall som det ovan. Än så länge är en sådan prioritetsordning mellan reglerna inte implementerad. I stället presenteras nu båda möjligheterna som fel, vilket allt som oftast också är den bästa metoden.

Många fel kan vara svåra att hitta på grund av flertydigheter som ibland inte går att avgöra med hjälp av kontexten; ett exempel är *en egen exekverings enhet* där ordet *exekverings* kan tolkas som ett genitivattribut. Ett annat problem är falska fel som råkar slinka med i vissa fall, till exempel *Den katt kvinnan såg var svart* där det utelämnade relativpronomenet ställer till problem. Ett annat exempel på falskt alarm vid kongruens är *Samtidigt utgör den nya hot mot de utrotningshotade arterna*. Arbete pågår med att förbättra reglerna med hänsyn till falska alarm och missade fel.

Resultat

Hela 97 procent av orden i en okänd text taggas korrekt, ett resultat som står sig väl i en internationell jämförelse. De okända orden taggas med 92 procent korrekt. Granskningsystemet i sin helhet (med ca 200 regler för ett tjugotal olika feltyper) granskar ungefär 1 000 ord per sekund inklusive taggning. Det är svårt att jämföra prestanda med andra system (se t.ex. Sågvall Hein 1998; Arppe 1999) eftersom resultaten sällan redovisas på ett jämförbart sätt, men vår uppskattning är att vi uppnått jämförelsevis hög prestanda.

Vi är fortfarande i arbete med att optimera prestanda och att förbättra reglerna i programmet. Preliminära tester med granskningsreglerna visar att vi kan hoppas på en täckning på över 50 % och en precision på över 90 % för kongruensfel och felaktiga särskrivningar. Hittills är resultaten mycket lovande (jfr t.ex. Birn 1999), men trots allt är det orealistiskt att tro att systemet kommer att nå full täckning och precision.

Diskussion

Föregående avsnitt har visat hur man utifrån kunskap om vissa feltypers realisationer kan definiera formella regler som kan användas för att detektera en del felaktigheter med datorns hjälp. Problemet är att en hel del fel också undgår upptäckt eller, än värre, presenteras som fel fastän de inte är det. Tyvärr är

denna språkteknologiska begränsning inte något som kommer att kunna botas i första taget. Tekniken och den underliggande språkbeskrivningen blir visserligen hela tiden bättre (se bara på svensk stavningskontroll som på senare år uppnått stora förbättringar) men vissa grundläggande problem är svåra att komma ifrån, vilket oundvikligen resulterar i falska alarm och oupptäckta fel (se t.ex. Kohut & Gorman 1995).

Olika användare har olika behov. Vissa vill framför allt effektivisera sitt skrivande, andra vill förbättra sitt skrivande. De användare som kanske har störst behov av hjälp med språket, till exempel andraspråksinlärare, har också sämst förutsättningar att avgöra när programmet ger missledande kommentarer (se t.ex. Öhrman 2000). Ingen vet egentligen vilka effekter en grammatikkontroll får på olika gruppers skrivande. Här behövs mer forskning som grund för bättre användaranpassade program. Därför vill vi utveckla ett lättanvänt och instruktivt gränssnitt för Granska och i praktiken testa programmet på användare för att undersöka vilka effekterna blir på språk och skrivförmåga (se Larsson 1998, Domeij 1998). Ett grafiskt gränssnitt för PC beräknas vara klart under våren 2000. Inledande användartester kommer att slutföras före årets slut.

Referenser

- Arppe, A. 2000: Developing a grammar checker for Swedish. I: T. Nordgård, ed. *Nodalida'99*, Trondheim, december 1999.
- Birn, J. 2000: Detecting Grammar Errors with Lingsoft's Swedish Grammar Checker. I: T. Nordgård, ed. *Nodalida'99*, Trondheim, december 1999.
- Carlberger, J. & Kann, V. 1999: Implementing an efficient part-of-speech tagger. I: *Software - Practice and Experience*, 29 (9), pp. 815-832.
- Domeij, R. 1997: Datorn och språkkriktigheten. I: O. Josephson, ed., *Svenskan och IT-samhället*. Hallgren & Fallgren, 1997.
- Domeij, R. 1998: Detecting, presenting and correcting low-level problems when editing with and without computer aids. I: *TEXT Technology*, vol 8, no. 1. Wright State University, Celina, USA.
- Domeij, R., Eklundh, K., Knutsson, O., Larsson, S. & Rex, Å. 1998: *Granskaprojektet 1996-1997*. Technical Report NADA, KTH.
- Domeij, R., Knutsson, O., Carlberger, J. & Kann, V. 2000: Granska - ett effektivt hybridssystem för kontroll av svensk grammatik. I: T. Nordgård, ed. *Nodalida'99*, Trondheim, december 1999.
- Kann, V., Domeij, R., Hollman, J., & Tillenius, M. 1999: Implementaion aspects and applications of a spelling correction algorithm. In: R. Koehler, L. Uhlirova, G. Wimmer, eds, *Text as a Linguistic Paradigm: Levels, Constituents, Constructs. Festschrift in honour of Ludek Hrebicek*.
- Knutsson O. 1996: *Agreement Error Detection*. Master's Thesis, Department of Linguistics, Stockholm University.
- Kohut & Gorman 1995: The effectiveness of leading grammar/style software in analysing business students' writing. I: *Journal of Technical Writing and Communication* pp. 341-361. July 1995.

- Kukich, K. 1992: Techniques for automatically correcting words in text. *ACM Computing Surveys*, Vol. 24, No. 4, pp. 377-439.
- Larsson, S. 1998: Interaktivitet och användbarhet vid datorstödd språkgranskning och redigering i en integrerad skrivmiljö. Exjobbssrapport TRITA-NA-E9833 (IPLab-150), Nada.
- Severinson Eklundh, K. 1995: Skrivmönster med ordbehandlare. I: *Språkvård*, nr 4, sid 11–18.
- Sågvall Hein, A. 1998: *A Chart-Based Framework for Grammar Checking. Initial studies*. Nodalida'98, Copenhagen, january 1998.
- Öhrman, L. 1998: *Felaktigt särskrivna sammansättningar*. C-uppsats i datorlingvistik, Institutionen för lingvistik, Stockholms Universitet.
- Öhrman, L. 2000: *Datorstödd språkgranskning och andraspråksinlärare*. D-uppsats i datorlingvistik, Institutionen för lingvistik, Stockholms Universitet.