LUND UNIVERSITY

Den finita satsen i små barns språk

Wikström, Åsa

2008

Document Version:
Förlagets slutgiltiga version

Link to publication

Citation for published version (APA):
Wikström, Å. (2008). *Den finita satsen i små barns språk*. (Nordlund; Vol. 28). Institutionen för nordiska språk, Lunds universitet.

Total number of authors:
1

General rights
Unless other specific re-use rights are stated the following general rights apply:
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.
• Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
• You may not further distribute the material or use it for any profit-making activity or commercial gain
• You may freely distribute the URL identifying the publication in the public portal

Read more about Creative commons licenses: https://creativecommons.org/licenses/

Take down policy
If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

LUND UNIVERSITY

PO Box 117
221 00 Lund
+46 46-222 00 00

Download date: 10. Oct. 2026

Nordlund 28
Småskrifter från Nordiska språk
vid Lunds universitet

Den finita satsen i små barns språk

Åsa Wikström

Nordlund 28
Småskrifter från Nordiska språk
vid Lunds universitet

ISSN 0281-5427

LUNDS UNIVERSITET

Lund 2008
Nordiska språk, Språk- och litteraturcentrum
Box 201
221 00 Lund

Den finita satsen i små barns språk

Åsa Wikström
Asa.Wikstrom@nordlund.lu.se

Lund 2008

© 2008 Åsa Wikström

Nordiska språk, Språk- och litteraturcentrum
Lunds universitet

Tryck: Media-Tryck, Lunds universitet, maj 2008

ISSN 0281–5427

Innehåll

1 Inledning **1**

1.1 Bakgrund 2

1.1.1 Språkuppbyggnad och språkförlust 2

1.1.2 Den generativa modellens minimalistiska version . . . 2

1.1.3 EPP i språkuppbyggnad och språkförlust 4

1.2 Ledföljd och obligatoriska led i tidig svenska 6

1.2.1 Ledföljd och obligatoriska led i finita satser 6

1.2.2 När finns den finita huvudsatsen? 8

1.3 Syfte 9

1.4 Avhandlingens disposition 10

2 Hannakorpusen 1.0 **13**

2.1 Bakgrund och syfte 13

2.2 Beskrivning av korpusen 15

2.3 Teknisk information 17

2.4 Transkribering 17

2.4.1 Att tolka det barnet säger 20

2.4.2 Begreppet *ord* 20

2.4.3 Begreppet *yttrande* 22

2.4.4 Upprepningar 26

2.4.5 Verb och verbmorfologi 29

2.5 Materialets representativitet 31

2.6 Översikt över Hannakorpusen 33

3 Manual för satsanalys i barnspråk **37**

3.1 Inledning 37

3.2 Syntax i samtal 40

3.2.1 Huvudsatser 40

3.2.2 Huvudsatsekvivalenter . . . 40
3.2.3 Tidigt barnspråk . . . 41
3.3 Analysprinciper . . . 44
3.3.1 Övergripande principer . . . 44
3.3.2 Hörbarhetsprincipen . . . 46
3.3.3 Undantag . . . 47
3.3.4 Satsens omfattning . . . 47
3.3.5 Utanför analysen . . . 49
3.3.6 Sammanfattning av analysprinciper . . . 50
3.3.7 Hur man avläser kodsträngarna . . . 51
3.4 Strängtyper . . . 52
3.4.1 € – meningsfragment . . . 52
3.4.2 $ – huvudsatser . . . 54
3.4.3 ¢– underordnade verbfraser . . . 55
3.5 Satsled . . . 57
3.5.1 S – subjekt . . . 57
3.5.2 F – finit verb . . . 59
3.5.3 V – verb . . . 61
3.5.4 O – objekt . . . 61
3.5.5 P – predikativ . . . 62
3.5.6 A – verbfrasadverbial . . . 62
3.5.7 I – satsadverbial . . . 63
3.5.8 L – verbpartikel . . . 64
3.5.9 W – frågeord . . . 65
3.5.10 Q – fråga/utelämnat v-frågeord . . . 65
3.5.11 U – bisatsinledare . . . 66
3.5.12 C – konjunktion . . . 66
3.5.13 + – utelämnat led . . . 67
3.5.14 N – obestämbart nominal . . . 68
3.5.15 R – rest/räkneord . . . 68
3.5.16 X – oanalyserbart . . . 69
3.6 Verbmorfologi . . . 69
3.7 Prepositionsfraser . . . 72
3.8 Slutkommentar . . . 73

4 Analys av utdata **75**
4.1 Språklig förmåga i stort . . . 76
4.1.1 Spontana imitationer . . . 76
4.1.2 Yttrandelängd . . . 78
4.1.3 Strängtyper . . . 81

4.2 Verb och verbmorfologi 84
4.3 Huvudsatser 90
4.3.1 Subjekt och verb 90
4.3.2 Interrogativa huvudsatser 94
4.3.3 Deklarativa huvudsatser 96
4.4 Underordnade verbfraser 104
4.5 Funktionsord 111
4.6 Den finita satsen hos Hanna 113

5 Avslutning **115**

Tabeller

2.1 CHAT-symboler 19
2.2 CHAT-symboler för prosodiska egenskaper 25
2.3 Hannakorpusen 34

3.1 Analyskoder 58

4.1 Upprepningar 77
4.2 Yttrandelängd 80
4.3 Antal strängar av varje typ 81
4.4 Antal och andel verb fram till 2;2 85
4.5 Verbformer 86
4.6 Huvudsatser 91
4.7 Andel pronomen som subjekt 92
4.8 Frågor 95
4.9 Deklarativa satser 97
4.10 Satsled i fundamentet 99
4.11 Underordnade verbfraser 105
4.12 Funktionsord 112

Förord

Mitt första och största tack går till min äldsta dotter, lätt maskerad till Hanna i denna avhandling. Tänk att du hade så mycket intressant att säga när du var helt liten! Mitt andra tack, också det det största, går till min andra dotter som dök upp när jag trodde mig veta hur saker och ting låg till och visade mig att det långt ifrån var så enkelt.

Utan mina handledare Christer Platzack och Lisa Christensen hade detta arbete aldrig blivit färdigt. Utan Christer, Erik Philip Sörensens fond för humaniora och genetik och Vetenskapsrådet hade jag inte ens kunnat påbörja arbetet. Tack Christer för ditt stöd, ditt engagemang och dina skarpsinniga kommentarer, och för att du inte lät mig ge upp. Tack Lisa för din förmåga att handleda med en lagom blandning av utmanande kritik och uppmuntrande entusiasm. Jag törs inte ens tänka på hur avhandlingen hade sett ut utan era insatser.

Tack alla som deltog i slutseminariet och kom med kloka synpunkter, särskilt förstås Jonas Granfeldt som gjorde ett imponerande jobb som opponent, men också Christian Waldmann och Henrik Rosenkvist som läste hela texten och hjälpte mig att se inkonsekvenser och obegripligheter. Nordiska språk kan vara världens bästa arbetsplats. Jag kan inte säga det säkert, men det skulle förvåna mig om det fanns ett bättre ställe. Tack alla kolleger, doktorander och kafferaster! Tack till språkinlärningsseminariet för inspirerande diskussioner om hur språkinlärning egentligen går till. Tack till Anna Gustafsson för uppiggande luncher och till mina rumskompisar Lena Lötmarker och Bo Wendt för seriösa och oseriösa samtal om språk och annat, och ett särskilt tack till Lena som tillhandahöll sitt exemplar av SAG.

Jag vill till slut också tacka mina föräldrar, Anne-Marie och Lars-Åke, som alltid har varit ett stort stöd, och alla andra släktingar och vänner som har ställt upp när tiden inte har räckt till. Och allra sist vill jag tacka min man Per för så mycket, att jag inte vet var jag ska börja. Så jag börjar inte – men du vet.

Kapitel 1

Inledning

År 1925 lade Gunnar Freudenthal fram en licentiatavhandling vid den filosofiska fakultetens humanistiska sektion vid Lunds universitet som behandlade skillnaderna mellan modersmålsutveckling och språkundervisning (Freudenthal, 1925). Syftet var främst språkdidaktiskt. Språkundervisningens metoder debatterades utifrån jämförelsen mellan ”å ena sidan det tvångslösa och omedelbara sätt, varpå barnet förvärvar sitt modersmål, å andra sidan skolkursens tunga arbete och torftiga behållning” (s. 5). Man var överens om att de olika språkinlärningssätten var väsensskilda men Freudenthal menade att man inte kunde nöja sig med det, utan att man måste fastslå vad som skiljer språkundervisningen i skolan från ”den effektbetingade utvecklingen” (förstaspråksinlärningen). För att kunna studera förstaspråksinlärningen tog Freudenthal fram ett eget barnspråksmaterial. Det består till största delen av nedskrivna spontana yttranden från en av hans söner under åldern 1;5 till 2;4 [1].

Drygt åttio år senare lägger jag fram en annan licentiatavhandling som behandlar modersmålets utveckling, och jag presenterar också ett nytt barnspråksmaterial. Det som för mig började som ett intresse för ledföljd och obligatoriska led i finita satser hos små barn blev snart till en undran över hur den finita satsen blir till. I svenska består finita satser i normalfall av ett finit verb på andra plats i satsen, ett i stort sett valfritt led på första plats och ett subjekt. Hur lär sig barnen det? Hur kan man avgöra om barnet vet det?

I den här avhandlingen ska jag undersöka om man utifrån spontandata kan avgöra när ett barn har tillägnat sig den finita satsen. Avhandlingen är en del i en större undersökning av obligatoriska led och ledföljd hos små barn. Jag beskriver en metod för att undersöka satser i små barns språk och jag presenterar

[1] 1 år och 5 månader till 2 år och 4 månader.

ett nytt longitudinellt barnspråksmaterial. I det följande ska jag kort beskriva den större undersökningen för att därefter ringa in syftet med föreliggande avhandling. Sist i kapitlet går jag igenom bokens disposition.

1.1 Bakgrund

1.1.1 Språkuppbyggnad och språkförlust

Inom språkinlärningsforskningen har man länge sysselsatt sig med frågor som rör felaktig placering av det finita verbet i V2-språk, förekomsten av satser med enbart infinit verb (så kallade *rotinfinitiver*), utelämning av subjekt i språk med subjektstvång och utelämning eller felaktig placering av v-frågeord i kvesitiva frågor. För svenska gäller att en del av dessa fenomen är vanligare i små barns förstaspråksinlärning och andra är vanligare i andraspråksinlärning, men att alla i någon utsträckning förekommer i inlärarspråk. Också barn med språkstörning, ofta kallad *SLI* (Specific language impairment[2]), visar samma sorts avvikelser (bland andra avvikelser, se t.ex. Håkansson och Hansson (2007)). Christer Platzack (2005, 2007) framför en hypotes om att dessa avvikelser kan beskrivas på ett samlat sätt för alla dessa grupper (i språkuppbyggnad) samt för Brocas afatiker (i språkförlust).

Projektet Språkuppbyggnad och Språkförlust fick 2001 medel från Erik Philph Sörensens fond för humaniora och genetik för att närmare undersöka Plazacks hypotes, och min uppgift i projektet är att undersöka det tidiga språket hos barn med typisk språkutveckling. I projektet avsattes också medel för att sammanställa ytterligare barnspråksmaterial, vilket har resulterat i Hannakorpusen som presenteras i den här avhandlingen.

Plazacks hypotes är formulerad inom den generativa grammatikens ramverk, och jag ska först mycket kortfattat gå igenom den generativa modellen för att sedan gå närmare in på Platzacks hypotes. Det fortsatta arbetet i avhandlingen kommer inte att vara beroende av den generativa begreppsapparaten, så den följande redogörelsen är starkt förenklad och avser endast att utgöra bakgrund för Platzacks hypotes.

1.1.2 Den generativa modellens minimalistiska version

Den generativa grammatiken utgår ifrån att mänskliga språk till del är ett resultat av ett medfött grammatiksystem som är oberoende av andra kognitiva

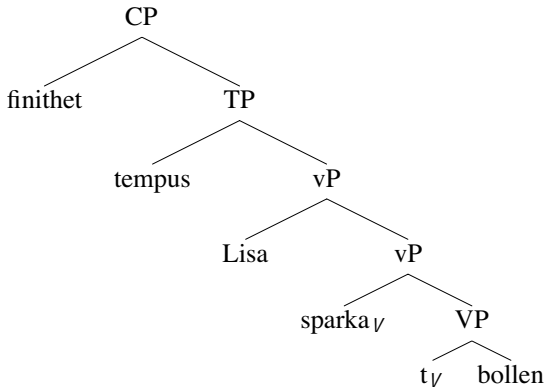
[2]Specifik språkstörning innebär att barnet har vissa svårigheter med språkinlärningen medan den övriga utvecklingen är opåverkad. Man har kunnat visa att språkstörningen i alla fall delvis har genetiska orsaker (Nettelbladt och Salameh, 2007).

förmågor. Språket är förstås också ett resultat av andra kognitiva system, och i det minimalistiska programmet vill man undersöka hur mycket av språket som bestäms av dessa andra system, som t.ex. begreppsapparat och minne, och vad som är specifikt för språket.

I den generativa grammatiken antar man att språkstrukturen är uppbyggd med hjälp av en generell princip – förbind. Den säger att man får sätta ihop något med något annat i binära kombinationer. Det betyder att grammatiken tar en byggsten (t.ex. ett ord eller ett morfem) och kombinerar det med andra byggstenar så att en fras bildas. Ett morfem som *sparka* kombineras t.ex. med en nominalfras *bollen* så att verbfrasen *sparka bollen* bildas. Nominalfrasen är i sin tur bildad av ett rotmorfem, som bär betydelsen *boll*, och särdrag som specificerar sådant som person, numerus och genus (som brukar sammanfattas med termen ϕ-särdrag) och species, i det här fallet särdrag som säger att det rör sig om *en* (utrum, singular) *specifik* (definit) boll. Verbfrasen kan sedan kombineras med en ny byggsten – t.ex. en annan nominalfras, som kan fungera som agent. I verbfrasen kombineras verbet alltså med sina argument genom att nya byggstenar förbinds in i strukturen.

För att göra en sats av verbfrasen måste man kombinera den med fler särdrag som t.ex. anger tempus och finithet, så att propositionen får en koppling till yttrandeögonblicket. I en sats måste det alltså finnas finithetssärdrag och tempussärdrag – man kallar de fraser eller funktionella noder som bär dessa särdrag för C respektive T. En förenklad trädmodell för en sats brukar se ut som i figur 1. VP står för den fras som bildas när verbet förbinds med sitt komplement, medan *v*P är nästa nivå där verbets yttre argument är infogat i strukturen. I figuren visas en ofärdig sats – en verbfras som ännu inte har förankrats i talarens här och nu.

(1)



Den grammatiska strukturen byggs i den här modellen alltså upp utifrån vilka särdrag som plockas ut från lexikon. Vissa särdrag måste stå i relation till andra särdrag för att få rätt tolkning. En sådan relation är den mellan subjektet och tempussärdragen. Det finns olika förslag på hur den relationen ska beskrivas. Platzack (2007) utgår från en idé av Pesetsky och Torrego (2001), som laborerar med otolkbara och tolkbara särdrag. Enligt den måste otolkbara särdrag matchas med tolkbara så att inga omatchade otolkbara särdrag finns kvar när grammatikprocessen är slutförd. När de semantiska och prosodiska gränssnitten ska avläsa resultatet av processen kan inga särdrag vara otolkade. För subjekt–tempus-relationen innebär det att det finns otolkbara personsärdrag (ϕ-särdrag) i tempusnoden som måste stå i relation till de tolkbara personsärdragen hos subjekts-DP:n för att det semantiska gränssnittet ska kunna avläsa särdragen rätt[3].

Relationen mellan särdragen visar sig ibland genom att hela särdragsknippet flyttar[4]. I svenska görs relationen mellan subjekt och tempus synlig genom att subjektet flyttar till tempusnoden, vilket helt enkelt innebär att knippet av särdrag, inklusive de fonetiska särdrag som också ingår, återigen förbinds in i strukturen, men den här gången med T. Den egenskap som gör att något måste flytta (eller återförbindas) finns hos särdragen och kallas EPP[5] och den är helt och hållet språkspecifik. Varför vissa särdrag bär på en sådan egenskap är något som vi ännu inte förstår oss på, menar Platzack, men EPP-begreppet kan användas för att beskriva ordföljsfenomen. I svenska finns alltså EPP-egenskaper hos särdrag i T så att subjektet måste flytta och återförbindas med T, och det är därför svenskan har en subjektsposition efter det finita verbet och obligatoriskt utsatt subjekt. Liknande EPP-egenskaper, som jag beskriver närmare nedan, gör också att svenskan har det finita verbet på andra plats och ett led i fundamentet.

1.1.3 EPP i språkuppbyggnad och språkförlust

Platzack (2007) menar att de grammatiska avvikelserna hos små barn, avancerade andraspråksinlärare, barn med SLI och vuxna med Brocas afasi kan

[3]Relationen är mer komplicerad än så i och med att man också antar det omvända, att det finns tolkbara tempussärdrag i T och otolkbara tempussärdrag i subjektet, men redovisningen här syftar bara till en överskådlig bild av hur syntaxen arbetar och jag försöker därför hålla det så enkelt som möjligt.

[4]Man använder ofta termen flyttning, trots att man antar att särdragen egentligen kopieras, och att kopian i den ursprungliga positionen i normalfall inte uttalas.

[5]Från början stod EPP för Extended Projection Principle, den funktionella frasen projicerar så att ett externt argument, alltså en fras, kan förbindas dit, men Pesetsky och Torrego (2001) använder termen även för huvudflytt.

betraktas på ett gemensamt sätt, närmare bestämt som ett problem med att automatiskt applicera kunskapen om var de språkspecifika EPP-egenskaperna finns. Språkbrukarna i dessa grupper utelämnar t.ex. oftare än andra subjektet i omgivningar där det är obligatoriskt, men det är inte på något sätt ett konsekvent drag i deras produktion. Ofta gör de rätt, dvs. de sätter ut subjekt på rätt plats i satsen. Det innebär, menar Platzack, att dessa språkbrukare kan antas ha samma inre grammatik som vuxna svenskar med svenska som förstaspråk, och att avvikelserna måste bero på något utanför det syntaktiska systemet. De har inte en annan kunskap om språket, en annan kompetens, utan problemet ligger i produktionen.

I en svensk sats finns EPP-egenskaper knutna till det otolkbara ϕ-draget i tempusnoden (T). Otolkbara drag anges med *u*, så det otolkbara draget i T betecknas uϕ^{EPP}. Detta söker neråt i strukturen efter ett matchande tolkbart ϕ-drag och hittar subjekts-DP:n, som på grund av EPP-egenskapen måste återförbindas med T. Det innebär att subjektet fogas till T-noden (man kallar positionen spec-TP). EPP-egenskapen medför dels att ett subjekt inte får utelämnas i svenskan, dels att subjektet står i den första positionen efter det finita verbet (om det inte står i fundamentet).

