

Manuskript för:

Nordand, Nordisk tidsskrift for andrespråksforskning, Årg. 3, Nr 1-2008

## **Konstruktioner som produkt och process – en studie av hur L1- och L2-talare utnyttjar ”det är”**

**Björn Hammarberg**

### **SAMMANFATTNING**

Språkliga konstruktioner kan ses som produkter eller processer. Artikeln utforskar samband mellan dessa aspekter och fokuserar på L1- och L2-talares bruk av svenska konstruktioner bildade med det starkt flerfunktionella syntaktiska fragmentet *det är*. En användningsbaserad modell för språkets utveckling läggs till grund, där särskilt också frekvensens roll uppmärksammas. Data hämtas från en korpus av konversationer med inlärare och infödda svenskar, som uppvisar en riklig användning av olika *det är*-konstruktioner, såväl med avseende på bredden i repertoaren som i fråga om frekvens. Inlärarna tillägnar sig successivt ett användningsmönster som i stort liknar de infödda talarnas. Vissa strukturer uppträder tidigare än andra; en karakteristisk utvecklingsprofil framträder. Sekvensen *det är* har en central roll i yttrandeproduktionen, där den ofta tjänar som en (ibland bara tentativ) satsinledare. Det finns indikationer på att *det är* tillägnas och fungerar som en formelsekvens med en stödjande roll i yttrandeplaneringen. Av speciellt intresse är de formuleringsbrott som ofta uppträder efter *det är*, med olika sätt att sedan fortsätta yttrandet. Dessa fall ger inblickar i det löpande formuleringsarbetet.

### **ABSTRACT**

Linguistic constructions can be viewed as products or processes. The paper explores connections between these aspects, focussing on L1 and L2 speakers' use of Swedish constructions formed with the highly polyfunctional syntactic fragment *det är* 'it is'. A usage-based model of language development is adopted, and the role of frequency of use is taken into account. The data for the study were extracted from a corpus of Swedish L1 and L2 conversational speech. This displays a rich use of different *det är* constructions, both in terms of the width of the repertoire and in terms of the frequency of use. The learners gradually acquire a pattern of usage which is largely similar to that of the native speakers, some structures occurring earlier than others; there appears a characteristic developmental profile. *Det är* has a prominent role in utterance production, often serving as a (sometimes only tentative) sentence introducer. There are indications that *det är* is acquired and functions as a formulaic sequence with a role in supporting utterance planning and execution. A particularly interesting phenomenon is the frequent occurrence of a formulation break after *det är*, with various ways of continuing the utterance; this provides insights into the ongoing formulation work.

# INTRODUKTION<sup>1</sup>

Språkliga konstruktioner låter sig studeras såväl i ett produktperspektiv som i ett processperspektiv. I det första fallet, det som är föremål för grammatiska beskrivningar, rör det sig kognitivt sett om mentala representationer av strukturer lagrade i långtidsminnet till förfogande för talare och lyssnare. I det andra fallet gäller det själva konstruerandet i realtid av yttranden med viss språklig struktur. Man kan ställa frågor dels om människors repertoar av konstruktioner, dels om hur dessa hanteras i formuleringsförloppet.

Under senare år har frekvensens roll i språkanvändningen och språktillägnet fått ökad uppmärksamhet. Olika forskare har diskuterat frekvensens betydelse för hur språklig struktur etableras och befästs hos individen (Bybee och Hopper 2001a; Ellis 2002). I en användningsbaserad modell (usage-based model; Langacker 1988, 1999; Kemmer och Barlow 1999) för hur språk utvecklas hos individen betonas att upprepat bruk av specifika språkliga uttryck leder till att de successivt blir befästa som kognitiva rutiner hos talaren. Talarens språkliga system är enligt denna teori användningsbetingat; det grundas, formas och omformas efterhand genom erfarenhet av konkreta ”usage events”, tillfällen då denne producerar eller tar emot och förstår språkliga uttryck. Frekvent användning främjar denna utveckling.

Frekvens har dock i sin tur en orsak. Det måste finnas något motiv eller några faktorer som leder till att ett språkligt uttryck används med en viss grad av relativ frekvens. Behovet att kommunicera vissa tankeinnehåll, eller det praktiska i att uttrycka språkliga budskap i viss form kan utgöra sådana faktorer. Om det aktuella språkliga uttrycket är en grammatisk konstruktion, så kan denna ha funktionella egenskaper som gör att den ofta kommer till användning i yttrandeformuleringar. Sambandet mellan funktionalitet och frekvens hos språkliga konstruktioner blir därmed intressant. Ett närmare studium av funktionerna hos högfrekventa språkliga strukturer kan ha en del att bidra till vår förståelse av hur språklig kompetens byggs upp.

Till detta kommer så processperspektivet på konstruktionen av yttranden. Ett studium av yttranden utifrån vad Lindström (2008: 43ff.) kallar *processyntax* belyser hur deras uppbyggnad hanteras i formuleringsprocessen. Språkliga uttrycks funktionella potential och sättet att använda dem i talförloppet hänger samman.

I det följande ska jag försöka demonstrera detta på ett område av svenskan genom att undersöka infödda talares och andraspråksinlärares användning av konstruktioner som bildas

med byggstenen *det är*. Predikationer av formen 'det är X' är oerhört frekventa i svenskan, i tal ännu mer än i skrift. I en ordräkning baserad på svenskt talkorpusmaterial utfaller *det* som det frekventaste ordet och *det är* som den mest frekventa tvåordskombinationen (Allwood 1999). Vad som också frapperar när man studerar användningen av sekvensen *det är*, är att den är en i hög grad flerfunktionell enhet i språket. Jag ska återkomma till varför jag betraktar *det är* som en enhet. Primärt kan emellertid den största delen av den grammatiska flerfunktionaliteten härledas till det pronomina elementet, *det*. Svenskan ingår typologiskt, tillsammans med de övriga germanska språken samt franskan och rätoromanskan, i den västeuropeiska gruppen av *platshållarspråk* (Beckman 1934; Hammarberg och Viberg 1977), vilket leder till ett frekvent bruk av neutrala pronomen och formella subjekt. *Det* uppträder i olika syntaktiska positioner, som subjekt, predikativ, objekt eller styrt av preposition, och kan kombineras med många olika verb. I anaforskt bruk kan det neutrala pronomenet *det* ofta användas för att på ett enkelt sätt syfta på komplexa enheter i den tidigare diskursen, genom att det möjliggör vag eller underspecificerad referens. Som formellt subjekt ingår det som platshållare i olika syntaktiska konstruktioner. En fallstudie av bruket av *det* i olika funktioner tyder på att det används frekvent även i inlärarspråk och tidigt uppträder i flera slags konstruktioner (Hammarberg 2000).

Den här studien koncentrerar sig således på en delmängd av *det*-användningar i och med att vi fokuserar enbart på yttranden bildade med sekvensen *det är*. Den förekommer i varierande kontexter, men särskilt ofta i satsens början, för att introducera en predikation av något slag. Detta ger *det är* en roll som ett flexibelt medel i den pågående formuleringsprocessen, eftersom det i många fall finns flera alternativa möjligheter att fullborda yttrandet efter *det är*, och det syntaktiska valet hur man vill fortsätta ofta fortfarande är öppet i detta läge. Det gör att det kan vara användbart att starta ett yttrande med *det är*, medan fortsättningen planeras. Som jag ska visa mer i detalj längre fram, finns det skäl att anta att *det är* fungerar som en enhet, d.v.s. som en *formelsekvens* (jfr t.ex. Wray och Perkins 2000) med en viktig roll att stödja den löpande yttrandeproduktionen.

I de följande avsnitten ska bruket av *det är* hos L1- och L2-talare av svenska undersökas närmare på grundval av en textkorpus av inspelade samtal. Huvudfrågorna är dessa:

I. I vilka grammatiska kontexter uppträder *det är* hos infödda talare och inlärare?

- II. Med vilken kvantitativ fördelning uppträder de olika *det är*-konstruktionerna hos infödda och inlärare och på olika stadier av inlärarspråk? Framträder en utvecklingsprofil?
- III. Vilken roll spelar *det är* i den löpande planeringen och utförandet av yttranden?
- IV. Kan man se en roll för *det är* i utvecklingen av inlärarspråket?

## DATAUNDERLAG

Data för den här studien är hämtade ur *ASU-korpusen*, som sammanställdes på 1990-talet vid Institutionen för lingvistik, Stockholms universitet. Den omfattar samtalstexter och skriftliga uppsatser insamlade från svenska och utländska studenter i Stockholm. Korpusen beskrivs utförligt av Hammarberg (2006). Den nu aktuella undersökningen utnyttjar det muntliga materialet från sex infödda talare och sex inlärare.

