

# Mening, Kommunikation, Förklaring

## Projektbeskrivning

### 1. Inledning

Den analytiska språkfilosofins centrala del är meningsteorin, dvs studiet av fundamentala frågor kring språklig mening (därutöver finns bl.a. talaktsteorin). Inom meningsteorin kan man i sin tur urskilja tre huvudområden. Dels har vi studiet av semantiken för speciella typer av uttryck, som egennamn ('Winston Churchill'), demonstrativer ('den där'), psykologiska kontexter ('Eva tror att ...'), logiska partiklar ('icke'), kvantifikatorer ('varannan björk'), osv.

Dels har vi också mer generella semantiska teorier, eller teoretiska ramverk, för beskrivning av formella (som aritmetikens) eller naturliga språk (som svenska). Hit kan vi räkna system som Davidsonska sanningsteorier (Davidson, Larson och Segal), Montaguegrammatik (Montague), Situationssemantik (Barwise och Perry) och DRT, Diskursrepresentationsteori (Kamp). Dessa olika system är tekniska ramverk med sina egna uppsättningar semantiska begrepp, dvs begrepp som har med mening att göra.

För att åstadkomma en allmän teoretisk förståelse av språklig mening måste man emellertid försöka göra reda för de semantiska begreppen själva. Detta är ambitionen för teorier inom den tredje huvudgruppen, de mer allmänfilosofiska teorierna om semantiska egenskaper. Somliga av dessa teorier är konstruktiva, somliga är destruktiva. Till de senare hör Wittgensteins diskussion av regelföljande, avsedd att underminera våra föreställningar om objektivitet hos regler och begrepp, liksom Quines diskussioner av analyticitet och översättning, avsedda att misskreditera traditionella idéer om meningsegenskaper som objektivt tillkommer språkliga uttryck.

Bland de konstruktiva meningsteorierna kan man skilja ut teorier som låter språkliga uttryck erhålla sin mening från mer grundläggande mentala tillstånd eller akter, som att tro något eller vilja något. Vissa teorier här, som hos Grice och Lewis, är i första hand koncentrerade på att redogöra för de komplicerade avsikter eller övertygelser som behövs för att göra yttranden meningsfulla, medan andra teorier, som hos Harman och Fodor, syftar till att förklara på vilket sätt de mentala tillstånden själva, och även delar av dem ("begrepp"), erhåller sin mening.

Den andra huvudgruppen bland de konstruktiva teorierna, i första hand representerad av Davidson och Dummett, grundas på tanken att språklig mening bestäms genom språklig interaktion, eller närmare bestämt genom villkor som måste vara uppfyllda för att språklig kommunikation skall vara möjlig. En gemensam tanke här är att det måste vara möjligt att *veta* vad andra talare menar (Davidson), eller att *veta* huruvida andra talare förstår ett språkligt uttryck (Dummett). Det är framför inom denna tradition man håller redogörelse för, eller förklaring av, språklig kommunikation som ett syfte för meningsteorier.

Det är denna sistnämnda tradition jag själv står närmast. Min uppfattning är att semantiska teorier har en funktion att fylla i en *förklaring* av det faktum att språklig

kommunikation i så hög grad lyckas. Med normala, förteoretiska mått mätt är en överväldigande stor andel av språkliga transaktioner framgångsrika. Med tanke på hur många olika tankar vi kan uttrycka, och faktiskt uttrycker, är denna framgångsfrekvens anmärkningsvärd. De grundtankar som är nya i mitt eget synsätt är två: för det första att denna förklaringsidé är själva *utgångspunkten* för teorin om språklig mening och för det andra att normer för förklaringsvärde hämtade från vetenskapsteoretiska förklaringsmodeller skall uppfyllas inom teorin om språklig kommunikation.

Vid första anblicken kan detta synas som en obetydlig variation av temat hos Davidsonsontraditionen, men i själva verket är den teoretiska ordningen omkastad. Detta visar sig i frågan om rättfärdigandet av *kompositionalitetsprincipen* för naturliga språk. Enligt min uppfattning rättfärdigas denna princip av att den kan tjäna till att förklara framgångsrik kommunikation. I Pagin (1999) har jag emellertid utförligt argumenterat för att kompositionalitetsprincipen för naturliga språk *inte* kan rättfärdigas inom ramen för en Davidsoniansk meningsteori, utan där helt enkelt måste postuleras. Skälet är i grunden detta: Davidson formulerar allmänna tolkningsprinciper som får tjäna till att *avgöra* om framgångsrik kommunikation föreligger eller inte, och dessa principer uppfylls bättre av icke-kompositionella tolkningar.

Davidson kan därför inte åberopa våra förteoretiska empiriska skäl för att tro att vi kommunicerar framgångsrikt: vi lyckas i hög grad samordna handlingar intersubjektivt (som att träffas på avtalad tid och plats), och denna samordning förklaras i sin tur bäst av att föregående språkliga transaktioner varit kommunikativt framgångsrika. Detta är utgångspunkten för mitt projekt. Jag har arbetat med dessa frågor av och till i nära tio år (med två arbeten publicerade och två under utgivning), och syftar nu till att delvis sammanfatta, delvis utveckla, tankarna i en monografi. Den preliminära dispositionen är denna:

1. Kommunikation och villkor för kommunikativ framgång
2. Idén om förklaring av framgång i språklig kommunikation
3. Kompositionalitet
4. Semantiska teorier och förklaringsmodeller
5. Alternativa förklaringar
6. Alternativa rättfärdiganden av kompositionalitetsprincipen
7. Föregivna skäl mot kompositionalitet för naturliga språk

I följande avsnitt ges utförliga beskrivningar av de planerade kapitlens tänkta innehåll.

## **2. Monografien**

### **2.1 Kommunikation och villkor för kommunikativ framgång**

För genomföra ett meningsteoretiskt program enligt vilket semantiska begrepp införs genom teorier som kan förklara framgång i språklig kommunikation, måste man först definiera vad som utgör kommunikativ framgång. Man måste också genomföra detta utan att åberopa semantiska egenskaper hos språkliga uttryck. Syftet med första kapitlet är att presentera och diskutera en sådan definition.

Mitt förslag kan sägas ansluta till klassiska uppfattningar, som hos John Locke (*An Essay Concerning Human Understanding*, bok III), Gottlob Frege ('Über Sinn und Bedeutung') och Claude Shannon (*The Mathematical Theory of Communication*). Även om Shannon inte behandlar semantiska relationer finns det en gemensam tanke: kommunikation lyckas när mottagarens kommunikationstillstånd motsvarar avsändarens.

Locke tänker sig att språkliga uttryck i första hand står för *idéer* i människans medvetande. För att kommunicera en idé använder sig talaren av ett språkligt uttryck som står för denna idé, och kommunikationen lyckas om detta språkliga uttryck står för *samma* idé i mottagarens medvetande. På motsvarande sätt menade Shannon att kommunikation lyckas om avsändarens meddelande, vilket kodats och förmedlats genom en kommunikationskanal, blir korrekt *reproducerat* på mottagarsidan. Hos Shannon är meddelanden informationstillstånd hos fysiska system, inte någonting mentalt eller semantiskt.

I bägge fallen krävs något slags *motsvarighet* mellan mottagarens tillstånd och avsändarens. I Lockes fall är det identitet mellan idéer, men detta är problematiskt, eftersom idéer hos Locke är tänkta att vara mentala och subjektiva. Om idéer verkligen är mentala, så kan två personer inte ha samma idé, lika lite som två personer kan ha samma förnimmelse. De kan däremot ha idéer om samma sak, dvs idéer med samma *innehåll*, men innehållet är självt inte något mentalt. Just detta var något som betonades av Frege. Frege skilde mellan *föreställningar* (*Vorstellungen*), vilka är subjektiva och privata, och innebörder (*Sinne, Gedanken*), som är objektiva och gemensamma för flera tänkare. Kommunikation kan lyckas, enligt Frege, därför att talare och mottagare kan förknippa *samma* innebörd med samma uttryck.

Det första kravet, på motsvarighet hos tillstånd, är gemensamt för all slags kommunikation, inklusive djurkommunikation. Vid ett varningsläte kan kommunikationen sägas ha lyckats om lätet hos mottagaren orsakar en känsla av rädsla, eller en flyktdisposition, som motsvarar avsändarens. Det andra kravet, på likhet mellan innehåll, är speciellt just för högre, framför allt språklig, kommunikation.

Kommunikation kan både lyckas och misslyckas. De ovannämnda kraven utgör villkor för att en kommunikativ händelse ska räknas som framgångsrik. För att definitionen skall vara fullständig krävs också att man anger villkor för att en viss (komplex) händelse skall vara *kommunikativ*. För mina syften är det viktigare att de första villkoren, framgångsvillkoren, är ordentligt utarbetade än att de senare, kommunikativitetsvillkoren, är det. Det är emellertid viktigt att ge ett alternativ till sådana teorier som t ex kräver mycket sofistikerade sk kommunikationsavsikter hos talaren för att ett yttrande alls ska räknas som kommunikativt. Det är nämligen orimligt att kräva att talare och mottagare oberoende av språkbehärskning är kapabla att tänka sofistikerade tankar om egna och andras tankar för att alls kunna kommunicera.

