Svenska-engelska tur och retur med Multra

Eva Forsbom och Stina Karlsson
{evafo, stinaka}@stp.ling.uu.se
Uppsala universitet

Maskinöversättning
Språkteknologiprogrammet
Lärare: Anna Sågvall Hein

HT 2000

Sammanfattning

I rapporten beskriver vi en inledande undersökning av möjligheten att återanvända lexikon och språkbeskrivning vid vändning av ett transferbaserat maskinöversättningssystem för översättning från svenska till engelska, Multra. Vi har använt en kort svensk text och en översättning av den i det befintliga systemet som studiematerial. För de ord som används i texten har vi vänt på lexikoningångarna i analys-, transfer- och genereringslexikonen, och därefter återanvänt eller anpassat den övriga språkbeskrivningen. Det visade sig att det gick lätt att vända lexikonen och att det skulle kunna göras till största delen automatiskt och att det var ganska enkelt att anpassa transferbeskrivningen. Språkbeskrivningarna i analys- och genereringssteget går i viss mån att återanvända, men i vilken grad och vad som behöver anpassas är ännu en öppen fråga.

1 Inledning

1.1 Uppgift

Denna uppgift är en del av kursen Fördjupningskurs i maskinöversättning och MATS-projektet[1]. Vår del gick ut på att se hur långt det går att komma med en naiv ansats till omvandling av den svensk-engelska versionen av maskinöversättningssystemet Multra till en engelsk-svensk version. Syftet har varit att undersöka i hur stor omfattning det går att utnyttja det befintliga materialet i den ursprungliga versionen, vad man behöver lägga till, ändra eller ta bort samt att uppskatta hur lång tid det skulle ta.

Eftersom vi hade begränsad tid till vårt förfogande fick vi hålla oss till en kort text. Vi har använt filen `scania.txt` (se bilaga A) som användes i en demoversion av Multra. Filen är på svenska, och vi har översätt den i MTA[2] till engelska (bilaga B) och använt resultatet som indata, med vissa modifieringar. Tidsbristen gjorde också att vi fick lov att koncentrera våra ingrepp på lexikal nivå och på nominalfrasnivå.

1.2 Bakgrund

Multra är ett transferbaserat maskinöversättningssystem, vilket innebär att det är uppdelat i tre komponenter: analys, transfer och generering. Det språk som används som källspråk vid översättningen beskrivs och analyseras i analyskomponenten, och det språk som är målspråk beskrivs i genereringskomponenten. För varje språkpar finns sedan översättningsförhållandet angivet i en beskrivning i transferkomponenten. Översättningen görs på meningsnivå och de enskilda orden överförs på lexemnivå. Multra är utvecklat för översättning från svenska till engelska och tyska.

Analysen görs med Uppsala Chart Processor (UCP), som är en chartparser skriven i det procedurella programmeringsspråket LISP. Det finns också en särskild UCP-formalism för språkbeskrivningar (närmare beskriven i Sågvall Hein 1983 och 1984, Ahrenberg 1984 och Dahllöf 1989). Man använder den för både fonologisk, morfologisk, lexikal och syntaktisk nivå, vilket är lite ovanligt. I och med det kan språkbeskrivningen innehålla ett flertal olika lexikon: t.ex. för tecken, affix, stammar, lemman och lexem. Formalismen innebär att beskrivningen blir en procedurell och deklarativ blandning som inte är helt lättigenomtränglig för noviser som oss. UCP har dock den fördelen att den arbetar både datadrivet och hypotesdrivet, vilket gör att den klarar mer än en rent deklarativ parser gör ännu. Resultatet av en meningsanalys är en hierarkisk beskrivning i form av en attribut-värdestruktur. I följande exempel visas analysen för meningen `Ta bort muttern.`:

[1] MATS står för Methodology and Application of a Translation System. Det är en förstudie om huruvida det går att använda Multra som bas för utveckling av ett kommersiellt maskinöversättningssystem. Webbadress: http://stp.ling.uu.se/mats.

[2] Multra Translator's Assistent, det gränssnitt till Multra som en översättare använder.

```
[PHR.CAT : CL
 MODE : IMP
 PRED : [OBJ.DIR : [NP.FEAT : #1[GENDER : UTR
                                 NUMB : SING
                                 DEF : DEF
                                 CASE : BASIC]
                    PHR.CAT : NP
                    DF : [NG.FEAT : #1
                          HEAD : [LEX : MUTTER.NN.1
                                  WORD.CAT : NOUN]
                          PHR.CAT : NG]]
         VERB : [LEX : TA.VB+BORT.PL.1
                 VERB.FEAT : [VSURF : +
                              INFF : IMP]]
         SUBJ : 2ND]
 SEP : [WORD.CAT : SEP
        LEX : STOP.SR.0]]
```

Transfern får denna struktur som indata och för över den till en struktur i samma format som motsvarar meningen på målspråket genom att följa sökvägen från den översta nivån och kontrollera de transferregler som gäller för det aktuella attributet, t.ex. `PRED`. Transferkomponenten är en unifieringsbaserad komponent, skriven i Prolog (närmare beskriven i Beskow 1993).

Reglerna för översättningsförhållandet är skrivna i ett lättöverskådligt, deklarativt, format, som i princip är reversibelt, även om nyckelorden `SOURCE` och `TARGET` anger att översättningen har en riktning. I regeln nedan anges t.ex. att källspråkets `PRED`-attribut har ett underattribut, `OBJ.DIR`, att målspråket också har det och att de står i översättningsrelation till varandra. Transfern fortsätter på sökvägen, ned till lexemnivå och dess attribut.

```
LABEL
  OBJ.DIR
SOURCE
  <* OBJ.DIR> = ?OBJ.DIR1
TARGET
  <* OBJ.DIR> = ?OBJ.DIR2
TRANSFER
  ?OBJ.DIR1 <=> ?OBJ.DIR2
```

Resultatet av transfern förs över till genereringskomponenten att arbeta vidare med, och i vårt exempel resulterar det i meningen `Remove nut`. Genereringskomponenten är också skriven i Prolog och fungerar med samma unifieringsbaserade metod som transferkomponenten, men de syntaktiska reglerna skrivs som omskrivningsregler som arbetar på attribut i sökvägen som måste uppfylla vissa restriktioner.

```
[SEP : [WORD.CAT : SEP
        LEX : STOP.SR.0]
 MODE : IMP
 PRED : [OBJ.DIR : [DF : [HEAD : [LEX : NUT.NN.0
                                  WORD.CAT : NOUN]
                          NG.FEAT : [NUMB : SING
                                     DEF : DEF]
                          PHR.CAT : NG]
                    NP.FEAT : [NUMB : SING
                               DEF : DEF]
                    PHR.CAT : NP]
         SUBJ : 2ND
         VERB : [LEX : REMOVE.VB.0
                 WORD.CAT : VERB
                 VERB.FEAT : [VSURF : +
                              INFF : IMP]]]
 PHR.CAT : CL]
```

Lexikoningångarna skrivs i ett eget format (närmare beskrivet i Beskow 1997a). Den morfologiska genereringen (se Beskow 1997b) sköts i en egen komponent med särskilda regelformat, som påminner om vanliga Prolog-predikat.