Inlärardelen av korpusen är longitudinellt anlagd. Informanterna deltog i universitetets preparandkurs i svenska för utländska studerande. Samtal mellan två svenska intervjuare och en inlärare åt gången spelades in fristående från kursen vid tio tillfällen mellan september 1990 och mars 1992. De omfattade olika språkliga aktiviteter såsom intervjuer, diskussioner, bildberättelser, beskrivningar av foton och föremål etc. Korpusen speglar en utveckling i svenska per individ från nybörjarstadiet då deltagarna nyligen anlät till Sverige och påbörjat kursen, till ett stadium då de studerade på svenska vid svenska högskolor. Den dokumenterar därmed framväxten av ett utvecklat inlärarspråk från tidigaste former. De deltagare som ingår i denna undersökning är en kvinna och fem män, i ålder mellan 19 och 22 år vid starten av perioden. Två hade kinesiska, två grekiska och två portugisiska som L1. Alla hade också L2-kunskaper i engelska och ett eller två ytterligare språk. Dessa inlärare kan i grova drag karakteriseras som *'semi-formella'* (då de fick sin språkliga input dels från kursen och dels från den svenska språkmiljön i Stockholm), *'kvalificerade'* (i den meningen att de alla hade gymnasieutbildning, erfarenhet av främmande språk och en stark instrumentell motivation att lära sig svenska för att kunna bedriva studier inom sina fackområden), och relativt *'snabba'* (i och med att de avancerade från nybörjarstadiet till eller näst intill den nivå som krävs för högskolestudier inom ett till två läsår). De sex inlärarnas muntliga produktion omfattar

sammanlagt ca 75 000 ord, vilket ger en genomsnittlig textlängd per informant om ca 12 500 ord.

Korpusdelen från svenska informanter konstruerades för att bli så jämförbar som möjligt med inlärarkorpusen med avseende på typ av informanter, innehåll och metoder för datainsamling; en skillnad är naturligtvis att den inte återspeglar en språkutveckling över tiden. Deltagarna var sex studenter i grundutbildning i icke-språkliga ämnen, varav fyra kvinnor och två män i ålder mellan 20 och 29 år. De talade en centralsvensk standardvarietet, sådan man hör i Stockholmsregionen; dock hade en av dem (Z5, nedan) spår av en norrbottnisk bakgrund. Fem samtal spelades in med liknande innehåll som för inlärarna. Textlängden för de sex svenska informanterna blev ca 77 000 ord, d.v.s. genomsnittligt ca 12 800 ord per person.

Ur den muntliga produktionen från dessa tolv personer har samtliga förekomster av sekvensen *det är* sökts fram och analyserats i sina kontexter. Detta ligger till grund för den kategoriindelning av *det är*-konstruktionerna som behandlas i nästa avsnitt. Grammatiskt besläktade sekvenser med *det + vara*, såsom variationer i tempus (*det var*, *det skulle vara*), inverterad ordföljd (*är det*) och mellanliggande element (*det inte är*, *det kanske är*) har däremot inte tagits med.

## ***DET ÄR I OLIKA KONTEXTER I INFÖDDAS OCH INLÄRARES TAL***

För att kunna se ett mönster i hur infödda talare och inlärare gör bruk av konstruktioner med *det är* ska vi först upprätta en taxonomi över de olika slags kontexter där *det är* uppträder i korpusen, och sedan se hur dessa strukturer fördelar sig kvantitativt hos infödda och inlärare. *Det är*-konstruktioner utgör då, som vi var inne på tidigare, ett specialfall av olika konstruktioner med det neutrala pronomenet *det*. (För utförliga beskrivningar av *det*-användningar, se grammatiska handböcker som Teleman m.fl. 1999; Thorell 1973; Holmes och Hinchliffe 2003; för formellt subjekt-konstruktioner också Sundman 1980; för bruket av *det* i inlärarspråk Hammarberg 2000.)

### **En taxonomi för konstruktioner med *det är***

De olika användningarna av *det är* skiljer sig på olika sätt med avseende på diskursfunktion och syntaktisk form. Det är praktiskt att sortera dem i tre huvudgrupper:

- A. fall där pronomenet *det* har någon form av anaforisk eller deiktisk referens (strukturerna 1–11 nedan);
- B. fall där pronomenet antingen är icke-referentiellt eller syftar framåt i yttrandet (strukturerna 12–16 nedan);
- C. fall där *det är* följs av någon form av syntaktiskt avbrott (strukturerna 17–19 nedan).

Exempel ur korpusen på de olika strukturtyper som faller inom dessa tre huvudgrupper har jag samlat i listorna 1–3. I mån av förekomst ges exempel från både de infödda informanterna och inlärarna. Exempelen ges här i den transkription som används i korpusen, d.v.s. en modifierad ortografisk form, där talspråksformer av orden återges i den mån talarna använder dem. Observera särskilt att standardortografins *det är* i exemplen återges som *de e*. Några specialtecken som används i transkriptionen förklaras i Appendix.

Inlärarna betecknas i korpusen som C1, C2, G2, G3, Q1 och Q2, där bokstaven anger modersmål: C för kinesiska, G för grekiska och Q för portugisiska. De svenska informanterna betecknas Z1–Z6. Källhänvisningar till exemplen ges som en identifikation av talare och inspelningstillfälle, t.ex. Z4-1 för talare Z4, tillfälle 1, C2-6 för talare C2, tillfälle 6. Siffran för tillfälle ger för inlärarna en antydning om det aktuella inlärningsstadiet.

----- Lista 1 ungefär här

Huvudgrupp A: Anaforiskt eller deiktiskt *det* (lista 1)

I fallen med anaforiskt eller deiktiskt *det* syftar pronomenet på något som i någon form är givet i den föregående diskursen eller i situationen. I alla fallen här utom det sista (struktur 11) introducerar *det är* en predikation av något slag, och den fortsatta satsen kan ha varierande form. I struktur 1 följs *det är* av en NP, som kan vara enkel eller komplex, definit eller indefinit, substantivisk eller pronominaliserad. I allmänhet fungerar *det* som subjekt, men ibland utgör det en framförställd predikativ, som i *nä de e han inte* (Z4-3). I den kvantitativa jämförelsen i nästa avsnitt räknar vi fall med en relativsats inbäddad i NP:n som en särskild undertyp. Andra slags konstituenten än NP uppträder efter *det är* i strukturerna 2–5.

De följande typerna utgör olika slags 'icke-kanoniska' fall. I struktur 6 är *det* prepositionsstyrt, och sekvensen *det är* uppkommer genom svenskans V2-ordföljd när *det* placeras först.

'Tough'-konstruktionen (struktur 7) har analyserats som att ett objekt till en infinitivfras har lyfts till subjekt, t.ex. *å säga de e hemskt svårt* → *de e hemskt svårt å säga*; se Sundman (1987: 114ff.). I den resulterande konstruktionen fungerar infinitiven som komplement till adjektivet: *svårt å säga*. Den anger i vilket avseende adjektivet är tillämpligt. Konstruktionen uppträder bara med vissa adjektiv. Majoriteten av våra exempel, från både infödda och inlärare, innehåller adjektivet *svårt* (ofta *tough* i motsvarande engelska konstruktion). 'Tough'-konstruktionen har jag valt att särskilja från typen *det är* + AdjP (struktur 2).

I struktur 8 ger ett spetsställt objektspronomen i kombination med svenskans V2-ordföljd upphov till sekvensen *det är*.

Vänsterdislokeringar med *det är*, som i struktur 9, är mycket vanliga i talspråk. Ett tematiskt element i satsen lyfts ut och ställs först och återupptas med pronomenet *det*, som får introducera satsens rematiska del. Den del som inleds med *det är* sammanfaller i formen med någon av strukturerna 1–8 ovan; här sammanför jag dock dislokationsfallen som en särskild strukturtyp.

Pseudoklyvningar med *det är* (struktur 10) är en särskild typ av vänsterdislokeringar, som jag också räknar separat. Här består den dislokerade delen av ett pronomen modifierat av en relativsats (*det som X*; *vad som X*). Pseudoklyvningar med ett återupptagande *det* är särskilt karakteristiska i informellt talspråk: *de som e svårast å måla de e ... snarare än de som e svårast å måla e ...*.

Till skillnad från de andra typerna i den här översikten fungerar *det är* i struktur 11 inte som introduktion till en predikation, utan uppträder i satsens slut där det fullbordar ett yttrande som vanligen redan är färdigplanerat på detta stadium.

Det faktum att det neutrala pronomenet kan referera på ett enkelt sätt till komplexa innehåll utan att vara onödigt explicit gör det mycket användbart i kommunikationen. Det neutrala anaforiska pronomenet *det* är ett mycket smidigt medel att skapa diskursreferens. Saken kan belysas om vi i anslutning till Brown och Yule (1983: 216ff.) antar att referens fungerar genom att upprätta en enhet i talarens och lyssnarens *diskursrepresentation*. Denna enhet, referenten, kan vara något föremål, eller kan utgöras av ett skeende, ett faktum, en proposition eller någon mer komplex information som kan härledas ur den föregående diskursen (Fraurud 1992). I några fall identifieras referenten explicit genom t.ex. ett nämnt substantiv, en infinitivfras eller en sats. I andra fall signaleras den åsyftade informationen inte uttryckligt i

den föregående kontexten, utan får härledas ur diskursen när den aktualiseras genom ett anaforiskt uttryck, som pronomenet *det*. Det vill säga, någonting i diskursen etableras som en enhet först när man refererar till det genom en anafor. Fraurud (2000) betecknar *det* i denna användning som en *hypostaseringsanafor*: ”tolkningsprocessen innebär ... att något som från början inte är introducerat som en diskursreferent s.a.s. ’förtingligas’ – hypostaseras – genom en nominal referens” (Fraurud 2000: 7). Anaforen *det* kan alltså hänföra sig till olika slags referenter, givna mer eller mindre explicit och i varierande syntaktisk form i kontexten, som i exemplen under (1), här citerade från Hammarberg (2000):

- (1) a. Han tycker om att dyka. Han är jätteintresserad av **det**.
- b. Jag jobbade på ett kafé ett halvår, men **det** blev alldeles för tråkigt.
- c. Känner du till den här boken? – Nej, **det** gör jag inte.
- d. **Det** [d.v.s. det som du har berättat nu] visste jag inte.