Det finns också en EPP-egenskap i det otolkbara tempusdraget i finithetsnoden (C) uτ^{EPP}. Det draget söker efter ett matchande tolkbart särdrag och finner då det tolkbara tempusdraget τ i T och det otolkbara draget i C elimineras. Men EPP-regeln säger att det måste finnas fonologiskt material i C, och det närmaste huvudet med fonetiska särdrag är det finita verbet, som då flyttas till C. Det innebär att det finita verbet står på andra plats i satsen, om den inte är fylld av en subjunktion, då det finita verbet står kvar i *v*P.

I C finns ett kantdrag (*edge feature*, ϵ^{EPP}), som är av en delvis annan typ. Det innebär att en fras måste finnas i förstapositionen. I deklarativa satser är det oftast satsens topik som återfinns först i satsen, i interrogativa satser ett v-frågeord eller en frågeoperator. Första positionen antas alltså alltid vara fylld, även om det inte alltid är med några fonetiska särdrag (Mörnsjö, 2002).

1.1.3.1 Hela strukturen finns

Platzack (2007) visar med exempel ur en och samma inspelning av ett barn, Sara vid 1;11, att hon har hela satsstrukturen klar för sig. Exempel 2 visar att hon är medveten om att det finns ett kantdrag i C så att ”nu” hamnar i första position, hon uppmärksammar uτ^{EPP} och placerar det finita verbet i C samt uϕ^{EPP} och placerar subjektet direkt efter det finita verbet i spec-TP.

(2) nu får pappa denna. Sara 1;11

Visserligen kan man inte vara säker på att subjektet inte står *v*P. Man antar att satsadverbial är infogade ovanför *v*P, så en liknande sats med satsadverbial kunde ha visat om subjektet har flyttat eller inte (”nu får pappa inte denna”). En sådan sats finns inte i just denna inspelning, men det finns andra skäl att anta att subjektet har flyttat ur *v*P.

1.1.3.2 Brott mot EPP

Det barnen har problem med, menar Platzack, är sådant som rör flyttningar till de funktionella projektionerna i satsen. I exemplen nedan (från Platzack, 2007) bryter barnen i tur och ordning mot $u\epsilon^{EPP}$ (i exempel 3 saknas topiken ”nu”, i exempel 4 är v-frågeordet inte utsatt), mot $u\tau^{EPP}$ (i exempel 5 har det finita verbet inte flyttat till C utan står kvar i vP och i exempel 6 saknas det finita verbet helt) och mot $u\phi^{EPP}$(i exempel 7 saknas subjektet). Jag har normaliserat stavningen i exemplen.

(3)	har pappa det.	Sara 1;11
(4)	gör den?	Kenneth 2;2
(5)	va Millan gjorde.	Markus 1;11.25
(6)	docka sitta där.	Sara 1;11
(7)	där ska stå.	Sara 1;11

Platzack visar därefter att satser med det finita verbet på tredje plats (V3-satser), rotinfinitiver och subjektslösa satser också är utmärkande för språket hos avancerade andraspråkinlärare, barn med SLI och Brocas’ afatiker. Detta gör att han kan beskriva de vanliga grammatiska avvikelserna i de olika språkgrupperna på ett enhetligt sätt. Enligt Platzack har de små barnen med typisk utveckling, SLI-barnen, afatikerna och de avancerade andraspråksinlärarna en inre grammatik som överensstämmer med målspråkets, men de har problem med att konsekvent uppmärksamma de flyttningar som är ett resultat av EPP-egenskaperna i vissa särdrag.

1.2 Ledföljd och obligatoriska led i tidig svenska

1.2.1 Ledföljd och obligatoriska led i finita satser

Projektet Språkuppbyggnad och språkförlust har som syfte att närmare undersöka Platzacks hypotes om EPP-relaterade problem. Det självklara sättet att

försöka att falsifiera hypotesen är att testa den mot ett brett material, bestående av olika grupper som talar olika språk. Varje språkbrukargrupp där avvikelserna kan beskrivas som en svårighet att automatisera kunskapen om EPP-egenskaperna korroborerar hypotesen. Hittar man däremot ett språk där de vanliga syntaktiska avvikelserna klart är orelaterade till EPP-egenskaper förlorar hypotesen i trovärdighet.

Ett annat sätt att testa hypotesen är att göra en grundlig felanalys av språket hos en individ ur någon av språkgrupperna, och se hur de avvikelser som hittas kan beskrivas. Om avvikelser som rör syntaxen (till skillnad från sådant som rör lexikal utveckling, semantik och pragmatik) inte kan beskrivas som EPP-relaterade förlorar hypotesen också i trovärdighet. I en felanalys av språket kan man utgå ifrån spontandata, men för att komma åt lågfrekventa avvikelser kan också eliciteringar och andra experimentella metoder komma i fråga.

Syntaktiska avvikelser kan vara avvikande placering av led eller utelämningar av obligatoriska led i finita satser. Platzack undersöker vissa led som är centrala i satsen – subjekt, finita verb och v-frågeled i fundamentet. Josefsson (2003b) redovisar en översikt över icke-målspråkliga strukturer i barns yttranden där även hon redogör för avvikande placering eller utelämningar av subjekt och finita verb, men också nämner utelämningar av prepositioner och subjunktioner. Josefsson och Håkansson (2000) föreslår också att framväxten av subjunktioner är kopplad till prepositionsbruket. I en bred undersökning av ledföljd och obligatoriska led i finita satser bör man få med så många som möjligt av dessa led, och den satsanalys jag ska beskriva motiveras inte bara av forskningsfrågan i den här avhandlingen utan också av detta kommande arbete.

Undersökningen fokuserar på *obligatoriska* led i *satsen*, vilket medför två viktiga avgränsningar. För det första analyserar jag inte obligatoriska led och ledföljd i nominalfrasen. Nominalfrasen är naturligtvis mycket central i det tidiga barnspråket, men den utgör också ett stort forskningsområde i sig självt och av tidsskäl har jag valt att helt utelämna den. För det andra analyserar jag inte led som aldrig ingår i satsstrukturer, dvs. interjektionsfraser, vokativfraser och konjunktioner utanför satsen. Interjektionsfraser och vokativfraser har en egen intern struktur och förekommer tidigt i barns språk. Men det är oklart om och i så fall hur de ska relateras till den finita satsen. För konjunktioner utanför satsen är det omöjligt att fastställa vad som utgör en obligatorisk syntaktisk kontext och det blir då svårt att undersöka eventuella utelämningar. Naturligtvis är interjektionsfraser, vokativfraser och konjunktioner utanför satsen intressanta när man undersöker barns grammatiska utveckling. Men de faller alltså utanför mina studier som har den finita satsen som huvudområde.

1.2.2 När finns den finita huvudsatsen?

Enligt EPP-hypotesen vet barnen att det finns en EPP-egenskap förknippad med ett visst särdrag, men de har inte automatiserat den fullt ut. Att det finns ett otoklbart tempussärdrag i C, dvs. att finithet har en relation till tempus, är givet av det språkliga systemet. Svenska barn identifierar också att denna relation ska vara synlig, dvs. att uτ^{EPP}bär på en EPP-egenskap, och barnen vet tidigt att det i typfallet är ett finit verb i C som gör denna relation synlig (se t.ex. Santelmann, 1995; Josefsson, 2003a). På samma sätt är relationen mellan tempus och person given av systemet, dvs. den relation som kallas uϕ^{EPP}. Barnen vet tidigt att också den relationen ska vara synlig (att det otolkbara personsärdraget i T bär på egenskapen EPP) och att det sker genom att subjektet är synligt i positionen spec-TP (ibid).

Frågan är hur tidigt barnen vet vilka relationer som ska vara synliga i satsen. Barn börjar förstås inte att tala med finita verb och subjekt. Det är en allmän iakttagelse att barn börjar med infinita verbformer. Det betyder att det finns en initial fas där vi inte kan veta om barnen har förstått om och hur relationen mellan finithet och tempus görs synlig, eftersom de inte använder finita verbformer. Innan vi vet att det finns finita verbformer och att de sitter i C blir det svårt att säga något om att kunskapen om EPP-egenskapen är automatiserad.

Barn börjar också med ett ord i taget. Det finns begränsningar för hur många ord de kan kombinera och processa i det tidigaste skedet. Därför är det naturligtvis inte konstigt att subjekten saknas i början. De kan knappast automatisera en kunskap om att subjektet alltid måste vara synligt om de bara kan processa ett ord i taget (jfr. Strömqvist och Ragnarsdóttir, 2000). Innan man börjar undersöka om kunskapen om att subjekten ska vara synliga är automatiserad bör man veta att barnen har kommit till en sådan språknivå att de generellt sett kan producera yttranden med subjekt.

Hypotesen om *automatisering* av kunskap förutsätter att kunskapen finns där. Innan barnet producerar finita huvudsatser är det inte meningsfullt att tala om saknade obligatoriska led. Det är inte heller självklart hur ledföljden i de allra tidigaste yttrandena ska betraktas. I förlängningen leder resonemanget in på frågan om små barns kompetens. Platzack (2007) utgår ifrån antagandet att de tidiga språkinlärarna har samma I-språk[6] som vuxna, i likhet med t.ex. Santelmann (1995) och Josefsson (2003a), vilket brukar benämnas *Full competence hypothesis*. Det är dock svårt att veta något om kompetensen när man analyserar en språkproduktion som är så begränsad, vilket är ett välkänt faktum. Och precis hur man kan förvänta sig att barnets allra tidigaste språkpro-

[6]I-språk innebär vårt interna språk, dvs. den omedvetna kunskap om språket som språkbrukarna har.

duktion ska se ut om man utgår ifrån att barnet har full kompetens från början, dvs. vad full kompetens egentligen innebär, är fortfarande en öppen fråga.

I Platzacks pilotstudie visar han med exempel att Sara har hela satsstrukturen i den första transkriberade inspelningen som finns för Sara. Men om man som jag har tillgång till inspelningar som sträcker sig längre tillbaks i barnets språkutveckling, hur kan man avgöra närifrån man kan börja att undersöka hypotesen? Hur kan man avgöra att barnet har etablerat den finita satsen?

Den finita satsen konstitueras av nexusförbindelsen mellan subjektet och predikatsledet, vars huvudord är ett finit verb (Teleman m.fl., 1999, IV, s. 3). En huvudsats är sats som inte fungerar som led i någon annan sats, och den känns igen på en viss struktur (ledföljd) och en viss prosodi. Den prototypiska huvudsatsen utgör också en egen språkhandling och har en textuell funktion (Teleman m.fl., 1999, IV, s. 674ff). Alla dessa egenskaper är relevanta om man vill identifiera framväxten av den finita huvudsatsen. Eftersom det finita verbet är centralt, är verbens morfologi också en viktig aspekt.

1.3 Syfte

Den här avhandlingen utgör en utgångspunkt för den vidare undersökningen av den finita satsen och Platzacks EPP-hypotes. Det första syftet här är att undersöka om jag utifrån transkriberad spontandata kan identifiera när ett barn har kunskap om vilka grundläggande egenskaper den finita huvudsatsen har i svenska.

För ändamålet har jag utformat en metod för att analysera satser och satsled i små barns språk. I analysen tar jag i huvudsak hänsyn till barnets syntaktiska kompetens, dvs. jag analyserar yttranden med avseende på satsled och ledföljd. Men i den syntaktiska analysen har jag också varit uppmärksam på barnets prosodiska kompetens och jag har trots viss brist på teoretisk underbyggnad inkluderat en del prosodiska egenskaper i analysen. Eftersom det finita verbet är förknippat med tempusböjning och verbmorfologi har jag också inkluderat en analys av den morfologiska kompetens hos barnet som är kopplad till verben. Däremot har jag inte tagit hänsyn till den pragmatiska sidan av yttrandena, dvs. satsernas språkhandlingar och textuella funktion. Att fastställa språkhandlingar i barnyttranden framstår som en mycket svår uppgift (annat än i de mest prototypiska fallen) och skulle kräva en insats som jag inte har haft utrymme för.

Analysmetoden är också utformad med syftet att ge mig svar på de frågor jag ställer mig i den större undersökningen av ledföljd och obligatoriska led i den finita satsen. Ett andra syfte blir därmed att beskriva, tillämpa och utvärdera

den metod som jag har utarbetat.

Analysen genomför jag på Hannakorpusen som är framtagen med delvis andra syften än att svara på den övergripande frågan i denna avhandling. Att presentera Hannakorpusen och att redogöra för de ställningstaganden jag har fått göra i sammanställningen av korpusen är ett tredje syfte i denna avhandling.

1.4 Avhandlingens disposition

Avhandlingen består av två delar som till stor del är oberoende av varandra. I den första delen, kapitel 2, presenterar jag en ny longitudinell barnspråkskorpus, Hannakorpusen, och redovisar en del metodiska frågor som man måste ta ställning till när man sammanställer en korpus av talat barnspråk. Hannakorpusen är framtagen med syftet att utöka mängden data över svenska enspråkiga förstspråksinlärare. Materialet består av 16 transkriberade halvtimmesinspelningar av min egen dotter Hanna mellan 1;6 och 2;10. Transkriptionerna omfattar ca 5000 barnyttranden som är länkade till ljud- eller videofiler. Tanken är att korpusen ska vara tillgänglig för andra forskare och kapitel 2 utgör materialets dokumentation. Kapitlet är därmed skrivet för att kunna stå på egna ben.

Avhandlingens andra del består av kapitlen 3, 4 och 5. I kapitel 3 beskriver jag analysmetoden. Större delen av kapitlet består av riktlinjer för hur yttranden, satser och satsmotsvarigheter, satsled och ledföljd ska analyseras och kapitlet mynnar ut i en manual för satsanalys i barnspråk. Medan den vidare undersökningen alltså sker inom ramen för den generativa grammatiken, baserar sig analysmetoden till största delen på en traditionell grammatikbeskrivning.

I kapitel 4 redogör jag för utfallet av analysen och diskuterar vad man kan dra för slutsatser. Utfallet av analysen visar att man kan få en god bild av hur den finita satsen växer fram hos Hanna och när den är etablerad. Men eftersom analysen ännu bara är utförd på ett enda barn får undersökningen betraktas som en pilotstudie, och det är en öppen fråga om analysen ger lika tydliga resultat för barn i allmänhet. Hela den transkriberade Hannakorpusen är analyserad och utfallet av analysen blir samtidigt en kartläggning över den del av Hannas språk som är dokumenterad och en utförlig beskrivning av hur olika satsrelaterade fenomen växer fram.

Utvärderingen av metoden sker delvis löpande i kapitel 4, då jag diskuterar vissa problem med analysmetoden, delvis mer systematiskt i det avslutande kapitel 5. I utvärderingen tar jag ställning till om analysen ger mig svar på följande frågor: 1) Kan jag med hjälp av analysen identifiera när barnet har

satsstrukturen klar för sig, så att det är möjligt att undersöka hypotesen om EPP-relaterade fel? 2) Kan jag undersöka alla de fenomen Platzack (2007) och Josefsson (2003b) undersöker? 3) Ger mig analysmetoden möjlighet att hitta andra systematiska satsrelaterade avvikelser som skulle kunna förekomma, så att jag kan undersöka om hypotesen om EPP-relaterade fel predicerar de avvikelser som finns i barnets språkutveckling?

Kapitel 2

Hannakorpusen 1.0

2.1 Bakgrund och syfte

Hannakorpusen är sammanställd inom projektet Språkuppbyggnad och språkförlust som finansierades av Erik Philip Sörensens fond för humaniora och genetik. Den huvudsakliga uppgiften i projektet var att undersöka satsrelaterade fenomen hos olika språkbrukare, men en del av projektresurserna avsattes också för ytterligare insamling av material. Detta har resulterat i den korpus som jag nu presenterar – Hannakorpusen.

Undersökningar av barnspråk grundar sig i iakttagelser av hur barn talar. Analyser av spontant tal är visserligen inte det enda sättet att undersöka barnspråk, men insamlade longitudinella material behövs om man ska kunna undersöka språkinlärningsprocessen. Med nya tekniker för att lagra talat språk har forskningen kunnat utvecklas. Barnspråksforskningen kunde inte riktigt ta fart förrän tekniken att spela in talat språk kom med bandspelaren under sent 50-tal (MacWhinney, 2006a, s. 7). Tekniken har sedan dess underlättat forskningen i allt högre grad. Nu är den så pass bra att vi kan ha stora korpusar av barnspråk med både text, ljud och bild, där varje yttrande är länkat till ljud- eller videofilen, så att man bara genom att klicka på ett transkriberat yttrande kan höra det. Med stora transkriberade material som underlag kan man dra säkrare slutsatser om vad barn säger och inte säger. Men inspelade transkriberade material rymmer också metodfrågor som man måste ta på allvar. Hur och när man väljer att spela in och vilka principer som ska ligga till grund för transkriptionen får följder för hur språket i materialet ser ut.

Metodfrågor som rör transkription av talspråk har diskuterats i olika forum, utifrån olika utgångspunkter (se t.ex. McDaniel m.fl., 1996; Byrman m.fl.,

1999). När det gäller barnspråk är CHILDES det mest kända forumet. CHILDES står för Child Language Data Exchange System och är ett löpande projekt som startades 1984 av Brian MacWhinney och Catherine Snow (MacWhinney, 2006a, s. 9). CHILDES består av tre delar, eller verktyg, som kompletterar varandra: För det första består det av CHAT-verktyget som är ett format och en teknik för transkription och kodning av talade texter. För det andra består det av CLAN-verktyget som är en samling dataprogram som kan göra olika automatiska beräkningar på språk som är transkriberat i CHAT-format. För det tredje består det av en databas av transkriberat barnspråksmaterial från en mängd olika språk som kan användas av olika forskare med olika syften. Till CHAT och CLAN finns utförliga manualer där metodfrågor lyfts fram och diskuteras. Det ska dock påpekas att CHAT-formatet ger utrymme för variationer och att CHAT-manualen inte ger några direktiv för hur man ska göra. Ansvaret för att utforma sin korpus efter genomtänkta principer lämnas helt till den som tar fram den.

Det finns totalt omkring 20 longitudinella korpora från enspråkiga svenska barn (Håkansson, 2003), vilket i och för sig är rätt mycket. Men många av materialen är framtagna för den egna forskningen och har inte gjorts tillgängliga för andra. Det betyder också att vi inte vet mer om dessa barn än det som beskrivs i de enskilda studierna. På CHILDES har transkriptioner av tre barns språk från Strömqvists och Richthoffs longitudinella korpus funnits lätt tillgängliga för alla slags undersökningar (Strömqvist m.fl., 1993; Richthoff, 2000). Nyligen har data från ytterligare två barn tillkommit. Göteborgskorpusen, som den populärt kallas, är mycket användbar, inte bara för att den är tillgänglig och maskinläsbar utan också för att den består av flera barn som är dokumenterade på samma sätt, efter samma genomtänkta principer. Man kan alltså på ett enkelt sätt göra undersökningar baserade på fler barn och därmed komma fram till resultat som är generaliserbara på ett mer övertygande sätt än om man bara utgår från ett enda barn (även om det också kan ge intressanta resultat).