Utan att här gå närmare in på detaljer tänker jag mig att villkoren för att vara en kommunikativ händelse skall anges dels i kausala termer, dels i termer av syfte. Kausalitet behövs, eftersom ett tillstånd hos mottagaren orsakas av ett tillstånd hos avsändaren via en *signal*. Avsändarens tillstånd orsakar signalen. Via mottagarens varseblivning av signalen orsakar denna mottagarens tillstånd. I det fall det är fråga om kommunikation av representationellt innehåll består dessa tillstånd i att tänka en viss tanke.

Denna kausala relation är emellertid inte tillräcklig. För att vi alls ska kunna tala om

framgång eller misslyckande måste det finnas ett visst *syfte* med signalen. Detta syfte kan vara avsändarens avsikter, men måste inte vara det. Inom djurkommunikation kan vi t ex säga att syftet med ett varningsläte är att varna de andra medlemmarna av flocken: detta är varningslätetets *funktion*. Avsändaren behöver inte vara medveten om denna funktion.

Kapitlet är tänkt att innehålla en diskussion om vilka villkor som måste ställas på kausalrelation och syfte för att en händelse ska räknas som kommunikativ. Utöver detta är det två större frågor som skall behandlas. Den första gäller valet att ange framgångsvillkor för språklig kommunikation i termer av en likhetsrelation mellan tankeinhåll. Även om detta ansluter till klassiska föreställningar är det en idé som strider mot dominerande föreställningar inom den analytiska filosofin. Michael Dummett har t ex kritiserat Freges idé utifrån kravet att vi måste kunna *veta* att vi förstår varandra: att förståelse föreligger måste därför vara verifierbart, och det är inte verifierbart för en talare vilken innebörd en annan talare förbinder med ett visst uttryck, eftersom detta är någonting privat för den andre talaren.

Jag kommer att trycka på att vi måste skilja mellan de krav man kan ställa på en kommunikationsmetod, som t ex ett helt språk, eller en översättningsmanual, och de krav man kan ställa för att en *enskild kommunikativ händelse* skall räknas som framgångsrik. Jag ska argumentera för att kravet på kunskap om att förståelse föreligger är för starkt när det gäller den enskilda händelsen. Det är inte för starkt på den övergripande nivån. Det är t ex rimligt att kräva att vi kan veta att vi *överlag* kommunicerar framgångsrikt på ett naturligt språk som svenska. Att detta krav är uppfyllt är emellertid förenligt med Fregeanska framgångsvillkor för den enskilda händelsen.

I en delvis annorlunda, delvis likartad riktning kan det invändas att kommunikativ framgång inte kan bestå i någonting mer än att talare och mottagare lyckas koordinera sina respektive handlingar. Detta skulle peka mot en mer behavioristisk uppfattning, som t ex förespråkats av W V Quine. Även här kommer jag att argumentera för att behavioristiska villkor är inadekvata för det enskilda fallet: beroende på störande faktorer, t ex felaktiga bakgrundsövertygelser, kan koordination av beteende misslyckas trots att en föregående kommunikation talarna emellan var lyckad. Likaså kan koordination lyckas trots att kommunikationen misslyckades, nämligen om flera felkällor tar ut varandra.

Det behavioristiska kravet kan däremot synas mer adekvat på det övergripande planet, vilket också är var Quine faktiskt tillämpat det, nämligen som krav för att en hel översättningsmanual skall räknas som riktig. Detta är emellertid också förenligt med Fregeanska framgångsvillkor i det enskilda fallet. Behavioristen kan å andra sidan inte, vilket jag skall argumentera för, ge rimliga villkor för det enskilda fallet, utan måste ge upp själva idén om framgång och misslyckande i en enskild kommunikativ händelse.

Resten av kapitlet skall ägnas åt frågan om likhetsrelationen mellan tankeinhåll. Även här har Frege ett försteg jämfört med senare filosofer som tagit det individuella kognitiva subjektet som utgångspunkt. Frege betonade att innehåll måste vara gemensamt för olika talare. Detta utesluter specificering av innehåll i termer av mentala eller neurofysiologiska tillstånd som i olika varianter varit populära (Harman, Block, Fodor, Schiffer). Tre frågor skall tas upp: för det första frågan om det begreppsliga ramverket för att karakterisera innehåll; för det andra frågan om relationen mellan det mentala till-

ståndet och innehållet; för det tredje frågan om vilken likhet som skall råda mellan talarers respektive innehåll för att framgångsvillkoret skall vara uppfyllt.

I det första fallet kommer jag att förespråka att innehåll karakteriseras i termer av *möjliga världar* som i något relevant avseende *passar* ett mental tillstånd. Användningen av ett sådant ramverk har två klara fördelar. Dels är kravet på intersubjektivitet uppfyllt, eftersom världarna är externa i förhållande till talarna och vi därför kan fråga om två talarers mentala tillstånd passar eller inte passar samma möjliga världar. Dels finns det inte något behov av att spekulera om *struktur* eller andra inre egenskaper hos de mentala tillstånden eller deras innehåll. Inre struktur bör inte bara postuleras eller antas.

I det andra fallet kan man tänka sig två stycken huvudalternativ för att specificera relationen att *passa* mellan ett tillstånd och en möjlig värld. Det ena är att utgå från ett *sanningsbegrepp* för mentala tillstånd. En möjlig värld passar då ett övertygelsetillstånd om detta tillstånd är *sant* med avseende på denna värld. Vi tänker oss i så fall detta semantiska begrepp om sanning hos övertygelsetillstånd som primitivt, med vars hjälp semantiska begrepp för språkliga uttryck senare skall förklaras. Det andra, som är det filosofiskt svårare alternativet, är att utgå från *dispositioner* hos talaren att under vissa externa omständigheter råka i vissa inre tillstånd. Det finns stora och kända problem med att definiera propositionellt innehåll utifrån sådan dispositioner (en form av detta är Fodors s k *disjunktionsproblem*). Jag tror att dessa problem går att lösa inom ramen för en teori om kommunikation, men en sådan ansats går utöver vad som kan rymmas inom kapitlet. Här kommer jag istället att preliminärt anta ett sanningsbegrepp.

Den tredje frågan rör kravet på likhet. Det enklaste är att helt enkelt kräva att talarens och lyssnarens respektive övertygelser skall vara sanna i exakt *samma* möjliga världar för att kommunikationen skall räknas som framgångsrik. Alternativen är att kräva en lägre grad av likhet. Min uppfattning är att en lägre grad av likhet är mest adekvat för en teori om kommunikation. Argumentationen för detta faller emellertid utanför ramen för detta förberedande kapitel. Preliminärt kommer jag därför att anta att identitetskrav.

## 2.2 Idén om förklaring av framgång i språklig kommunikation

Hur kommer det sig att kommunikation lyckas? Frågan kan ställas och besvaras för den enskilda kommunikativa händelsen. Talaren A tänker att *morötter innehåller karotein*. A faller yttrandet  $u$ , av satsen  $s = \text{'morötter innehåller karotein'}$ , vilket uppfattas av B, vilket leder till att B också tänker tanken att morötter innehåller karotein, dvs en tanke med samma innehåll. Givet att yttrandet hade ett kommunikativt syfte har vi här ett fall av framgångsrik kommunikation. Vi vill nu veta varför det lyckades.

Såsom framgång har definierats är det inte uteslutet kommunikationen lyckades av en slump. Det enda som krävs är att B:s tanke faktiskt har samma innehåll som A:s. Slumpbegreppet är emellertid problematiskt. När vi intuitivt säger att det lyckades av en slump menar vi inte att det inte finns någon förklaring till varför den lyckades. Det skulle kunna vara så att B av helt oberoende orsaker tänkte på morötter och karotein, och utan att riktigt uppfatta A:s yttrande fick för sig att A sa att morötter innehåller karotein, mest för att denna tanke helt enkelt låg närmast till hands i B:s associationsbanor. Det finns alltså en förklaring, men i en viss bemärkelse var det ändå en slump att tankeinnehållen sammanföll.

Den relevanta bemärkelsen tycks vara att en förklaring av ett visst *slag* saknas, nämligen en förklaring enligt vilken yttranden av satsen *s tillförlitligt* ger kommunikativ framgång. Kommunikationen ovan kunde lyckats på det sätt som beskrevs även om B inte förstod svenska, och den lyckade i så fall bara på grund av vad B råkade tänka på vid tillfället. I motsats till detta gäller att en talare med god kompetens i svenska har en *disposition* att tänka tanken *att morötter innehåller karotein* som en reaktion på observationen av ett yttrande av satsen *s*. Likaså har talaren en disposition att yttra denna sats, eller någon liknande, när talaren önskar säga *att morötter innehåller karotein*, dvs när denna tanke på typiskt sätt ger upphov till ett yttrande. Nu kan vi förklara varför en viss kommunikativ händelse lyckades genom att hänvisa till sådana dispositioner hos talaren och lyssnaren. Denna förklaring åberopar faktorer som ger tillförlitlig framgång, dvs i hög grad oberoende av egenheter hos yttrandekontexten.

Hur dispositionsbegreppet skall förstås och hur det bäst uppfattas inom ramen för en förklaring av en enskild händelse är något som behöver diskuteras, men som jag inte tar upp här. Att talare av svenska språket har sådana dispositioner ser jag som okontroversiellt. Men liknande dispositioner är operativa vid djurkommunikation, som dispositionen hos den gröna markattan i östafrika att avge ett tvåstavigt hostljud vid åsynen av en annalkande örn, och dispositionen hos mottagarna att tro, eller åtminstone bete sig som om dom trodde, att en örn närmar sig, när de hör en medlem utstöta detta ljud (se Cheney och Seyfarth 1990).