Ett exempel med *det är* ur korpusen är (2):

- (2) (Samtalsämne: Att diskutera ett foto av en historieberättare och en grupp lyssnande människor i en afrikansk miljö. E är intervjuaren, I är informanten.)  
E va får du för associationer när du ser den där bilden? du kanske tänker på paralleller.  
I en dagisgrupp tänker ja på.  
E jaha.  
I **de** e dom enda tillfällena tycker ja som som man samlar en grupp människor å berättar en historia. (Z2-5)

Här syftar *de* tillbaka på *dagisgrupp*, men inte på gruppen som sådan. Snarare är det tillfället och händelserna då en sådan grupp samlas, som hypostaseras som referent. Att talaren avser detta, gör hon tydligt i den fortsatta satsen genom *tillfällena* och de aktiviteter som hon nämner.

En viktig egenskap hos anaforisk referens är att den kan vara mer eller mindre starkt *underspecificerad*, d.v.s. enheten man refererar till kan vara bara vagt utpekad genom anaforen. Underspecificerad referens är framgångsrik om lyssnaren kan upprätta en diskursrepresentation som motsvarar talarens tillräckligt väl. Den här möjligheten att variera och reducera graden av precision i syftningen är en värdefull tillgång i språket. Å ena sidan



kan kontexten ofta bidra till att tillräckligt identifiera en referent; å andra sidan är stor precision ofta inte nödvändig eller ens önskvärd i sammanhanget. Ett illustrativt exempel är (3):

- (3) han % hyra % en del av v- villa me dom andra. och % vi laga mat tillsammans.  
% **de e** rolit tror ja. (C1-6)

Här kan man tolka det som talaren tyckte var 'roligt' på varierande sätt som 'att laga mat', 'att laga mat tillsammans', 'att vi lagade mat tillsammans', eller till och med 'att vi lagade mat tillsammans i hans villa'. Poängen med yttrandet *de e rolit tror ja* är snarast att uttrycka en attityd till det talaren varit med om, och graden av precision i referensen är inte det viktiga här.

När det gäller deiktisk referens, alltså när referenten ska sökas i den yttre situationen, är det snarast ännu mer uppenbart att referenten sällan är ett explicit utpekad föremål, utan oftast måste hypostaseras ur situationen med ett spelrum för vag referens.

Även om exemplen i lista 1 i de flesta fall är återgivna utan sin diskurs- eller situationskontext, kan man lätt inse att de här möjligheterna till hypostasering och utnyttjande av underspecificering är viktiga förutsättningar som gynnar användningen av *det* och *det är* i dessa konstruktionstyper.

----- Lista 2 ungefär här

Huvudgrupp B: Icke-refererande eller kataforiskt *det* (lista 2)

Det kan vara knepigt att skilja olika slags konstruktioner med icke-refererande eller kataforiskt *det*, i synnerhet som detta *det* genomgående har en likartad funktion som syntaktisk platshållare.

I struktur 12 uppträder *det är* i opersonlig konstruktion. Ett icke-refererande *det* krävs som subjekt vid opersonliga predikat av skilda slag (se Teleman m.fl. 1999, vol. 4: 58ff.). Många olika verb kan kombineras med opersonligt *det*. *Det är* uppträder med adjektiv (ex. a-b), particip eller adverbial (ex. c).

I sin undersökning av existentialsatser och liknande konstruktioner i svenskan listar Sundman (1980) följande kriterier för att identifiera en existentialkonstruktion, vår struktur 13: (i) ett tomt *det* i subjektposition, (ii) en NP som logiskt subjekt och (iii) ingen predikativ.

Hon framhåller också (1980: 19f.) att det logiska subjektet uppträder i objektsposition, att det logiska subjektet är indefinit, att det logiska subjektet i de flesta fall kan ersätta *det*, och att ett rumsadverb *här*, *där* ibland kan tjäna som platshållare i stället för ett formellt subjekt.

Funktionen hos denna konstruktion är att uttrycka det logiska subjektet som rematiskt. Den används typiskt för att introducera nya referenter i diskursen (Ekerot 1979: 96f.). Om det logiska subjektet innefattar en relativsats, som i (ex. a), så innehåller denna ny information. Detta särskiljer existentialkonstruktionen från satsklyvningskonstruktionen, som vi presenterar som struktur 16 nedan. Fall där det logiska subjektet innefattar en relativsats tas upp som en särskild undertyp i den kvantitativa översikten i nästa avsnitt.

I konstruktionen med ett anteciperande *det* (struktur 14) följs *det är* av ett predikativt adjektiv, och *det* föregriper ett postponerat led som utgör satsens logiska subjekt. Det postponerade leDET kan ha formen av en infinitivfras (ex. a) eller en bisats (ex. b).

Vid högerdislokeringar (struktur 15) uppkommer sekvensen *det är* som introducerande uttryck när ett neutralt singular subjekt i en predikativ sats med kopulan *är* flyttas till slutposition och efterlämnar pronomenet *det* som platshållare: ***de där e ju så himla omfattande*** → ***de e ju så himla omfattande de där***.

Satsklyvningar med inledande *det är* (struktur 16) bildas genom att man lyfter ut ett led från grundsatsen, introducerar det med *det är* och gör resten av satsen till en relativsats. Vanligtvis fungerar dessa satsklyvningar som emfatiska framhävningar av det utbrutna leDET, som blir betonat.

I fallen med icke-refererande eller kataforiskt *det* är en annan mekanism verksam än vid anaforisk och deiktisk referens. Det man talar om anges i yttrandet som följer, och det introducerande *det*, även om det bara är en tom platshållare, fungerar som en framåtppekare. Observera att det råder en samhörighet mellan bakåtppekande (anaforiska) och framåtppekande *det*-konstruktioner, en sorts alternering i sättet att ge information. En anaforisk konstruktion (särskilt en med vag referens) kan ofta ändras till en framåtppekande genom att det som åsyftas preciseras i den predikation som följer. Den här flexibiliteten illustreras i (4), som är ett modifierat exempel ur en dialog med inläraren C1:

(4) (Intervjuarens fråga:)

tyckte du att de va lätt att förstå henne?

(Alternativa svar:)

a. de e lättare än att prata

b. de e lättare att förstå än att prata

I (4a) måste *de* tolkas som ett anaforiskt pronomen, som syftar tillbaka på intervjuarens *att förstå*. Men om denna information sägs ut senare i satsen, som i (4b), så kommer *de* i stället att fungera som ett framåtpekande formellt subjekt. Observera att en framåtpekande *det*-konstruktion i regel är mer explicit än den motsvarande anaforiska konstruktionen. Valet mellan en bakåt- och en framåtpekande konstruktion är därmed ett viktigt sätt att reglera graden av explicititet i fall som dessa och hantera balansen mellan sparsamhet och tydlighet i formuleringen.

----- Lista 3 ungefär här

### Huvudgrupp C: Brott i formuleringen (lista 3)

Olika slags avbrott i den löpande formuleringen är en naturlig del av yttrandeproduktionen. Det är typiskt för spontant tal, där meningarna oftast inte är fullständigt planerade när de börjar yttras. Sådana avbrott påverkar satsstrukturen på varierande sätt.

En typisk plats i yttranden där brott ofta uppstår är efter ett inledande *det är*. Före brottet kan också komma ett obetonat modalt adverb inom samma prosodiska enhet: *det är ju, det är väl, det är liksom, det är inte* etc., och dessa räknar vi också in här.

Genom att ta fasta på vad som följer efter brottet ska vi skilja mellan fall av *upprepning*, *omformulering* och *avbrutet yttrande*, typer som gradvis innebär en mindre eller mer radikal inverkan på det aktuella yttrandet. Dessa avbrottsfenomen ger oss en viss inblick i den löpande formuleringsprocessen genom att de vittnar om tvekan eller omplanering på kritiska ställen i yttrandet. Blotta förekomsten av en paus, fylld eller ofylld, räknar vi emellertid inte, även om detta också helt klart är ett tvekansfenomen. De ställen i yttranden där pauser uppträder verkar variera mer och inte vara så typiskt förknippade med *det är*.

I struktur 17 upprepas *det är* (+ ModAdv), med eller utan paus emellan, innan yttrandet går vidare. Brottet kan tolkas som ett ögonblick av fortsatt planering av ett ännu inte fullt planerat yttrande. Även om ingen omformulering är märkbar i den yttrade meningen, så kan naturligtvis en latent förändring av det tänkta yttrandet ha ägt rum. Observera att det är fragmentet *före* brottet som vi räknar här; den efterföljande delen som inleds med *det är* kategoriseras för sig och räknas inom den strukturtyp under A eller B ovan som är tillämplig.