Hur många barn behöver man kartlägga för att resultatet ska vara representativt för svenskt barnspråk i allmänhet? Jag vet inte svaret på den frågan, men jag anar att fem är för få, och att inte ens 20 räcker så långt. Med hjälp av andra metoder och med tvärsnittsmaterial vet vi mer. Men de longitudienlla studier som finns av svenska barns språkinlärning täcker bara en handfull olika barn, och varje studie som tillkommer bidrar därför till att bilden av hur svenska barn lär sig sitt modersmål blir fullständigare. Om den allmänt tillgängliga databasen med svenska barn blir större kan också fler studier av olika karaktär göras, utan att varje projekt ska behöva börja med den kostsamma processen att sammanställa material. Att bidra med ett nytt svenskt barnspråksmaterial är

därmed ett syfte i sig i denna avhandling.

Ytterligare ett syfte med denna del av avhandlingen är att lyfta fram några metodfrågor som var och en som transkriberar svenskt barnspråk måste ta ställning till. De flesta barnspråksstudier presenteras i artiklar i tidskrifter eller samlingsvolymer, och där finns sällan utrymme för att i detalj redovisa vilka metodiska överväganden man har gjort. Två större svenska barnspråkskorpusar finns beskrivna i litteraturen: Söderberghs Stockholmskorpus (Söderbergh, 1975; Lange och Larsson, 1977) och Strömqvists och Richthoffs Göteborgskorpus (Strömqvist m.fl., 1993; Richthoff, 2000). I beskrivningarna av båda dessa korpusar redovisas en del av de metodfrågor som jag har förhållit mig till i mitt arbete. Det finns också en Lundakorpus som innehåller material från förstaspråksinlärare med typisk språkutveckling, tidiga andraspråkinlärare och barn med språkstörning (Hansson m.fl., 1999). Korpusen har genererat många studier, och i några artiklar betonas vikten av en metodisk medvetenhet särskilt (Hansson m.fl., 1999; Nettelbladt, 1994). Jag har dock inte funnit någon sammanhållen redovisning av hur de själva har valt att göra i de metodfrågor som jag tar upp.

I detta kapitel går jag igenom de metodiska överväganden jag har fått göra i transkriptionsarbetet och jag redovisar också vilka lösningar jag har valt. Jag diskuterar begreppet *ord* och *yttrande*, samt hur man ska betrakta upprepningar av olika slag. Jag redovisar också problem som uppstår när man transkriberar verbmorfologi – problem som är viktiga att vara medveten om när man använder transkriptionen som underlag för slutsatser om barns grammatiska utveckling.

Att sammanställa en talspråkskorpus innebär ett arbete utan ände. Det finns alltid något att förbättra eller att vidareutveckla. Hannakorpusen publiceras nu i version 1.0.A (videoversionen) och 1.0.B (audioversionen), men tanken är att nya versioner ska kunna bli möjliga i framtiden. Versionnumret finns i transkriptionfilernas inledande avsnitt. Målet är att B-versionen så småningom ska ingå i CHILDES-databasen, men den finns tills vidare tillgänglig via mig.

2.2 Beskrivning av korpusen

Hannakorpusen omfattar 22 videoinspelningar av ett barn som tillägnar sig svenska som modersmål i en enspråkig omgivning. Inspelningarna är gjorda under perioden september 2002 till juni 2004 när barnet är mellan 1;1.10 och 2;10.8[1]. Jag har i stort sett spelat in en gång i månaden, en halvtimme åt gången. Av de 22 inspelningarna är de 16 mellan 1;6.18 och 2;10.8 transkriberade

[1] 2;10.8 utläses 2 år 10 månader och 8 dagar.

och länkade till video- eller audiofilen. Den transkriberade korpusen består av totalt ca 40 000 ord (varav 14 500 från Hanna) och 11 000 yttranden (5 000 från Hanna)[2]. I mitt eget arbete har jag också haft nytta av en del sporadiska anteckningar, som dock ännu inte finns tillgängliga för andra.

Hanna är min egen dotter, något som naturligtvis har fått konsekvenser för hur materialet ser ut. Å ena sidan är det en stor fördel att arbeta med ett barn vars språk man känner väl och har alla förutsättningar för att förstå. Det är ofta nödvändigt att ha gemensamma referensramar för att kunna förstå vad barnet vill säga. Å andra sidan finns alltid risken att man som förälder övertolkar barnens yttranden och tror sig förstå mer än vad man faktiskt gör. Är man dessutom den som ska använda materialet i sin forskning finns kanske en viss risk att man styr samtalet under inspelningen.

Demuth (1996) betonar vikten av att planera datainsamlandet noggrant så att materialet blir så optimalt som möjligt. Innan man börjar att spela in ska man utifrån både kortsiktiga och långsiktiga forskningsfrågor bestämma sig för sådana saker som inspelningsfrekvens, interaktionspart, inspelningssituationer och inspelningsutrustning. För min del började jag med inspelningarna innan jag ens hade hunnit börja på projektet, och hann inte ta ställning till dessa frågor. Det innebär naturligtvis att metoden när det gäller själva inspelningarna är otillfredställande, men materialet är ändå tillräckligt bra för att vara användbart.

Hanna är första barnet till universitetsutbildade föräldrar. Pappan talar en sydsvensk dialekt och jag talar en mellansvensk dialekt, ibland med inslag av dalmål. Hanna började i förskolan vid ett års ålder, och gick under hela den inspelade perioden på en förskola med tre skånska barnskötare/förskollärare och en grupp barn som till största delen bestod av svenska enspråkiga barn. Hanna började tidigt att tala och talade också mycket och gärna med dem hon kände. I möte med främmande kunde hon dock vara lite återhållsam. I interaktion och lek var hon utåtriktad och social. När Hanna var 2;1 fick hon en lillasyster, vilket hon reagerade på med viss misstänksamhet och en period av något dämpat humör.

Inspelningarna är inte arrangerade på något utarbetat sätt. I de flesta fall deltar jag och pappan – den ena filmar medan den andra samtalar med Hanna. Inspelningssituationerna varierar, men vi har för det mesta spelat in i vårt vardagsrum där vi leker, bygger med lego, leker med ett dockhus, ritar och läser böcker. Det finns också några inspelningar från måltider. Eftersom Hanna talade mycket behövde vi inte driva på samtalet eller anordna några särskilda situationer för att få henne att säga något. I de senare inspelningarna deltog också lillasyster. Hon bidrog inte till den språkliga interaktionen i så hög grad,

[2]Alla slags ord och yttranden är inräknade, även otolkbara ord och minimala responser.

men däremot till det allmänna kaoset och en del av interaktionen har gått förlorad i barnskrik.

Sist i detta kapitel finns en tabell där jag har sammanställt det transkriberade materialet och kortfattat beskrivit inspelningarnas innehåll.

2.3 Teknisk information

Inspelningarna är gjorda med en digital videokamera, Canon MV400i. Filmerna har sedan importerats till iMovie och finns säkerhetskopierade på DVD. För att kunna arbeta med filmerna har det varit nödvändigt att komprimera dem (bilden är komprimerad, men inte ljudet), och det har jag gjort enligt de anvisningar som fanns på CHILDES webbplats år 2003 (MacWhinney, 2000), med komprimeringsprogrammet Sorensen Video 3 i ca 10 minuter långa avsnitt.

Inspelningarna är transkriberade i CHAT-format (MacWhinney, 2006a) och länkade till video- eller ljudfiler och man kan alltså lyssna på ett valfritt yttrande genom att klicka på det i CHAT-filen[3].

En effekt av de bristande förberedelserna är att jag inte använde en extern mikrofon vid inspelningarna. Vissa delar av inspelningarna är dessutom ljudförorenade, dvs. brus från diskmaskinen, barnskrik, legoskrammel eller andra ljudkällor stör materialet. Från många håll betonas det hur viktigt det är med en bra ljudupptagning (MacWhinney, 2006a; Demuth, 1996). Bedömningen av om ett ord bara är reducerat eller helt utelämnat kan vara helt beroende av den tekniska utrustningen. När det t.ex. gäller rotinfinitiver har Dye (2005) visat att man med bättre inspelningsutrustning och -metod får en högre andel verb i infinitiv med utfyllnadsstavelser före infinitivformen, sk *fillers*, som hon menar är reducerade hjälpverb. Också när det gäller vuxnas tal är en god ljudkvalitet nödvändig för att man ska kunna urskilja delarna i snabbt tal.

Hannamaterialet har alltså vissa tekniska brister. Men det finns ändå en hel del hörbart och tydligt språk att analysera, och en fördel med materialet är att det är länkat till ljud- eller videofilen, så att det är enkelt att lyssna på varje yttrande och avgöra om det är användbart eller inte. Att länka materialet är ett tidskrävande arbete, men det gör att en större del av materialet blir användbart.

2.4 Transkribering

Inspelningarna har först transkriberats och länkats utifrån videofilerna av tre assistenter och därefter kontrollerats av mig. Jag har kunnat fylla i med det

[3]Genom att öppna filmerna direkt i QuickTime och utnyttja resurserna där kan man också lyssna på inspelningarna i lägre hastighet för att lättare kunna urskilja ljud.

som är svårt att förstå för en utomstående och har dessutom arbetat med att likrikta transkriptionerna för att uppnå en hög grad av konsekvens genom hela materialet. Det är dock oerhört tidskrävande att redigera transkriptioner och jag har genomgående proiriterat konsekvens i återgivandet av barnspråket och inte varit lika noggrann när det gäller vuxenspråket.

När man undersöker språkliga fenomen utifrån ett skrivet material måste man kunna lita på att transkriptionen speglar barnets språk, och inte transkriberarens uppfattning av hur språket ska vara. De assistenter som jag har arbetat tillsammans med har alla varit studenter från kandidatkursen i svenska och har haft goda kunskaper om språkinlärning. De har varit medvetna om att transkriptionen ska ligga så nära språket i inspelningen som möjligt. Samtidigt innebär en transkription alltid att man måste göra en tolkning av det man hör och då uppstår det vissa problem. I det följande kommer jag att gå igenom vad man använder för ledtrådar i tolkningen, hur ord återges i transkriptionerna, vad som menas med ett yttrande och hur man ska se på upprepningar och omtagningar av olika slag. Eftersom ett av mina huvudintressen har varit tillägnandet av finithet finns också ett avsnitt om transkriberingen av verbmorfologi.

När man sammanställer ett material som ska vara tillgängligt för andra forskare kan man nog utgå ifrån att många kommer att använda det enbart i den skrivna formen och utan att läsa dokumentation och transkriptionsöverväganden. Jag har därför tyckt att det var viktigt att transkriptionerna speglar språket på ett sådant sätt att man får en god bild av det utan att läsa anvisningarna.

Det är vanligt att man kontrollerar sina transkriptioner genom att låta två personer transkribera samma inspelning oberoende av varandra och sedan jämföra avskrifterna (Demuth, 1996; Nettelbladt, 1994). Ju högre grad av överensstämmelse, desto pålitligare transkriptioner. Jag har inte kunnat genomföra en sådan kontroll av materialet i den version som föreligger, men hoppas att kunna kontrollera en del av materialet till nästa version. Jag hoppas på att den som använder korpusen också drar nytta av att transkriptionerna är länkade till själva yttrandet och att tveksamma fall då kan utvärderas av forskaren själv. Därmed är det inte nödvändigtvis *transkriptionen* som måste ligga till grund för undersökningar, utan det talade språket i inspelningen som kan undersökas.

I CHAT skrivs yttrandena in på en så kallad *main tier* som inleds med en asterisk och tre versaler som betecknar namnet på talaren: *CHI för barnet, *MOT och *FAT för mamma och pappa i Hannakorpusen. Man har sedan möjlighet att lägga till flera rader som hör ihop med just det yttrandet, så kallade *dependent tiers*[4]. Dessa inleds med ett %-tecken och tre gemener som betecknar vilken information som finns i raden. Det finns en rad fördefinierade *tiers*

[4]Ordet *tier* är svåröversatt, men liknar kanske mest 'skikt' eller 'nivå'. Jag använder här den engelska termen oöversatt.

för olika syften (%pho för information om fonetik och fonologi, %sit för information om situationen osv.), men det står var och en fritt att skapa sina egna koder och *dependent tiers*. Jag har valt att begränsa mig till en %com-rad (kommentarrad) för uppgifter av alla olika slag [5].

I yttrandena används versaler enbart för att inleda egennamn (och för pronomenet *I* på engelska). Yttrandet avslutas med ett mellanslag och därefter en punkt eller annat skiljetecken. Jag ger en översikt över de vanligaste CHAT-symbolerna i tabell 2.1. Tecknen inom hakparenteser hänför sig till närmast föregående ord eller till orden inom vinkelparenteser. Mer utförliga förklaringar och anvisningar finns i CHAT-manualen (MacWhinney, 2006a).

Tabell 2.1: CHAT-symboler

Symbol	Förklaring	Exempel
xx	ohörbart ord	
xxx	flera ohörbara ord	
[/]	reparation	FAT: <har vi>[//] ska [/] ska ha en skorsten på taket också ?
[//]	reparation med omformulering	
[/-]	reparation med nystart	MOT:<oj , va du>[/-] ska vi bygga högt ?
#	paus i ett yttrande	FAT: du # ska vi läsa en bok .
[>]	nästa talare överlappar	FAT: <ska vi lä>[>] +/. CHI: xx [<] läsa ka.
[<]	föregående talare överlappar	
+/.	avbrutet yttrande	
+...	yttrande som dör ut	MOT: om du vill så kan ja bygga +...
+,	fortsättning på ett tidigare yttrande	CHI: kan du hjälpa mej bygga ? CHI: +, föster [: fönster] .
[?]	osäker transkription	pejta [?]
[=? text]	alternativ transkription	vända <på den> [=? foten]
[: text]	normalortografisk översättning	luglo [: lego]
&=text	extralingvistiska händelser	&=laugh, &=imit:motor
[=! text]	paralingvistiska kommentarer	MOT: ja [=! whispers]
[!]	föregående ord, eller materialet inom hakar, är betonat	CHI: däj [!] e biten .

[5] I undersökningen kodar jag dessutom yttrandena för syntax (se kapitel 3) i en %syn-rad.

2.4.1 Att tolka det barnet säger

Att avkoda vad någon annan säger innebär alltid en gissning. När vuxna samtalar är vi ganska bra på att gissa eftersom vi följer samma konventioner när det gäller vilka ord och ordkombinationer man använder, relevansen av ett yttrande och så vidare. Svårigheten att avkoda små barns yttranden ligger i att man inte riktigt vet vilka konventioner barnet har hunnit lära sig. Ofta kan man nog gissa ganska bra med ledning av det språkliga uttrycket i kombination med den icke-språkliga kontexten (vad barnet gör när hon fäller yttrandet) och den vuxna samtalspartens tolkning i yttrandeögonblicket. När man lyssnar på inspelat språk går viss ljudinformation förlorad och den som lyssnade på språket på plats har därför ofta tolkningsföreträde. Å andra sidan kan den som lyssnar och tittar på inspelningarna göra det på ett mycket mer koncentrerat sätt och har möjlighet att spela upp samma sekvens flera gånger. Det innebär att den som i efterhand tolkar yttrandena i vissa fall kan göra bättre gissningar om vad barnet menar.

Den vuxna samtalspartnen upprepar ofta det som barnet säger, kanske särskilt i inspelningssituationer. Richthoff (2000, s. 33) menar att om barnet accepterar den vuxnas tolkning och går vidare i samtalet är det rimligt att anta att yttrandet har fått rätt tolkning. I min genomgång av Hannakorpusen har jag vid ett flertal tillfällen i efterhand förstått yttranden som missförstods i samtalet. Hanna accepterar dock ofta den misslyckade upprepningen och går vidare i samtalet utifrån den. Det är möjligt att hon inte ser feltolkningarna som misslyckade upprepningar, utan som nya initiativ i samtalet.

(8) CHI: de ka ja katta [: kasta] på löv^en .
FAT: ska du klättra upp på lövet ?
%com: missförstår förmodligen 'katta'
CHI: mm .
FAT: va ska du göra där uppe då ?
CHI: kasta fotbollen . 2;1

Barn är ofta samarbetsvilliga och man kan inte alltid utgå ifrån att de protesterar mot felaktiga tolkningar. Trots det utgör förstås den vuxnas upprepning av barnets yttrande för det mesta en pålitlig tolkning.

2.4.2 Begreppet *ord*

CHILDES-systemet bygger på att man ska kunna göra automatiska beräkningar av språket i CHAT-filerna med hjälp av de CLAN-kommandon som finns.

Man ska t.ex. kunna göra frekvensundersökningar av olika ord, konkordanser osv. För vissa språk finns också morfo-syntaktiska analysredskap (MOR-analyser) som automatiskt analyserar ordens morfologi och markerar ordens ordklasstillhörighet (dock ej för svenska än). För att detta ska kunna fungera krävs det att transkriptionen så långt som möjligt skrivs med standardiserad ortografi.

När man lyssnar på riktigt små barns språk känns det ofta som en grov övertolkning att skriva orden med vanlig ortografi, eftersom avvikelserna är alltför stora och dessutom alldeles för intressanta att bortse ifrån. Lösningen i CHILDES-systemet är att yttrandena skrivs med standardiserad ortografi men kombineras med en fonetisk transkription (med IPA), där den faktiska ljudbilden framgår. En sådan lösning är naturligtvis bra, men också kostsam. Eftersom det inte heller finns någon MOR-analys för svenska, vilket innebär att materialet ändå måste taggas för hand, valde jag bort ett sådant dubbelt system.

Anvisningarna till assistenterna har varit att transkribera så nära det som sägs som möjligt, men att använda normal ortografi. Det betyder att ett ord som *ljuger* som barnet kanske uttalar [jʉːgɔ], skrivs *ljugå*. Det här är i linje med Göteborgskorpusen – Richthoff (2000) kallar det ”icke disambiguerat tal”. I någon mån är alltså orden normaliserade så att de ska bli lättare att läsa både för maskiner och människor. Ord som i snabbt tal dras samman (t.ex. *kan inte* blir *kante*) skrivs som två ord (*kan nte*) för att det ska bli lättare att söka ut lexem, men den talspråksnära transkriptionen innebär att det ändå är svårt. För att hitta alla yttranden med satsnegationen *inte* måste man ha fantasi nog att hitta de alternativa uttalssätt som skulle kunna finnas: ”ente”, ”itte”, ”nte” och kanske till och med ”te”. Ett påtagligt problem uppstod i analysfasen med ordet ”ja” som står för två högfrekventa men mycket olika lexem, interjektionen *ja* och det personliga pronomenet *jag* (se avsnitt 4.1.2). I Göteborgskorpusen löstes en del av de problem som uppstod med det uttalsnära transkriptionssättet med så kallade *lexikonfiler* (Richthoff, 2000, s. 31), något som i förlängningen också kan bli aktuellt att skapa för Hannakorpusen.

Alla ljud som är urskiljbara skrivs ut; *xx* används för mummel och ohörbart material. Ljud som inte är ord (skrik, joller, härmningar av djurläten osv) inleds med ett &-tecken och räknas inte som ord av CLAN-programmen. Men i de tidiga inspelningarna är det inte lätt att veta när Hanna säger något eller bara låter. De automatiska ordbaserade beräkningar i CLAN får alltså användas med försiktighet i de första filerna (t.ex. MLU(word)[6]).