1.3 Metod

Vi har genomfört undersökningen i tre delar som motsvarar komponenterna i Multra:

- Analysen: Vi har studerat vilka lexikoningångar och regler i den befintliga svenska grammatiken som går att återanvända som de är, vilka som går att justera samt vilka som behöver bytas ut och mot vad. Dessutom ville vi se i hur stor utsträckning transferlexikonet och det engelska genereringslexikonet går att använda för analyslexikonet.
- Transfern: Vi har kontrollerat om de nuvarande transferreglerna och -lexikonet är riktningsberoende eller om det går att vända på dem.
- Genereringen: Vi har precis som för analysen undersökt om det går att omvandla transferlexikonet och det svenska analyslexikonet till svenskt genereringslexikon. Eftersom formaten i genererings- och analysstegen på syntaktisk nivå skiljer sig så pass mycket åt har vi inte använt oss av de svenska analysreglerna för de svenska genereringsreglerna, utan i den mån vi hunnit med ändrat i de engelska genereringsreglerna istället.

Vi skapade en egen ren MDT-miljö[3] (`mdtense`) där vi använder bara våra grammatikbeskrivningar (Filerna finns lokalt på ~evafo/MT/ense). Sedan samlade vi alla lexikoningångar som var aktuella för vår text på ett ställe, så att vi kunde se om och hur man kunde modifiera dem automatiskt. Därefter vände vi på transferlexikonet, bytte utifrån detta ut lexikoningångarna i genererings- respektive analyslexikonet samt "utväxlade" den morfologiska informationen mellan analys- och genereringsdelen. När vi kommit så långt arbetade vi parallellt med testning av fraser och meningar av olika svårighetsgrad genom systemet

[3]Multra Developer's Tool, det gränssnitt till Multra som utvecklare använder. Vi fick hjälp med detta (och mycket mer) av Per Starbäck.

och ändringar av analys- och genereringsregler för att se hur mycket vi lyckades klara av.

1.4 Hypoteser

Innan vi började vår undersökning hade vi följande hypoteser:

- Analysen: Vi trodde att det går att utnyttja en del av reglerna eftersom de båda språken har relativt lika struktur och är närbesläktade, men att en del skulle behöva ändras. Vi visste t.ex. redan att vi behövde ändra från omvänd till rak ordföljd i huvudsatser, stryka kongruensreglerna för nominalfraser och lägga till kongruens mellan person och verb samt lägga till regler för `do`-omskrivning. Däremot trodde vi inte att det behövdes några stora förändringar för lexikonet, förutom att överföra det till UCP-formatet samt att göra något med morfologin.

- Transfern: Vi antog att det går att återanvända många transferregler. Överföringsregler skulle det bara gå att föra över och transferrelationsregler skulle det bara gå att föra över (med vissa tillägg och strykningar), medan elimineringsregler måste elimineras. Även för de lexikala transferreglerna hade vi stora förhoppningar om att bara kunna vända på dem, dvs. att byta plats på informationen under `SOURCE` och `TARGET`.

- Genereringen: Vi trodde att det till viss mån går att utnyttja analyslexikonet med vissa modifieringar. Det borde inte vara helt omöjligt att överföra eller lägga till morfologisk information. För reglerna hade vi inga större förhoppningar.

2 Genomförande

Vår undersökning påverkades naturligt nog av den svenska text vi använde som grunddata. Den är ett utdrag ur en servicemanual för lastbilar. Dess ordklassfördelning är ganska typisk för teknisk text, med en övervikt för nomen, även om det är osedvanligt mycket nomen i just denna text i och med den inledande tabelluppställningen. Ordklassfördelningen[4] visas i tabell 1. Den översatta engelska texten hade ungefär samma fördelning.

NN	VB	AB	PP	NL	SN	AL	AV	CN
62	13	10	6	4	3	2	1	1

Tabell 1: Ordklassfördelning för de 102 typorden i den svenska texten.

Eftersom det är en instruktionstext är meningarna till övervägande del i imperativ. Det var lämpligt i denna inledande studie, eftersom svenskan och engelskan har samma grundläggande ordföljd i meningar med imperativ.

[4] Texten hade totalt 323 token varav 65 skiljetecken och 72 sifferuttryck som var desamma i båda texterna. Om man inte räknar med dem återstår 186 token uppdelade på 102 typord. Texten hade 6 partikelverb, där partiklarna står både som partiklar och adverb i analyslexikonet. I tabellen är partiklarna upptagna under adverben. Förkortningarna står för nomen, verb, adverb, preposition, numeral, subjunktion, artikel, adjektiv och konjunktion.

2.1 Generering till analys

Beskrivningen av svenskan i UCP är inte bara tänkt att användas för analys av källtext vid översättning, utan den är utvecklad för allsidigt bruk, vilket gör den betydligt mer detaljerad och täckande än den engelska beskrivningen i genereringskomponenten. Det, tillsammans med det blandade procedurella och deklarativa formatet, gör också att den är svårare att sätta sig in i. Den version vi fick att arbeta med bestod t.ex. av ett flertal filer som beskrev olika syntaxnivåer, och ett flertal filer med olika lexikon, ibland med dubblettregler och olika formatering i reglerna som orsakade en del förvirring. Det var inte heller den senaste versionen av den morfologiska beskrivningen, även om den i princip är densamma, så våra resultat kanske inte alltid är helt rättvisande i detalj för den aktuella versionen.

2.1.1 Morfologi

Den morfologiska analysen i UCP sker till största del med hjälp av mönsterord för olika böjningsmönster som anropas från stamlexikonet med hjälp av regeln `pattern.<ord>`, som i exemplet nedan:

```
(define sve.dic-entry "kåp"
 #u <& gd lem>:=:'kåpa.nn,pattern.flicka;)
```

Mönsterordsregeln kan lägga till stammen i charten, kontrollera och unifiera vissa attributvärden och sedan anropa ytterligare regler med de värden som gäller för just det mönstret. I exemplet nedan anropas regeln `noun` med värdena för mönstret `flicka`, dvs. pluralsuffix, genus och sluttecken (som används för val av speciessuffix).

```
(define sve.gram-entry pattern.flicka
            #u assign.infl('flicka),
               assign.dicstem,
               (advance,
                <* char>='a,
                <& gd numb>:=:'sing//
                <& numb>:=:'plur),
               noun('-or,'utr,'vowel);
)
```

Suffixen är definierade i ett affixlexikon, som i exemplet nedan, där det senare slås upp av respektive regel.

```
(define numb-entry or  #u plursuff('-or);
)
```

För engelskan, som kräver mycket mindre morfologisk analys än svenskan, i synnerhet för substantiven, verkade det vara ett smidigt beskrivningssätt, där antalet mönsterord kan hållas nere rejält.