I struktur 18 är omstruktureringen av det löpande yttrandet tydlig. Avbrottet för omformulering markeras i transkriptionen med ”’”.

I struktur 19 uppträder *det är* i en position för att inleda en predikation, men fortsättningen uteblir. Talaren startar med den allmänna inledaren *det är*, men misslyckas med att producera antingen ett relevant innehåll eller en passande formulering. Yttrandet ändras inte heller, som i struktur 18, utan *det är* lämnas hängande. Många gånger kan sådana fall tjäna som turavslutande signaler.

## **Kvantitativ och longitudinell fördelning hos talarna. Utvecklingsaspekter**

De följande tabellerna 1–4 visar i översikt hur de olika *det är*-konstruktionerna fördelar sig hos de infödda talarna och inlärarna.

----- Tabell 1 ungefär här

Om vi först ser på frekvensfördelningen hos *de infödda informanterna* (tabell 1), så kan vi notera att bilden är tämligen likartad för de sex informanterna: de frekventare strukturerna är relativt rikligt företrädda hos alla sex, och i stort sett är det bara de med låg total frekvens (mindre än 1%) som inte uppträder hos alla. Detta visar på ett stabilt distributionsmönster för denna grupp av infödda talare. Till de frekventaste strukturtyperna hör 1:1 (*det är* + NP), 2 (*det är* + Adjektiv), 14 (Anteciperingsbrott) och 17–19 (*det är* följt av formuleringsbrott). Särskilt de två första har ju en vid tillämplighet i och med att det handlar om mycket centrala satsstrukturer i språket, där *det är* kan kombineras med ett stort antal olika substantiv och adjektiv, och hög frekvens i dessa fall bör därför inte vara oväntad. Det faktum att typerna 17–18 och i viss mån 19 är högfrekventa hos de infödda informanterna bekräftar att sådana avbrottsfall är generella och vanliga fenomen i tal och inte är specifika för inlärare. Till de mest lågfrekventa strukturerna hör 4 (*det är* + Infinitiv), 6 (*det är* med Strandad preposition), 8 (Objekts-*det*), 9 (*det är* vid Vänsterdislokering), 10 (*det är* vid Pseudoklyvning) och 15 (*det är* vid Högerdislokering).

----- Tabell 2 ungefär här

De sammanlagda frekvenserna för de infödda har också lagts in som jämförelse i översikten över *inlärarnas* frekvensdata (tabell 2). I huvudsak uppvisar frekvenssummorna

för inlärarna en liknande fördelning som hos de infödda, med strukturerna 1:1, 2, 14, 17 och 18 som de mest framträdande.

Det faktum att struktur 2, *det är* + Adjektiv, är ännu mer frekvent hos inlärarna (26,6%, mot 17,4% för de infödda), beror troligen på att inlärarna under samtalen har haft en större benägenhet att yttra beskrivande och värderande omdömen. Hos dem finner man också ett rikligt bruk av beskrivande och värderande adjektiv, som *dyr, rolig, bra, svår, möjlig, viktig* och många andra.

Struktur 9, *det är* vid Vänsterdislokering, är vanligare hos inlärarna än hos de infödda (2,4% gentemot 0,5%). Å andra sidan utnyttjas Satsklyvning (struktur 16) mindre av inlärarna än av de infödda (2,0% gentemot 8,2%). Även *det är* i Existentialkonstruktion (struktur 13:1–2) används mindre av inlärarna, något som till en del kan förklaras av en viss överrepresentation på tidiga stadier av *det finns* i kontexter där det är naturligt för infödda att säga *det är*.

En översikt av utvecklingen av de olika strukturerna hos inlärarna kan vi få om vi betraktar *den kvantitativa profilen över tid* för varje struktur. Ett exempel på en sådan longitudinell översikt för en individuell inlärare, C1, ges i tabell 3. Den visar dels hur tidigt de olika strukturerna börjar uppträda, dels frekvensfördelningen över tid. Liknande profiler har erhållits för de andra fem inlärarna, men av utrymmesskäl återges de inte separat här. I tabell 4 har värdena lagts ihop till en kombinerad utvecklingsprofil för gruppen av sex inlärare.

----- Tabell 3 och 4 ungefär här

Som tabell 4 utvisar, utvecklas flerfunktionaliteten hos *det är* gradvis i inlärarnas tal. På det rena nybörjarstadiet, vid tillfälle 1, uppträder *det är* bara i enkla satser följt av nominalfras (struktur 1:1 och existentialkonstruktionen 13:1). Men redan vid tillfälle 2 framträder 12 av de 21 strukturerna (undertyperna av 1 och 13 då räknade separat). Vid tillfälle 6 förekommer 17 strukturer, och i slutet av observationsperioden, vid tillfälle 10, 20 strukturer. Både anaforiska och icke-referentiella användningar av *det är* tidigt representerade. Inlärarnas totala användning av *det är* ökar med tiden, vilket kan ses i tabellens nedersta rad. *Det är* blir efterhand till och med mer frekvent hos inlärarna än hos de infödda informanterna. Andelen *det är* av inlärarnas totala textmassa uppgår till 2,2% jämfört med 1,8% hos de infödda. Medelvärdet för de sista fem inspelningstillfällena är så högt som 2,4%, d.v.s. i medeltal en förekomst av *det är* per 42 ord i texten. Det är tydligt att konstruktioner med *det är* är

favoriserade strukturer hos inlärarna, och att repertoaren av olika sådana konstruktioner växer ut snabbt.

Formuleringsbrott efter *det är* (strukturerna 17–19) uppträder regelbundet hos alla inlärarna, och som tabell 4 visar, ökar de snarare än avtar med högre språkfärdighet. Detta stödjer slutsatsen att dessa strukturer är normala diskursfenomen; tydligen uppnår inlärarna efterhand stadier av yttrandekomplexitet och flyt där dessa drag utgör en reguljär del av hur yttranden byggs upp.

Även om konstruktioner med *det är* allmänt visar sig vara favoriserade hos inlärarna, är det samtidigt påfallande skillnader mellan sådana strukturer som uppträder tidigt och sådana som börjar tas i bruk först senare, om de alls dyker upp. Strukturer med anaforskt *det + är + NP*, *Adj*, *Adv* eller *PP* (strukturerna 1:1, 2 och 3) kommer således tidigare än t.ex. *det är + Infinitiv* (str. 4), *Strandad preposition* (str. 6), *Objekts-det* (str. 8, inga exempel), *Pseudoklyvning* (str.10) eller *terminalt det är* (str. 11). Bland strukturerna med icke-refererande eller kataforiskt *det* kommer *Opersonlig*, *Existentiell* och *Anteciperande konstruktion* (str. 12, 13:1 och 14) tidigare än *Högerdislokering* och *Satsklyvning* (str. 15 och 16). Sådana skillnader i utvecklingsgången blir särskilt tydliga om vi jämför nära besläktade strukturer med varandra, såsom 1:1 med 1:2, 9 med 10 eller 13:1 med 13:2, strukturer som skiljer sig åt väsentligen genom frånvaron eller närvaron av en inbäddad relativsats. Det ger – inte oväntat – belägg för att graden av syntaktisk komplexitet har betydelse för hur tidigt en struktur tillägnas.

Jämförelsen av tidiga och sena strukturer i tabell 4 utvisar också ett samband mellan total frekvens för en struktur och tidpunkten då den dyker upp. Tendensen är att de strukturer som har visat sig mindre frekventa hos de infödda och inlärarna också uppträder sent i inlärnarnas produktion. Märk att det här sambandet inte kan förklaras som en automatisk effekt av att sent uppdykande strukturer bara hinner användas under kortare tid och därför får lägre frekvenssumma. Två saker talar emot det. Dels kan vi se i utvecklingsprofilen att de lågfrekventa strukturerna tenderar att vara lågfrekventa även då de väl börjat användas. Dels visar också frekvenssummorna för de infödda (tabell 1) att dessa strukturer även där är lågfrekventa. Korrelationen lågfrekvent struktur – sent inlärd tycks således vara reell. Skillnaden mellan tidiga och sena strukturer kan ses också i skillnaden i det antal inlärare som har börjat använda strukturen, såsom framgår av en jämförelse med tabell 2. Under det att de strukturer som noterades som tidiga används av alla inlärarna, kan de som visade sig vara sena beläggas bara hos en del.

## **DET ÄR I YTTRANDEPRODUKTIONEN**

De föregående avsnitten har behandlat användningen av *det är* grammatiskt, kvantitativt och i ett utvecklingsperspektiv. Frågan är nu hur *det är* fungerar i formuleringsförloppet. Vi ska försöka se de olika aspekterna i sitt inbördes sammanhang och belysa de beröringspunkter som finns mellan den grammatiska flexibiliteten, den höga frekvensen, det snabba tillägnandet och den roll som *det är* spelar i yttrandeprocessen i realtid.