[6]MLU, *mean length of utterance* beräknas egentligen på antalet morfem som ingår i ett yttrande, och man rekommenderas att ange morfemgränser i transkriptionen. Om man inte har gjort det kan man istället beräkna MLU på ord (MLU(word)) vilket i ett språk som inte har så rik morfologi kommer att ge ett jämförbart värde.

Några av assistenterna har använt parentestecken för fonem som faller bort i snabbt tal, t.ex. *(s)ka*. Detta förekommer bara i vuxentalet, men är inte konsekvent genomfört. Ord som avviker mer är skrivna inom hakparenteser, dvs. *g'in [: gå in]*, både hos Hanna och de vuxna. Den notationen används också när det inte går att förstå av kontexten vad som är målformen, t.ex. för att skriva ut Hannas egna ord för saker (*luglo [: lego]*). CLAN-programmen analyserar i default-läget ersättningssträngen inom hakparenteser.

Tecknet för osäker transkription används när det som sägs tolkas som ett ord som dock är obegripligt, eller tolkas som ett känt ord men med stor osäkerhet. I några fall där jag har kunnat tolka ett yttrande men insett att det är obegripligt för andra har jag glossat yttrandet i kommentarraden.

2.4.3 Begreppet *yttrande*

Ett samtal byggs upp av yttranden från de personer som deltar. När vi samtalar identifierar vi oftast yttrandet utan svårighet, men när man ska transkribera ett samtal inser man snart att det inte är självklart hur yttranden ska avgränsas.

Själva termen *yttrande* är vag och inom samtalsanalysen talar man hellre om *samtalsturer* och *turkonstruktionsenheter*, TKE (se t.ex. Karlsson, 2006). En samtalstur består av det en talare säger tills någon annan talare tar över. Men turen kan vara uppbyggd av flera TKE:er. En TKE avslutas med en *turbytesplats*, alltså en slutpunkt där en annan talare har möjlighet att komma in i samtalet. Turbytesplatsen identifieras i normalfallet med hjälp av såväl syntaktiska som prosodiska och pragmatiska medel, även om det inte alltid är fallet att alla tre signalerna sammanfaller. En TKE kan alltså bestå av en sats, men också en fras eller en komplex mening.

Sedan 70-talet har det varit vanligt att mäta den grammatiska utvecklingen hos små barn med måttet MLU, dvs. hur många morfem per yttrande barnet i genomsnitt har vid en given tidpunkt. CHAT-manualen (MacWhinney, 2006a, s. 54ff) betonar mycket starkt att var och en bör ta ställning till vad man avser med enheterna *yttrande*, *ord* och *morfem*, även om man inte är beroende av ett MLU-värde för sin forskning, och att det är upp till varje forskare att utifrån sina syften definiera enheterna. MLU-måttet går tillbaks på Roger Brown och hans forskarteam under sent 60-tal och tidigt 70-tal, och beskrivs t.ex. i Brown (1973), s. 54. Beskrivningen handlar dock nästan bara om vad som ska räknas som ett *morfem* och problematiseringen av begreppet *yttrande* är märkligt frånvarande. Inte heller CHAT-manualen ger någon ledning till hur man ska avgränsa ett yttrande, även om resonemanget och exemplen snarast ger intryck av att man bäst avgränsar dem med hjälp av syntaktiska medel, dvs. i satser. CHAT-manualen fastställer att varje yttrande ska avgränsas av

en delimiter, alltså en punkt, ett frågetecken eller ett utropstecken, dvs. tecken som används för självständiga språkhandlingar, och i manualen står att ”CHAT requires that there be only one utterance on each main line” (MacWhinney, 2006a, s. 56).

Om man vill använda MLU-måttet som jämförelsemått måste man naturligtvis veta vad man menar med ett yttrande och veta att man tillämpar samma principer för yttrandeindelning i olika korpusar. Att detta ändå inte har diskuterats så mycket är förmodligen för att MLU-måttet bara anses kunna säga något upp till en längd av 4 morfem per yttrande (Brown, 1973, s. 54), dvs. under den period när barnens yttranden inte är så utbyggda.

Men också bortsett från MLU-måttet behöver man dela in transkriptionen i yttranden och det är då två saker som framstår som viktiga. Det första är att man ska kunna beskriva de kriterier man använder i yttrandeindelningen, så att man har en möjlighet att vara konsekvent genom hela materialet. Det andra är att det är önskvärt att yttrandeindelningen inte avviker alltför mycket från andra korpusar som man kan få anledning att jämföra sitt material med.

De korpusar som kan komma att bli intressanta att jämföra Hannamaterialet med är Stockholmskorpusen (Söderbergh, 1975; Lange och Larsson, 1977), Santelmann och Platzacks korpus (Santelmann, 1995) och Göteborgskorpusen (Richthoff, 2000). I Stockholmskorpusen avgränsas yttranden av den prosodiska konturen, så de kan bestå av ett eller flera ord eller en eller flera satser. Santelmann (1995, s. 91) anger att hon i stort sett beräknar MLU enligt kriterierna i Lange och Larsson (1973), men går inte närmare in på begreppet *yttrande*. I dessa studier kommer dock barnen sällan upp till yttrandelängder som överstiger 4 morfem. Eftersom jag inte har tillgång till materialen i maskinläsbar form kan jag inte heller på ett enkelt sätt undersöka vilka principer transkriberingen har följt mer i detalj.

I Göteborgskorpusen definieras yttrandet ”med utgångspunkt från det talade språkets intonationsmönster, pausering, andningsuppehåll o.s.v.” (Richthoff, 2000, s. 23). Man går inte närmare in på vad det innebär, men formuleringen ger intrycket att möjliga turbytesplatser ligger till grund för indelningen även här. En enkel kontroll med hjälp av CLAN-kommandot MAXWD, som söker ut det längsta ordet eller yttrandet i en fil, tyder på att det stämmer. Det är i alla fall klart att yttrandena inte är indelade i huvudsatser. Några exempel presenteras nedan[7].

(9) MOT: ska han bo här [!] tror du eller bor han där [!] eller bor han i vattnet eller +... bel27_23.cha

[7] I de återgivna exemplen har morfemindelningstecknen uteslutits för att läsbarheten ska öka.

(10) HAR: så hade vi bara två stycken å sen ämh dog den andra å sen dog den andra å sen dog en ingen mer f att vi hade ingen mer fö a va de alldeles synd å så å [/] å s f att pappa tj öh han skötte inte dom f att äh han va så dåli på å sköta dom han gedde öh chfå himla äh b han had han hade bara en [!] bröbit <å j> [/] å ge dom # bara [!] en bröbet.
har46_23.cha

(11) MAR: den har du satt fast å den har xxx å den har du satt fa den har a satt fast å den har du satt den har ja ja satt fast å den har du satt fast.
mar33_29.cha

En yttrandeindelning på grundval av TKE:er verkar ligga i linje med både Stockholms- och Göteborgsmaterialen. Det ger mig också ett sätt att dela in yttranden som är intuitivt, men samtidigt beskrivbart så att indelningen kan bli konsekvent. Yttrandena i Hannamaterialet avgränsas alltså på grundval av TKE:er. Det betyder att ett yttrande tar slut där jag bedömer att det finns en möjlig turbytesplats. Interjektioner i förfältet räknas alltid in i yttrandena om de ingår i samma prosodiska kontur, men skrivs oftast som egna yttranden om de avslutas med en möjlig turbytesplats.

Oftast sammanfaller den syntaktiska satsen (den utvidgade satsen, se s. 41) med den möjliga och faktiska turbytesplatsen, exempel 12. Men några barnyttranden består, som i exempel 13, av flera satser. Att exemplet inte utgör två skilda TKE:er framgår av att den andra satsen kommer så snabbt att det inte finns utrymme för någon annan att komma in med en samtalstur. Barnyttranden som består av flera huvudsatser är dock få i materialet – de utgör endast 22 av Hannas 5000 yttranden.

(12) CHI: nu kan man äta . 2;2

(13) CHI: kan du hjälpa mej -’ det va ett hårt smör .
(Tecknet -’ anger frågeprosodi, se nedan.) 2;5

Varje yttrande i CHAT måste alltså avslutas med en *delimiter*: en punkt, ett frågetecken eller ett utropstecken. Om yttrandet utgörs av endast en prosodisk enhet kan man låta delimitern ange enhetens språkhandling. En vanlig punkt anger påståendeintonation, ett utropstecken utropsintonation och ett frågetecken står för frågeintonation i ett yttrande (och inte för att yttrandet är syntaktiskt frågeformat). Ett oavslutat yttrande med fortsättningston får antingen delimitern +... (yttrandet dör ut) eller +/. (yttrandet avbryts av någon annan), motsvarande oavslutade yttranden med frågande intonation får +..? och +/?. Om man behöver markera prosodi inuti ett komplext yttrande kan man använda *Nonfinal Tone Markers*: -_ markerar fallande ton, -’ stigande (se exempel 13) och

-, fortsättningston. Vid svårbedömda fall använder jag ett komma för att skilja olika prosodiska enheter åt inom ett yttrande. En översikt av prosodiska tecken ges i tabell 2.2.

Tabell 2.2: CHAT-symboler för prosodiska egenskaper

	Avslutar yttrande	Avslutar avbrutet yttrande	Anger prosodi inuti ett yttrande
Påstående-intonation	.	+... +/.	-_
Neutral			,
Utrop	!		
Fråge-intonation	?	+..? +/.	-’

Det finns dock fortfarande en del problematiska fall. Ett är när talaren verkar avsluta en TKE men i stället väljer att fortsätta den. Då uppstår en liten paus som är svårbedömd. I exempel 14 är prepositionsfrasen påhängd i efterhand, så att den första delen får avslutningsprosodi, men fortsättningen följer så nära inpå att jag ändå har bedömt det som en sammanhållen TKE. Hade en annan talare kommit in med en tur eller uppbackning, hade yttrandet dock skrivits som två, jämför exempel 15. Att det andra yttrandet hör samman med det tidigare anges med tecknet plus och komma +,.

(14) CHI: nu ka pappa stå -_ på taket. 2;2

(15) CHI: kan du hjälpa mej bygga ?
FAT: ja okej .
CHI: +, föster [: fönster] . 2;1

I vissa fall kan alltså samtalspartens återkopplingsfrekvens påverka yttrandeindelningen. Eftersom det inte är beskrivet i tidigare korpusar hur man har hanterat detta, vet jag inte om det påverkar jämförbarheten. I Hannamaterialet förekommer symbolen +, 14 gånger i barnyttrandena.

Det kan också vara svårt att avgöra hur pass fristående inledande led är. Det förekommer ofta, särskilt i barnspråket, att en sats föregås av ett led som sedan upprepas i satsen: ”Nu! Nu kan du ta senap!”. I talspråk förekommer också ofta att topiker lyfts ut ur satsen och bildar en utgångspunkt för satsen (jfr. Ekerot, 1979): ”MOT: nä , **igår** -_ **ute** -_ # så kom en luftballong” (2;0). Det är ofta svårt att avgöra om ledet utgör en egen TKE, och konstruktionerna kan ha transkriberats på olika sätt i korpusen.

Andra problematiska fall utgör yttranden som består av upprepade fraser, oftast nomnialfraser eller interjektionsfraser. Det förekommer främst i de tidigare inspelningarna och CHAT-manualen diskuterar alternativa transkriberingssätt för detta (MacWhinney, 2006a, s. 54f). I Hannakorpusen är dessa upprepningar oftast i skrivna ett och samma yttrande med komma mellan (exempel 16), men jag har inte varit så noggrann med indelningen av yttranden som saknar verb och det kan finnas variation. Såvitt jag kan utläsa av Brown 1973 (s. 54) gör också han så, och därmed både Söderbergh och Santelmann. Det är oklart hur dessa yttranden ser ut i Göteborgskorpusen.

(16) CHI: ä binnå , bennå [: pennor] , dom bannå .
CHI: manne ma -, tack , tack [!] . 1;6

Ibland är det omöjligt att avgöra om en ordström utgör en eller flera syntaktiska eller prosodiska enheter, särskilt i de tidiga inspelningarna. I de fallen har jag angett i en kommentar att segmenteringen är oklar.

(17) CHI: pada den hästen [=? hä den] boken läsa.
%com: oklar segmentering 1;9

Rim, ramsor och sånger har inte delats upp i yttranden på ett konsekvent sätt och räknas inte med i analysen av syntaxen. De markeras inte heller på annat sätt och utgör därför en möjlig felkälla till MLU-måttet.

Avslutningsvis kan man konstatera att begreppet *yttrande* inte är särskilt väldefinierat i tidigare svenska barnspråksmaterial och att det därför krävs en del försiktighet om man vill använda MLU-måttet för jämförelser mellan olika barn. Däremot är MLU användbart för att för ett och samma barn mäta den språkliga utvecklingen om man kan lita på att transkriptionen är tillräckligt konsekvent. I Hannamaterialet finns några felkällor, men de är knappast så stora att de påverkar MLU-måttet i någon större utsträckning.

2.4.4 Upprepningar

Ett återkommande problem när man är intresserad av hur ofta vissa syntaktiska strukturer förekommer i barnspråket är hur man ska göra med upprepningar. Om ett barn upprepar *dä bil* åtta gånger utan att bli avbruten, kan det verka logiskt att bara räkna en förekomst och betrakta de andra sju som ekon. Men å andra sidan ingår alla åtta yttrandena i barnets språkproduktion, och därför är de egentligen lika intressanta vart och ett (jfr också Westergaard, 2005, s. 125). Ofta skiljer de sig också åt på något litet sätt, kanske bara i prosodin.

Jag har valt att bara markera två typer av upprepningar i transkriptionerna: självinitierade reparationer inom yttrandet och barnyttranden som är upprepningar av närmast föregående talare.

2.4.4.1 Självinitierade reparationer inom yttrandet

Direkta lagningar av ett yttrande för att rätta till något som har blivit fel, antingen med fel ord eller fel prosodi, brukar kallas självinitierade reparationer (se t.ex. Norrby 2004, efter Schlegloff et al 1977). De förekommer också när man blir störd av en annan talare och behöver börja om från början, inte hittar rätt ord osv. Jag markerar bara reparationer inom yttranden. Dessa är vanliga också i vuxenspråket och transkriberade med de konventioner som finns i CHAT.

(18) CHI: <dä Hanna>[/] dä [!] Hanna ritat . 1;11

(19) CHI: nu ä pap(xx) [//]<mammas th>[//] pap(xx) [//]
mammas tur stå ph [/] på tomet [: tornet] . 2;2

I dessa yttranden rättar man genast till det som blev fel. På prosodin hör man att det första ledet inte är avslutat, och ofta kommer det andra ledet snabbt in. CLAN-kommandot MLU avläser tecknen [/] och [//] så att det närmast föregående ordet, eller orden inom hakar, räknas bort från yttrandelängden [8]. Yttrandet i exempel 18 består alltså av tre ord. Med den här notationstekniken blir det lättare att tolka yttrandena rätt vid läsningen av transkriptionerna och det ger förmodligen en bättre bild av barnets förmåga när man räknar yttrandelängd. Om man vill säga något om barnets förmåga att sätta samman satser bör man räkna bort det som barnet hakar upp sig på. Det verkar troligare att Hanna i exempel 19 visar att hon har förmågan att sätta samman satser med 7 ord än med 12 ord.

Observera att det bara är reparationer som markeras med tecknen [/] och [//]. Andra typer av upprepningar och svårtolkade fall separeras med komman eller inte alls (jämför exempel 16 ovan).

2.4.4.2 Upprepningar av närmast föregående talare

I den språkliga analysen av yttranden är det nödvändigt att veta vilka som är direkta upprepningar av vuxeninput och jag har därför märkt upp de spontana imitationerna. Det är dock uppenbart att många upprepningar inte bara är en imitation av det givna, utan i många fall gör barnet något med ordet eller orden

[8] MLU-kommandot räknar alltså automatiskt bort alla reparationer (MacWhinney, 2006b, s. 106) till skillnad från de flesta andra CLAN-kommandon som räknar med dem. Man kan välja att låta kommandot räkna med alla reparationer genom att inkludera flaggan +r6 (ibid, s. 144).

innan det återger dem (jfr. Lust m.fl., 1996). Jag har därför valt att skilja på upprepningar som är ordagranna, som i kommentarraden har fått beteckningen 'imit', och upprepningar som på något sätt är kreativa som har fått beteckningen 'rep'. Observera att verbet "äta" i exempel 21 skiljer sig i både ändelsevokal och ordaccent (se avsnitt 2.4.5).

(20) MOT: Annika gråter .
CHI: gåte .
%com: imit 1;10

(21) MOT: skulle hon äta kanske ?
CHI: än # äte kanske .
%com: rep 1;10

Jag återkommer till de spontana imitationerna i avsnitt 4.1.1 eftersom det visade sig att de bidrog mer till kartläggningen av Hannas tidiga språk än jag hade väntat.

2.4.4.3 Övriga upprepningar

Övriga upprepningar markeras inte på något särskilt sätt, och vilka hänsyn man tar till dem beror på vad man vill undersöka. Jag kan bara konstatera att barn upprepar sig mycket. Upprepningar av hela prosodiska fraser skrivs på separata rader i CHAT-filen och räknas därmed som fullvärdiga yttranden. Även hela yttranden kan ju vara reparationer av ett tidigare yttrande som blev fel på något sätt. Skillnaden är att man gör det i efterhand, när yttrandet redan är fullbordat.

Inte alla upprepningar består dock av reparationer. Ofta upprepar barnet sina yttranden tills det har fått bekräftelse på att samtalspartnern har hört vad de säger, eller för att de helt enkelt säger det de tänker på tills de börjar tänka på något annat. I praktiken kan man ofta inte skilja på dessa upprepningar och reparationer som består av hela yttranden, vilket också är en anledning till att jag inte har markerat dem. Upprepningarna innehåller kanske reparationer i väntan på bekräftelse. Båda sorterna förekommer också i det vuxna talet, om än inte i lika stor omfattning.

(22) CHI: kugga katten .
CHI: bygga katt # ten .
CHI: bygga katt .
CHI: bygga katten , nen . 1;8

Det förekommer också upprepningar med andra funktioner. En typ används för att verbalt illustrera ett skeende, antingen ett man berättar om eller ett man leker fram.

(23) CHI: <gupp [: gå upp], gupp , gupp>[= description] på din arm .
(Hanna leker med en docka som går upp för pappas arm.) 2;1

Strömqvist (1984) kallar detta *ludic speech act*, dvs. en lekfull talhandling som barnen använder för att skapa en låtsasvärld, och som karakteriseras av reduplikation av verb på ett monotont sätt. Jag är inte säker på vad denna typ av upprepning säger om barnets yttrandelängd, men det är inte frågan om reparationer och inte om separata yttranden. Därför skrivs de samlat i ett yttrande men markeras med [= description] så att det framgår att det är en särskild sorts upprepning.