De oböjliga ordklasserna (PP, CN, SN, NL, AL, AB/PL[5]) fick varsitt mönsterord. Verben har vi inte hunnit undersöka, och de har än så länge bara fått ett gemensamt mönsterord för att analysen skulle fungera. Där finns det en hel del att göra, med engelskans flora av oregelbundna verb. Adjektiv har vi inte behandlat alls, eftersom vi bara hade ett i texten, men där behövs ingen

[5] Vi hade bara oböjliga adverb i texten, och inga ordningstal.

information om genus eller numerus, och inte i de regler som behandlar dem heller. Däremot har vi kunnat återanvända information om pluralböjning av substantiv från den engelska morfologiska genereringskomponenten. Där angavs fyra böjningsmönster: standardfallet (`-s`), ord som slutar på s-haltiga ljud (`-es`), ord som slutar på `y` (`-ies`) och ett ”oregelbundet” (`teeth`). (För varje annan oregelbunden pluralböjning får man i nästa steg skapa ett eget mönsterord, eller ange informationen direkt i lexikoningången, men det kommer antagligen inte att behövas särskilt många.) Utifrån den angivna informationen skapade vi fyra mönsterord där vi angav respektive pluralböjning, plus ett för ord som inte förekommer i plural och ett för ord som bara förekommer i plural. Dessutom lade vi till respektive pluralsuffix i affixlexikonet.

Eftersom engelska substantiv inte böjs efter genus men däremot påverkar verbböjning genom vilket värde för person det har så bytte vi attributen `GENDER` och `PERSON` rakt av (även i de regler där genus förekom på något sätt).

I den svenska beskrivningen anropar vanligen mönsterordsreglerna pluralregler som i sin tur anropar speciesregler som i sin tur anropar kasusregler. Egentligen borde vi ha strukit speciesreglerna och anropat kasusreglerna direkt från pluralreglerna, samt även strukit slutteckenargumentet i anropet till `noun`, men det får anstå till ett senare skede. Vi har heller inte fört in engelska genitivsuffix, eftersom vi bara hade `of`-genitiv i texten. Det borde dock vara ett ganska enkelt jobb.

Vändningen av den morfologiska beskrivningen går knappast att utföra automatiskt, eller ens halvautomatiskt. Däremot skulle man kunna bestämma sig för ett standardmönsterord per ordklass och ange det vid vändningen av lexikonet (jfr avsnitt 2.1.2). För substantiv skulle man även kunna välja mellan tre standardmönsterord utifrån slutljud.

2.1.2 Lexikon

När vi skulle byta ut orden i analyslexikonet utgick vi från transferlexikonet och kunde sedan slå upp ordet i genereringslexikonet och hämta ytterligare information om så behövdes. I analysdelen finns det tre lexikon på ordnivå: för stam, lemma och lexem. Så för varje ord i transferlexikonet måste vi byta ut ordet på tre ställen. Formatet på lexemen i transfer- och analyslexikonet var inte exakt lika, vilket gör att man alltid måste göra en enkel konvertering (se avsnitt 2.2.1). Eftersom orden överfördes på lexemnivå i transferlexikonet verkade det lämpligast att börja med att söka upp lexemet i analyslexikonet och gå vidare därifrån. Ingångarna i de tre analyslexikonen hade i huvudsak följande format:

```
(define sve.dic-entry "och" #u <& gd lem>:=:'och.cn, pattern.och;)
(define lemma och.cn (lex och.cn.1))
(define lexeme och.cn.1)
```

Vi började med att byta ut lexemet i lexemingången. Därefter kunde vi genom att söka på lexemformen hitta lemmat, byta ut det och till sist använda lemmat till att söka upp stammen och byta ut den. Under vår studie kom vi fram till lämpliga mönsterord allteftersom, men om det ska genomföras med automatik kan man som föreslagits i avsnitt 2.1.1 använda standardmönsterord och utnyttja den information som är tillgänglig från genereringsdelen. I genereringslexikonet finns även information om huruvida ordet börjar med konsonantiskt eller vokaliskt ljud. Vid genereringen behövs det för val av variantform för obestämd artikel, `a` respektive `an`. Den informationen är i och för sig inte strikt nödvändig vid analys, men det vore bra om man kunde få med den så att den går att använda i en eventuell grammatikkontrollkomponent.

För vissa lexem- och lemmaingångar fanns mer information än i exemplet ovan. Verb hade t.ex. regeln `verbaction` och nomen kunde ha regeln `nounaction` som båda angav vilka komplement som hörde till ordet. Reglerna gäller på högre syntaktiska nivåer och kan inte anges automatiskt på denna nivå, utan det måste gås igenom för hand. På lexikal nivå räcker det sedan med att eventuellt byta, ta bort eller lägga till en regel. För adverb kunde det t.ex. finnas information om att det rörde sig om ett sätts-, rums- eller tidsadverbial. Sådan information behöver antagligen inte ändras annat än i undantagsfall.

Artiklarna saknade helt lemma- och lexemingång i den svenska beskrivningen. De sköts i stället i mönsterordsregeln. Eftersom det bara rör sig om två ord är det enklast att sköta det för hand. Det fungerade för våra behov alldeles utmärkt att bara sätta in de engelska artiklarna (och göra en till smärre ändring).

Engelska partikelverb måste också skötas manuellt. I genereringslexikonet är de angivna som ett flerordsuttryck vilket som helst. Verbböjningen sköts i den morfologiska komponenten direkt. Vid bytet i analyslexikonet fungerar det bra ända fram till stamlexikoningången, där flerordsuttrycket måste delas upp eftersom verbet går att böja. Det fungerar fortfarande för verbet, men däremot är det inte säkert att det finns en översättningsrelation för partikeln. Motsvarande gäller för andra flerordsuttryck med inre böjning, t.ex. `have to`. Information om alternerande stam saknas också för oregelbundna ord som `be`. Den finns visserligen i den morfologiska genereringskomponenten, men är knappast automatiskt utdragbar.

Ingångarna i stamlexikonet hänvisade ibland till två eller flera lemman när den tekniska stammen utgjorde en homograf. Det hände också att flera engelska lemman fick samsas om en stam där det inte var så för de svenska motsvarigheterna.

```
(define sve.dic-entry sätt
  #u (<& gd lem> :=: 'sätta.vb, pattern.sätta
     //<& gd lem> :=: 'sätt.nn, pattern.bord);)
```

Ett sätt att lösa båda problemen skulle kunna vara att förbearbeta den svenska filen genom att dela upp sådana ingångar i två (eller flera) så att varje stamingång bara hänvisar till ett lemma. När alla engelska ord är insatta kan man sedan sortera stammarna och sammanfoga de ingångar som har gemensam stam.

```
(define sve.dic-entry sätt
  #u <& gd lem> :=: 'sätta.vb, pattern.sätta;)
(define sve.dic-entry sätt
  #u <& gd lem> :=: 'sätt.nn, pattern.bord;)
```

Eftersom lexikonen var uppdelade på flera filer, bl.a. ett kärnvokabulär och domänspecifika lexikon, med lite olika sorteringsvarianter, skulle man även behöva förbearbeta dem innan det skulle gå att byta ut lexikonet automatiskt. För att inte förlora informationen om vilka ord som tillhör en viss domän skulle man kunna ha ett attribut för detta i lexemingången som sedan filtrerades bort i analysen. Då skulle man kunna samla alla ingångar i samma fil och sedan bryta ut dem när bytet var klart[6].

Vid vändningen av transferlexikonet fick en del engelska ord flera svenska översättningar (se avsnitt 2.2), vilket gjorde att vi för undersökningen bara kunde välja ett alternativ. Det ledde i sin tur att vi inte kunde analysera alla meningarna i testtexten. Men det borde inte utgöra något problem för bytet i analyslexikonet eftersom det måste lösas redan vid vändningen av transferlexikonet.