I nästa steg är vi därför intresserade av en förklaring till varför olika individers dispositioner motsvarar varandra på så sätt att kommunikation lyckas. Även detta kan vara en slump. Det vore t ex en slump om A och B talade olika språk, och just satsen 'morötter innehåller karotein' är den enda satsen som finns på bägge språken och dessutom har samma mening. Även här är vi intresserade av en förklaring i en starkare mening, en förklaring som visar på varför vi kunde *förvänta* oss att de olika individerna hade samma eller motsvarande dispositioner, därför att den visar på faktorer som tillförlitligt ger upphov till likhet i detta avseende.

I markattornas fall är nästa steg att ge en *kausal* förklaring till dispositionens existens. Vi får åberopa dels genetiska faktorer, dels en inlärningshistoria, läte för läte i deras repertoar av kommunikativa uttryck. Att dispositioner stämmer överens mellan individer förklaras då av tillräckligt överlappande genuppsättning samt av kausal påverkan inom gruppen för utvecklande av dispositioner hos unga individer.

Här skiljer sig emellertid den mänskliga talaren. Den mänskliga talaren kan uttrycka en oöverskådligt stor mängd tankar och är kapabel att tolka en oöverskådligt stor mängd språkliga uttryck. Speciellt är den mänskliga talaren kapabel att uttrycka *nya* tankar, dvs finna ett språkligt uttryck för ett tankeinnehåll som talaren själv tidigare aldrig tänkt, och kapabel att tolka *nya* uttryck, dvs uttryck han tidigare inte använt, hört eller tilldelat mening. Och framför allt är det vanligt att vi *lyckas* i kommunikation där både tankeinnehåll och språkligt uttryck är nya för både talare och lyssnare. Detta är vanligt, och det är anmärkningsvärt. Denna framgång kan *inte* på ett tillfredsställande sätt förklaras med hänvisning till inlärning sats för sats.

Explananda, dvs fakta som skall förklaras, utgörs nu av ett extremt stort antal fakta om intersubjektiv likhet ifråga om språkliga dispositioner, inklusive dispositioner som

ännu inte blivit manifesterade. A har en disposition att yttra satsen  $s$  för att kommunicera tanken att  $p$ , och B har dispositionen att tänka tanken att  $p$  som ett resultat av att uppfatta ett yttrande av satsen  $s$ , för ett mycket stort antal satser  $s$  och propositioner  $p$ , inklusive sådana som aldrig tidigare yttrats eller tänkts.

Vi kan beskriva den typiska situationen som en situation med ett *kommunikationsproblem*. Talaren vill kommunicera ett visst tankeinhåll och har ett mycket stort antal möjliga språkliga uttryck till sitt förfogande. Uppgiften är att välja ett uttryck som är adekvat för lyssnaren. Lyssnaren å sin sida uppfattar det språkliga uttrycket och har ett mycket stort antal möjliga innehåll att välja mellan. Uppgiften är att välja samma innehåll som talaren tänkt från början. Hur kommer det sig att talare av naturliga språk i så hög grad löser detta problem, och dessutom i de allra flesta fall med automatik, utan att ens vara medvetna om någon uppgift som behöver lösas? Givet de stora urvalen är chansen att lyckas av en slump försumbar. Vi kan av empiriska skäl utesluta möjligheten att kommunikation regelbundet lyckas på grund enbart av telepati, intuition, empati eller kontextuella ledtrådar.

För att förklara detta vill vi ha en allmän teori. Denna teori skall alltså kunna förklara den intersubjektiva likheten ifråga om dispositioner. Hypotesen är att en sådan teori är *semantisk*. (*meningshypotesen*). Denna hypotes går alltså ut på att vi med hjälp av hänvisning till *relationer mellan språk och värld* kan formulera en teori som kan förklara likheter ifråga om dispositioner.

### 2.3 Kompositionalitet

Detta kapitel är avsett att behandla begreppen *kompositionalitet* respektive *stark kompositionalitet*. Historiskt tycks den första moderna formuleringen av *kompositionalitetsprincipen* ha givits av Rudolf Carnap i *Meaning and Necessity* 1947. Carnap kallade den med rätta *Freges princip*. Visserligen formulerade inte Frege principen på modernt sätt, men den rimligaste tolkningen av flera passager hos Frege (se nästa kapitelbeskrivning) är att han formulerade en starkare idé än vanlig kompositionalitet. En vanlig modern formulering av kompositionalitetsprincipen är denna:

(FK) Meningen hos ett sammansatt språkligt uttryck är bestämt av meningen hos dess delar och sammansättningsformen.

En numera vanlig, men likafullt kontroversiell, tanke är att det är denna kompositionalitet hos naturliga språk som gör att vi kan förstå nya satser: vi känner meningen hos den nya satsens delar, vi känner den semantiska funktionen hos sammansättningsformen, och tack vare dessa kunskaper kan vi förstå en sats vi aldrig tidigare mött. Senare (kap 6) kommer jag att kritisera detta och andra argument, trots att jag håller med om argumentens slutsats: naturliga språk har en kompositionell semantik.

En kompositionell semantisk teori för ett visst språk är en teori i vilken man kan härleda meningen hos ett sammansatt uttryck utifrån de ingående uttryckens mening samt sammansättningsformen. De första kompositionella semantiska teorierna gavs för formella, matematiska språk, genom Gottlob Frege på 1890-talet, och senare framför allt genom Alfred Tarski på 1930-talet. Steget att tillämpa de formella semantiska idéerna på

*naturliga* språk—i första hand engelska—togs på 1960-talet genom Donald Davidson och Richard Montague. Davidson använde sig av Tarskis formella arbeten för att ge vad som kommit att kallas sanningsteoretisk semantik för enklare delar av engelska. Montague, å andra sidan, använde sig av ett flertal existerande logisk-matematiska teorier: universell algebra, typteori, kategorialgrammatik samt s k möjlig-världssemantik. Med hjälp av dessa tekniker kunde Montague utveckla teorier för större fragment av engelska än Davidson med tekniskt enklare (men filosofiskt bättre motiverade) medel klarade av.

Montagues ramverk passar bättre för en abstrakt beskrivning av kompositionalitet. Grunden för detta är en tillämpning av en gren av matematik som kallas universell algebra. En algebra består i enklaste form av en mängd element  $M$  samt en mängd operationer  $O$  på dessa element, och brukar anges som paret  $M, O$ . Montagues idé var att betrakta språkets syntax och språkets semantik som separata algebror. Den syntaktiska algebran har språkets enkla uttryck som grundelement, och genom operationer bildas nya mer komplexa element ur enklare element. T ex har vi en operation vi kan kalla  $F_1$  som bildar en sats ur en nominalfras samt en intransitiv finit verbfras. Tillämpat på argumenten 'Per' och 'springer' erhåller vi då  $F_1('Per', 'springer')$ ='Per springer'. Tanken är nu att ett naturligt språks hela syntax kan beskrivas som en syntaktisk algebra, med ett stort men ändligt antal enkla uttryck samt ett antal syntaktiska (eller snarare syntaktiska och morfologiska) operationer på mängden av uttryck.

Mot detta svarar en semantisk algebra, dvs en algebra av sådana entiteter som utgör *semantiska värden* till språkets uttryck. Typiskt är det semantiska värdet hos ett egenamn som 'Peter Pagin' den person som är bärare av namnet. Det semantiska värdet för en fullständig sats (eventuellt givet en viss yttrandekontext) är ett sanningsvärde, värdet SANN eller värdet FALSK. Det semantiska värdet för en verb som 'springer' kan vi, i Freges och Montagues efterföljd, betrakta som en funktion, närmare bestämt som en funktion från objekt till sanningsvärden (jag bortser här från behovet att ta hänsyn till tidpunkt). Låt oss kalla denna funktion 'SPRINGER'. Då gäller att SPRINGER är en funktion som ger värdet SANN för argumentet Per om, och endast om, Per springer. Om Per springer gäller alltså att  $SPRINGER(Per)$ =SANN. För att fullborda bilden har vi en tillämpningsfunktion APP, sådan att  $APP(SPINGER, Per)$ =SPRINGER(Per) (det är lite missvisande att använda samma uttryck 'SPRINGER' på bägge platserna, men jag bortser från det).

Nu kan vi introducera kompositionalitet genom att *relatera* den syntaktiska och den semantiska algebran. Det krav som Montague formulerade var att det ska finnas en meningsfunktion  $H$  från uttryck till deras semantiska värden, och att denna funktion skall vara den s k *homomorfism*, dvs en strukturbevarande avbildning. Homomorfismkravet innebär att det till varje syntaktisk funktion  $F$  från uttryck till uttryck skall finnas en semantisk funktion  $G$  från värden till värden sådana att

$$H(F(x_1, \dots, x_n)) = G(H(x_1), \dots, H(x_n))$$

där  $x_1, \dots, x_n$  är de uttryck som utgör argument för funktionen  $F$ .