Sådana här illustrationer av leken består ibland av hela avslutade yttranden. Också de får markören [= description] på det första yttrandet.

(24) CHI: gå fån ne i tappan [= description].
CHI: gå få upp i tapparen .
CHI: gå ne fån tappan .
CHI: å gå upp ifån tappan . 2;1
(Hanna leker med en docka som går upp och ner i (ifrån) en trappa.)

2.4.5 Verb och verbmorfologi

Vid min genomgång av transkriptionerna har jag varit extra uppmärksam på verbmorfologi av olika slag, vilket innebär att jag har lyssnat igenom alla verb upprepade gånger. Jag har dock bara använt mig av auditiv analys i transkriptionsarbetet[9].

I den största verbklassen i svenska, konjugation 1, faller det mesta av morfologin bort; presens- och preteritumformerna låter ofta (men långt ifrån alltid) likadana som infinitiven. Också supinum-t:et kan falla bort, så att inte heller den formen kan skiljas från infinitiven. I många andra högfrekventa verb som *ha*, *få* och *gå* är infinitiven svår att skilja från presensformen.

(25) MOT: titta grodan spruta(r) vatten . 1;9

(26) MOT: fastna(de) den ? 1;9

(27) MOT: oj , uj nu ramla(de) han ner . 1;9

men

(28) MOT: oj oj ramlade du ? 1;9

(29) MOT: e de Nelly som har bajsat ?
CHI: Elly bajsat . 1;9

[9]För en diskussion av auditiv och instrumentell fonetik, se Bruce (1998), s. 59f.

De dentala ändelserna är inte svåra att identifiera om de finns där, så preteritum- och supinumformer är angivna som sådana med ganska stor säkerhet. Presens-r-et är däremot svårare att höra. I de flesta fall har jag inte ens försökt avgöra om det finns en presensändelse på verben i konjugation 1. I min uppsvenska dialekt framgår det dock ibland indirekt att det finns där när den följande konsonanten blir en supradental (*kastar sig* blir [kʰastaʂɛj]), och i vissa barnyttranden är r-et framträdande. I de fallen är ändelsen utskriven, men generellt gäller att man inte kan lita på transkriptionen av presensmorfologin för just dessa verb.

Som tur är kommer många av barnets tidiga verb från klasser som har en presensform som tydligare skiljer sig från den infinita (*sitta, bygga, läsa, äta, sova, komma, ligga*). Presensformerna av dessa verb har akut accent och slutar dessutom på -er (som kan låta som -ä/-ö i uppsvenska eller -å i skånska), som *sitter*, *bygger*, *läser* osv. Varje sådant verb har jag lyssnat på extra noga, för att jag ska kunna säga om ändelsen är ett -a eller ett -ä/-ö/-å(r). Närvaron eller frånvaron av ett -r är dock ofta omöjlig att avgöra, och man får ta transkriberingen av slut-r som ytterst osäker även här.

Accenten är ett tydligare tecken på presens i dessa verb, även om den är inte heller alltid är så lätt att avgöra. I något fall har accenten inte stämt överens med ändelsen och då har jag angivit det i en hakparentes. I de fall jag har varit osäker anges en alternativ transkription. Verbändelse med -ä/-ö/-å(r) innebär alltså att verbet också har akut accent, om inte något annat anges. På samma sätt innebär en verbändelse på -a att accenten är grav, om inte annat anges.

(30) MOT: den sover .
CHI: sovå dä .
CHI: sovår .
MOT: ja . 1;9

(31) CHI: sovä [=? sova] .
CHI: sova lite . 1;9

(32) CHI: dä Mak mälla [= accent 1]
mål: där Max smälla/smäller 1;10

I de tre första inspelningarna (1;6–1;8) har det varit mycket svårt att avgöra accent och ändelsekvalitet. Den iakttagelsen får jag bekräftad av ett par undersökningar av Germund Kadin och Olle Engstrand (Engstrand och Kadin, 2004; Kadin och Engstrand, 2005) som visar att svenska barn vid 18 månader inte producerar tvåstaviga ord med tydlig akut accent i någon större utsträckning, vilket de däremot gör vid 24 månader. Förmågan att producera skillnaden mellan akut och grav accent blir alltså hörbar någonstans mellan 1;6 och 2;0.

Bedömningar av prodsodiska egenskaper är svåra, och det är vanligt att man också i dessa fall låter flera personer oberoende av varandra transkribera

samma material för att sedan mäta överenstämmelsen mellan de olika transkriptionerna, sk *interbedömarreabilitet* (Bruce, 1998). En sådan procedur är ännu inte genomförd för Hannakorpusen, men planeras till nästa version. Också en maskinell analys av orden, t.ex. i ett analysprogram som PRAAT, kan komma att ge stöd åt accentbedömningen i en framtida version av korpusen.

Jag har också varit uppmärksam på närvaron av *är*, både som kopula och befintlighetsverb. Det är naturligtvis svårt att avgöra om det finns ett kopulaverb efter t.ex. *där*, även i vuxenspråk. I transkriptionen finns olika grader av säkerhet angivet: från ett säkert särskrivet verb, ett möjligt verb (*däˆe*), omöjligt att avgöra (*dä:*) till med viss säkerhet inget verb (*dä*).

(33) CHI: dä e blå . 1;9

(34) CHI: däˆe Elli , älgen . 1;9

(35) CHI: dä: nollen . 1;9

(36) CHI: dä mus . 1;9

Till sist har jag också varit uppmärksam på om det finns *fillers* som kan vara reducerade hjälpverb (även fillers för subjektspronomen) (Dye, 2005), dvs. ljud som inte nödvändigtvis har en särskild fonetisk form, men uppträder i de positioner där man kunde väntat sig ett hjälpverb (eller ett subjekt). I de tidiga inspelningarna finns många olika utfyllnadsljud, så det är svårt att säga med säkerhet vad som är vad. Men jag har strävat efter att i alla fall skriva ut de ljud som jag kan höra.

(37) CHI: se ligga dä . 1;9

(38) CHI: e sovår . 1;9

2.5 Materialets representativitet

Den typ av longitudinell data, som Hannakorpusen utgör, är mycket vanlig inom barnspråksforskningen. Det material som finns att tillgå på CHILDES består till stor del av data som är insamlade genom sporadiska inspelningar. Analyser av spontant tal har många fördelar, men man måste också vara medveten om de felkällor som kan gömma sig i både metod och material (jfr t.ex. Stromswold, 1996; Tomasello och Stahl, 2004; Rowland och Fletcher, 2006). I det följande ska jag nämna två faktorer som man bör vara uppmärksam på i Hannakorpusen. För det första spelar inspelningssituationerna roll för språket i korpusen, för det andra får man komma ihåg att korpusens densitet är låg.

I den vanligaste inspelningssituationen befinner sig barnet och den vuxna samtalsparten i barnets sfär, ofta i en leksituation där barnet får styra över vad som ska hända och vad man ska göra. Den vuxne försöker vara så följsam som möjligt. Det resulterar i att man pratar om det som finns på plats och det man gör då och där och bara en en begränsad del av barnets språk dokumenteras. Det leder t.ex. förmodligen till att presensformerna dominerar och att man i analysen underskattar barnets förmåga att tala om det förflutna. Även om den vuxne försöker att uppmuntra barnet att tala om händelser i det förflutna är den pågående leken ofta intressantare. Om man vill fånga barnets olika tempusformer borde man kanske snarare dokumentera övergångssituationer, t.ex. hämtning vid dagis (Christensen 2003, s. 36).

Inspelningssituationen påverkar också det språk som riktas till barnen, vilket kan exemplifieras med antalet barnriktade imperativformer i Hannakorpusen. I Hannakorpusen är vuxenspråket undersökt i två inspelningar. I inspelning 2;0 leker Hanna med sin mamma i sitt lekrum och till största delen är det Hanna som styr över situationen. Inspelning 2;4 följer Hanna när hon hjälper sin pappa att förbereda kvällsmaten och under familjens måltid, och då är det till största delen de vuxna som styr. Vem det är som styr situationen visar sig tydligt i antalet imperativformer i vuxenspråket: i 2;0 är de 8 och i 2;4 29 (se tabell 4.5 i avsnitt 4.2). De flesta inspelningar, inte bara i Hannakorpusen, är gjorda i leksituationer där barnet får styra och man bör därför vara uppmärksam på att den input man kan undersöka inte i alla avseenden är representativ för barnets totala input.

Inspelningarna utgör bara stickprov av barnets språk. I en halvtimmeslång inspelning av Hanna 1;9 finns det 411 yttranden från Hanna. Om man antar att hon under en vanlig halvtimme inte får riktigt lika mycket uppmärksamhet från sina föräldrar och därför pratar hälften så mycket, och att hon är vaken 10 timmar om dygnet betyder det att 0,3 % av alla hennes yttranden under en månad finns dokumenterade. Det är därmed klart att man inte kan uttala sig om lågfrekventa fenomen i Hannas språkutveckling. Att ett fenomen inte finns i inspelningarna kan antingen bero på att det var kortvarigt och inspelningarna skedde med för långa mellanrum så att man missade det, eller på att fenomenet är lågfrekvent (jfr Demuth, 1996). Hannas verbpartikelplacering är ett bra exempel på det senare. I kompletterande dagboksanteckningar finns noteringar om att Hanna vid 2 1/2 års ålder påfallande ofta placerar verbpartikeln efter objektet (”kan du sätta den fast?”), men i det dokumenterade språket finns bara ett par förekomster av felplacerade partiklar.

Det kan till sist vara på sin plats att påminna om att en longitudinell korpus som bara har dokumenterat ett barn möjligen endast kan säga något om hur just det barnet lär sig språk. Hanna var tidig i sin språkutveckling. I motsats

till sin yngre syster hade hon från början ett uttal som låg nära målspråket och använde inte särskilt många onomatopoetiska och egna ord. Hennes fonetiska träffsäkerhet gör henne förhållandevis lätt att förstå och hennes språk tacksamt att analysera. Hannas språkutveckling verkar likna bl.a. Emblas (Lange och Larsson, 1973) och Markus (Strömqvist m.fl., 1993; Richthoff, 2000). Hannas lillasyster Nina däremot, som inte är dokumenterad på samma systematiska sätt, hade många onomatopoetiska och egna ord (många egna tecknade också) och var mycket svårare att förstå. Hon talade tidigt i långa meningar utan att man kunde urskilja orden. Det tycktes mig också som om hennes språk innehöll fler avvikelser längre upp i åldern, även om jag inte har haft möjlighet att studera det i detalj.

Man brukar skilja mellan två strategier i språkinlärning som Peters (1977) kallar *analytic* och *gestalt* (se också Nelson, 1981)[10]. Gestaltstrategin gör språket svårare att förstå vilket kan leda till att man måste välja bort dessa barn när man ska undersöka det allra tidigaste barnspråket. Ett av barnen i Göteborgskorpusen, Anton, är t.ex. inte inkluderad i någon undersökning av tidig grammatik som jag känner till, förmodligen eftersom det är så tidskrävande att tolka vad han säger. Här finns ett metodiskt hinder som är svårt att komma över – för att kunna analysera någons språk måste man förstå det och då väljer man naturligtvis att titta på de barn som är begripliga tidigt. Det innebär dock också att man bara kan uttala sig om de begripliga barnens tidiga språkutveckling, och det är inte på förhand givet att de mer obegripliga barnens utveckling ser likadan ut.

Man måste alltså vara medveten om att man generaliserar utifrån sparsamma och något ensidiga data. Man kan fråga sig hur väl ett material med så låg densitet representerar ett svenskt barns språkutveckling. Vi får dock konstatera att vi inte har så många svenska barnspråksmaterial än så länge. Att transkribera barnspråk och på annat sätt insamla och färdigställa material är arbetskrävande och därmed kostsamt. Undersökningar av spontandata ger dock ett bra underlag när man vill precisera frågeställningar som man sedan kan undersöka med andra metoder.

2.6 Översikt över Hannakorpusen

I tabell 2.3 finns en översikt över alla inspelningar. I tabellen finns information om ålder (anges med år;månader.dagar), deltagare (m=mamma, p=pappa,

[10] Andra termer som har använts är *referential* å ena sidan och *expressive* eller *holistic* å andra. Demuth (2001) föreslår att skillnaden mellan de olika strategierna har att göra med var i den *prosodiska hierarkin* barnet befinner sig.

s=syster), aktivitet och antalet ord och yttranden samt MLU(word) för de transkriberade inspelningarna[11]. I anmärkningsfältet står kommentarer om sådant som har påverkat inspelningen och i aktivitetsfältet beskrivs inspelningssituationen. I fältet för *Språklig kommentar* försöker jag ge en översiktlig bild av Hannas språk. I kapitel 4 ges en mer heltäckande beskrivning av den språkutveckling som kan knytas till satsen.

Tabell 2.3: Hannakorpusen

Ålder	*Transkription*	*Deltagare*	*Anmärkning*
	Antal yttranden	*Antal ord*	*MLU(w)*
	Aktivitet		
	Språklig kommentar		

1;6.18	Ha1y06m.cha	m, p	
	317 yttranden	532 ord	MLU(w) = 1,7
	lägger saker i lådor, läser, tittar på djurbilder och namnger djur, ritar		
	en- och tvåordsyttranden, endast infinita verb		

1;7.16	Ha1y07m.cha	m, p	Hanna är sjuk
	337 yttranden	621 ord	MLU(w) = 1,8
	läser, tittar på djurbilder och namnger djur, bygger med lego		
	mest tvåordsyttranden, men enstaka yttranden med verb som har två argument, oklara verbändelser, enstaka bestämda slutartiklar i nominalfraser		

1;8.15	Ha1y08m.cha	m, p	
	338 yttranden	640 ord	MLU(w) = 1,9
	bygger med lego, gömmer och tar fram djur, läser, tittar på sig själv i kameran		
	vissa treordsyttranden och enstaka längre men svårtolkade yttranden, oklara verbändelser, övergeneraliserar bestämd slutartikel i nominalfraser		

1;9.15	Ha1y09m.cha	m	
	411 yttranden	935 ord	MLU(w) = 2,3
	läser, tittar på djurbilder, leker med bilar, djuren ska sova		
	en del finita verb, antydan till hjälpverb, satsprosodi och annan intressant prosodi, många spontana imitationer, imiterar också prosodi		

[11] MLU(w)-måttet är framräknat på enklast möjliga sätt med CLAN-kommandot ”mlu” och är bara användbart som ett grovt mått på meningslängden i de olika inspelningarna.

1;10.17	Ha1y10m.cha	m, p	
	385 yttranden	808 ord	MLU(w) = 2,1
	Hanna vill gå ut, läser, bygger med lego		
	några hjälpverb, sammansatta målspråkliga och icke-målspråkliga nominal, en del avvikande ordaccenter, många spontana imitationer		

1;11.14	Ha1y11m.cha	m, p	
	289 yttranden	628 ord	MLU(w) = 2,2
	ritar och pratar om färger, grannbarnen knackar på och Hanna vill gå ut och leka med dem, hämtar och packar ner saker, läser om rymdraketer		
	språket börjar likna vuxenspråket, få svårtolkade yttranden, fler satser, inga avvikande nominalfraser, få imitationer		

2;0.11	Ha2y00m.cha	m	
	382 yttranden	1036 ord	MLU(w) = 2,7
	stoppar klossar i hål, leker med dockhuset, bygger med lego		
	många satser med frågeprosodi, många hjälpverb, några underordnade verbfraser utan subjunktioner		

2;1.16	Ha2y01m.cha	p, (m)	
	418 yttranden	1048 ord	MLU(w) = 2,5
	leker med dockhuset, bygger med klossar, kastar boll, läser		
	svårtolkad frågeprosodi, lek med ord, många attribut, genitiver med -s		

2;2.14	Ha2y02m.cha	m	
	358 yttranden	1314 ord	MLU(w) = 3,7
	spelar munspel, leker med dockhuset och klossar, pratar om pengar och att handla glass		
	relativsatser, supinum, många inverterade deklarativer		

2;3.14	Ha2y03m.cha	m, (p), (s)	bullrig inspelning, Hanna har fingrarna i munnen mycket
	311 yttranden	955 ord	MLU(w) = 3,1
	Hanna har just kommit hemfrån dagis och äter mellanmål, läser böcker, leker med dockan		
	bisatser, tydliga kopula, blandar ”han” och ”hon”		

2;4.17	Ha2y04m.cha	p, (m)	bullrig inspelning, Hanna talar otydligt
	275 yttranden	880 ord	MLU(w) = 3,2
	Hanna hjälper pappa att laga mat och duka bordet, familjen äter, pratar om vad som hänt på dagis		
	för att-satser (fatt. . .), lek med ord		

2;5.22	Ha2y05m.cha	p, m, (s)	bra inspelning med tydligt tal
	287 yttranden	834 ord	MLU(w) = 2,9
	Hanna har just kommit hem från dagis och familjen äter mellanmål och pratar om vad som har hänt och vad som ska hända		
	ordentliga bisatser		

2;6.21	Ha2y06m.cha	m, p, (s)	
	296 yttranden	966 ord	MLU(w) = 3,3
	leker med bokstäver, leker med kameran		
	fullständiga bisatser, övergeneraliserar ”hon”, hittar på egna ord		

2;7.19	Ha2y07m.cha	m, (p), (s)	otydligt tal
	260 yttranden	880 ord	MLU(w) = 3,4
	spelar memory, läser bok, leker tåg		
	utbrytningskonstruktioner och många andra ”de e . . . ”, bisatsordföljd		

2;8.26	Ha2y08m.cha	m, (p)	Hanna är förkyld och hes
	335 yttranden	1375 ord	MLU(w) = 4,1
	ritar och skriver bokstäver, läser, pratar om dåtiden		
	negationsflytt, frågebisatser, satsflätor, variation i konjunktionsutsättning i samordningar, övergeneraliseringar av dentalsuffix i preteritum		

2;10.08	Ha2y10m.cha	p, m, (s)	
	251 yttranden	961 ord	MLU(w) = 3,8
	bakar scones, bollar med studsboll, fikar och pratar om farmor och farfar		
	verbkedjor, när-satser, överdrivna r-ljud, lek med ord, nonsensramsor		

Kapitel 3

Manual för satsanalys i barnspråk

3.1 Inledning

När man undersöker barnspråk är det naturligtvis en stor fördel att materialet är transkriberat och maskinläsbart. Bara att ha materialet i CHAT-format gör att man automatiskt kan få frekvensordlistor, beräkna genomsnittlig yttrandelängd och t.ex. söka ut ord och kombinationer av ord med hjälp av CLAN-kommandot *combo*. I exempel 39 finns alla satser med pronomenet *mej* producerat av Hanna i inspelningarna fram till och med 2;0.

(39) combo +smej

```
————————————————————-
File "Ha1y08m.cha": line 715.
CHI: hjäpa mej .
————————————————————-
File "Ha1y08m.cha": line 730.
CHI: hjäpa mej .
————————————————————-
File "Ha1y08m.cha": line 735.
CHI: xx tä hjäpa mej .
————————————————————-
File "Ha2y00m.cha": line 342.
CHI: mm # mamma , kan hjälpa mej !
————————————————————-
```

File "Ha2y00m.cha": line 1180.
CHI: den e liten mej .
————————————-
File "Ha2y00m.cha": line 1183.
CHI: stor mej .