2.1.3 Syntax

Vi hann tyvärr inte gå in så mycket på de mer syntaktiska reglerna. Det finns flera strukturer som behöver åtgärdas, t.ex. att engelskan har rak ordföljd i huvudsatser. Eftersom analysen i stort bygger på Diderichsens fältschema för svenska satser (se exempelvis Holm & Larsson 1980) där konstituenterna inte beskrivs hierarkiskt på samma sätt som i vanliga frasstrukturgrammatiker så kunde det kanske gå att byta plats på subjektsfältet och fältet för det finita verbet så att fältordningen blir som i tabell 2.

FUND	(SUBJ)	PRED1	(ADV1)	(PRED2)	(OBJ/PF)	(ADV2)
The boy	+	will	maybe	be	a big boy	soon

Tabell 2: Möjligt fältschema för engelskan.

Det borde fungera med att bara byta plats på motsvarande regler i följande regel för deklarativa satser:

```
(define sve.gram-entry cl.decl
 #u      <& gd phr.cat> :=: 'cl,
         not <* sent.frag>,
         fundament,
         <& gd type> :=: 'main,
         advance(v.fin),
         advance(subj),
         adverbials1,
         do(<& reg v1.lem :verbcomp>),
         do(<& gd pred verb lex :verbaction>),
         adverbials2,
         adv.in.pva,
         pva.used,
         end.of.sent,
         store*;
  #! phr.cat;
)
```

Annars har vi, som nämnts, bytt ut attributet `GENDER` mot `PERSON` i alla regler där det förekom, oavsett om det ska vara där eller inte, bara för att analysen skulle fungera. Vi lyckades inte analysera obestämda nominalfraser (och det fungerade inte med motsvarande svenska uttryck i den ursprungliga miljön heller), och därigenom gick det inte att klara transfern eller genereringen.

[6] Man skulle även kunna använda andra terminologihanteringsattribut för lexemen, t.ex. för projekt, källa, definition och kommentar, som går att använda för filtrering vid översättning. Det förutsätter förstås att det går att välja lexikon i MTA.

Om det beror på våra ändringar eller på det ursprungliga systemet har vi inte haft möjlighet att kontrollera under den tid vi hade till förfogande.

Vi har också bytt ut eventuella svenska lexem som förekom i de komplementregler som gällde orden från texten mot motsvarande engelska lexem (`noun-action`, `verbaction`, `verbcomp`). Det är förmodligen inte alls tillräckligt, utan flera måste nog skrivas om helt. Systemet måste också klara konstruktioner som `of`-genitiv och `do`-omskrivning.

2.2 Vändning av transfer

Vid en vändning av Multra torde inte transfersteget vara något problem. De flesta regler är lätta att vända och i den mån något måste åtgärdas för hand är formalismen lätt att sätta sig in i.

2.2.1 Lexikon

Transfersteget använder sig av ett lexikon där varje engelskt ord med sin svenska översättning har en egen lexikoningång som ser ut som regeln nedan.

```
LABEL
  BOLT
SOURCE
  <* LEX> = BOLT.NN.0
  <* WORD.CAT> = NOUN
TARGET
  <* LEX> = SKRUV.NN.1
  <* WORD.CAT> = NOUN
TRANSFER
```

Programmet söker upp det engelska ordet och går vidare med den svenska motsvarigheten.

Detta lexikon gick relativt lätt att vända på. Vändningen borde dessutom kunna hanteras automatiskt eftersom den i många fall endast innebär ett byte av lexemen under `SOURCE` respektive `TARGET`. Vissa modifieringar måste dock göras p.g.a. programmets struktur. Lexikoningångar där orden är särskrivna anges t.ex. med understreck i `TARGET`, men måste ha punkter i `SOURCE`:

```
LABEL
  DREV
SOURCE
  <* LEX> = DREV.NN.3
  <* WORD.CAT> = NOUN
TARGET
  <* LEX> =  DIVE_GEAR.NN.0
  <* WORD.CAT> = NOUN
TRANSFER

LABEL
  DIVE.GEAR
SOURCE
  <* LEX> =  DIVE.GEAR.NN.0
  <* WORD.CAT> = NOUN
TARGET
  <* LEX> = DREV.NN.3
```

```
  <* WORD.CAT> = NOUN
TRANSFER
```

Något mer komplicerade fall var verb och prepositioner, där endast `SOURCE`-formen var försedd med information om t.ex. ordklass. Var denna knuten till just lexemet i fråga eller hörde det mer ihop med `SOURCE`? I detta fall valde vi att låta informationen stå kvar under `SOURCE` tills vidare. Konjunktionerna var än mer komplicerade, då information om ordklass där förekom till synes sporadiskt både i `SOURCE` och `TARGET`. Där valde vi att lägga in informationen på båda ställena i samtliga ingångar. Beträffande dessa lexikoningångar måste man bestämma sig för ett slags mall, med de olika särdrag som behövs, innan man kan vända på lexikonet maskinellt.

Vissa lexikoningångar var särskilt problematiska. Det gällde främst flerordsuttryck, som t.ex. `på plats`, där det blev knepigt att lista ut vilka olika särdrag som behövdes i `SOURCE` respektive `TARGET`. Partikelverben, som till en början såg problematiska ut, hanterar vi på så vis att man i transferlexikonet helt enkelt slår ihop verbet och partikeln till ett enda lexem och sedan endast definierar detta lexem i genereringslexikonet. Regeln nedan är en sådan regel.

```
LABEL
  SET.UP
SOURCE
  <* LEX> = SET.VB+UP.PL.0
TARGET
  <* LEX> = SÄTTA_UPP.VB.0
  <* WORD.CAT> = VERB
TRANSFER
```

För vissa engelska ord finns det två eller flera svenska översättningar. Dessa ingångar har vi lämnat utan åtgärd tills vidare. Detta problem borde dock enkelt kunna rättas till genom att man anger vilket lexemnummer de olika orden har, t.ex. `cover.NN.1`, `cover.NN.2` respektive `cover.NN.3` (med översättningarna `lock/lucka/kåpa`), på samma sätt som i den svenska versionen.

Vid en vändning av Multra är det lämpligt att börja med transferlexikonet. Från detta kan nämligen information hämtas till både analys- och genereringssteget. Om transferlexikonet då redan vänts och bearbetats så att dubbletter försetts med olika lexemnummer är det korrekt information som sprids vidare till de övriga delarna och dessa behöver då inte gås igenom separat.

Sammantaget kan sägas att transferlexikonet är en av de delar i Multra som lättast låter sig ändras för att användas från engelska till svenska.

2.2.2 Syntax

Precis som vi antog är det möjligt att återanvända en hel del av transferreglerna.

Många regler är av typen

```
LABEL
 X
SOURCE
 <* X> = ?X
TARGET
 <* X> = ?X
TRANSFER
```

och kan stå kvar precis som de är, eftersom de är helt riktningsoberoende. Detta gäller rena överföringsregler som för över t.ex. numerus från källspråket till målspråket.

I vissa regler står två element från `SOURCE` respektive `TARGET` i transferrelation till varandra och ska sålunda inte översättas rakt av. Själva reglerna borde dock kunna återanvändas precis som de är, eftersom reglerna bara innehåller variabler och ska föra dem vidare oavsett vilket språk som kommer in först.