I Levelts (1989) inflytelserika modell över hur talade yttranden produceras utgör principen om ”tillväxande processning” (*incremental processing*) en grundsten. Principen, som lanserades av Kempen och Hoenkamp (1987), innebär att olika bitar av det blivande yttrandet utarbetas parallellt medan det hela planeras och efterhand byggs ihop till en sats. I modellen tänks ett yttrande under tillblivelse stegvis genomlöpa komponenter för konceptualisering, formulering och artikulation, men dessa tre delprocesser tänks arbeta fortlöpande och parallellt så att fragment av det tänkta yttrandet levereras vidare från en komponent till nästa att börja arbeta med under tiden som processen fortsätter. Så snart början till ett yttrande är klart kan det uttalas medan fortsättningen planeras. Detta möjliggör en snabb och flytande talproduktion i och med att yttrandet kan komma igång medan talaren planerar hur det ska fullbordas. Här är det klart att det finns ett värde i att ha tillgång till användbara uttryck att starta yttranden med. Det faktum att *det är* är så frekvent som introduktion till olika slags predikationer talar för att detta uttryck fungerar som en standardsekvens för att starta med och därmed som ett stöd i den pågående satsplaneringen.

De kvantitativa översikterna i det förra avsnittet visade på ett utvecklingsmönster för strukturer med *det är*, grundat på en flitig användning av sådana strukturer. Vad som händer här kan ses inom ramen för en *användningsbaserad språkmodell* (*usage-based model*; Langacker 1988, 1999; Kemmer och Barlow 1999), som jag nämnde inledningsvis. Enligt denna modell är talarens språkliga system grundat i tillfällena av konkret användande, ”usage events”. Språkanvändarnas mentala representationer av språkliga enheter uppstår som resultat av erfarenhet och blir successivt *befästa* (*entrenched*) som *kognitiva rutiner* hos individen. Ur specifika enskilda användningar i meningsfulla sammanhang abstraheras mer generella representationer, såsom fonem, morfem och syntaktiska mönster. Enligt en användningsbaserad språkuppfattning spelar *frekvensen* en central roll i denna process, i och med att den ackumulerade erfarenheten av upprepad användning av ett språkligt element eller

mönster successivt leder till att det blir kognitivt befäst hos språkanvändaren (Bybee och Hopper 2001a, b; Ellis 2002). Användningarna har också en dubbelriktad och interaktionell roll i språktilläggnandeprocessen. Samtidigt som yttrandena är produkter av talarens språkliga system är de också input till andra personers språkssystem, liksom till talarens eget. Även erfarenheten av att tolka inkommande yttranden har en effekt. Detta innebär med andra ord att det kognitiva befästandet av språklig struktur får näring av både reception och egen produktion av tal.

Observationen tidigare, att de strukturer med *det är* som är mest frekventa hos infödda och inlärare också används tidigast av inlärarna, kan nu tolkas som en dubbel effekt av funktionalitet och frekvens. Funktionaliteten leder till frekvent behov och tät användning. Det faktum att vissa strukturtyper uppträder frekvent hos de infödda informanterna gör det rimligt att anta att också våra inlärare möter frekvent användning av dessa strukturer i sin språkliga input från infödda talare. Frekvent input bör leda till tidigt tillägnande. Samtidigt befrämjar också inlärarens egen användning av dessa strukturer en rutinisering av dem. De blir allt lättare att processa i tal, och därmed bekvämare att använda.

Våra data tyder på att *det är* utgör en byggbit som rutiniseras som en sammanhängande sekvens. Det faktum att *det är* uppträder som en gemensam, ständigt återkommande inledning till predikationer med en mångfald olika konstruktioner gör denna tolkning rimlig. Den tydligaste indikationen på att *det är* fungerar som en enhet i yttrandeproduktionen är fallen där det följs av ett formuleringsbrott, ett fenomen som jag strax ska återkomma till. Ellis (1996) diskuterar den roll språkliga sekvenser har i inläringen och fastslår att mycket av språktilläggnandet består i inläring av sekvenser, och att abstrakt grammatisk kunskap kommer av analys av sekvensinformation (Ellis 1996: 91). Inlärarens korttidsminne analyserar regelbundet återkommande sekvenser av tal, som integreras i långtidsminnet och blir åtkomliga därifrån på ett automatiskt sätt. En central process här är *chunking*, kombinationen av mindre element till sammanhängande block som lagras som sådana i minnet (Field 2004: 60f.). *Chunking* ”is the development of permanent sets of associative connections in long-term memory and is the process that underlies the attainment of automaticity and fluency in language” (Ellis 1996: 107).

I en studie där Axelsson (1988) bland annat undersökte förekomsten av s.k. *ramar* i svenskt talat inlärarspråk fann hon att den mest frekventa ramen hos en grupp vuxna inlärare var *de e* följt av en adjektiv- eller nominalfras. En ram definieras då som ”ett yttrande bestående av en fast enhet med tillägg av en kreativ del” (Axelsson 1988: 191). Sekvensen *de*



*e* är här den fasta enheten och den följande frasen den kreativa delen. Motsvarigheten i vårt fall är den dominerande frekvensen av strukturtyperna 1 och 2, som vi såg i tabell 2.

Sett i relation till den löpande planeringen och produktionen av yttranden finns det således skäl att anta att *det är* fungerar som en *formelsekvens*. Begreppet formelsekvens (eng. *formulaic sequence, formula, prefab* etc. – termerna varierar) definieras vanligen i litteraturen på basis av hur sådana enheter tänks vara åtkomliga i talproduktionen. Enligt en ofta citerad definition hos Wray (1999: 214) är en formulaic sequence ”stored and retrieved whole from memory at the time of use, rather than being subject to generation or analysis by the language grammar”. Formler fungerar således som planeringsenheter i talproduktionen och uppträder som automatiserade chunks. Användningen av sådana enheter utgör holistiska processer som samverkar med kreativa processer i kodning och avkodning av yttranden. Som Wray och Perkins (2000: 11) formulerar det: ”... the best deal in communicative language processing is achieved by the establishment of a suitable balance between creative and holistic processes. The advantage of the creative system is the freedom to produce or decode the unexpected. The advantage of the holistic system is economy of effort when dealing with the expected”.

Till skillnad från de flesta formeluttryck som har behandlats i litteraturen, såsom olika slag av lexikala och pragmatiska fraser, är *det är* ett exempel på en *grammatisk* formelsekvens. Erman och Warren (2000: 41) karakteriserar grammatiska prefabs, till skillnad från lexikala och pragmatiska prefabs, som ”intralinguistic text-forming items rather than units with extralinguistic reference”. (För indelningar och kategoriseringar av typer av formeluttryck och deras funktioner, se också Raupach 1984; Weinert 1985; Wray 1999, 2002; Wray och Perkins 2000; Forsberg 2006.)

En indikation på att en sekvens har formelkaraktär är också dess fonetiska form, som kan avslöja att talaren avkodar och producerar uttrycket som ett samlat block. Formelsekvenser tenderar att uttalas sammandraget eller reducerat. *Det är* uppvisar ett kontinuum av varianter i talet: [de: e:, de e, de, də, d<sup>ə</sup>], där en mer eller mindre reducerad form är vanlig. (ASU-korpusens transkription använder *de e* som en grov representation av detta fonetiska kontinuum.) Koherens är obligatoriskt, d.v.s. ingen paus är möjlig mellan *det* och *är*. Uppreppningsformer där *det* följs av *det är* kan således ha formen *de de e*, eller *de = de e* (där ”=” betecknar en ofylld paus), eller *de % de e* (där ”%” är ett pausfyllnadsljud), men inte *\*de de = e* eller *\*de de % e*. Här uppträder *de e* som ett block i en återupptagande funktion efter ett formuleringsbrott.

De olika fallen av formuleringsbrott efter *det är*, alltså strukturerna 17, 18 eller 19 ovan, bidrar till att belysa hur konstruktioner väljs och tar form i formuleringsprocessen. Som vi

såg, uppstår upprepningar, omformuleringar eller helt avbrutna yttranden, beroende på hur väl talaren lyckas producera en fortsättning som passar ihop med startsekvensen *det är*. Det är typiska drag i löpande diskurs, då talaren söker efter formuleringar. Exempel (5) illustrerar hur en inlärare använder satsinitialt *det är* på ett rutinmässigt sätt under det att han arbetar ut ett svar på en fråga:

(5) B ja. hur går de till å läsa arkitektur i i / på kth?

I ja. **de e de e** väldigt spännande verkligen men också mycke jobbit. **de e** mycke mycke jobb. **de e** / man sitter där från nie till fem (B: MH) varje da. (B: MHM) och **de e** mycke intensivt kurs. kursen e mycke intensivt. (Q1-10)

Omformuleringar av olika slag (struktur 18) är särskilt intressanta, eftersom de belyser hur mer utvecklade yttranden successivt byggs upp med hjälp av en ospecifik startformel *det är*, och på så sätt ger inblickar i den gradvisa processen av konceptualisering och formulering.

Detta kan illustreras med några exempel på olika omformuleringsfall från infödda och inlärare. Det finns ett visst spelrum för att variera *det är*-formeln så att den också kan innefatta ett obetonat modalt adverb inom samma prosodiska enhet före brottstället (*det är ju, det är väl, det är liksom* etc.). Det är skäl att anta att även dessa sekvenser fungerar som formelartade block.

Omformuleringen kan variera från minimal till mer radikal. Den kan till exempel bestå i bara ett skifte av tempus, som i (6a), eller aktionsart, som i (6b).