Om man som jag är intresserad av den finita satsens egenskaper, såsom ledföljd och obligatoriska satsled, hjälper det emellertid inte långt att söka på ord, utan man måste gå igenom materialet och analysera det yttrande för yttrande. Genom att tagga materialet kan man söka i det mer effektivt, och med en generell taggning kan man få svar på de frågor som uppkommer efterhand, utan att man behöver gå tillbaka och gå igenom hela materialet igen.

I CHILDES-systemet är det möjligt att göra en automatisk analys och taggning av morfo-syntaxen i yttrandena (en sk MOR-analys), men det redskapet finns inte för svenskt material än. Programmet utför heller ingen satsanalys, utan en ett-till-ett-analys som för varje ord anger ordklass och böjningsmorfem. I projekten Talsyntax och Skrivsyntax användes en utförlig analysmetod, där både satsled på alla nivåer och ordklasser och frastyper märktes upp (Loman och Jörgensen, 1971; Teleman, 1974), och en sådan metod hade varit idealisk för mina syften. Nackdelen är att bearbetningen av materialet blir alldeles för tidskrävande. Jag har därför valt att göra en ganska enkel analysmodell som jag har kunnat ta fram på egen hand, under den tidsrymd jag har haft till förfogande.

Min analysmodell är utarbetad så att man på ett enkelt och ganska snabbt sätt ska kunna koda alla yttranden med hjälp av de verktyg som finns i CHAT och CLAN. I korthet innebär det att jag för varje yttrande anger vilka satser (eller satsmotsvarigheter) det innehåller samt en sträng av koder som talar om vilka primära satsled som finns i dem och i vilken ordning satsleden kommer. Däremot anger jag inte vad satsleden utgörs av, dvs. jag gör varken ordklass- eller frastypsanalys i kodningsskedet. Jag inför en separat kod för verbens morfologi, eftersom det är en viktig aspekt när man ska avgöra verbets finithet. För att fånga alla de systematiska avvikelser som (Josefsson, 2003b, s. 156) redovisar, för jag också in en kod för närvaro eller frånvaro av prepositioner. Kodenrna fogas in i en rad under yttrandet i transkriptionsfilen, en *dependent tier* som inleds med %syn.

Med hjälp av den syntaktiska analysen och kodningen i CHAT-format kan man med de CLAN-program som finns söka ut alla yttranden som har en viss ordföljd, och med den morfologiska analysen kan man oberoende av satsanalysen undersöka det morfologiska verbsystemet hos barnet. I exempel 40 finns alla Hannas yttranden med topikaliserat objekt, tillsammans med de två föregående yttrandena, från inspelningen då hon är 2;0.

(40) combo +t*CHI +t%syn +s$OFS* -w2 Ha2y00m.cha

File "Ha2y00m.cha": line 303.
MOT: ska vi göra de ?
CHI: mm .
CHI: den ha ja sett huset .
%syn: $OFSVO M:basf M:supm
————————————————-
File "Ha2y00m.cha": line 374.
MOT: ska han ligga stora sängen istället ?
CHI: mm .
CHI: de kan hon .
%syn: $OFS M:ipre
————————————————-
File "Ha2y00m.cha": line 1012.
MOT: får den plats där bättre [>] .
CHI: mm [<] .
CHI: de kan den få de .
%syn: $OFSVX M:ipre M:basf

Om det material man undersökt är länkat till ljud- eller videofilen kan man i den utdatafil som sökningen genererar också se och höra filmsekvensen av yttrandet genom att klicka på det. CHAT- och CLAN-verktygen gör det alltså mycket enkelt att utvärdera det talade språket. Det finns då möjlighet att också inkludera svåranalyserade yttranden. Små barns yttranden har ofta flera möjliga tolkningar vilket leder till att man får utesluta många osäkra yttranden. Men i de svåranalyserade yttrandena kan det finnas intressanta mönster som man då missar. Enligt metoden föregår analysen minst i två faser. I den första fasen taggar man materialet enligt den mest troliga tolkningen och i den andra fasen ska man på ett enkelt sätt kunna söka fram alla yttranden som är relevanta för det fenomen man undersöker, så att man ska kunna göra en andra samlad bedömning av dem.

Den analysmetod jag använder är utformad för att jag utifrån *en* sortering ska kunna få svar på mina två frågor i syftesbeskrivningen. Den ska å ena sidan göra det möjligt att identifiera *när* den finita satsen är etablerad hos barnet och den ska å andra sidan göra det möjligt för mig att närstudera ledföljd och obligatoriska led i satsen efter det att den finita satsen är etablerad. I detta kapitel ska jag beskriva analysen i detalj. Innan dess ska jag dock gå igenom syntax i samtal och diskutera några saker som jag måste ta ställning till för att kunna analysera tidigt barnspråk.

3.2 Syntax i samtal

Ett *yttrande* är, som jag nämnde i föregående kapitel (2.4.3), något som identifieras med fler medel än syntaxen. Mitt huvudintresse omfattar den finita satsen, så jag analyserar egentligen inte *yttranden*, utan främst *satser*. I vardagligt talat språk består dock många ord- och fraskombinationer av något som är mindre än en sats, som SAG (Svenska Akademiens Grammatik, Teleman m.fl., 1999) kallar *huvudsatsekvivalenter*. I detta avsnitt ska jag gå igenom SAGs beskrivning av dessa båda begrepp och därefter diskutera hur man ska betrakta de tidigaste barnyttrandena. I avsnitt 3.4 redogör jag sedan för mina analysbegrepp.

3.2.1 Huvudsatser

SAG beskriver den utbyggda satsen med följande schema:

(41) SAG IV, s. 6

Utvidgad sats				
		Inre sats		
Förfält	Initialfält	Mittfält	Slutfält	Efterfält
Ja,	i morgon	kan hon nog	vara med,	eller hur?
I morgon,	då	kan hon nog	vara med,	din syster.

I förfältet och efterfältet hittar man dels annex av olika slag, dels vokativer och interjektioner som är prosodiskt knutna till satsen. Annexen utgör satsdelar i satsen och har antingen ett dubblerat led i den inre satsen eller är lösare knutna till satsen (fria annex) (SAG IV, s. 438). Den inre satsen är välundersökt och består av ett antal satsled med en prototypisk ordning.

3.2.2 Huvudsatsekvivalenter

Långt ifrån alla yttranden i ett samtal ser ut att ha full satsstruktur. Vissa yttranden saknar verb och andra också subjekt, t.ex. svar på frågor, utrop eller rubriker.

(42) Hur har maten varit?
–Bra.

(43) Jag först!

(44) Latinet tillbaka! Nybörjarkurs redan till våren.

Loman och Jörgensen (1971) skiljer på *meningsfragment* och *huvudsatser*, medan SAG använder termen *huvudsatsekvivalenter* i stället för meningsfragment. Samma namnskifte ser man också på engelska; *fragments* benämns i nyare litteratur ofta *nonsententials* (Progovac m.fl., 2006). Namnskiftet indikerar ett intresse för att undersöka dessa självständiga meningar som består av led som i vanliga fall fungerar som satsled (SAG IV, s. 806), men som inte ser ut att uttalas i fullständiga satser.

Huvudsatsekvivalenter (SAG IV, s. 806) är meningar som består av satsled som kan ingå i huvudsatser och som är syntaktiskt avgränsade utan att ha prototypisk huvudsatsstruktur. De skiljs från ellipser (SAG IV, s. 945) som är fulla huvudsatsstrukturer, men med vissa led, som dock är lätta att supplera, outtalade. I huvudsatsekvivalenter är leden naturligtvis också syntaktiskt relaterade till varandra, men när man ska göra primär satslösning kan man få svårt att avgöra vilka satsdelar som kommer till uttryck i huvudsatsekvivalenten. I ett yttrande som *Javisst ja, mina handskar!* (SAG IV, s. 811) är det svårt att svara på vilken satsdel *mina handskar* utgör. SAG skriver: ”I de flesta fall kan en huvudsatsekvivalent visserligen mer eller mindre naturligt parafraseras med en huvudsats där den fras som utgör ekvivalenten ingår som primärt satsled (vanligast predikativ eller objekt), men parafrasen kan åtminstone lexikalt göras på på flera olika sätt, och det finns inte någonting i kontext eller talsituation som utpekar en unik supplering. Därigenom skiljer sig huvudsatsekvivalenterna från de konstruktioner vilka i denna grammatik redovisas som ellipser (Ellips § 1). Många uttryck är dock gränsfall och kunde redovisas antingen som ellipser eller som huvudsatsekvivalenter”.

Traditionellt har huvudsatsekvivalenterna betraktats som elliptiska satser och i generativ terminologi som fulla CP-strukturer (Progovac m.fl., 2006, s.1). Men interjektionsfraser och vokativfraser kan knappast beskrivas som satser, även om de också har bestämda strukturer, så det kan inte vara uteslutet med strukturer som är något annat än satser i språket. Man kan tänka sig att huvudsatsekvivalenterna ”egentligen” är hela satsstrukturer, med vissa led utelämnade, men de kan också beskrivas på annat sätt som hos Barton (2006) och Progovac (2006). Hur man än betraktar huvudsatsekvivalenter i vuxenspråket blir nästa fråga om det tidiga barnspråket kan beskrivas på samma sätt.

3.2.3 Tidigt barnspråk

I små barns språk blir frågan om vad man ska betrakta som primära satsled aktuell. Se t.ex. på rotinfinitiven i exempel 45.

(45) CHI: pappa bygga den åsså . 1;10

Det har från flera håll hävdats att satser som denna, utan finit verb, ändå utgör fulla satsstrukturer, men med ett hjälpverb som inte realiseras fonetiskt (t.ex. Josefsson, 2002; Dye, 2005), dvs. ellipser. I exempel 45 ovan finns alltså ett ”ska” även om det inte hörs, vilket fint passar in i kontexten när det gäller satsens modalitet och tempus. Satsen kan alltså analyseras som en huvudsats och man vågar analysera ”pappa” som subjekt i den huvudsatsen, trots att det inte finns något finit verb. Samma resonemang håller för kopulalösa barnsatser som ”mamma arg” eller ”nallen trött”. ”Mamma” och ”nallen” utgör tydliga predikationsbaser som tillskrivs något, ett predikativ. Att det finita verbet saknas är inte konstigt; det är obetonat och i många omgivningar nästan ohörbart.

Men det finns också andra typer av rotinfinitiver, t.ex. sådana som klart refererar till en ospecificerad dåtid. Exempel 46 är hämtat från Christensens barnspråkskorpus (Christensen, 2003, ex. 18):

(46) Amma [: Hjalmar] åka buss Hjalmar 1;8
(på kvällen, när Hjalmar har åkt buss för första gången tidigare samma dag)

Detta exempel kan inte på samma sätt som exempel 45 betraktas som en elliptisk sats där man enkelt kan supplera det finita verbet. Här används en infinit form av verbet där man i målspråket skulle ha haft en finit form. Denna sats saknar alltså ett finit verb, men det verkar ändå rätt naturligt att betrakta ”Hjalmar” som subjekt, eftersom han är den som åker och ledet befinner sig på något som skulle kunna vara subjektsplatsen. Det finns dock också exempel där man inte kan säga något om subjektsstatusen utifrån placeringen, men ändå kan analysera verbets argumentstruktur. I exempel 47 finns en form av verbet som ser ut som en infinitiv. Agenten, som verkar naturligast att göra till subjekt, är dock placerad mellan det infinita verbet och objektet, en plats som subjekt aldrig intar i vuxen svenska.

(47) CHI: Hanna, läsa Hanna den Mak [: den Maxboken]. 1;8
(Hanna pekar på en bok som hon vill läsa själv.)

Alla rotinfinitiver kan alltså inte självklart beskrivas som elliptiska satser, men de är ändå mer lika huvudsatser än de huvudsatsekvivalenter som beskrevs ovan, eftersom de har ett verb och en analyserbar argumentstruktur.

Frågan om vad man kan klassa som subjekt uppstår också i den mycket vanliga benämningsleken, som är en vanlig sysselsättning särskilt i inspelningssituationer. När man läser böcker eller tittar på bilder tillsammans med barnet genererar det ofta en sorts lek som går ut på att barnet benämner det hon ser på bilderna, ofta med enkla nominal, men också i kombination med ytterligare

ord. Frågan här är om det är meningsfullt att försöka ge orden någon satsdelsbeteckning i dessa benämningar (jfr också Bohnacker (2003) och däri citerade verk) eller om man kan identifiera en argumentstruktur. När barnet i en tidig inspelning tittar på en bild av en katt i en bok och kombinerar ihop ”dä” och ”katt” skulle den rimligaste översättningen kanske vara ”Det där är en katt”, dvs. ”dä” skulle vara subjekt och ”katt” ett predikativ. Men man kan också tänka sig att satsen ska tolkas ”Där är en katt” som svar på frågan ”Var är en katt?”: Då skulle ”dä” vara adverbial och ”katt” subjekt. Ett tredje alternativ är att det inte alls finns några satsdelar i yttrandet eftersom det inte utgör en sats, utan något annat.

När det gällde huvudsatsekvivalenten ”Javisst ja, mina handskar” ovan menar SAG att funktionen är att aktualisera en bekant referens eller ett bekant sakförhållande (SAG IV, s. 811). Det påminner just om barnens benämningslek. Brown (1973) kallar det *nomination* eller *demonstrative and entity* och noterar att nominalen ofta föregås av något, antingen ett pronomen (en/ett, de(n/t)) eller ett obestämbart schwaaktigt ljud, eller något som ser ut som lokativer (där, här). Brown påpekar dock att det inte är äkta lokativer, utan snarare demonstrativer. Han menar att frågan ”Var är x” inte är en äkta lokativ fråga, utan snarare betyder ”Visa mig att du vet vad en x är för något”. Benämningen kan också föregås av ”titta”, som då inte har några verbegenskaper utan bara fungerar demonstrativt (jfr. Lange och Larsson, 1973).

De allra första tecknen på syntax kommer redan i ettordsstadiet, vilket Lois Bloom kunde visa i boken *One word at a Time* (1973). Orden som till en början används för att namnge objekt och händelser börjar under den andra hälften av det andra året att även användas för att referera till relationer. ”Daddy” tillsammans med en pekande gest mot pappans portfölj namnger inte objektet utan fångar den possessiva relationen mellan pappan och objektet. Därefter kan barnet börja att sätta samman två ord och göra relationerna explicita, och det är först då, i tvåordsstadiet, vi kan börja undersöka hur barnets syntax ser ut, dvs. hur hon kombinerar orden. Då får vi ”de första glimtarna av grammatik”, skriver Tom Roeper i boken *The Prism of Grammar*(s. 50 ff), genom att den grundläggande strukturen blir synlig. Han menar att man kan visa att strukturen blir asymmetrisk när två element kombineras, så att det ena elementet dominerar det andra. Kort tid därefter kan barnet också skapa nya strukturer genom rekursion – genom att återigen sätta samman element som i sin tur består av sammansatta element.

Barnens tidiga tvåordssatser har visat sig uttrycka en begränsad mängd semantiska relationer (Brown, 1973; Roeper, 2007), varav benämningen, Browns *nomination*, är en. Det är dock oklart hur dessa tvåordsyttranden ska relateras till syntaxen i vuxenspråket. Kan man beskriva dem som satser, dvs. utgör de

CP-strukturer i den generativa terminologin? Eller utgör de en sorts satsekvivalenter som ska beskrivas på annat sätt (t.ex. som i Potts och Roeper (2006))? I min undersökning intresserar jag i huvudsak för den finita satsen och när man kan säga att den finns hos barnet. Perspektivet utgår alltså från vuxenspråket och jag kommer att behöva pressa in dessa tidiga ett- och tvåordsyttranden i den kategoriindelning som jag kan göra för vuxenspråket. Det betyder dock att jag måste försöka avgöra om de tidiga barnyttrandena är att jämföra med satser eller huvudsatsekvivalenter. På samma sätt kan man kanske inte tala om *subjekt* om man inte har en finit sats att utgå ifrån. Men jag vill också undersöka de tidiga förekomsterna av det som skulle kunna tolkas som subjekt, även om det inte har alla de subjektsegenskaper som krävs i målspråket. Jag får kanske anledning att återkomma till de tidiga barnspråksyttrandena och hur de ska tolkas, men för att kunna tagga upp barnspråkskorpora på ett enhetligt sätt behöver jag kunna använda mig av grammatiska termer något friare.

3.3 Analysprinciper

3.3.1 Övergripande principer

Huvudsyftet med den analysmetod som presenteras här är att med den traditionella grammatiken som grund tagga upp ledföljen i satsen. Det betyder som sagt att det till största delen är den *inre satsen* i SAGs utbyggda sats som analyseras. Alla yttranden består inte av satser, men jag tar med alla yttranden som innehåller led som skulle kunna fungera som satsled i analysen. Analysen bygger till största delen på SAG (Teleman m.fl., 1999), men jag har också haft god hjälp av Teleman (1974) (hädanefter Mamban). Varje sats eller satsmotsvarighet översätts till en sträng av bokstäver, där varje bokstav symboliserar en satsdel. Varje sträng får sedan ett inledande tecken som inte är en bokstav, för att man enkelt ska kunna skilja bokstavssträngarna från vanliga ord i utsökningarna. S står för subjekt, F för finit verb och O för objekt, och $ betyder (i stort sett) huvudsats. Satsen ”Lammet äter välling” får alltså koden $SFO.

Satsledsbeteckningen utgörs av en enda bokstav så att bokstävernas plats i strängen representerar deras placering i (den linjära) satsen. Därmed kan man t.ex. söka ut alla satser som har det finita verbet på andra plats ($_F*) [1]. Taggningen är gjord med de verktyg som finns i CLAN (MacWhinney, 2006b). Mitt sätt att utforma kodningssystemet medförde att det blev otympligt med alltför många olika taggar, och jag fick därför försöka att begränsa antalet. Taggning-

[1] I sökströngen förekommer jokertecknen understreck, som står för *ett* valfritt tecken, och asterisken, som står för noll, ett eller flera valfria tecken.

en kan därmed ses som ett ganska grovt filter som man kan använda för att söka fram en mängd yttranden som sedan måste analyseras vidare.

Den huvudsakliga analysen av ledföljd i satser och satsmotsvarigheter kompletteras av en separat kod för verbens morfologi och en separat kod för frånvaron och närvaron av prepositioner.

Analysen utgår ifrån att barnets yttranden kan beskrivas på samma sätt som vuxenspråkets motsvarande yttrande skulle ha beskrivits, men eventuellt med utelämningar av vissa led och morfem. Det betyder inte att det är så man måste betrakta det tidiga språket, men jag har valt att ha en enhetlig princip för taggning av hela korpusen. Så småningom blir barnets yttranden målspråkliga och mitt syfte är att kunna undersöka hur det går till genom att söka ut alla relevanta yttranden som leder fram till en viss typ. Jag kallar den här principen att ”tagga stort” och den innebär t.ex. att holofraser som ”E de?” taggas som om de vore analyserade enheter hos barnet, i det här fallet som finit verb och subjekt. De tidigaste förekomsterna av ”E de?” är med all sannolikhet inte syntaktiskt segmenterade av barnet, men någon gång under inspelningsperioden förstår hon att ”e” är ett finit verb och att ”de” i det här fallet fungerar som subjekt (och att kvesitiva frågor av det här slaget också kräver ett v-frågeord). Om man i efterhand kan undersöka alla frågor kan man kanske få en uppfattning om hur barnet räknade ut detta.