Så här långt bör det inte vara några som helst problem att automatiskt gå igenom reglerna och vända dem. När man kommer till de regler som behandlar t.ex. särdrag hos nominalfraser blir det dock genast mer komplicerat. Dessa regler bör hanteras för hand, då det gäller att avgöra vilka särdrag som är relevanta för det aktuella språket, vilka som bör läggas till och vilka som måste strykas. Det borde dock gå relativt lätt att vända även dessa regler.

Då vi inte hunnit testa längre meningar kan vi inte säga något om de regler som hanterar strukturer på en högre nivå, men vi kan inte se att de skulle medföra några stora problem.

2.3 Analys till generering

Vid en vändning av Multra kommer det svenska språket inte längre att hanteras av analysen, utan istället bli föremål för genereringssteget. Eftersom analys och generering sköts av två olika system, UCP respektive Multra, finns det mycket stora skillnader i deras struktur. Vår teori var att det inte skulle vara helt enkelt att göra om genereringen så att den passade för svenska istället för engelska, men att man ändå skulle kunna utnyttja den information som fanns i den svenska analysen.

Ju närmare lexikal nivå man kommer, desto lättare är det att modifiera de olika reglerna. Vi har dock inte gett oss på några av reglerna som hanterar de högre nivåerna. Vi kan föreställa oss att dessa måste skrivas om helt för att passa svenskans struktur.

2.3.1 Morfologi

Vissa delar av de engelska genereringsreglerna går att återanvända vid skapandet av de svenska. Så är t.ex. fallet med de lexikala insättningsreglerna. För substantiven räcker det med att utöka regeln lite, så att fler särdrag skickas med till den morfologiska delen. De särdrag vi lagt till är `PATTERN` (mönsterord) och `DEF` (species). Även kasus bör läggas till.

Eftersom engelska och svenska är så pass olika morfologiskt sett går det inte att använda samma typ av uppläggning för dem båda. Den engelska böjningen för substantiv består t.ex. i stort sett bara av ändelserna `-s` och `-es`. I detta fall kan den morfologiska komponenten bestå av ett fåtal regler som kontrollerar vilken bokstav stammen slutar på. Ord som slutar på vokal eller ett fåtal andra bokstäver tilldelas ändelsen `-es`, övriga får `-s`.

För svenskans del blir det hela mer komplicerat, eftersom orden böjs på ett betydligt mer omfattande sätt. Ett system med mönsterord blir nödvändigt. Detta får byggas upp från grunden, men man kan få hjälp med innehållet från analysdelens mönsterord och deras ändelser. Vi har valt att ha fyra (vore fler om kasus lagts in) predikat för varje mönsterord, som anropas från ett gemensamt predikat. Varje enskilt predikat kontrollerar vilka särdrag ordet har och lägger till ändelser som matchar dessa:

(anropspredikat:)

```
noun(NUMB, PATTERN, DEF) => X becomes Y+Z when
pattern(PATTERN, NUMB, DEF,X,Y,Z).
```

(mönsterord:)

```
pattern(stol, 'SING', 'INDEF',X, Y, Z):-
Y=X, Z="".
pattern(stol, 'SING', 'DEF', X, Y, Z):-
Y=X, Z="en".
pattern(stol, 'PLUR', 'INDEF', X, Y, Z):-
Y=X, Z="ar".
pattern(stol, 'PLUR', 'DEF', X, Y, Z):-
Y=X, Z="arna".
```

Vissa ord måste även modifieras. Detta gäller t.ex. ord med mönstret `nyckel`, vars 'e' faller bort i pluralformerna. Från predikatet sker då ett anrop till en funktion som helt enkelt plockar bort den näst sista bokstaven i ordet (samt tredje sista bokstaven i ord med mönsterordet `nummer`). För ord som `flicka`, `gosse` och `kypare` har vi gjort en lösning som bygger på att ordets stam är av typen `flick/goss/kypar`. Ordet `låsringstång` får omljud och har därför ett separat predikat.

Vårt program klarar av att generera samtliga nomenformer utom kasusformerna korrekt.

2.3.2 Lexikon

Vi hade tänkt oss att man skulle kunna skapa ett svenskt genereringslexikon med utgångspunkt i analyslexikonet. Vid en närmare titt på det hela kom vi dock fram till att det är mycket enklare att helt enkelt kopiera informationen under `TARGET` i transferlexikonet. Regeln

```
LABEL
  DIVE.GEAR
SOURCE
  <* LEX> =  DIVE.GEAR.NN.0
  <* WORD.CAT> = NOUN
TARGET
  <* LEX> = DREV.NN.3
  <* WORD.CAT> = NOUN
TRANSFER
```

ger följande lexikoningång i genereringslexikonet:

```
WORD
 DREV.NN.3 :
  <* WORD.CAT> = NOUN
  <* PLT> = -
  <* STEM> = drev
  <* PATTERN> = bord
```

Denna extraktion borde inte vara så svår att göra automatiskt. Det som krävs är en grundstruktur för varje typ av lexikoningång, i detta fall nomen.

Denna kan sedan fyllas på undan för undan. Först plockar man alltså ut det svenska lexemet från transferlexikonet.

I lexikoningången ska även en stam anges, vilken kan fås fram genom att ta lemmats namn och göra om det till små bokstäver. (Vi har valt att låta verbets stam stå i imperativ för att förenkla för oss i den morfologiska genereringen.)

Till skillnad från det engelska genereringslexikonet har vi även lagt in mönsterord i varje nomen-lexikoningång. Detta behövs vid den morfologiska genereringen, för att rätt ändelser ska knytas till de olika orden. Mönsterorden kan lätt hämtas från analysdelen. Vid vidare arbete med verben byts det sätt vi använt (med imperativformer) lämpligen ut mot användning av mönsterord.

Därefter är lexikoningången färdig!

2.3.3 Syntax

Vi har endast hunnit koncentrera oss på de lägre nivåerna, främst lexikonnivån. Under arbetet har vi kanske snuddat lite vid de mer övergripande reglerna. Vi har dock endast modifierat två av dessa (för att inte ett nomen i bestämd form singular ska få en artikel som i engelskan) och kan inte säga något om hur dessa regler kommer att bete sig när man vill översätta fler och längre meningar från engelska till svenska. Vi kan dock tänka oss att relativt mycket måste skrivas om för att passa svenskans struktur.

3 Resultat

Det resultat vi fått av våra undersökningar stämmer relativt bra överens med de hypoteser vi ställde upp innan vi började. Kanske var vi lite väl optimistiska då, men i stora drag har de ändringar vi fått göra inte avvikit från de vi förutsatte att vi skulle behöva göra. Nedan följer en punktvis redogörelse för vad resultatet blev av vårt arbete med de olika delarna.

3.1 Analysen

Vändningen av den morfologiska beskrivningen går knappast att utföra automatiskt, eller ens halvautomatiskt. Man skulle dock kunna bestämma sig för ett eller flera standardmönsterord per ordklass och ange det vid vändningen av lexikonet (jfr avsnitt 2.1.2). För nomen skulle man även kunna välja mellan tre standardmönsterord utifrån slutljud.