Om vi nu tillämpar denna idé på vårt exempel, så tänker vi oss Per utgör det semantiska värdet för termen 'Per', dvs  $H('Per')$ =Per. Likaså tänker vi oss att  $H('springer')$ =SPRINGER. Till sist kan vi tänka oss APP-funktionen som en funktion som semantiskt motsvarar den syntaktiska operationen att kombinera en nominalfras med en

verbfras. Vi erhåller då

$$\begin{aligned} H(F_1(\text{'Per'}, \text{'springer'})) &= \text{APP}(H(\text{'Per'}), H(\text{'springer'})) \\ &= \text{APP}(\text{Per}, \text{SPRINGER}) \\ &= \text{SPRINGER}(\text{Per}) \end{aligned}$$

vilket i sin tur är värdet SANN, om nu Per springer. Detta kan synas trivialt, eftersom vi använder nästan samma termer i teoriformuleringen för att tala om språkliga uttryck som för att tala om deras semantiska värden, men intrycket av trivialitet skulle snart försvinna om de språkliga uttrycken hörde till ett främmande språk och konstruktionen ifråga var syntaktiskt eller semantiskt mer komplicerad. Framställningen behöver också korrigeras för att ta hänsyn till syntaktiska och semantiska tvetydigheter, men jag går inte in på det här.

Nu är det ju inte särskilt rimligt att tänka sig ett sanningsvärde som *meningen* hos en deklarativ sats, och därför komplicerar Montague bilden genom att införa *intensioner* i teorin. Utan att närmare gå in på detaljer kan vi säga att slutresultatet genom detta inte längre är ett sanningsvärde utan en viss intension, närmare bestämt en funktion från *möjliga världar* till sanningsvärden, vilket motsvarar vad man inom traditionen kallat en *proposition*. Tanken är att en proposition karakteriseras just genom den mängd av möjliga världar där den är sann. Propositionen *att Per springer* är sann vid exakt varje möjlig där Per springer.

Den sk Montaguegrammatiken har tillämpats av olika forskare på ett stort antal fragment av naturlig engelska (Dag Westerståhl och jag själv utvecklade en variant av den som kan hantera vissa annars problematiska anaforiska konstruktioner; se Pagin och Westerståhl 1994). Teoretiskt har det algebraiska ramverket för den utvecklats av de holländska logikerna Theo Jansen och Herman Hendriks. I min egen tillämpning kommer jag istället att använda mig av ett enklare algebraiskt ramverk som formulerats av Wilfrid Hodges och vidareutvecklats av Dag Westerståhl. Dessa formella bakgrunder kommer att beskrivas relativt detaljerat i kapitlet.

Mitt eget bidrag till den formella teorin utgår från beskrivningen av kommunikationsproblemet i förra avsnittet. Så som många har tänkt erbjuder kompositionalitet en lösning, men vid närmare eftertanke är det en lösning endast på *ena halvan* av problemet, mottagarens problem, dvs problemet hur man ska hitta ett propositionellt innehåll givet ett viss språkligt uttryck. Men kompositionalitet hjälper inte talaren att hitta det lämpliga *uttrycket* som lyssnaren sedan effektivt kan tolka. Vi kan likna situationen vid ett översättningsproblem. En svensk och en engelsman konverserar med hjälp av ett lexikon. Detta är emellertid endast engelsk-svenskt. De saknar svensk-engelskt lexikon. Givet vad *engelsmannen* säger kan talarna genom den alfabetiska ordningen effektivt hitta ordöversättningar i lexikonet, men de kan *inte effektivt* hitta en översättning av vad *svensken* säger. De måste helt enkelt leta igenom lexikonet ända till dess att de hittar det svenska ordet ifråga, utan att ha någon aning om var i lexikonet det finns. Om vi lämnar lexikonanalogin och tänker på användningen av hela satser, är situationen ännu värre, eftersom antalet satser är så mycket större (och potentiellt oändligt).

Frågan är vad som krävs för en lösning av denna del av problemet. Vad som skulle ge en lösning är detta: komplexa semantiska värden har en unik komposition. Om det är fal-

let att *tankeinhåll* är unikt uppbyggda, dvs så att talaren som tänker en tanke har tillgång till detta tankeinhålls konstituent, och dess sammansättningsform, så kan talaren effektivt hitta de enkla språkliga uttrycken och sedan effektivt bygga upp satsen utifrån dessa delar. Vi kommer i så fall att ha kompositionalitet i motsatt riktning, eftersom det då finns en homomorfism från den semantiska algebran till (en del av) den syntaktiska.

Om det finns en semantisk egenskap hos naturliga språk som hänger samman med existensen av en sådan unik komposition, så har vi också separat empiriskt stöd. Detta stöd kommer från observation av vissa *substitutionsegenskaper*. Redan hos Frege förekommer formulering av substitutionsprinciper, och Carnap tänker sig en substitutionsversion (SK) av kompositionalitetsprincipen som ekvivalent med funktionsversionen (FK). Kortfattat kan vi formulera det så här:

(SK) Substitution av synonyma uttryck ger synonyma resultat.

Dvs, om vi har en sats  $A$  som innehåller ett uttryck  $e$ , och ersätter  $e$  med ett annat uttryck  $e'$  som är synonymt med  $e$ , dvs har samma mening, så får vi en ny sats  $A'$  som är synonym med  $A$ . Detta är nära besläktat med idén att meningen hos satsen beror på meningen hos delarna och sammansättningsformen. Det är också så, vilket Hodges visat, att under antagande av två extra principer, så är FK och SK ekvivalenta.

Poängen nu är att det finns en motsvarande substitutionsprincip, vilken jag kallat principen om *invers* kompositionalitet:

(ISK) Substitution av icke-synonyma uttryck ger icke-synonyma resultat.

Dvs, om vi har en sats  $A$  som innehåller ett uttryck  $e$ , och ersätter  $e$  med ett annat uttryck  $e'$  som *inte* är synonymt med  $e$ , så får vi en ny sats  $A'$  som *inte* är synonym med  $A$ . (ISK) förefaller vara *sann* för naturliga språk. Dessutom har jag formellt bevisat att, givet några extra villkor, invers kompositionalitet *implicerar* unik komposition av innehåll/semantiska värden. Det är därför meningsfullt tala om icke-primitiva innehåll som *sammansatta av delar*, och anta principen:

(IFK) Uttrycket för ett sammansatt innehåll är bestämt av uttrycken för dess delar samt sammansättningsformen.

Kombinationen av kompositionalitet invers kompositionalitet kallar jag *stark kompositionalitet*. Stark kompositionalitet ger därmed en fullständig lösning på kommunikationsproblemet, och jag tänker mig att det också är vad Frege hade i tankarna i följande citat från uppsatsen 'Tankefogning', som publicerades 1918:

Det är förvånansvärt vad språket åstadkommer, då det med några få stavelser kan uttrycka oöverskådligt många tankar, så att till och med en tanke som för första gången fattats av en jordevarelse kan erhålla en språklig dräkt, i vilken tanken kan uppfattas av en annan för vilken den är helt ny. Detta vore inte möjligt om vi inte i tanken kunde urskilja delar, vilka motsvarade delar av satsen, så att satsens uppbyggnad kan tjäna som bild av tankens.

## 2.4 Semantiska teorier och förklaringsmodeller

Nästa steg är att se på formuleringen av semantiska teorier. Jag avser att behandla en Davidsonteori och en Montagueteori (och i mån av tid och utrymme någons diskurssemantisk teori). I princip är det två saker som skall åstadkommas i detta kapitel.

För det första ska jag visa att teorierna är kompositionella respektive inverst kompositionella. I det förra fallet gäller det främst att visa att Davidsonteorier, som har ett Tarski-format snarare än ett algebraiskt format, uppfyller kravet på kompositionalitet. Till detta kan vanliga modellteoretiskt formulerad semantik användas för att visa att de nödvändiga likheterna finns. I det senare fallet gäller det att visa att teorin tillåter härledningar i *motsatt* riktning mot normala tillämpningar. Att detta faktiskt är fallet för en Davidsonteori illustrerar jag strax.

För det andra gäller det att visa hur sådana starkt kompositionella teorier, tillsammans med extra principer, kan uppfylla villkor ställda av olika *vetenskapliga förklaringsmodeller* för att utgöra acceptabla förklaring av de explananda vi angivit, nämligen intersubjektiv likhet ifråga om dispositioner.

För att illustrera tillvägagångssättet antar vi en viss engelsk sats:

(1) Charles loves Lisa

Denna har tre komponenter, ett relationsuttryck—det transitiva verbet 'loves'—samt två egennamn, 'Charles' och 'Lisa'. Låt oss ställa upp en Davidsonteori **T** för att hantera satsen (1). Först har vi två *referensaxiom* för egennamnen:

$R_1$  Referenten till 'Charles' är Kalle.

$R_2$  Referenten till 'Lisa' är Lisa.

Sedan har vi ett *satisfieringsaxiom* för verbet 'loves'

L 'loves' är sann om  $a$  och  $b$  om, och endast om,  $a$  älskar  $b$ .

(en mer korrekt version skulle formuleras i termer av *tillordningsfunktioner* av värden till variabler). Sedan säger vi om satsformen

K En sats av formen  $e_1 R e_2$  är sann om, och endast om,  $R$  är sann om referenten till  $e_1$  och referenten till  $e_2$ .