Men det är alltså inte så att en klassificering som t.ex. bisats i min taggning betyder att barnet har analyserat, förstått sig på och producerat en bisats på samma sätt som en vuxen talare. Det kan vara så, men det återstår att undersöka. Tanken är inte att man ska använda taggningen för att räkna fram statistik över barnets språkförmåga. Taggningen ska göra det möjligt att på ett enkelt sätt leta fram de satser (strängar) som har de egenskaper man är ute efter så att man kan testa den hypotes man har om ett visst fenomen. Den statistik, som man kan få fram, kan man dock förstås använda för att upptäcka fenomen som är intressanta att studera vidare.

En annan sida av ”tagga stort”-begreppet är att också yttranden som bara är delvis tolkbara tas med i analysen i taggningsskedet. De taggas efter den mest rimliga tolkningen. Om de kan bidra med någon information avgörs i ett senare skede. Yttranden som är helt otolkbara får dock ingen kod.

Ytterligare en övergripande princip är att verben innehar en särställning i analysen. Alla målspråkliga satser innehåller minst ett verb, och det är (de tematiska) verbens argumentstruktur som ger oss de andra leden i satsen. Jag är alltså inte bara intresserad av satsen, utan av verbenheten också. Det innebär att alla yttranden som innehåller verb taggas som satsstrukturer, även om både strukturen och funktionen kan diskuteras.

Min ambition har varit att låta analysen ligga så nära barnspråket som möj-

ligt och att analysera det som faktiskt finns där. Man kan dock fråga sig vad det betyder att något ”finns där”. När man t.ex. talar om fundamentellips av ett i sammanhanget tydligt subjekt kan man hävda att det finns där, men inte uttalas. Det finns mycket som pekar på att det finns i vår språkliga representation av satsen. Överlag kan tystnad i en viss kontext bära på lika mycket information som ljud. Jag får därför definiera att något ”finns där” om det hörs, dvs. uttalas (om en fras bär fonetiska särdrag i en särdragsbaserad teori) och utelämnade subjekt märks därmed inte upp. Jag måste göra vissa avsteg från denna princip och jag ska snart gå igenom dem. Först måste jag dock ägna ett kort avsnitt åt vad det innebär att något hörs.

3.3.2 Hörbarhetsprincipen

Språkets ljudegenskaper kan delas upp i tre huvudklasser: vokaler, konsonanter och prosodi (Bruce, 1998, s. 9ff). Vokaler och konsonanter utgör betydelseskiljande segment i ljudsekvensen medan de prosodiska egenskaperna är suprasegmentella (har räckvidd över flera segment) och har flera funktioner. De prosodiska signalerna hörs tydligt men eftersom de inte är segmentella kan de inte på samma självklara sätt inordnas i strängar på samma nivå som ord och satsled. Det är dessutom svårt att analysera prosodi i naturligt tal (jfr Bruce 1998, kap. 4), vilket också Loman och Jörgensen konstaterade i talsyntaxprojektet 1971. Prosodin är en viktig faktor i språkinlärning, det medger alla. Men information om prosodi är ganska sällsynt i de svenska barnspråksstudier som har presenterats, även om det finns några undantag, t.ex. Peters och Strömqvist (1996), Strömqvist m.fl. (2001) och Bohnacker (2003). Hänsyn till prosodin tas naturligtvis i alla de studier som arbetar nära ljud- och videomaterialet, på så sätt att man tolkar yttrandena utifrån hur de låter och inte bara utgår från den transkriberade strängen. Typiskt urskiljer man då intonation som skiljer frågor från påståenden.

I Göteborgskorpusen har dock ”[i]nga prosodiska markeringar eller detaljerade markeringar av ton- eller röstläge [...] gjorts varken i barnens yttranden eller de vuxnas” (Richthoff 2000, s. 22). I barnyttrandena i korpusen är inte ens fråge- eller utropsprosodi utmärkt med skiljetecken, på grund av svårigheter att avgöra den verkliga funktionen hos yttrandena (ibid). Josefsson (2003a), som undersöker ledföljd hos tre barn varav två från Göteborgskorpusen, skiljer inte heller på verbinitiala satser som har interrogativ eller deklarativ prosodi, utan räknar in alla satser med det finita verbet först och subjekt sedan i samma kategori. Hon nämner dock att satstypen bl.a. används med frågefunktion. I Josefsson (2003b) noterar hon att vissa V3-satser skulle kunna vara satser med fritt annex, något som prosodiska markeringar (eller avlyssning) kunde

ha avgjort. När det gäller barns förmåga att skifta objekt påpekar Josefsson att förmågan att härma satsernas intonation förbättrar barnens möjligheter att upprepa satser med skiftade objekt (”Ser du mej inte?”), och hon drar slutsatsen att ”intonation patterns are important for children acquiring syntactic patterns” (ibid s. 181).

Prosodi är alltså en viktig faktor och definitivt något som hörs, men det är svårt att handskas med den. I Hannakorpusen är vissa prosodiska signaler transkriberade, men utan någon oberoende kontroll. En av fördelarna med Hannakorpusen är dock att den är länkad till ljud- eller videofilerna så att det alltid är lätt att bedöma yttranden utifrån hur de låter[2]. På så sätt tar jag hänsyn till prosodin i min taggning, men den teoretiska underbyggnaden brister i min analys. Frågeprosodi är t.ex. något som hörs och taggas upp (med Q sist, se s. 65). Man kan också höra att en v-frågeordsinledd fras är en bisats och inte en V3-sats, och jag kan därmed tagga den därefter, även om jag inte kan ge en god beskrivning av hur det låter[3].

3.3.3 Undantag

Jag gör några avsteg från principen att bara tagga det som hörs. Vissa fenomen som jag vill undersöka går annars inte att söka fram. De tidigaste kvesitiva frågorna saknar ofta v-frågeord (jfr ”E de?” ovan) vilket gör att ledföljden ser likadan ut som för en rogativ fråga, som ju har det finita verbet först (”Sover hon?” $FSQ). Om det utelämnade v-frågeordet inte märks upp går det inte att skilja frågetyperna åt. Att det är skillnad på dem framgår oftast av just prosodin. Jag taggar alltså frånvaron av v-frågeord (med Q, se s. 65). Jag taggar också upp frånvaron av prepositioner i den kompletterande prepositionsanalysen. Indirekt taggar jag också frånvaron av finit verb, genom att jag klassificerar vissa yttranden som satser trots att de saknar finit verb (se s. 54).

3.3.4 Satsens omfattning

Att avgränsa yttranden är, som jag har diskuterat i kapitel 2, inte oproblematiskt, eftersom de syntaktiska, prosodiska och pragmatiska avgränsningarna inte alltid sammanfaller. Ett problem utgör satser som består av flera TKE:er, som ibland har blivit transkriberade över flera rader. Det händer t.ex. när en annan

[2] Arbete med att länka delar av Göteborgskorpusen pågår nu också.

[3] Några exempel från Lundin (1987) kan belysa detta (s. 199): R: ”Berätta om den nye Martin för mig.” Embla: ”Var han bor?” (2;1), dvs. Ska jag berätta var han bor?; S: ”Ja titta.” Ask: ”Mm. Vad han gjort på Jesper.” (3;6).

talare kommer in med en samtalstur, eller en minimal respons (+, markerar att intonationen visar att yttrandet är en fortsättning på ett tidigare yttrande).

(48) CHI: kan du hjälpa mej bygga ?
FAT: ja okej .
CHI: +, föster [: fönster] . 2;1

I exempel 48 sträcker sig huvudsatsen till och med ”fönster” och man kan hävda att ”fönster” utgör objekt till verbet ”bygga”. Men den prosodiska strukturen säger något annat. Ordet ”bygga” uttalas med en avslutande frågeprosodi och satsen är fullbordad redan där. Därefter bestämmer sig talaren för att bygga på den redan avslutade satsen med ett komplement till verbet, och i det nya yttrandet innefattas också innehållet från det föregående. Men det betyder inte att innehållet från det andra yttrandet nödvändigtvis innefattas i det första yttrandet.

Loman och Jörgensen (1971) väljer att arbeta med *makrosyntagmer* och låter den syntaktiska satsen bli så stor som möjligt. De låter därmed de syntaktiska kriterierna gå före de prosodiska i fall som dessa och analyserar hela den *syntaktiska* satsen (makrosyntagmen) som en enhet[4]. Jag använder dock inte makrosyntagmbegreppet, utan utgår, i fall som exempel 48, ifrån hur yttrandet är avgränsat i transkriptionen. Som jag redovisade i avsnitt 2.4.3, kan yttrandeindelningen bero på samtalspartens interaktion, och därför är tillvägagångssättet inte helt konsekvent. I exemplet ovan har jag allstå inte analyserat ”fönster” som objekt, utan som satsekvivalent. Som jag dock också redovisade rör det sig om 14 barnyttranden som är skrivna som skilda yttranden, och de kan lätt sökas fram.

Ett annat problem utgör led som står som prosodiskt avskilda led utanför den inre satsen, men skulle kunna betraktas som satsled i extraposition. I exempel 49 ser vi hur MOT förtydligar sig genom att placera adverbialen utanför satsen i egna prosodiska fraser (markerade med fetstil), utan att det finns turbytesplatser mellan dem.

(49) MOT: ja: [<], de flög en luftballong precis här över gården igår .
CHI: ä uppe , min tak .
MOT: nä u [//] nä , **igår** -_ **ute** -_ # så kom <en luftballong> [>] . 2;0

Det är vanligt med yttranden som dessa och de är svårbedömda (jfr. 2.4.3). Om leden står i förfältet räknar jag inte in dem i strängen. Undantag gör jag

[4]En av anledningarna till att Loman och Jörgensen låter den syntaktiska analysen vara primär är att en systematisk beskrivning av prosodin inte var utarbetad. De avsåg att i ett senare skede jämföra prosodiska och syntaktiska gränser i spontant talspråk (Loman och Jörgensen, 1971, s. 12).

för led som följs av adjunktionellt *så* så att båda leden då ingår i satssträngen och bildar en V3-sträng (se exempel 49 och 50). Det förekommer också andra målspråkliga V3-satser med *kanske* och *bara*.

(50) CHI: nä(r) du skruˆde upp de , så måste ja skruva på .
%syn: $AAFSVL ¢USFLO M:sptm 2;10

Jag lägger däremot inte ner så mycket möda på att försöka avgöra om led befinner sig i slutfältet eller efterfältet. Om leden kan klassificeras som satsled tar jag med dem i strängen. Interjektionsfraser och vokativfraser lämnas alltså oanalyserade, medan efterställda annex ingår i strängen.

3.3.5 Utanför analysen

Element som aldrig ingår i den syntaktiska satsen, konjunktioner, interjektionsfraser och vokativfraser, lämnas oanalyserade och saknar helt %syn-rad. Jag går inte heller in i nominalfrasen och analyserar ledföljd och obligatoriska led i den.

Interjektionsfraserna kan vara problematiska, då gränsdragningen mot huvudsatsekvivalenter och ellipser kan vara svår. Jag har valt att klassificera alla yttranden som innehåller verb som satssträngar, även lexikaliserade interjektionsliknande fraser, eftersom de på ett enkelt sätt kan byggas ut till en produktiv fras (sov så gott → nalle sover så gott), medan fraser utan verb, där kopplingen till en huvudsatsmotsvarighet är mer långsökt, lämnas utanför analysen (gonatt → jag önskar nalle en god natt (?)). Nedan har jag listat de vanligaste förekomsterna av interjektionsfraser, vokativfraser och konjunktioner i Hannas språk i inspelningarna.

interjektioner		*vokativer*	*konjunktioner*
ja	hejdå	mamma	å (och)
nej	gonatt	pappa	men
mm	tack	Hanna	fast
hoppsan	varsågod	du	så
oj	så, sådär, sådärja		

Det multifunktionella *så* tolkas som interjektion när det står självständigt och har samma betydelse som *sådärja*, men som adverbial i andra fall. Se s. 62. Lange och Larsson (1973) räknar också *titta* till interjektionerna i de fall då ordet används isolerat för att rikta uppmärksamheten på något (eller till kategorin DEIX tillsammans med andra demonstrativa element i samband med

benämningar). Jag har dock valt att tagga alla potentiella verb som verb oavsett funktion.

Det finns fall då det är svårt att skilja vokativer från subjekt. Prosodin är inte en helt entydig ledtråd, men oftast ganska klar. I exempel 51 nedan har jag med hjälp av prosodin tolkat *mamma* som en vokativfras. Tolkningen är redan gjord i transkriptionen, men vid taggningen har jag alltid genomfört en ny analys av tvetydiga yttranden.

(51) CHI: mm # mamma , kan hjälpa mej !
%syn: $FVO M:ipre M:infv 2;0

Det är dock inte bara prosodin som avgör tolkningen av exempel 51 utan ledtrådar finns på flera håll. Hanna vänder sig direkt till mamman med sitt yttrande, så det är tydligt att mamma utgör andra person för Hanna. Vid 2;0 använder Hanna pronomen för att referera till andra person och har lämnat den nominala stilen. I inspelningen säger Hanna ”kan du” flera gånger, men aldrig ”mamma” eller ”pappa” för att referera till mamma och pappa som andra person (förutom i vokativer)[5]. Detta illustrerar också hur jag låter många faktorer bidra till den syntaktiska tolkningen av ett enskilt yttrande.

3.3.6 Sammanfattning av analysprinciper

Mina övergripande principer sammanfattas här:

1. Alla yttranden som innehåller led som kan fungera som satsled får en kodsträng. Interjektionsfraser, vokativfraser samt konjunktioner utanför satsen får ingen kod.

2. Yttranden som är delvis tolkbara taggas enligt den mest rimliga tolkningen. Yttranden som är helt otolkbara får ingen kod.

3. Varje led i kodsträngen består av endast en bokstav, och bokstavssträngarna inleds av tecken som inte är bokstäver.

4. Verbmorfologi och prepositionsbruk taggas med separata koder.

5. Jag ”taggar stort”, dvs. det kan ibland se ut som om jag tillskriver barnet större språklig kompetens än det kanske har när jag analyserar barnets yttranden som de skulle ha formulerats i vuxenspråket.

[5]Hon använder däremot ”Hanna” två gånger för att referera till sig själv, men båda gångerna refererar hon till bilden av sig själv på kameraskärmen. I alla övriga självreferenser säger hon ”jag”.

6. Jag märker upp de led som finns uttalade i yttrandet, förutom utelämnade v-frågeord och prepositioner i obligatoriska kontexter som får en kod fast de inte uttalas. Prosodiska signaler märks upp när jag kan tolka dem, framförallt i frågor och sam- och underordnade verbfraser.

I tabell 3.1 på s. 58 finns en översikt över de koder som jag har använt. I de avsnitt som följer går jag igenom koderna en och en och beskriver hur jag har använt dem och vilka avgränsningar jag har fått göra. Först ska jag förklara hur man ska avläsa %syn-raden, för att underlätta läsningen av exemplen.

3.3.7 Hur man avläser kodsträngarna

(52) CHI: ja kan sitta på andra sidan .
%syn: $SFVA M:ipre M:infv PP:prep 2;0

(53) MOT: om han inte kan böja på knäna så måste de gö ont .
%syn: $AAFSVO M:ipre M:infv ¢USIFVA M:ipre M:basf PP:prep 2;0

(54) CHI: den -' , kan den ligga da: ?
%syn: €NQ $FSVAQ M:ipre M:infv 2;0

I exempel 52 finns en huvudsats och det markeras med det inledande tecknet $. Huvudsatsen består av satsleden subjekt (S), finit verb (F), infinit verb (V) och (rums)adverbial (A) i den ordningen, så hela strängen blir $SFVA. Verben i strängen består av först (längst till vänster) en icke-suffigerad presensform (M:ipre) och därefter en suffigerad infinitivform (M:infv)[6]. I yttrandet finns en prepositionsfras som innehåller en preposition. I exemplet utgör prepositionsfrasen ett primärt satsled, men det är inte nödvändigt för att prepositionsförekomsten ska märkas ut. Utsatta och utelämnade prepositioner på alla nivåer i satsen märks ut.

I exempel 53 består yttrandet också av en huvudsats, $. Den inleds av ett (omständighets)adverbial i annex, (A – bisatsen "om han inte kan böja på knäna"), följt av ett adjunktionellt *så* (A), och därfter finit verb, subjekt, infinit verb och objekt, FSVO. Verben i huvudsatsen består av en icke-suffigerad presensform (M:ipre) och en suffigerad infinitiv (M:infv). I yttrandet finns en underordnad verbfras (bisatsen) som ger en sträng med inledaren ¢. Bisatsen inleds med en subjunktion, U, och följs av subjekt, satsadverbial (I), finit verb, infinit verb och (objekt)adverbial: USIFVA. Verben i den strängen utgörs av en icke-suffigerad presensform (M:ipre) och en basform (M:basf). I den underordnade verbfrasen finns också en prepostionsfras som innehåller en preposition,

[6] Angående terminologin, se avsnitt 3.6.

PP:prep. Bisatsen i exemplet fungerar som ett primärt satsled i den överordnade satsen, men det är inte ett nödvändigt kriterium. Underordnade verbfraser på alla nivåer märks ut.

I exempel 54, till sist, består yttrandet av en verblös satsekvivalent först, så den första strängen inleds med €. Satsekvivalenten består av ett obestämbart nominal (N) som bär på en frågande prosodi (Q). Därefter kommer en huvudsats, $, som består av finit verb, subjekt, infinit verb och adverbial. Satsen har frågande prosodi vilket markeras med ett Q sist i strängen: $FSVAQ. Verbformer som ovan.

3.4 Strängtyper

Jag har valt att skilja på tre sorters strängtyper: 1) meningsfragment, som motsvarar SAGs *huvudsatsekvivalenter*, med den skillnaden att alla huvudsatsekvivalenter som innehåller verb klassas som huvudsatser; 2) huvudsatser, som i den här klassificeringen också omfattar alla verbfraser som inte är underordnade; 3) underordnade verbfraser, dvs. bisatser och infinitivfraser[7]. Det enda som blir över efter denna kategoriindelning är interjektionsfraser, vokativfraser, konjunktioner utanför satsen samt helt otolkbara yttranden.

3.4.1 € – meningsfragment

Yttranden som inte innehåller verb, och inte på ett enkelt sätt kan suppleras med ett verb, inleds med tecknet €. Jag kallar dem *meningsfragment* för att inte blanda ihop dem med SAGs huvudsatsekvivalenter. I vuxenspråket innebär det att eurotecknet inleder fraser (enledade eller flerledade) som inte på ett självklart sätt kan betraktas som huvudsatser med utelämnat led. Exempel på sådana satsmotsvarigheter är ”Inte nu” och ”Upp igen” (som taggas €IA respektive €AI).