I tabell 3 visas automatiseringsgraden för att byta ut de olika ordklassernas respektive lexemingångar i analyslexikonet. För de flesta ingångarna går det att automatisera, enligt följande:

Kopiera: Det innebär att omformatera lexemet från transferlexikonet till UCP-formatet, söka upp lexemingången och ersätta den. Sedan måste det engelska lexemet och det svenska lemmat lagras i variabler till nästa steg: bytet av lemman, och vidare på liknande sätt. En del ingångar innehåller ytterligare information som kan behöva åtgärdas manuellt senare (se avsnitt 2.1.2), men den rör snarare syntaktiska och semantiska hänseenden.

Manuellt: Det innebär i detta fall att det saknades lexemingångar för artiklarna. De sköttes istället av en regel på högre nivå (se avsnitt 2.1.2).

	NN	VB	AB	PP	NL	SN	AL	CN
Kopiera	59	11	5	7	4	3		1
Manuellt							3	

Tabell 3: Enkelhet vid vändning för lexem- och lemmalexikon i analyssteget.

I tabell 4 visas automatiseringsgraden för att byta ut de olika ordklassernas respektive stamingångar i analyslexikonet. För de flesta ingångarna går det att automatisera, enligt följande:

Kopiera: Det innebär att omformatera lemmat till stam, genom att t.ex. stryka en slutvokal eller liknande som för `dolly`. För att det ska fungera måste man antagligen för- och efterbearbeta lexikonfilerna till en del, men det kan ske automatiskt (se avsnitt 2.1.2).

Lätt mall: Det innebär i detta fall att ett engelskt verb motsvarar ett svenskt partikelverb, t.ex. `remove` och `ta bort`.

Manuellt: Det innebär att det kanske saknas relevant information i transfer- och genereringslexikonen, t.ex. för engelska flerordsuttryck med inre böjning som `have to` eller ord med alternerande stammar.

	NN	VB	AB	PP	NL	SN	AL	CN
Kopiera	59	5	5	7	4	3		1
Lätt mall		3						
Manuellt		3					3	

Tabell 4: Enkelhet vid vändning för stamlexikonet i analyssteget.

Vi har inte hunnit göra något vettigt med de syntaktiska reglerna, men vi har fortfarande vissa förhoppningar om att det går att behålla åtminstone grundstrukturen i de flesta reglerna. Grundtanken är att det är mer som är likt än som skiljer.

3.2 Transfern

De transferregler som Multra använder sig av såg vid en första anblick ut att vara i stort sett reversibla. En närmare titt på dem förstärkte denna teori, med undantag för vissa regler som var mer eller mindre riktningsberoende. Trots att vi i stort sett endast bytte plats på informationen under `SOURCE` och `TARGET` i de olika reglerna, samt i vissa fall lät den riktningsberoende informationen stå kvar på sitt ordinarie ställe, fick vi ett bra resultat när vi testade att översätta olika ord och kortare fraser. Detta visar att det åtminstone på lexikal nivå är lätt att vända på översättningssystemet. I tabell 5 redovisas de olika typer av ingrepp som kan behövas för att vända på programmet, samt antalet regler för varje typ. De kortfattade rubrikerna för de olika processerna förklaras mer i detalj nedan.

Kopiera: Dessa regler behöver endast kopieras och kan alltså stå kvar precis som de är. Detta gäller t.ex regler som för över särdrag från källspråk till målspråk (se avsnitt 2.2.2).

Vända: Dessa regler kan enkelt vändas på så sätt att man byter plats på informationen under `SOURCE` och `TARGET`. Inga modifieringar behövs.

Lätt mall: Dessa regler behöver, utöver att vändas, modifieras något. Modifieringarna kan enkelt göras automatiskt enligt en lätt mall. Exempel på regler av det här slaget är de särskrivna ord som har understreck i `TARGET` och punkter i `SOURCE` (se avsnitt 2.2.1).

Svårare mall: Dessa regler skiljer sig från punkten ovan i det att den modifiering som sker av varje regel rör lite större saker än interpunktion. Istället kan det t.ex. handla om att verb och partikel slås ihop (se avsnitt 2.2.1).

Manuellt: Dessa regler är lite svårare och kan inte hanteras automatiskt utan måste åtgärdas manuellt. Regler i denna grupp är ofta flerordsuttryck som t.ex. `i universalstativ`.

	Kopiera	Vända	Lätt mall	Svårare mall	Manuellt
Lexikon	0	63	41	5	3
Regler	73	12	0	0	2

Tabell 5: Enkelhet vid vändning i transfersteget.

Som synes i tabellen är flertalet regler enkla eller till och med mycket enkla att vända på. Transfersteget skulle ta absolut kortast tid att omvandla till en engelsk - svensk version av Multra.

3.3 Genereringen

Vår teori var att det inte skulle vara helt enkelt att göra om genereringen så att den passade för svenska istället för engelska, men att man ändå skulle kunna utnyttja den information som fanns i den svenska analysen. Det visade sig att det i verkligheten var enklare att plocka den lexikala informationen från transferlexikonet. De mönsterord som behövs för den svenska morfologin hämtades dock från analysen. Utöver detta var det dock svårt att hitta något som kunde hjälpa oss med modifieringen av de engelska genereringsreglerna.

Hittills har vi endast lyckats översätta ett fåtal meningar och en ordlista, men åtminstone för denna begränsade mängd har vi fått ett bra resultat.

I tabell 6 visas lexikonreglernas fördelning på de operationer som måste utföras på reglerna för att de ska fungera för översättning från engelska till svenska. De kortfattade rubrikerna förklaras nedan.

Kopiera: Dessa regler kan kopieras och användas precis som de är. Detta gäller t.ex. regler för interpunktion.

Ta lexem: Dessa regler kan fås fram relativt enkelt genom att man endast plockar ett lexem från transferlexikonet och gör om det till små bokstäver för att få fram stammen. Inga övriga modifieringar eller tillägg behövs.

Ändra enl.mall: För dessa regler gäller att man utöver att ta fram stammen på samma sätt som i föregående stycke även modifierar regeln på ett eller annat sätt. Det kan t.ex. röra sig om tillägg av mönsterord som hämtas från analyssteget.

	Kopiera	Ta lexem	Ändra enl. mall
Lexikon	5	26	82

Tabell 6: Enkelhet vid vändning i genereringssteget.

Som tabellen visar innebär genereringslexikonet inga större problem för en vändning av Multra. Vi hittade inga regler alls som behöver hanteras manuellt. Det som krävs är alltså att man skriver ett program som skapar lexikoningångar enligt en mall man definierar. Denna del av vändningen tar alltså i idealfallet inte längre tid än det tar att skriva programmet.

När man kommer in på morfologi och syntax blir det dock mer komplicerat. Här måste nog det mesta skötas manuellt. I morfologifilen måste samtliga mönsterord läggas in och samtliga syntaxregler måste synas för att se hur mycket som måste skrivas om från början. Vi kan därför inte redovisa någon tabell som visar hur lätt det skulle kunna vara att vända på just detta avsnitt. Det beror helt på hur insatta i programmet de personer som utför vändningen är, samt hur stor del av reglerna som går att återanvända.