(6) a. så att **de e / de va** därför ja reagera tror ja. (Z4-2)

b. vi snackar tillsammans skojar tillsammans. (B: JA) **de e / de blir** kul när vi jobbar tillsammans. (G3-10)

Många omformuleringar visar hur argumentstrukturen i satsen gradvis byggs upp. Talaren startar med ett provisoriskt *det är* och utarbetar satsen sedan med ett annat verb och/eller ett mer specifikt subjekt i form av ett substantiv eller pronomen. I (7) ges några exempel.

(7) a. men på nåt sätt / **de e % / de känns** inte som nåra problem. (Z4-1)

b. på sommarn (E: JA) jobba ja på danderyd sjukhus (E: MHM, B: JA) som städa / städerska. å ja = ja tycker **de e / de hjälpte** mycke me svenska. ja började prata svenska då. (C1-10)

- c. han sa att **de e / musik e** för hög. (C2-10)
- d. jo men **de e / ja e** väldigt intresserad av ganska mycke. (Z4-1)
- e. I ja tåget.  
E ja.  
I men lastbile+ **de e också / man kan kombinera** tåg å lastbil (C2-8)

Andra vanliga sätt att modifiera *det är*-satsen är att topikaliserar något icke-subjekt, som i (8a–b), eller att konstruera meningen med en överordnad sats, *jag tror/tycker/anser/antar* etc., som i (8c).

- (8) a. ja **de e / ibland = e de** = orättvist tror ja. (C2-9)
- b. ja tycker att **de e / på nå sätt / ur moralisk synvinkel så e de väl** ganska så förkastligt (Z1-3)
- c. ja precis. **de e / ja anser att de e** så. (G3-8)

Särskilt i fall av mer radikala nyformuleringar av satsen märks det att väsentliga delar av det blivande meddelandet inte var konceptualiserade vid tidpunkten för formuleringsbrottet. Ibland förkastar talaren helt en satsplanering med *det är*, som t.ex. i (7e), men ofta verkar det finnas en tendens att vilja få ihop en sats med denna byggbit bevarad eller något modifierad (*är det, jag är, det blir, det känns* etc.). Den första, fragmentariska satsplanen tycks ha kvar en viss styrande verkan.

Variationen mellan bakåtpekande (anatorisk) och framåtpekande *det*-konstruktion och den därmed förbundna regleringen av ekonomi och precision i syftningen, som diskuterades ovan i samband med exempel (4), kan också fungera som ett syntaktiskt skifte under löpande satsproduktion. Ett sådant exempel (fast med *det var*) är (9), som visar spår av hur det kan ha gått till i formuleringsförloppet:

- (9) % på gymnasiet va **de** nån dikt vi läste. nej **de** va en sångtext = (B: MHM)  
= som ja nu inte kommer ihåg = hur den va på svenska heller.  
men **de** va lite roligt = å läsa igenom den. (Z2-5)

Den första meningen i (28) är en existentialkonstruktion med formellt *de*. I den andra meningen syftar *de* anatoriskt på *dikt*. Sekvensen *nej de va en sångtext* är uppenbarligen tillräckligt tydlig, och talaren kan fortsätta och lägga ny information till *sångtext*. Observera

pausen i den tredje meningen. Avsnittet *men de va lite roligt* skulle fungera som en anaforisk konstruktion om meningen slutade där. Med fortsättningen *å läsa igenom den* skiftar *de* till att bli ett framåtppekande formellt subjekt. Det förefaller här som om talaren under hand märker att den första delen får en alltför vag syftning i sammanhanget och bygger ut meningen med en precisering.

## SLUTORD

Det syntaktiska fragmentet *det är* har visat sig ha en speciell roll i språket genom sin höga frekvens, sin mångsidiga användbarhet i olika konstruktioner, sin formelartade karaktär och sin funktion som (ibland tentativ och provisorisk) satsintroduktör i yttrandeproduktionen. Sekvensen *det är* hör till de element i språket som lätt tillägnas. Uppenbarligen sammanhänger detta både med dess funktionella användbarhet och dess bekvämlighet att processa. Inlärarna i undersökningen tillägnade sig på kort tid ett användningsmönster för olika typer av *det är*-konstruktioner, som både i fråga om repertoar och frekvensfördelning i stort liknade användningsmönstret hos de infödda informanterna. Det framträdde också en karakteristisk distribution av hur tidigt de olika konstruktionstyperna dök upp i inlärarnas tal. En utgångspunkt för tolkningen av resultaten har varit att framväxten av det språkliga systemet grundas i konkreta ”usage events”, där lyssnaren/talaren tar emot och producerar yttranden. Jag har sökt visa något av hur konstruktionsprocessen och konstruktionsprodukten hänger samman, och hur valet av konstruktioner i talarens tillgängliga repertoar manifesteras som val (och ibland ändrade val) i yttrandeproduktionen.

Det kan tyckas märkligt att inlärarna så beredvilligt tillägnar sig ett så komplext mönster av syntaktiska konstruktioner som det som förknippas med det neutrala pronomenet *det* och satsfragmentet *det är*. Men ett sådant intryck får tillskrivas vad vi kan beteckna som ”grammatikerns illusion”. Det som ter sig som ett intrikat system för den beskrivande lingvisten kan utgöra en enkel och effektiv mekanism för språkanvändaren och inläraren i talsituationen. *Det* och *det är* är fonologiskt och morfologiskt okomplicerade element, som möjliggör dels bekväm referens till olika diskursreferenter, dels bildandet av satser med opersonliga predikat och en rad textlingvistiskt motiverade formellt subjekt-konstruktioner med bevarande av svenskans V2-ordföljd och platshållarkrav. Det som är komplexitet i beskrivningen och kategoriseringen för grammatikern som analyserar konstruktioner är inte

nödvändigtvis formuleringsproblem för användaren. De exempel och kvantitativa utvecklingsprofiler vi sett ovan tyder på att det neutrala pronomenet *det* och formeln *det är* snarare är tillgångar för inläraren i språkanvändningen och gynnar upptagandet av de berörda konstruktionerna i inlärarspråket. *Det är* har visat sig vara ett brukbart standardelement i formuleringsarbetet. Att tillägna sig sådana standardelement som uttrycksmedel är en viktig del av språkutvecklingen.

## **NOT**

- - - - (Not 1 infogas här)

## REFERENSER

- Allwood, Jens 1999: *Talspråksfrekvenser: frekvenser för ord och kollokationer i svenskt tal- och skriftspråk*. (Gothenburg Papers in Theoretical Linguistics S 21.) Göteborgs universitet, Institutionen för lingvistik.
- Axelsson, Monica 1988: Helfraser och ramar – ett viktigt inslag på andraspråksinlärares väg mot ett kreativt regelsystem. S. 191-202 i *Första symposiet om svenska som andraspråk. Vol. I: Föredrag om språk, språkinläring och interaktion*, red. Kenneth Hyltenstam och Inger Lindberg. Stockholms universitet, Centrum för tvåspråkighetsforskning.
- Bartning, Inge och Björn Hammarberg 2007: The functions of a high-frequency collocation in native and learner discourse: the case of French *c'est* and Swedish *det är*. S. 1–45 i *IRAL, International Review of Applied Linguistics in Language Teaching* 45 (1).
- Beckman, Natanael 1934: Västeuropeisk syntax. Några nybildningar i nordiska och andra västeuropeiska språk. I *Göteborgs högskolas årsskrift* 40. Göteborg.
- Brown, Gillian och George Yule 1983: *Discourse Analysis*. Cambridge: CUP.
- Bybee, Joan och Paul Hopper (red.) 2001a: *Frequency and the emergence of linguistic structure*. Amsterdam/Philadelphia : John Benjamins.
- Bybee, Joan och Paul Hopper 2001b: Introduction to frequency and the emergence of linguistic structure. S. 1-24 i *Frequency and the emergence of linguistic structure*, red. Joan Bybee och Paul Hopper Amsterdam/Philadelphia : John Benjamins.
- Ekerot, Lars-Johan 1979: Syntax och informationsstruktur. S. 79-108 i *Svenska i invandarperspektiv*, red. Kenneth Hyltenstam. Lund: Liber Läromedel.
- Ellis, Nick 1996: Sequencing in SLA: Phonological memory, chunking, and points of order. S. 91–126 i *Studies in Second Language Acquisition* 18.
- Ellis, Nick 2002: Frequency effects in language processing. S. 143-188 i *Studies in Second Language Acquisition* 24.
- Erman, Britt och Beatrice Warren 2000: The idiom principle and the open choice principle. S. 29–62 i *Text* 20.
- Field, John 2004: *Psycholinguistics : The Key Concepts*. London och New York : Routledge.
- Forsberg, Fanny 2006: *Le langage préfabriqué en français parlé L2. Étude acquisitionnelle et comparative*. Diss. Stockholms universitet, Institutionen för franska, italienska och klassiska språk.