Principerna  $R_1$ ,  $R_2$ , L och K, tillsammans med vissa logiskt giltiga slutledningsformer, utgör den semantiska teorin **T**. Tillsammans med **T** har vi två bryggprinciper mellan semantik och disposition:

D premiss: det följer ur teori **T** att satsen  $s$  är sann om, och endast om,  $p$   
slutsats a): talaren är disponerad att, när hon vill kommunicera att  $p$ , yttra  $s$   
slutsats b): lyssnaren är disponerad att, när hon hör  $s$  yttras, tänka att  $p$

Låt oss först koncentrera oss på b). Vi vill härleda slutsatsen att lyssnaren B är dispone-

rad att tänka att Kalle älskar Lisa när hon hör satsen (1). Först tillämpar vi principen K på satsen och erhåller resultatet

- (2) 'Charles loves Lisa' är sann om, och endast om, 'loves' är sann om referenten till 'Charles' och referenten till 'Lisa'

(genom den predikatlogiska regeln för universell instantiering). Sedan tillämpar vi satisfieringsaxiomet L för verbet 'loves' på detta resultat och erhåller

- (3) 'Charles loves Lisa' är sann om, och endast om, referenten till 'Charles' älskar referenten till 'Lisa'

(genom en satslogiskt tillåten substitution av en sats för en därmed ekvivalent sats). Därefter tillämpar vi referensaxiomen  $R_1$  och  $R_2$  för de bägge namnen, och får

- (4) 'Charles loves Lisa' är sann om, och endast om, Kalle älskar Lisa

(genom predikatlogiskt tillåten substitution enligt Leibniz lag). Nu har vi erhållit en semantisk beskrivning av satsen 'Charles loves Lisa'. På den kan vi nu tillämpa bryggprincipen D och erhålla slutsatsen

- (5) Lyssnaren är disponerad att, när hon hör satsen 'Charles loves Lisa', tänka att Kalle älskar Lisa

(eftersom satsen (4) följer ur teorin **T** har vi rätt att använda detta som premiss enligt regel D, för att erhålla slutsatsen (5) enligt variant b)).

Detta enkla exempel ger en glimt av hur vi kan *förklara* varför en talare av engelska har en disposition att förstå en ny sats. Vi associerar objekt och egenskaper med enskilda språkliga uttryck, vi formulerar principer för sammansättningsformer, och med hjälp av dessa tillordningar kan vi härleda beskrivningar av sammansatta uttryck, inklusive beskrivningar av nya satser. Dessa beskrivningar är semantiska beskrivningar, och från dessa semantiska beskrivningar kan man sedan dra slutsatser om språkanvändarnas dispositioner att tala och förstå. Om sedan dessa slutsatser också blir systematiskt riktiga, för en viss talare och för de nya satser hon konfronteras med, så kan vi korrekt *förutsäga* hur talaren kommer att förstå en ny sats, och därmed också *förklara* hur det kommer sig att talaren är disponerad på detta sätt för att förstå nya satser.

I härledningen ovan utgår vi från satsen 'Charles love Lisa', och arbetar oss fram till teoremet (4) genom att successivt tillämpa teorins principer K, L och R. Detta visar att det är möjligt för en lyssnare att effektivt hitta ett innehåll att para ihop med satsen. Omvänt kan vi emellertid också utgå från innehållet *att Kalle älskar Lisa*, och därur successivt arbeta oss fram till samma teorem (4). Fört erhåller vi

- (6) referenten till 'Charles' älskar referenten till 'Lisa' omm Kalle älskar Lisa

genom referensaxiomen  $R_1$  och  $R_2$ . Därefter tillämpar vi satisfieringsaxiomet L, och

erhåller

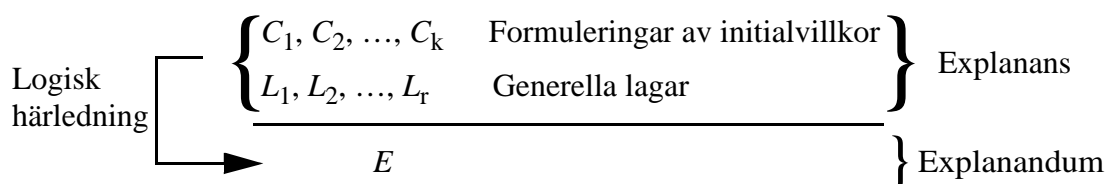
- (7) 'loves' är sann om referenten till 'Charles' och referenten till 'Lisa' omm Kalle älskar Lisa.

Ur (7) erhåller vi till sist återigen (4) genom att tillämpa principen K och substituera på vänster sida. Denna alternativa härledning av (4) visar hur det är möjligt för talaren att effektivt hitta ett uttryck för att kommunicera innehållet att Kalle älskar Lisa. Detta motiverar nu en tillämpning av a)-versionen av principen D, för att sluta oss till talarens disposition. En Davidsonteorin av detta slag uppfyller därför villkoret för att vara *starkt* kompositionell. Teorin kan, tillsammans med principen D, förklara existensen både av talarens och av mottagarens dispositioner. Här är mina egna bidrag dels modellen att den semantiska teorins uppgift är att på detta sätt ge systematiska härledningar av dispositionsutsagor, dels att påvisa existensen av den andra härledningsformen samt observationen om dess koppling till talarens kommunikationsuppgift.

Väsentligen förklarar vi alltså kommunikativ framgång genom påståendet att samma starkt kompositionella semantiska teori är sann om både talare och lyssnare, varför den tillsammans med en dispositionsprincip D kan förklara varför dispositioner för enskilda satsen hos talare och lyssnare stämmer överens. I nästa steg kan man sedan fråga efter hur det kommer sig att samma teori gäller för bägge språkanvändarna, och här får man bli hänvisa till inlärning, dvs en kausal förklaring av hur vissa kognitiva tillstånd hos språkanvändarna uppkommit, men detta är inte nödvändigt att genomföra för att dispositionsteorin **T+D** själv ska ha det avsedda förklaringsvärdet.

Nästa steg är att visa att en sådan teori om språkliga dispositioner, vilken innehåller den semantiska teorin som en del, kan sägas förklara existensen av enskilda språkliga dispositioner i enlighet med etablerade vetenskapsfilosofiska modeller för vetenskapliga förklaringar. Mina planer är att undersöka förklaringsvärdet i förhållande till fyra olika förklaringsmodeller: Gustav Hempels deduktivt-nomologiska/induktivt-statistiska modell, Philip Kitchers unifieringsmodell, Wesley Salmon's modell för statistisk relevans samt James Woodward's invariansmodell. De två första har jag tittat närmare på. I samtliga fall gäller det att förklara, inte en enskild händelse, utan en disposition, vilket kan jämföras med ett lagmässigt samband, eller approximativt lagmässigt samband.

När det gäller Hempel är det alltså frågan om den generella varianten av den deduktivt-nomologiska modellen (DN-modellen). Schematiskt framställs den deduktivt-nomologiska modellen (DN-modellen) så här:



Eftersom det är frågan om den generella varianten av DN-modellen är explanandum själv en generell utsaga, nämligen påståendet att en talare är disponerad att yttra en viss

sats, givet vissa tankar, eller tänka vissa tankar, givet yttrandet av en viss sats.

Den språkliga dispositionsteorin (dvs **T+D**) uppfyller Hempels förklaringsvillkor när det gäller härledning av dispositionsutsagor, så när som på det faktum att principen **D** självt inte kan betraktas som en *strikt lag*, och av två skäl: dels för att formuleringen av den kommer att innehålla referens till enskilda talare/lyssnare, dels för att den inte garanterat saknar undantag. Vi kan säga att den approximerar en **DN**-förklaring. Å andra sidan kan den sägas strikt uppfylla Hempels modell för *induktivt-statistiska* förklaringar (**IS**-modellen). Även **IS**-modellen har formen av en slutledning, där explanans utgör premisser och explanandum slutsats. Här krävs dock inte strikta lagar, utan det räcker med statistiska lagar, och dessa behöver enligt Hempel inte heller ange exakta numeriska sannolikheter. Enligt **IS**-modellen ges explanandum ett *induktivt* stöd av explanans snarare än ett deduktivt. **IS**-modellens villkor kan sägas vara uppfyllda av **T+D**, eftersom det där räcker att **D**-principen ger *hög sannolikhet* för en viss disposition, givet ett teorem i den semantiska teorin, och detta villkor kan utan tvekan vara uppfyllt även om störande faktorer leder till att kommunikation misslyckas, och även om begränsade kognitiva performansresurser hos språkanvändarna leder till att de i vissa sammanhang är disponerade att misslyckas med tolkning eller formulering.

En annan förklaringsmodell vars villkor är uppfyllda är Philip Kitchers unifieringsmodell, där den centrala idén är att teoretiska utsagor blir *unifierade* genom att kunna härledas som specialfall av mer generella teoretiska utsagor. Förklaringskraften beror där inte enbart på den enskilda härledningen, utan på det mer generella faktum att ett större antal sanna teoretiska utsagor (i detta fall de enskilda dispositionsutsagorna) kan härledas på ett enhetligt sätt. Jag går inte in på detaljerna.

En viktig observation här är att skillnaderna mellan de två härledningsformerna i teorin **T**, och därmed skillnaden mellan att förklara talarens och att förklara lyssnarens dispositioner, *inte* spelar någon roll för dessa förklaringsmodeller. I detta avseende kan dessa modeller därför betraktas som alltför svaga och det återstår att utreda hur dessa skulle kunna stärkas på ett generellt sätt för att härbärgera de distinktioner som kommunikationsförklaringar kräver.