3.4 Testkörningar

Vi har testkört hela den text vi använt som indatatext. En lista på de meningar/ord som klaras av finns i bilaga C. En längre testkörning av en mening (med halv spårning) samt ett ord finns i bilaga D

4 Nästa steg

Om man ska fortsätta på den inslagna linjen bör man lägga ner den mesta tiden på den engelska språkbeskrivningen som nu är betydligt mindre omfattande än den svenska. Vårt syfte har mest varit att studera hur mycket som går att återanvända, men man måste ju även ha i åtanke vad som kan göras för att underlätta en vändning i det pågående arbetet med att skala upp det svensk-engelska systemet.

4.1 Automatisering

Vändningen av stora delar av framför allt lexikonen i de olika delarna kan med största sannolikhet hanteras automatiskt. Under avsnitt 2 ovan har vi beskrivit hur detta skulle kunna göras för de olika avsnitten. Det som krävs är att man skriver program som utför de olika växlingarna och tilläggen/strykningarna. Som synes i tabellerna 5 och 6 tror vi att 194 av 199 (transfersteget) respektive 113 av 113 (genereringssteget) av våra regler kan vändas på automatisk väg.

4.2 Språkbeskrivning

Om man ska använda systemet med mönsterord även för engelskan bör en uppsättning sådana tas fram för de ordklasser som böjs. Den engelska beskrivningen måste också i annat bli mer finkornig. En mer detaljerad nivå av lemma- och framför allt lexembeteckningar behövs för att välja bland möjliga alternativa översättningar. Det kanske också kan vara en god idé att skapa attribut för terminologihantering, t.ex. för domän och projekt, och även göra så att det går att filtrera fram rätt ord på rätt plats. Det gäller naturligtvis även för svenskan.

Det krävs också en hel del arbete med den syntaktiska beskrivningen av engelskan i analysen.

Vad gäller transfern så vore det intressant att se om de förändringar vi gjort av reglerna räcker även för längre och svårare meningar. Vi kan i nuläget inte se hur man skulle kunna gå vidare med denna bit.

För genereringsdelen finns det mycket kvar att undersöka. Vi tycker att vi gjort en bra lösning för substantiven, men man skulle där även behöva beakta kasus för att kunna generera samtliga former. Nästa steg när det gäller morfologin vore att gå vidare och lägga in mönsterord även för verb och andra ordklasser. Om man sedan går vidare från den lexikala nivån återstår mycket jobb. Där gäller det främst att undersöka om reglerna alls går att modifiera eller om det helt enkelt är lättast att börja om från grunden.

5 Sammanfattning

Vi har i vårt projekt försökt se hur långt det går att komma med en naiv ansats till omvandling av den svensk-engelska versionen av Multra till en engelsk-svensk version. Syftet var att undersöka i hur stor omfattning det går att utnyttja det befintliga materialet i den ursprungliga versionen, vad man behöver lägga till, ändra eller ta bort.

Multra kan delas upp i tre delar: analys, transfer och generering. För var och en av dessa delar finns regler som hanterar lexikon och syntax och för analys och generering finns även morfologiska regler. Vi började med att studera den lexikala nivån. För samtliga delar gäller att lexikonet är relativt lätt att vända på, ibland med vissa modifieringar. Lexikonreglerna är också de som i störst utsträckning borde gå att vända på maskinellt.

De morfologiska reglerna innebär mer problematik, då svenskan och engelskan har olika strukturer. Det går inte heller att få någon större hjälp från analyssteget när man ska skapa genereringsregler och vice versa eftersom analys och generering använder sig av olika formalismer: UCP respektive Multra. De morfologiska reglerna måste därför skrivas om så att de passar strukturen hos det språk de ska hantera.

Vi har egentligen inte kommit så långt att vi på allvar gett oss i kast med syntaxen och kan därför inte säga något exakt om hur lång tid det skulle kunna ta att vända på Multra. Det är helt beroende av hur mycket som måste skrivas om från början och hur lång tid det tar att skriva program för/manuellt hantera lexikonen.

Referenser

[1] Lars Ahrenberg. De grammatiska beskrivningarna i SVE.UCP. I Sågvall Hein [8], s. 1–13.

[2] Björn Beskow. *Unification-Based Transfer in Machine Translation*, 1993. Reports from Uppsala University, Department of Linguistics, RUUL #24.

[3] Björn Beskow. *Generation in MULTRA*, 1997a. Uppsala University. Department of Linguistics.

[4] Björn Beskow. *Morphology in MULTRA*, 1997b. Uppsala University. Department of Linguistics.

[5] Mats Dahllöf. *En lexikonorienterad parser för svenska.* Magisteruppsats, Göteborgs universitet, 1989. (http://stp.ling.uu.se/~matsd/pub/ex.ps).

[6] Lars Holm och Kent Larsson. *Svenska meningar.* Ord och stil 7, Språkvårdssamfundets skrifter. Studentlitteratur, Lund, 2 uppl., 1980.

[7] Anna Sågvall Hein. *A Parser for Swedish. Status Report for SVE.UCP*, 1983. UCDL-R-83-1, Center for Computational Linguistics, Uppsala University.

[8] Anna Sågvall Hein, red. *Föredrag vid De nordiska datalingvistikdagarna 1983*, UCDL-R-84-1, Center for Computational Linguistics, Uppsala University, 1984.

[9] Anna Sågvall Hein. Regelaktivering i en parser för svenska (SVE.UCP). I Sågvall Hein 1984 [8], s. 187–199.

A Svensk grundtext

HUVUDVÄXELDEL

Isärtagning, växellådshus

1. Växellådshus
4. Cyl. pinne
8. Sidolock
9. Packning
10. Skruv
11. Brickfjäder
12. Oljesil
13. Gummislang
21. Lock
22. Tätningsring
23. O-ring
24. Skruv
25. Brickfjäder
26. V-ring
27. Bärhylsa
28. Skruv
29. Bricka
37. Växlingshus
39. Spiralspännstift
40. Packning
41. Skruv
42. Skruv
60. Packning

Specialverktyg

Fig Nummer Benämning
1 79 046 Dorn
2 87 172 Avdragare
3 87 792 Lyftverktyg
4 98 261 Fixtur
5 98 765 Spårhylsa
6 99 154 Stöddorn
-99 001 Adapter
-99 003 Hydraulcylinder
-99 004 Hydraulpump
-99 012 Spindel

Förmedlingsverktyg

587 511 Avdragare
587 312 Låsringstång

1. Ta bort växlingsförarhuset.

2. Ta bort kopplingsservomekanismen
   och kopplingshävarmen.

3. Ta bort luckan för kraftuttagshuset.

4. Sätt på lyftverktyget 87792 på
   växlingsförarhusets plats och
   fixtur 98261 på kraftuttagshuslockets plats.

5. Sätt upp växellådan i
   universalstativ.

6. Ta bort urkopplingslagret.

7. Ta bort växlingshuset.

8. Ta bort planetväxeldelen, se sidan
   40.

9. Ta bort bärhylsans fyra
   fästskruvar, ta bort bärhylsan och
   tätningsringen.

10. Ta bort de fyra skruvarna för
    locket och kopplingshävarmen. Ta
    bort kopplingshävarmen och de sju
    skruvarna för kopplingskåpan. Ta
    bort locket och kopplingskåpan
    samtidigt. Lägg kåpan på arbetsbänken
    med pumpaxeln vänd uppåt för att
    undvika skador.