- Fraurud, Kari 1992: *Processing Noun Phrases in Natural Discourse*. Diss. Stockholms universitet, Institutionen för lingvistik.
- Fraurud, Kari 2000: Demonstrativer i svensk sakprosa. S. 5-28 i *Denna – den här – den där. Om demonstrativer i tvärspråklig belysning. En minnesskrift till Elsie Wijk-Andersson*. [ASLA Information, 26 (2)]. Uppsala: ASLA.
- Hammarberg, Björn 2000: A polyfunctional word in native usage and L2 acquisition: the Swedish pronoun 'det'. S. 103-129 i *Kontraster i språk*, red. Johan Falk, Gunnar Magnusson, Gunnel Melchers and Barbro Nilsson. Stockholm: AWE International.
- Hammarberg, Björn 2006: *Introduktion till ASU-korpusen, en longitudinell muntlig och skriftlig textkorpus av vuxna inlärares svenska med en motsvarande del från infödda svenskar*. Stockholms universitet, Institutionen för lingvistik.
- Hammarberg, Björn and Åke Viberg 1977: The place-holder constraint, language typology, and the teaching of Swedish to immigrants. S. 106–163 i *Studia Linguistica* 21.
- Holmes, Philip och Ian Hinchliffe (2003). *Swedish: A comprehensive grammar*. 2:a uppl. New York: Routledge
- Kemmer, Suzanne och Michael Barlow 1999: Introduction: A usage-based conception of language. S. vii-xxvii i *Usage-Based Models of Language*, red. Michael Barlow och Suzanne Kemmer. Stanford: CSLI Publications, Center for the Study of Language and Information.
- Kempen, Gerard och Edward Hoenkamp 1987: Incremental sentence generation: implications for the structure of a syntactic processor. S. 201-258 i *Proceedings of the Ninth International Conference on Computational Linguistics*, red. Ján Horecký. Amsterdam: North-Holland.
- Langacker, Ronald W. 1988: A usage-based model. S. 127-161 i *Topics in Cognitive Linguistics*, red. Brygida Rudzka-Ostyn. (Current Issues in Linguistic Theory 50.) Amsterdam: Benjamins.
- Langacker, Ronald W. 1999: A dynamic usage-based model. S. 1-63 i *Usage-Based Models of Language*, red. Michael Barlow and Suzanne Kemmer. Stanford: CSLI Publications, Center for the Study of Language and Information.
- Levelt, Willem J.M. 1989: *Speaking : From intention to articulation*. Cambridge MA: The MIT Press.
- Lindström, Jan 2008: *Tur och ordning. Introduktion till svensk samtalsgrammatik*. Norstedts Akademiska Förlag.

- Raupach, Manfred 1984: *Formulae in Second Language Speech Production*. S. 114-147 i *Second Language Productions*, red. Hans W. Dechert, Dorothea Möhle och Manfred Raupach. Tübingen: Gunter Narr.
- Sundman, Marketta 1980: *Existentialkonstruktionen i svenskan*. Publications of the Research Institute of the Åbo Akademi Foundation, 57. Åbo.
- Sundman, Marketta 1987: *Subjektval och diates i svenskan*. Åbo: Åbo Akademis förlag.
- Teleman, Ulf, Staffan Hellberg och Erik Andersson (1999). *Svenska Akademiens Grammatik*. Vol. 1-4. Stockholm: Norstedts.
- Thorell, Olof 1973: *Svensk grammatik*. Stockholm: Esselte Studium.
- Weinert, Regina 1995: The role of formulaic language in second language acquisition: A review. S. 180-205 i *Applied Linguistics* 16.
- Wray, Alison 1999: Formulaic language in learners and native speakers. S. 213-231 i *Language Teaching* 32.
- Wray, Alison 2002: *Formulaic Language and the Lexicon*. Cambridge: Cambridge University Press.
- Wray, Alison och Michael R. Perkins 2000: The functions of formulaic language: an intergrated model. S 1-28 i *Language and Communication* 20.



## APPENDIX

### Nyckel till källhänvisningar och transkription

Källhänvisningar till informant och inspelningstillfälle ges inom parentes efter varje exempel; t.ex. Z4-1 = informant Z4, tillfälle 1; C2-6 = informant C2, tillfälle 6. Informanter med bokstaven Z avser infödda svenska talare; C, G och Q avser inlärare med kinesiska, grekiska resp. portugisiska som L1.

Bokstäverna I, B och E i vänsterkolumnen i vissa exempel är talaridentifikationer som är medtagna om textutdraget innehåller turer från fler än en talare. I betecknar här alltid den aktuella informanten, och B och E är intervjuarna. Om ingen talaridentifikation står utsatt, är talaren informanten.

Exemplen återger ASU-korpusens transkriptionspraxis, med talspråksformer där sådana förekommer (t.ex. *de e* för standardortografins *det är*). Några specialtecken används:

- = Tom paus
- % Pausfyllare (ersätter bokstavs-beteckningar som ”eh”, ”öh” etc.)
- Avbrutet ord
- + Efter en morfologiskt otydlig form
- / Syntaktiskt avbruten sekvens, eller avbrott för omplanering eller reparation. Används inte om identiska former upprepas.
- ? Efter avslutad direkt fråga
- . Efter avslutad mening som inte är fråga
- < > Omger text på annat språk än svenska
- ( ) Omger uppbackningar från samtalspartnern

Lista 1. Strukturtyper med *det är* med exempel ur ASU-korpusen.

Huvudgrupp A: Anaforiskt eller deiktiskt *det*.

Struktur 1. *Det är* + Nominalfras

- a. **de e** saker som ja funderar ganska mycke över. (Z4-1)
- b. **de e** en nyhet ti mej (E: JA) tror ja. (C2-6)

Struktur 2. *Det är* + Adjektivfras

- a. (Samtalsämne: mjältbrandsbakterier kan förbli aktiva i flera år)  
å **de e** ju lite otrevligt kanske. (Z1-4)
- b. (Samtalsämne: hur man spelar in en film)  
% **de e** inte så lätt. (G2-4)

Struktur 3. *Det är* + Adverbiell fras / Prepositionsfras

- a. **de e** så himla längesen. (Z6-5)
- b. **de e** på kvällen (C1-3)

Struktur 4. *Det är* + Infinitivfras

- a. nu finns de ju nästan bara en väg ut. å **de e** ju å anfalla. (Z2-3)
- b. (Samtalsämne: om informanten har svårigheter att skriva svenska)  
... för ja läser teknik. (B: JA) så **de e** bara lösa problem. (B: JAHA MH)  
s- man behöver inte skriva uppsats. (C2-10)

Struktur 5. *Det är* + Bisats

- a. **de e** ju bara va ja tror att de heter. (Z2-5)
- b. de har ett hål. - - - **de e** för att man kan hänga de på en krok (Q1-7)

Struktur 6. *Det är* + Strandad preposition

- de e** ja helt övertygad om. (Z1-3)

Struktur 7. 'Tough'-konstruktion (*det är* + Adjektiv+ Infinitivkomplement)

- a. (Samtalsämne: vad som händer på en bild)  
**de e** hemskt svårt å säja. (Z5-5)
- b. (Samtalsämne: att beskriva ett handtag)  
**de e** lätt att = att hålla. (Q1-6)

Struktur 8. Objekts-*det* lyft från bisatsposition

- de e** ja inte säker på att afrikanerna tycker (Z4-1)

Struktur 9. *Det är* vid Vänsterdislokering

- a. astrologi **de e** ju en ny ny sorts religion eller va man ska säja. (Z2-5)
- b. femti år **de e** inte så kul att fylla. (Q1-7)

Struktur 10. *Det är* i Pseudoklyvnings-konstruktion

- a. de som e svårast å måla **de e** typ ansikten å händer å sånt. (Z4-4)
- b. (Samtalsämne: betydelsen av begreppet "idiot" i grekiska)  
de som inte e sta- statligt eller offentligt **de e** idiotiskt. (G3-10)

Struktur 11. *Det är* i terminal position i bisats

- a. nu kan man till å me se hur långt **de e** / tjugi centimeter långt. (Z3-4)
- b. ja vet inte riktigt hur lång tid **de e**. (Q2-10)

Lista 2. Strukturtyper med *det är* med exempel ur ASU-korpusen.

Huvudgrupp B: Icke-referentiellt eller kataforiskt *det*.

Struktur 12. Opersonlig konstruktion

- a. hon pekar på klockan å visar att **de e** sent. (Z3-3)
- b. **de e** för varmt där. (Q1-10)
- c. **de e** dags att kliva upp. (Z5-5)

Struktur 13. Existentialkonstruktion

- a. **de e** ju fortfarande barn som blir misshandlade (Z2-4)
- b. **de e** skillnader mellan generation. (C1-10)

Struktur 14. Anteciperande *det*

- a. de e klart att **de e** bra å kunna andra språk också. (Z3-5)
- b. för de **de e** viktigt hur man presentera ett projekt (Q1-10)

Struktur 15. *Det är* vid Högerdislokering

- a. **de e** ju så himla omfattande de där. (Z6-5)
- b. ja tycker **de e** så sjukt just de här att dom tillverkar ruskia vapen (Z3-3)

Struktur 16. *Det är* vid Satsklyvning

- a. **de e** i köket man använder den. (Z1-2)
- b. ja ja kan tänka mej att **de e** därför som han fortfarande lever. (Z1-3)

Lista 3. Strukturtyper med *det är* med exempel ur ASU-korpusen.

Huvudgrupp C: Brott i formuleringen.