Den meningsteoretiska tanken kan nu sammanfattas så här: Vad det innebär för en språklig term att *referera* till ett objekt, och vad det innebär för ett predikat att *satisfieras* av objekt med en viss egenskap, eller av objekt som står i en viss relation, är just att vara relaterad till ett objekt, respektive en viss egenskap, i en semantisk teori **T** som tillsammans med dispositionsprincipen förklarar kommunikativ framgång. Det är min centrala tanke.

## 2.5 Alternativa förklaringar

I och med kapitel fyra är bokens konstruktiva del väsentligen fullbordad. De resterande delarna planeras ha blandat konstruktiv och polemisk karaktär. I femte kapitlet avser jag att diskutera dels alternativa formella teorier som delar av dispositionsteorier, dels två alternativa uppfattningar om vad som krävs för förklaring av kommunikation.

De alternativa formella teorierna är av två slag. Dels har vi möjligheten till andra, förklaringsmässigt *likvärdiga* kompositionella semantiska teorier. Om det finns sådana kommer vi att ha en slags underbestämning av semantisk teori, och möjligheten även s k

obestämdhet, liknande sådan som försvarats av Davidson. Jag avser att argumentera för att underbestämningen blir mycket mindre i min teori (på grund av att intuitiva skillnader i tankeinhåll kan accepteras i min teori men inte i Davidsons), samt för att graden av underbestämning minskar ju högre teoretisk nivå det är på den kommunikation som ska förklaras. Det krävs t ex mer för att kommunicera framgångsrikt med universella utsagor och modala utsagor än med enkla singulära utsagor som (1), och det minskar möjligheten att framgång kan förklaras lika väl med olika teorier.

Dels har vi möjligheten till systematiska men icke-kompositionella teorier. Dessa faller återigen i två undergrupper. Å ena sidan har vi semantiskt systematiska, "nästan kompositionella" teorier, som versioner av Hans Kamps DRT och Jaakko Hintikkas spelteoretiska semantik. Båda dessa kan emellertid säga uppfylla kompositionalitetsprincipen, om dess förmerling tolkas på ett lite svagare sätt än det vanliga. Dessutom har strikt kompositionella varianter av dessa teorier formulerats, av Henk Zeevat respektive Wilfrid Hodges. Teorier som dessa utgör därför inte substantiellt icke-kompositionella alternativ.

Å andra sidan har vi också systematiska icke-semantiska teorier (man kan t ex säga att meningen hos ett syntaktiskt komplext uttryck är bestämd av dess delar och sammansättningsformen, inte av delarnas *mening*). Min avsikt här är att försöka visa att, givet ett naturligt mått på förklaringsvärde, de kompositionella teoriernas värde är klart högre. Här har jag endast haft preliminära idéer som återstår att utveckla.

När det gäller alternativa uppfattningar finns det två som jag avser diskutera. Den ena idén kräver mer än vad jag föreslagit för förklaringskraft, nämligen att talaren/lyssnaren skall ha någon slags *kunskap* om den riktiga semantiska teorin. Eftersom man inte kan kräva explicit teoretisk kunskap brukar man kräva någon form av implicit, eller *tyst* kunskap. Ett sådant krav är oförenligt med min teori, eftersom den semantiska teorins riktighet i så fall måste vara bestämd *oberoende* av dess förklaringsvärde (annars kan inte talarnas kunskap om dess riktighet, oavsett hur implicit den är, krävas som ett villkor för att den ska ha förklaringsvärde). Tankar i denna riktning har förts fram av Michael Dummett, Gareth Evans, Crispin Wright och Martin Davies. Väsentligen är min invändning att förklaringsvärde kan etableras utan hänvisning till tyst kunskap, så som visats i avsnitt fyra. Dessutom avser jag att argumentera för att kravet på tyst kunskap leder till fler problem än det löser.

Den andra invändningen är formulerad av Stephen Schiffer, och går ut på att framgångsrik kommunikation kan förklaras utan hänvisning till semantik för språkliga uttryck överhuvud taget. Det enda som krävs enligt Schiffer är möjligheten till en systematisk mental översättningsmekanism hos talare och lyssnare, mellan mentala representationer av språkliga uttryck och mentala representationer av innehåll. Min invändning mot detta resonemang är det inte uppfyller Freges villkor att innehåll måste vara *gemensamma* för att vi ska kunna tala om framgångsrik kommunikation. Så länge vi endast talar om mentala representationer av innehåll och inte om innehållen själva har det gemensamma utelämnats. Ingen sådan teori kan ge en *fullständig* förklaring. På sin höjd kan den bidra med delar som behövs i en sådan.

## 2.6 Alternativa rättfärdiganden av kompositionalitetsprincipen

Detta kapitel kommer att utgöra en utveckling och komplettering av resonemang jag utför i Pagin 1999. Idéer som förts fram som skäl för att språk måste vara kompositionala innefattar tanken att kompositionalitet behövs för att språk skall vara möjliga att *lära sig* (Davidson), för att man ska kunna förstå nya satser (vanlig uppfattning), för att kunna producera nya satser (Chomsky), samt tanken att språk är systematiska (i en viss bestämd bemärkelse) och att detta förklaras av att de är kompositionella (Fodor). I samtliga fall, kommer jag att argumentera, är det antingen så att alltför starka antaganden görs, eller så att argumentet mer eller mindre implicit förutsätter framgångsrik kommunikation med nya satser, varför det i så fall kollapsar till ett kommunikationsargument.

Att detta är fallet är kanske mest uppenbart hos Chomsky, eftersom argumentet egentligen inte bara är att vi kan generera nya satser, utan att vi kan generera nya satser som kan *förstås* av en lyssnare. Det är alltså nogräknat endast en formuleringsvariant av kommunikationsargumentet.

Både inlärningsargumentet och argumentet om nya satser, om de tolkas strängt, bygger på ett mycket starkt antagande, nämligen antagandet att språkliga satser som ingen ännu använt redan *har* en välbestämd mening. Detta är ett antagande som är så starkt att argumentet för kompositionalitet blir mycket svagt om inte antagandet ges ett oberoende rättfärdigande. Varför kan vi inte lika gärna säga att språkliga uttryck *ges* en mening när de väl används, och att de inte har någon dessförinnan? Detta alternativ måste på något sätt uteslutas. Problemet här är att det enda rimliga rättfärdigandet av påståendet att satser har en välbestämd mening i förväg *just är* att språket har en kompositionell semantik. I brist på ett gott oberoende rättfärdigande blir argumentet därför antingen cirkulärt eller alldeles för svagt.

Systematicitetsargumentet kräver en lite mer komplicerad argumentation, som jag inte ska försöka presentera här. Till detta kan läggas det förslag jag argumenterat mot i Pagin 1999, nämligen att Davidsons idéer om radikal tolkning skulle kunna rättfärdiga kompositionalitet, och här blir min argumentation densamma som tidigare.

## 2.7 Föregivna skäl mot kompositionalitet hos naturliga språk

Invändningar mot påståendet att naturliga språk har en kompositionell semantik kan delas in i fem grupper. Den första typen av invändning går ut på att kompositionalitetsprincipen är *empiriskt tom*. Att vara empiriskt tom skulle innebära att det inte gör någon empirisk skillnad om ett språk är kompositionellt eller inte. I flera fall har argument av denna typ underbyggts med ett formellt argument för att visa att vilken semantik som helst för ett språk, dvs även en icke-kompositionell semantik, i någon viss mening kan ligga till grund för eller omformuleras till en kompositionell semantik. Slutsatsen är nu att alla språk kan kallas kompositionala, oavsett vilken semantik de har. Alltså gör det ingen empirisk skillnad.

De formella argumenten (som hos Jansen och Zadrozny) är oantastliga, men slutsatserna som dragits ur dem är för starka. Det enda som är korrekt är att det inte är *tillräckligt* för en semantik att formellt uppfylla kompositionalitetsvillkoret för att också vara empiriskt intressant. Att detta är fallet kan enkelt visas, eftersom en semantik för att språk L automatiskt är kompositionell om det inte finns några synonymipar i L. För att

hitta ett motexempel mot kompositionalitetsprincipen krävs nämligen att två synonyma uttryck  $e$  och  $e'$  kan bytas ut i en kontext  $A$ , så att resultatet av substitutionen,  $A'$ , inte är synonym med  $A$ . Om det inte finns några synonyma uttryck  $e$  och  $e'$  finns heller inga motexempel, och principen kan betraktas som giltig för  $L$ , även om det är helt omöjligt att förstå nya komplexa  $L$ -uttryck utan att först få dem förklarade för sig. För att kompositionalitet skall vara förklaringsmässigt relevant krävs att mening hos komplexa uttryck är bestämda på ett visst *enhetligt sätt*. Vilken meningsfunktion som helst duger inte, men att så är fallet utelämnas ur normala *definitioner* av kompositionalitetsbegreppet. Det är just detta som de formella argumenten utnyttjar. Deras resultat är helt förenliga med att vi kan behöva *adekvat* kompositionell semantik i våra teorier om naturliga språk. Liknande saker har påpekats, t ex i Westerståhl 1998. Vilken empirisk roll kompositionalitetsprincipen spelar för naturliga språk försöker jag förklara just i de fyra första kapitlen.