11. Ta bort oljepumpen (3 skruvar med
    brickfjädrar). Oljepumpen är en
    rotorpump med fem kuggar på drevet
    och sex på rotorn.

12. Skilj locket från kopplingskåpan
    genom att försiktigt knacka ut
    locket framåt. Ta bort O-ringen
    och tätningsringarna i locket.

13. Ta bort muttern till solhjulet med
    spårhylsa 98 765.

    Obs! Om inte mutterdragare
    används måste två växlar läggas i!

B Engelsk grundtext

MAIN GEARBOX

Dismantling, gearbox housing

1. Gearbox housing
4. Dowel
8. Side cover
9. Gasket
10. Bolt
11. Cup washer
12. Oil strainer
13. Rubber hose
21. Cover
22. Seal
23. O-ring
24. Bolt
25. Cup washer
26. V-ring
27. Support sleeve
28. Bolt
29. Washer
37. Selector housing
39. Spiral spin
40. Gasket
41. Bolt
42. Bolt
60. Gasket

Special tool

Fig No Name
1 79 046 Drift
2 87 172 Puller
3 87 792 Lifting tool
4 98 261 Fixture
5 98 765 Socket
6 99 154 Dolly
-99 001 Adapter
-99 003 Hydraulic cylinder
-99 004 Hydraulic pump
-99 012 Spindle

Intermediary tool

587 511 Puller
587 312 Retaining ring plier

1. Remove gear selector housing.

2. Remove clutch servo mechanism
   and clutch lever.

3. Remove cover of the power take-off housing.

4. Attach lifting tool 87792 in position of gear
   selector housing and fixture 98261 in position
   of power take-off housing cover.

5. Set up gearbox on universal stand.

6. Remove declutching bearing.

7. Remove selector housing.

8. Remove planetary gear unit, see page 40.

9. Remove the four bolts of the support sleeve,
   remove support sleeve and remove seal.

10. Remove the four bolts for the cover and for
    the clutch lever. Remove clutch lever and the seven
    bolts for the clutch cover. Remove cover and clutch
    cover at the same time. Place cover on the work bench
    with the pump shaft turned up to avoid damages.

11. Remove oil pump (remove 3 bolts with cup washers).
    The oil pump is a rotor pump with five teeth on
    the dive gear and with six teeth on the rotor.

12. Separate cover from the clutch cover by gently tapping
    the cover to the front. Remove o-ring and seals
    in the cover.

13. Remove the nut of the sun wheel with socket 98 765.

    Note. If a nut runner is not used two gears have to be engaged.

C Vad programmet klarar av att översätta

Indata

```
Remove the gear selector housing.

Remove the clutch servo mechanism and the clutch lever.

Remove the clutch servo mechanism and the rubber hose.

Set up the gearbox on the universal stand.

Remove the declutching bearing.

Remove the selector housing.

Remove 4 nuts of the support sleeve, remove the support sleeve and remove the seal.

Remove the cover and the clutch cover at the same time.

Place the cover on the work bench.

Remove the oil pump (remove 3 bolts with cup washers).
```

Utdata

```
Ta bort växlingsförarhuset!

Ta bort kopplingsservomekanismen och kopplingshävarmen!

Ta bort kopplingsservomekanismen och gummislangen!

Sätt upp växellådan i universalstativet!

Ta bort urkopplingslagret!

Ta bort växlingshuset!

Ta bort 4 muttrar av bärhylsan, ta bort bärhylsan och ta bort tätningsringen!

Ta bort locket och kopplingskåpan samtidigt!

Lägg locket på arbetsbänken!

Ta bort oljepumpen (ta bort 3 skruvar med brickfjädrar).
```

D Testkörning

```
Testkörning ord: "the retaining ring pliers"
Resultat: Låsringstängerna

Parser:

[PHR.CAT : NP
 NP.FEAT : [DEF : DEF
            NUMB : PLUR
            PERSON : 3
            CASE : BASIC]
 DET : [LEX : THE.AL.0
        WORD.CAT : ART]
 DF : [NG.FEAT : [PERSON : 3
                  NUMB : PLUR
                  DEF : INDEF
                  CASE : BASIC]
       HEAD : [LEX : RETAINING.RING.PLIER.NN.0
               WORD.CAT : NOUN]
       PHR.CAT : NG]]


Transfern gav strukturen

[NP.FEAT : [DEF : DEF
            NUMB : PLUR
            CASE : BASIC]
 DF : [NG.FEAT : [DEF : DEF
                  NUMB : PLUR]
       HEAD : [LEX : LÅSRINGSTÅNG.NN.X
               WORD.CAT : NOUN]
       PHR.CAT : NG]
 PHR.CAT : NP]
```

Testkörning mening: "Place the bolt on the work bench." med halv spårning.

Resultat: Lägg skruven på arbetsbänken! (OBS utropstecknet!)

Parser:

```
[PHR.CAT : CL
 MODE : IMP
 PRED : [OBJ.DIR : [NP.FEAT : [PERSON : 3
                               NUMB : SING
                               CASE : BASIC
                               DEF : DEF]
                    PHR.CAT : NP
                    DET : [LEX : THE.AL.0
                           WORD.CAT : ART]
                    DF : [NG.FEAT : [PERSON : 3
                                     NUMB : SING
                                     DEF : INDEF
                                     CASE : BASIC]
                          HEAD : [LEX : BOLT.NN.0
                                  WORD.CAT : NOUN]
                          PHR.CAT : NG]]
         OBJ.LOC : [RECT : [NP.FEAT : [PERSON : 3
                                       NUMB : SING
                                       CASE : BASIC
                                       DEF : DEF]
                            PHR.CAT : NP
                            DET : [LEX : THE.AL.0
                                   WORD.CAT : ART]
                            DF : [NG.FEAT : [PERSON : 3
                                             NUMB : SING
                                             DEF : INDEF
                                             CASE : BASIC]
                                  HEAD : [LEX : WORK.BENCH.NN.0
                                          WORD.CAT : NOUN]
                                  PHR.CAT : NG]]
                    PREP : [LEX : ON.PP.0]
                    PHR.CAT : PP]
         VERB : [LEX : PLACE.VB.0
                 VERB.FEAT : [VSURF : +
                              INFF : IMP]]
         SUBJ : 2ND]
 SEP : [WORD.CAT : SEP
        LEX : STOP.SR.0]]
```

```
Transfer:

Applying Rule STOP
Applying Rule MODE
Applying Rule PRED
   Applying Rule OBJ.DIR
      Applying Rule NP.FEATMEDDET
         Applying Rule BOLT
      Applying Rule PHR.CAT
   Applying Rule OBJ.LOC
      Applying Rule PREP
         Applying Rule ON
      Applying Rule RECT
         Applying Rule NP.FEATMEDDET
            Applying Rule WORK.BENCH
         Applying Rule PHR.CAT
      Applying Rule PHR.CAT
   Applying Rule SUBJ
      Applying Rule IMP.SUBJ
   Applying Rule VERB
      Applying Rule PLACE
      Applying Rule VERB.FEAT
Applying Rule PHR.CAT
Success!


Generering:

Applying Rule CL
   Applying Rule PRED2a
      Applying Rule NG1
      Applying Rule PP
         Applying Rule NP.NG.DEF
            Applying Rule NG1
Success!
```