Struktur 17. *Det är* upprepat

- a. ja **de e ju** = de e ju ganska självklart i såna situationer på nå sätt. (Z1-4)
- b. nej **de e de e** bra att åka bort tycker ja (Q1-10)

Struktur 18. *Det är* + omformulering

- a. ja tycker att **de e** / på nå sätt / ur moralisk synvinkel så e de väl ganska så förkastligt (Z1-3)
- b. så / = ja **de e** / ibland = e de = orättvist tror ja. (C2-9)

Struktur 19. *Det är* + avbrutet yttrande

- c. ja. de beror ju på / man kan ju argumentera för å emot i dom flesta frågor så att **de e** / (Z1-4)
- d. ja dom bara gå hemifrån å låter alla ljus på. å (B: JA) **de e** / (Q1-9)

Tabell 1. Kvantitativ fördelning av *det är* hos 6 infödda talare.

Strukturer	Infödda talare						Summa	%
	Z1	Z2	Z3	Z4	Z5	Z6		
1:1 <i>det är</i> + NP, ej Rel	36	54	45	60	53	52	300	21.2
1:2 <i>det är</i> + NP, med Rel	4	18	5	4	13	5	49	3.5
2 <i>det är</i> + Adj	30	38	30	59	30	59	246	17.4
3 <i>det är</i> + Adv/PP	9	10	6	8	10	15	58	4.1
4 <i>det är</i> + Infinitiv	-	3	1	1	-	-	5	0.4
5 <i>det är</i> + Bisats	4	6	11	6	1	1	29	2.1
6 <i>det är</i> med Strandad prep	1	1	-	-	1	-	3	0.2
7 'Tough'-konstruktion	6	2	2	2	5	2	19	1.3
8 Obj <i>det</i> lyft från bisats	-	1	-	1	-	-	2	0.1
9 <i>det är</i> med Vänsterdislok	-	2	2	1	1	1	7	0.5
10 Pseudoklyvning	-	-	-	1	2	-	3	0.2
11 <i>det är</i> terminalt i Bisats	6	3	2	8	-	1	20	1.4
12 Opersonlig konstruktion	3	2	2	2	4	1	14	1.0
13:1 Existentialkonstr, ej Rel	8	16	5	15	16	17	77	5.4
13:2 Existentialkonstr, m Rel	5	19	17	9	2	4	56	4.0
14 Anteciperande	18	22	22	22	18	12	114	8.1
15 Högerdislokering	1	-	1	1	1	3	7	0.5
16 Satsklyvning	12	23	15	10	31	25	116	8.2
17 <i>det är</i> upprepat	9	18	13	21	28	5	94	6.7
18 <i>det är</i> + omformulering	25	26	10	44	28	15	148	10.5
19 <i>det är</i> + övergivet yttrande	10	8	11	3	7	8	47	3.3
<b>Summa <i>det är</i></b>	187	272	200	278	251	226	1414	100
<b>Textlängd (antal ord)</b>	12140	13562	10239	11532	15397	14367	77237	
<b>% <i>det är</i> av textlängden</b>	1.5	2.0	2.0	2.4	1.6	1.6	1.8	

Tabell 2. Kvantitativ fördelning av *det är* hos 6 inlärare.

Strukturer	Inlärare						Inlärare		Infödda	
	C1	C2	G2	G3	Q1	Q2	Summa	%	Summa	%
1:1 <i>det är</i> + NP, ej Rel	33	109	52	64	48	48	354	21.3	300	21.2
1:2 <i>det är</i> + NP, m Rel	6	2	5	15	10	4	42	2.5	49	3.5
2 <i>det är</i> + Adj	72	152	47	53	66	52	442	26.6	246	17.4
3 <i>det är</i> + Adv/PP	15	17	16	8	12	2	70	4.2	58	4.1
4 <i>det är</i> + Infinitiv	-	7	2	-	4	-	13	0.8	5	0.4
5 <i>det är</i> + Bisats	2	3	2	7	2	2	18	1.1	29	2.1
6 <i>det är</i> med Str. prep	-	-	-	1	-	-	1	0.06	3	0.2
7 'Tough'-konstruktion	-	-	-	1	9	2	12	0.7	19	1.3
8 Obj <i>det</i> lyft från bis	-	-	-	-	-	-	-	-	2	0.1
9 <i>det är</i> m Vänsterdisl	3	10	4	6	6	10	39	2.4	7	0.5
10 Pseudoklyvning	1	-	-	1	-	-	2	0.1	3	0.2
11 <i>det är</i> terminalt i bis	-	2	2	1	-	1	6	0.4	20	1.4
12 Opersonlig konstr	10	2	5	2	9	4	32	1.9	14	1.0
13:1 Exist-kstr, ej Rel	10	15	1	4	18	5	53	3.2	77	5.4
13:2 Exist-kstr, m Rel	-	2	5	4	4	1	16	1.0	56	4.0
14 Anteciperande	20	26	34	17	44	4	145	8.6	114	8.1
15 Högerdislokering	-	2	2	1	-	-	5	0.3	7	0.5
16 Satsklyvning	1	-	6	11	12	4	34	2.0	116	8.2
17 <i>det är</i> upprepat	21	16	20	46	58	30	191	11.5	94	6.7
18 <i>det är</i> + omformul	19	42	16	18	36	7	138	8.3	148	10.5
19 <i>det är</i> + övergiv. yttr	5	9	9	5	16	4	48	2.9	47	3.3
<b>Summa <i>det är</i></b>	216	416	228	265	354	180	1661	100	1414	100
<b>Textlängd (antal ord)</b>	11952	15077	13344	13678	12968	8363	75382		77237	
<b>% <i>det är</i> av textlängden</b>	1.8	2.8	1.7	1.9	2.7	2.2	2.2		1.8	

Tabell 3. Individuell longitudinell profil för inläraren C1.

Strukturer	Inspelningstillfällen										Summa 1-10
	1	2	3	4	5	6	7	8	9	10	
1:1 <i>det är</i> + NP, ej Rel	2	3	2	3		7	8		3	5	33
1:2 <i>det är</i> + NP, med Rel		1				2		1	1	1	6
2 <i>det är</i> + Adj		7		5	8	20	3	8	12	9	72
3 <i>det är</i> + Adv/PP			3			7		1	1	3	15
4 <i>det är</i> + Infinitiv											-
5 <i>det är</i> + Bisats							1		1		2
6 <i>det är</i> med Strandad prep											-
7 'Tough'-konstruktion											-
8 Obj <i>det</i> lyft från bisats											-
9 <i>det är</i> med Vänsterdislokering				1		1	1				3
10 Pseudoklyvning										1	1
11 <i>det är</i> terminalt i Bisats											-
12 Opersonlig konstruktion			6			1	1	2			10
13:1 Existentialkonstr, ej Rel				1	1	3		1		4	10
13:2 Existentialkonstr, med Rel											-
14 Anteciperande						2	2	6	3	7	20
15 Högerdislokering											-
16 Satsklyvning									1		1
17 <i>det är</i> upprepat		1	2	1		4	3	4	2	4	21
18 <i>det är</i> + omformulering			1	1	2	2	2	2	5	4	19
19 <i>det är</i> + övergivet yttrande		1				3	1				5
<b>Summa <i>det är</i></b>	2	13	14	12	11	52	22	25	29	38	218

Tabell 4. Longitudinell utvecklingsprofil, gruppfrekvenser för 6 inlärare.

Strukturer	Inspelningstillfällen										Summa 1-10
	1	2	3	4	5	6	7	8	9	10	
1:1 <i>det är</i> + NP, ej Rel	18	30	10	26	14	50	58	26	37	85	354
1:2 <i>det är</i> + NP, med Rel		2		2	2	3	12	2	5	14	42
2 <i>det är</i> + Adj		21	7	34	47	73	58	38	75	89	442
3 <i>det är</i> + Adv/PP		5	3		1	14	11	7	15	14	70
4 <i>det är</i> + Infinitiv						3	1		2	7	13
5 <i>det är</i> + Bisats		1		2		1	2	1	5	6	18
6 <i>det är</i> med Strandad prep										1	1
7 'Tough'-konstruktion				2	3	2	1	1		3	12
8 Obj <i>det</i> lyft från bisats											-
9 <i>det är</i> med Vänsterdislokering		2		3	4	6	5	3	7	9	39
10 Pseudoklyvning										2	2
11 <i>det är</i> terminalt i Bisats						1		2		3	6
12 Opersonlig konstruktion		1	8	3		2	2	3	2	11	32
13:1 Existentialkonstr, ej Rel	8	1	5	5	2	9	1	2	2	18	53
13:2 Existentialkonstr, med Rel							1		6	9	16
14 Anteciperande		2	3	4	13	13	16	21	28	45	145
15 Högerdislokering					2	1				2	5
16 Satsklyvning				1		2	4	2	9	16	34
17 <i>det är</i> upprepat	3	12	7	25	20	19	22	17	22	44	191
18 <i>det är</i> + omformulering		1	6	5	13	10	24	12	31	36	138
19 <i>det är</i> + övergivet yttrande	1	6		2	5	7	11	3	3	10	48
<b>Summa <i>det är</i></b>	30	84	49	114	126	216	229	140	249	424	1661
<b>Textlängd (antal ord)</b>	2787	4239	5056	6505	5660	7735	8311	8241	9109	17744	75382
<b>% <i>det är</i> av textlängden</b>	1.0	2.0	1.0	1.8	2.2	2.8	2.8	1.7	2.7	2.4	2.2