Den andra typen av invändning går ut på att vi kan redogöra för kommunikation utan hänvisning till kompositionell semantik. Den ena undertypen här är t ex som hos Jaakko Hintikka, enligt vilken vi visserligen har en systematisk semantik, men en som inte är kompositionell i normal, strikt mening. Detta har jag diskuterat i kapitel fem.

Den andra undertypen är den enligt vilken vi inte behöver någon systematisk strukturell semantik överhuvud taget. Allt som behövs är pragmatiska principer, möjligen tillsammans med grammatik och ordmening: semantik för samansättningsformer behövs inte. Invändningar av detta slag underbyggs ofta med hänvisning till hur vi på ett kreativt sätt lyckas förstå yttranden i fall där kompositionalitet inte kan förklara våra prestationer.

Att normala språkanvändare har en psykologisk-pragmatisk förmåga att kommunicera är uppenbart riktigt, och att vi använder den dagligen för att tala och förstå är oundvikligt, liksom det faktum att den hjälper oss att förstå där kompositionell semantik inte kan förklara kommunikativ framgång. Men detta är fullt förenligt med att pragmatisk förmåga *enbart* inte räcker för att förklara förståelse (t ex vid läsning av dagstidningsartiklar). För att ge en fullständig förklaring av vår kommunikativa förmåga räcker pragmatiken inte till, utan en semantisk teori krävs.

En variant av denna invändning går ut på att vår förmåga att tala och tolka inte består i att följa vissa bestämda *regler* som den semantiska teorin formulerar. Men denna invändning bygger på ett missförstånd om hur relationen mellan teori och språkanvändning måste vara beskaffad. I min redogörelse i de fyra första kapitlen talas ingenstans om hur språkanvändare följer semantiska regler. Teorin är deskriptiv till skillnad från normativ, och de semantiska principerna formuleras för att *förklara* talarnas dispositioner till att tala och tolka, inte för att utgöra principer talarna måste *känna till* eller medvetet försöka *följa*. Teorin beskriver ytterst komplexa regelbundenheter i talarnas språkliga dispositioner, men säger ingenting om vilka kognitiva förmågor eller mekanismer som underliggör dessa dispositioner. Framför allt sägs ingenting om regler. Med detta visas också att invändningar mot kompositionalitet grundade på Wittgensteins överväganden om regelföljande inte har någon direkt relevans (se Pagin *under utgivning*).

Den tredje invändningen går ut på att kompositionella teorier är falska, därför att de meningsentiteter, eller semantiska värden, som krävs för att göra den semantiska teorin sann, helt enkelt inte existerar. Stephen Schiffer har argumenterat för att semantiken för psykologiska kontexter, som i

(8) Nils tror att barnläkare är snälla

kräver existensen av s k *presentationssätt* (först föreslagna av Frege), men att inga existerande entiteter kan spela den roll som krävs för att *vara* presentationssätt. Alltså existerar inte presentationssätt, och alltså finns ingen korrekt kompositionell semantik för satser som (8). Mitt svar på detta består i att ge en alternativ teori om vilka villkor en semantik för satser som (8) måste uppfylla. Se nedan.

En fjärde invändning går ut på att det sätt mening faktiskt bestäms för uttryck i naturliga språk är oförenligt med kompositionalitet, nämligen för att meningsbestämningen är holistisk. Att en rimlig form av semantisk holism är förenlig med kompositionalitet har jag emellertid utförligt argumenterat för i Pagin 1997.

Den femte invändningen, till sist, går ut på att peka på *motexempel* mot kompositionalitetsprincipen för naturliga språk. Dessa är huvudsakligen av två slag, ett strikt och ett icke-strikt. Det icke-strikta slaget är av följande allmänna karaktär: vi har ett komplext uttryck av ett visst synbarligen syntaktiskt välbestämt slag, bestående av delar med känd mening, och det komplexa uttryckets delar följer *inte* ur delarnas mening och sammansättningsformen. Exempel på sådana uttryck, som anförts som motexempel, är idiom, som t ex 'kila runt hörnet' (i dess idiomatiska mening). I en viss bemärkelse är exemplen korrekta, vilket också visar sig i att man just inte kan förstå uttryckens idiomatiska mening bara genom förståelse av delar och sammansättning (lägg märke till att detta talar för, inte emot, kopplingen mellan kompositionalitet och förståelse). Men därur följer inte att det inte är möjligt att *utvidga* en kompositionell teori till att omfatta idiom, utan att den därmed upphör att vara kompositionell. Detta är möjligt, och olika metoder för att göra detta har föreslagits och diskuterats i Westerståhl (*under utgivning*).

Den strikta typen av motexempel avser visa hur substitution av synonyma uttryck kan leda till icke-synonyma resultat. Huvudexemplet är just satser som (8). Det har hävdats att sådana satser utgör motexempel mot kompositionalitetsprincipen, eftersom vi kan byta ut ett uttryck mot ett synonymt uttryck utan att bevara sanningsvärdet. Det kan vara så att Nils tror att ordet 'pediatriker', som faktiskt är synonymt med 'barnläkare', i själva verket står för aggressiva cykelreparatörer. Då kan, enligt invändningen mot kompositionalitet, satsen

(9) Nils tror att pediatriker är snälla

vara falsk, trots att satsen (8) själv är sann, och trots att vi erhållit (9) ur (8) genom substitution av synonyma uttryck. Lösningen på detta problem är följande. För det första är det helt enkelt rimligt att hävda att om (8) är sann, så är (9) också sann, även om Nils t ex svarar nekande på *frågan* 'Är pediatriker snälla?'. Det är *talaren* av satsen (9) som använder ordet 'pediatriker', med dess vanliga mening, för att säga något om vad Nils tror, och att då Nils inte vet vad 'pediatriker' betyder spelar ingen roll.

Om detta åter bemöts av invändningen att det *visst* spelar roll hur Nils förstår ordet 'pediatriker', så måste slutsatsen bli att den talare som *använder* satsen (9) i enlighet med denna invändning avser att *andra* egenskaper hos ordet 'pediatriker' än just dess mening är relevanta för att bestämma dess sanningsvillkor. T ex kan det vara relevant vad Nils tror att 'pediatriker' betyder. Men detta innebär i sin tur att en analys av (9), såsom

använd på detta sätt, visar att ordet 'pediatriker' egentligen förekommer på två sätt, dels på vanligt sätt, där det helt enkelt uttrycker sin vanliga mening, dels underförstått också citerat, på ett sätt analogt med 'Giorgione' i Quines berömda exempel:

(10) Giorgione was so called because of his size

vilken betyder detsamma som

(11) Giorgione was called 'Giorgione' because of his size.

Att vi inte utan vidare kan byta ut 'Giorgione' i (10) mot något annat namn på samma person förklaras av analysen (11), enligt vilken ordet förekommer både som använt på vanligt sätt och som citerat, fastän detta inte syns på ytan. Så måste också analysen bli av (9), såsom använd i enlighet med invändningen.

Den grundläggande svårigheten med psykologiska satser som (8) och (9) är att det inte finns någon enhetlig användning av dem. Istället finns en familj av *liknande* användningar. Det är också en familj utan skarpa och definitiva gränser. Varje enskild användning kan ges en kompositionell analys, men ingen enskild analys kan hantera alla tänkbara användningar av sådana satser. Så i ett avseende är det möjligt, och i ett annat omöjligt, att ge en kompositionell teori för psykologiska satser. Att det i ett avseende är omöjligt visar inte på att kompositionalitetsprincipen är falsk, endast på att ingen enskild kompositionell teori kan redogöra för all mänsklig språkanvändning. Mänsklig språkanvändning är *bland annat* genuint kreativ, dvs skapar nya innebörder som inte tidigare var systematiskt förberedda. I sådan fall krävs också en kreativ insats av lyssnaren för att kommunikationen ska lyckas. Ju mer kreativitet som krävs, desto större är emellertid risken för missförstånd.

Jag har idéer för en generell teori om psykologiska satser som består i en enkel metod att *generera* en semantisk analys utifrån de substitutionsvillkor talaren avsett för ord som förekommer efter det psykologiska verbet. Det är emellertid oklart om det finns utrymme för den i kapitlet, eller i boken (psykologiska satser har jag tidigare arbetat med i Pagin 1992).

### 3. Arbetsplan

Planen är enkel. Jag avser att skriva kapitel 1-3 under det första året, kapitel 4-6 under det andra året, samt det sjunde kapitlet, tillsammans med färdigställande av monografin, under det tredje.

### 4. Egna förarbeten samt omnämnda skrifter

Preliminära tankegångar har presenterats som föredrag, relaterade till: kap 1 vid Uppsala universitet i september 2000 och vid Tjeckiska vetenskapsakademien, Prag, i december 2000; kap 2 och 5 (Schiffer) vid Rutgers University i oktober 1999; kap 3 på kompositionalitetssymposiet vid *European Summer School of Logic, Language and Information* i Aix-en-Provence i augusti 1997, samt vid Uppsala universitet i maj 2000; kap 4 och 5 (tyst kunskap) vid University of Edinburgh i maj 1998. Som bakgrund till de mer tek-

niska delarna i kapitel 3 och 4 finns också avslutade arbeten tillsammans med Dag Westerståhl (Pagin och Westerståhl 1993, 1994). En mindre del av diskussionen planerad för kap 6 finns utförd i Pagin 1999. Vidare finns frågor för diskussion i kap 7 behandlade i Pagin 1997, vad gäller holism, samt i Pagin (under utgivning), vad gäller regelföljande.

Preliminära texter utöver dessa, relaterade till senare års föredrag, är bl a 'Compositional explanation and tacit knowledge', 'Frege, communication and strong compositionality' samt 'When does communication succeed?'

*Egna omnämnda skrifter* (för referenser, se förteckning över publicerade arbeten)

Pagin, P, 1992, 'Names in and out of thought'.

Pagin, P, 1997, 'Is holism compatible with compositionality?'

Pagin, P, 1999, 'Radical interpretation and compositional structure.'

Pagin, P, *under utgivning*, 'Rule-following, compositionality and the normativity of meaning'.

Pagin, P, *under utgivning 2*, 'Mening hos yttranden', *Filosofisk Tidskrift*.

Pagin, P och Westerståhl, D, 1993, 'Predicate logic with flexibly binding operators and natural language semantics'.

Pagin, P och Westerståhl, D, 1994, 'Flexibly binding operators and Montague grammar'.

## 5. Allmänna meriter

Ifråga om allmänna meriter kan nämnas att jag är docent sedan 1992 och att jag 1999 ansökt om befordran till professor. En sakkunnig har yttrat sig positivt, en har ännu inte yttrat sig.

## 6. Kostnader och forskningsinformation

Kostnaden för projektet är utgörs av lön, omkostnadsavgifter samt mindre summor årligen för material och konferensresor.

Informationen om forskningen avses i första hand ske genom internationell publicering av monografin (Oxford University Press har uttryckt intresse för mitt arbete), i andra hand genom facktidskriftartiklar, populära artiklar (som Pagin *under utgivning 2*) och föredrag (som t ex föredraget 'Vad är språk?' vid *Språkårets* föreläsningsserie *På Tal om Språk*).

## **Bilaga: Övriga litteraturhänvisningar**

- Barwise, J and Perry, J, 1983, *Situations and Attitudes*, MIT Press, Cambridge, Mass.
- Block, N, 1986: 'Advertisement for a semantics for psychology', *Midwest Studies in Philosophy* 10, pp. 615-78.
- Carnap, R, 1947, *Meaning and Necessity*, Chicago University Press, Chicago.
- Cheney, D L and Seyfath R M, *How Monkeys See the World*, University of Chicago Press, Chicago 1990.
- Chomsky, N, 1971, 'Topics in the theory of generative grammar', i Searle (red): *The Philosophy of Language*, Oxford University Press, Oxford.
- Davidson, D, 1984, *Inquiries into Truth and Interpretation*, Clarendon Press, Oxford.
- Davies, M, 1981, *Meaning, Quantification, Necessity*, Routledge and Kegan Paul, London.
- Dummett, M, 1976, 'What is a theory of meaning (II)' i G Evans och J McDowell J (red), *Truth and Meaning*, Oxford University Press, Oxford.
- Evans, G. 1980, 'Semantic theory and tacit knowledge', i Holtzman and Leich (red), *Wittgenstein: to Follow a Rule*, Routledge and Kegan Paul, London.
- Fodor, J, 1992, *A Theory of Content and Other Essays*, MIT Press, Cambridge, Mass.
- Frege, G, 1892, 'Über Sinn und Bedeutung', *Zeitschrift für Philosophie und Philosophische Kritik*, NF 100, sid 25-50.
- Frege, G, 1918, 'Gedankengefüge', *Beiträge zur Philosophie des deutschen Idealismus* 3:36-51.
- Grice, P, 1989, *Studies in the Ways of Words*, Harvard University Press, Cambridge, Mass.
- Harman, G, 1987, 'Conceptual Role Semantics', i E LePore (red): *New Directions in Semantics*, Academic Press, London.
- Hempel, C G, 1965, *Aspects of Scientific Explanation*, Free Press, New York, NY.
- Hendriks, H, 2001, 'Compositionality and model-theoretic interpretation', *Journal of Logic, Language and Information* 10, sid 29-48.
- Hodges, W, 2001, 'Formal features of compositionality', *Journal of Logic, Language and Information* 10: 7-28 (gästredaktörer Peter Pagin och Dag Westerståhl).
- Jansen, T, 1986, *Foundations and Applications of Montague Grammar*, Centrum voor Wiskunde en Informatica.
- Jansen, T, 1996, 'Compositionality', i J van Benthem och A ter Meulen (red), *Handbook of Logic and Language*, Elsevier, Amsterdam, sid 417-473.
- Salmon, W, 1984, *Scientific Explanation and the Causal Structure of the World*, Princeton University Press, Princeton, NJ.
- Kamp H, 1993, *From Discourse to Logic. Introduction to Modeltheoretic Semantics of Natural Language, Formal Logic and Discourse Representation Theory*, Kluwer, Dordrecht.
- Kitcher, P, 1989, 'Explanatory unification and causal structure', i P Kitcher och W Salmon (red), *Scientific Explanation, Minnesota Studies in the Philosophy of Science*, XIII, University of Minnesota Press, Minneapolis.

- Larson, R och Segal, G, 1995, *Knowledge of Meaning. An Introduction to Semantic Theory*, MIT Press, Cambridge, Mass.
- Lewis, D, 1969, *Convention. A Philosophical Study*, Harvard University Press, Cambridge, Mass.
- Locke, J, 1689, *An Essay Concerning Human Understanding*.
- Montague, R, 1974, *Formal Philosophy*, R H Thomason (red), Yale University Press, New Haven.
- Quine, W V O, 1953, *From a Logical Point of View*, Harvard University Press, Cambridge, Mass.
- Quine, W V O, 1960, *Word and Object*, MIT Press, Cambridge, Mass.
- Schiffer, S, 1987, *Remnants of Meaning*, MIT Press, Cambridge, Mass.
- Shannon, C E and Weaver, W, 1949, *The Mathematical Theory of Communication*, University of Illinois Press, Chicago, Ill.
- Westerståhl, D, 1998, 'On mathematical proofs of the vacuity of compositionality', *Linguistics and Philosophy* 21, sid 635-43.
- Westerståhl, D, *under utgivning*, 'On the compositionality of idioms' under utgivning i D Barker-Plummer, D Beaver, J van Benthem och P Scotto di Luzio (red), *Proceedings of LLC8*, CSLI Publications, Stanford, CA.
- Wittgenstein, L, 1953, *Philosophical Investigations*, Basil Blackwell, Oxford.
- Woodward, J, 2000, 'Explanation and invariance in the special sciences', *British Journal for the Philosophy of Science* 51, sid 197-254.
- Wright, C, 1987, 'Theories of meaning and speaker's knowledge', i Wright, *Realism, Meaning and Truth*, Basil Blackwell, Oxford.
- Zadrozny, W, 1994, 'From compositional to systematic semantics', *Linguistics and Philosophy* 17, sid 329-342.
- Zeevat, H, 1989, 'A compositional approach to discourse representation theory', *Linguistics and Philosophy* 12, sid 95-131.

## Viktigaste egna publicerade arbeten

1. *Ideas for a Theory of Rules*, doktorsavhandling, 1987.
2. 'Names in and out of thought', *Philosophical Studies* 66:27-51, 1992.
3. 'Predicate logic with flexibly binding operators and natural language semantics', med Dag Westerståhl, *Journal of Logic, Language and Information* 2:89-128, 1993.
4. 'Knowledge of Proofs', *Topoi* 13:93-100, 1994. Under återutgivning i D Jacquette (red), *Philosophy of Mathematics*, Blackwell.
5. 'Flexible variable-binding and Montague Grammar', med Dag Westerståhl, i P Dekker och M Stokhof (red), *Proceedings of the 9th Amsterdam Colloquium*, Amsterdam, 519-525, 1994.
6. 'Is compositionality compatible with holism?', *Mind and Language* 12:11-33, 1997.
7. 'Bivalence: meaning theory vs metaphysics', *Theoria* 64:157-186, 1998.
8. 'Radical interpretation and compositional structure', i Zeglen (red) *Discussions with Donald Davidson; Truth, Meaning and Knowledge*, Routledge 1999. Även i polsk översättning i Zeglen (red), *Dyskusje Z Donaldem Davidsonem: O Prawd-  
zie, Jezyku I Umysle*, Lublin 1996.
9. 'Rules of meaning and practical reasoning', med Kathrin Glüer, *Synthese* 117: 207-27, 1999.
10. 'Publicness and indeterminacy', i P Kotatko och A Orenstein (red), *Knowledge, Language and Logic: Questions for Quine*, Boston Studies in the Philosophy of Science, Kluwer, Dordrecht, 2000.
11. 'Att förstå vad andra menar; frågeställningar kring kommunikation och språklig mening', i K Dahlbäck (red), *Att Förstå Det Mänskliga*, Natur och Kultur, Stockholm 2000.
12. 'Sensation terms', *Dialectica* 54:177-199, 2000.
13. 'A Quinean definition of synonymy', under utgivning i *Erkenntnis* 2001.
14. 'Semantic triangulation', under utgivning i P Kotatko, P Pagin och G Segal (red), *Interpreting Davidson*, CSLI Publications, Stanford, CA, 2001.
15. 'Rule-following, compositionality and the normativity of meaning', under utgivning i D Prawitz (red), *Meaning and Interpretation*, Vitterhetsakademiens skriftserie.