Alla yttranden som innehåller verb får någon av de två andra inledarna, även om de skulle vara mindre än en sats. Det betyder att målspråkliga huvudsatsekvivalenter som består av verb i infinit form som svarar på en fråga eller ett annat initiativ i min analys taggas som huvudsatser.

(55) MOT: Vad ska vi nu då?
CHI: Åka till farmor och farfar.
%syn: $VA 2;10

[7]Med underordnade verbfraser avser jag inte infinita verbfraser till ett finit modalt eller temporalt hjälpverb.

Genom att konsultera sin egen språkkänsla kan man förhållandevis oproblematiskt identifiera meningsfragmenten. Men när man ska tolka yttranden i barnspråket blir det svårare att avgöra vad som motsvarar en sats och vad som är något annat, och klassificeringen blir ibland något stipulativ. Tvåordsyttranden som skulle kunna sakna en finit form av *vara* betraktas som satser (”Mamma (är) arg”). Jag har dock efter viss tvekan valt att inleda benämningar (nominations) med tecknet €. Det gäller alltså konstruktioner med ett demonstrativt element och ett nominal eller ett adjektiv som förkommer när man gemensamt med barnet tittar på något och barnet berättar vad sakerna heter (”dä katt”, ”dä röd” osv.). Att jag inte betraktar dessa konstruktioner som satser beror på att de verkar vara en egen genre. I de tidiga inspelningarna förekommer många benämningar av slaget ”dä N”, men i de senare inspelningarna förändras benämningsleken så att Hanna hellre beskriver vad som händer i situationen man betraktar, ”Katten sover” eller ”Hon har en katt” (jfr Bohnacker, 2003, s. 242).

Nästa fråga är hur man ska tolka det demonstrativa elementet i dessa tvåordsbenämningar. Det enklaste hade kanske varit att låta dem utgöra en egen kategori, i likhet med Lange och Larssons DEIX. Men jag är inte villig att lägga till ytterligare en kod i mitt system, särskilt som alla förekomster av *dä* och *den* enkelt kan sökas ut i texten. Så jag väljer att gå på den enkla principen att det som är transkriberat som *dä*/*da* utgör ett adverbial och det som är transkriberat som *den*/*de* utgör ett nominalt led, och *titta* är alltid ett verb. Ibland när det har varit svårt att avgöra om det finns ett finalt -n eller inte (”dä” eller ”den”) har jag valt att tagga det som adverbial. Jag låter också den icke-språkliga kontexten spela en avgörande roll i analysen av dessa yttranden. I ett sammanhang när man letar efter katten, och barnet säger ”dä katt” när hon hittar den betraktar jag ”dä” som ett lokativt adverbial och ”katten” som predikationsbasen och därmed subjekt (se vidare avsnitt 3.5.1). Under arbetets gång visade sig denna indelning vara svår och förhållandet mellan benämningar och satser med ”dä” utvärderas i det kapitel 4.

Jag har valt att tagga alla ettordsyttranden, eller snarare enledsyttranden, med inledaren €, med undantag av dem som består av ett ensamt verb som får inledaren $. Ibland frestas man att tolka led i ettordsyttranden som subjekt, som i exempel 56, men eftersom yttrandet består av två enskilda prosodiska fraser gör jag inte det (observera att båda de prosodiska fraserna bedöms ingå i samma TKE). Även kongruerande predikativ inleds med €, även om det är tydligt att de kongruerar med ett outtalat nominal.

(56) FAT: va e det som e borta ?
CHI: bo:ta -_ kaå [: kaka] !
%syn: €A €N
FAT: e kakan borta ? 1;8

(57) MOT: aj aj aj , gör ont på Lisa .
MOT: ont på Max .
CHI: letha .
%syn: €P
MOT: dom e lessna . 1;10

Tvåordsyttranden som består av något som ser ut som ett led (”stor nalle”, ”den gröna”, ”annan bil”, ”mer kaka”) taggas också som €N (nominal) och kan alltså inte utan vidare sökas ut i mitt system (men man kan söka ut alla €N först och därefter alla yttranden som innehåller två ord). Jag låter alltså ordföljden få en ganska avgörande roll genom att ”Stor nalle” taggas som en benämning (€N) medan ”Nalle stor” som en sats med utelämnad kopula ($SP). Yttranden där samma led upprepas i nära följd får bara en markering (”häst , häst” taggas med strängen €N en gång).

Strängarna som inleds av € består alltså dels av vuxenlika huvudsatsekvivalenter utan verb, dels av tidiga ett- och tvåordssatser utan verb, där jag inte har anledning att tro att en full satsstruktur ligger bakom.

3.4.2 $ – huvudsatser

En huvudsats är enligt SAG en sats som inte fungerar som led i någon annan sats (SAG IV, s. 674). Den prototypiska huvudsatsen har fa-ordföljd (det finita verbet före adverbialet) och uttrycker en språkhandling. Jag använder begreppet huvudsats för prototypiska huvudsatser, men även för vissa satsmotsvarigheter som inte ingår som led i någon annan sats. Satsmotsvarigheter som uttrycker en språkhandling där jag misstänker att den målspråkliga varianten har ett finit verb och där jag med någon grad av säkerhet kan fastställa vilka satsled de ingående delarna utgör inleds också med $-tecknet. Det gäller å ena sidan så kallade *rotinfinitiver*, dvs. satser med enbart infinit verb (se s. 41). Exempelmening 58 nedan, och exemepel 55 ovan, klassificerar jag utan vidare som huvudsatser, även om det saknas led för att den ska få fullständig huvudsatsstruktur. Men även rotinfinitiver som inte kan beskrivas som ellipser (t.ex. 46 och 47, s. 41) analyseras som huvudsatser.

(58) CHI: de vila min kna .
%syn: $SVA M:basf PP:noll
MOT: ska den vila på ditt knä ?
%syn: $FSVAQ M:ipre M:basf PP:prep 2;0

Å andra sidan gäller det kopulalösa satser i barnspråket, eller snarare satser som saknar presens av *vara*. I Josefssons översikt över Sarakorpusen (Josefsson, 2003b, s. 161) förekommer satser med utelämnat *är* i hela den undersökta

perioden (fram till 2;11) även om de blir färre med tiden. För att jag ska kunna genomföra en konsekvent taggning genom hela materialet klassificerar jag alla yttranden som skulle bli målspråkliga med ett utsatt *är* som satser (t.ex. ”mamma arg” och ”nallen trött”). För de allra tidigaste yttrandena är den klassificeringen dock inte helt optimal.

Det betyder att jag analyserar led som primära satsdelar som om det hade funnits ett finit verb närvarande och utan att de har alla egenskaper som satsdelarna har i en fullständig sats. Strängar som inleds med dollartecknet innehåller antingen ett verb eller kan lätt suppleras med ett verb.

Eftersom mitt huvudintresse är ledföljd i satsen räknar jag inte expressiva satser till huvudsatser, trots att de är självständiga. Den formmässiga likheten med bisatser gör att jag räknar dem till ¢-strängarna (se nedan).

3.4.3 ¢– underordnade verbfraser

Centtecknet inleder till största delen underordnade verbfraser. De underordnande verbfraserna kan finnas på olika nivåer i satsstrukturen. De kan utgöra primära eller sekundära satsdelar, eller stå självständigt. Det framgår inte av kodningen vilken satsedel den utgör, utan ¢-strängen följer bara på $-strängen (matrissatsen kommer alltid först). ¢ märker ut bisatser (exempel 59) och infinitivfraser (exempel 60), expressiva satser som har bisatsform men inte är underordnade (SAG IV, s. 759) (exempel 61) samt satssammandragningar (den del i samordnade huvudsatser som har led strukna under identitet, exempel 62). Till de underordnade verbfraserna räknas även omkategoriserade anföringssatser (exempel 63, Mamban, s. 188)[8].

(59) CHI: ja måste titta denna flyge himle .
%syn: $SFVA M:ipre M:basf ¢SFA M:spre PP:noll 2;0

(60) CHI: nu flickans tur # sitta .
%syn: $AS ¢V M:infv 2;0

(61) MOT: vilka rammeldockor de här e .
%syn: ¢WSF M:ipre 2;0

(62) FAT: ska du gå bort till mamma å spela in ?
%syn: $FSVLAQ M:ipre M:basf PP:prep ¢++CVLQ M:basf 2;0

(63) CHI: den dä [!] slut # mena ja .
%syn: $SPA ¢FS 2;4

[8] SAG IV, s. 844 betraktar den anförande meningen (”mena ja” i exempel 63) som överordnad, medan Mamban betraktar den som underordnad. Jag följer Mamban i det här fallet.

De expressiva satserna utgör huvudsatser i den mening att de inte är underordnade (SAG VI, s. 759ff). Till formen liknar de dock bisatsen i hög grad. De har obligatoriskt en inledare som, med undantag av *sådan*, liknar vanliga bisatsinledare (*att*, *som*, *om* eller *så*, *vad*, *vilka*, *vilken*) och de har obligatoriskt af-ordföljd (se exempel 64–67). Grunden för denna undersökning är ledföljd, och därför placerar jag de expressiva satserna i samma grupp som bisatser, även om de skiljer sig från bisatserna i både prosodi och funktion.

(64) Att det aldrig kan sluta regna!

(65) *Det aldrig kan sluta regna!

(66) *Att det kan aldrig sluta regna!

(67) *Det kan aldrig sluta regna! (ej expressiv, men ok som proposition)

Många andra ¢-strängar står också som egna yttranden. De kan ingå som satsled i ett tidigare eget yttrande eller någon annans yttrande, eller i ett underförstått yttrande, men det markeras inte särskilt.

(68) CHI: ja , till kojven !
%syn: €A PP:prep
FAT: mm .
CHI: (s)om ja tatt fam .
%syn: ¢USVL 2;4

Alla kombinationer där den troliga motsvarigheten i vuxenspråket är en underordnad verbfras märks upp; analysen görs alltså från ett målspråksperspektiv. En del av de tidigaste underordnade verbfraserna verkar vara helhetsinlärda hos Hanna, men om man ska kunna se hur och när de blir produktiva måste de märkas upp. Hannas första att-sats (eller infinitivfras) består av något som liknar en reducerad variant av konstruktionen ”Det är bättre/bäst att...”.

(69) CHI: m , bätte sitte [=? sitta] golvet .
%syn: $PS ¢VA M:oanb PP:noll[9] 2;0

Det är inte troligt att ”bättre” i yttrandet i exempel 69 är ett adverbial (?Bättre sitter jag på golvet), utan satsen ligger närmare den vanliga konstruktionen ”Det är bättre att...”. Jag kan inte veta att det är så yttrandet ska tolkas, men det är den tolkning som ligger närmast min egen språkkänsla. (När jag i efterhand har undersökt dessa förmodat underordnade verbfraser verkar gissningen stämma, se avsnitt 4.4.)

[9]Jag har analyserat satsen som ’Det är bättre att jag sitter på golvet’, dvs. ’att jag sitter på golvet’ är egentligt subjekt=S och ’bättre’ är subjektiv predikativ=P. Bisatsen består av ett svårtolkat verb=V och ett rumsadverbial=A, dvs. ¢VA.

I satsflätor får det satsled som har lyfts ur bisatsen samma satsdelsbeteckning som den resterande bisatsen, oftast objekt (exempel 70). Det betyder att satsdelarna blir dubblerade i strängen. Vad det utlyfta ledet har för funktion i bisatsen anges inte.

(70) CHI: mm , den tycker ja e jättegod .
%syn: $OFSO ¢FP

Det kan vara svårt att avgöra om det är två huvudsatser samordnade med konjunktion, eller en huvudsats med en bisats i fundamentet, som har en ohörbar bisatsinledare. I barnyttrandena taggas de då som underordnade verbfraser enligt principen ”tagga stort”. I vuxenyttrandena har jag däremot gått på transkriptionens förstaalternativ. I exemplet ”MOT: man [=? om man] lyfter upp armarna lite, så kan den spela sen” (2;0) är ”man” förstaalternativet, vilket gör den första satsen till en huvudsats.

3.5 Satsled

Jag försöker använda så få satsdelsbeteckningar som möjligt, eftersom mitt kodningssystem blir för svårarbetat med för många koder på samma nivå. Dessutom kan det vara en fördel att inte finkategorisera för mycket från början, eftersom det är lättare att dela upp sökresultat i efterhand istället för att behöva slå ihop flera kategorier i utsökningar.

3.5.1 S – subjekt

Både formellt och egentligt subjekt markeras med S (exempel 71). Det betyder att en sats kan innehålla flera S. Också dubblerade subjekt i annex kan göra att det finns flera subjekt i satsen (exempel 72).

(71) MOT: e de 'skräp på golvet ?
%syn: $FSSAQ M:ipre PP:prep 2;0

(72) MOT: mm # e de mamman de där ?
%syn: $FSPSQ M:ipre 2;0

Det är inte alltid lätt att skilja mellan egentliga subjekt och nominala predikativ. Det finns många ”Det är...”-konstruktioner i svenskan och jag har inte haft möjlighet att fördjupa mig i problemet. Jag har försökt hålla mig till Mambans enkla tumregel att egentliga subjekt ska kunna ersätta formella subjekt ”utan att satsen blir språkvidrig och utan att satsens innebörd väsentligen ändras”

Tabell 3.1: Analyskoder

Symbol	Förklaring
	Strängtyper
€	verblösa meningsfragment
$	huvudsatser
¢	underordnade verbfraser
	Satsdelar
S	subjekt
F	finit verb
V	verb
O	objekt
P	predikativ
A	verbfrasadverbial
I	satsadverbial
L	partikel
W	frågeord
Q	fråga/utelämnat v-frågeord
C	konjunktion
U	subjunktion
+	utelämnat led
N	obestämbart nominal
R	rest/räkneord
X	oanalyserbart
M	*Verbmorfologi*
M:basf	basform
M:infv	suffigerad infinitiv
M:impv	imperativ
M:ipre	icke-suffigerad presens
M:spre	suffigerad presens
M:supm	supinum
M:iptm	icke-suffigerad preteritum
M:sptm	suffigerad preteritum
M:stam	verbstam
M:oanb	oanalyserbar form
M:ssss	s-form
PP	*Prepositionsfras*
PP:noll	utelämnad preposition
PP:prep	utsatt preposition

(s. 42). Det har lett till att satsen i exempel 73 analyseras som en sats med formellt och egentligt subjekt, trots att nominalfrasen är definit. Till skillnad från exempel 72 verkar ”det”-et inte vara referentiellt (det är helt obetonat). Det liknar mest en utbrytning som saknar subjunktion och finit verb i utbrytningsbisatsen.

(73) MOT: de e köket [!] där inne va ?
%syn: $SFSAWQ M:ipre

Satser med ”det finns...” tolkas också som satser med formellt och egentligt subjekt (Mamban s. 44). Frågan ”Vad är det?” förekommer ofta som misstänkt holofras ”e de?” men jag analyserar ändå ”de” som subjekt, enligt principen att ”tagga stort”.

3.5.2 F – finit verb

En av de saker man gärna diskuterar i språkinlärningssammanhang är skillnaden mellan hur finita verb och infinita verb hanteras av inläraren. För att kunna analysera det måste man veta vilka verb som är finita och vilka som är infinita. Svenskans verbmorfologi är inte optimal för detta, eftersom vi har en stor mängd verb där skillnaden mellan de olika böjningsformerna knappt är hörbar. Mest utmärker sig den stora mängd verb som har stamslut på vokal (se avsnitt 3.6). När man vill analysera satsdelar och ledföljd och få en överblick över ledföljdsutvecklingen är det emellertid bra om alla verb är definierade som antingen finita eller infinita.

Man kan också använda syntaktiska kriterier för att avgöra om ett verb är finit, i och med att endast finita verb i normalfallet kan föregå ett subjekt eller ett satsadverbial. Att svenska barn tidigt skiljer på finita och infinita verb genom att de placerar finita verb före negationen och infinita efter har tidigare visats av Lange och Larsson (1977), Strömqvist m.fl. (2001) och Santelmann (1995). Waldmann (2008) bekräftar detta. Men i det allra tidigaste barnspråket är det ändå vanskligt att utgå från syntaktisk placering för att avgöra finithet, dels för att det inte finns så många negationer och andra satsadverbial, dels för att det ändå hos de tidigare undersökta barnen finns enstaka tidiga exempel på negation före finita verb och efter infinita verb. Dessutom finns det hos Hanna fram till 1;8 ganska många exempel på morfologiskt säkra infinitiver som står före subjektet. Placering utgör alltså inte något pålitligt kriterium i de första inspelningarna. Principen att ”tagga stort” riskerar i det läget medföra direkt felaktiga analyser och jag valde här i stället att analysera mer försiktigt. För Hannas del beslutade jag mig under taggningen av inspelning 2;0 att hon be-

härskade finita verb (se kapitel 4), och använde därför syntaktiska kriterier för att avgöra finithet från och med 2;0.

För att kunna avgöra verbets status måste man tolka det utifrån barnets allmänna språknivå och yttrandets kontext. Delvis därför har jag kompletterat den syntaktiska analysen med en morfologisk analys, så att man ska kunna jämföra satser som jag har tolkat som finita, med satser som också har morfologiskt finita verb. I satsanalysen använder jag mig av följande kriterier för att avgöra om ett verb är finit:

1. Verb som visar upp morfologisk finithet, dvs. verb i presens eller preteritum, får alltid taggen F, oavsett placering i satsen och språknivå.

2. Det finita systemet verkar ha etablerats hos Hanna vid 2;0, så från och med inspelning 2;0 taggas också syntaktiskt säkra finita verb, dvs. verb som följs av antingen ett subjekt eller ett satsadverbial, med taggen F. Före 2;0 tolkas de som infinita och får taggen V.

3. Hjälpverben *få* och *ha*, som oftast inte visar upp morfologisk finithet, taggas som F om de följs av ett verb i infinitiv.

4. Andra verb som inte har en morfologiskt säker finit form (basformer, se s. 69) tolkas fram till 2;0 som infinita (med taggen V) oavsett placering och funktion. Från och med 2;0 görs en bedömning utifrån kontexten om verbet mest sannolikt motsvarar en presensform eller preteritumform, då det får taggen F, eller en rotinfinitiv som får taggen V. Denna bedömning är i många fall inte så svårt att göra, eftersom jag kan se vad som händer när yttrandets fälls. I de fall man kan tänka sig flera alternativa uttryck taggas verbet med V.

5. Imperativa verb taggas som F, förutom verbet *titta* som nästan alltid liknar en interjektion mer än ett verb. Eftersom det förekommer så ofta hade det varit svårt att urskilja när de riktiga finita verben kommer om också *titta* hade fått taggen F. Därför taggas *titta* som V.

6. I transkriptionen har jag angivit presens av *vara* med olika grad av säkerhet (se avsnitt 2.4.5). I satsanalysen överförs den bedömningen på så sätt att säkra särskrivna verb och möjliga verb (ˆe) analyseras som finit verb.

Taggen F är alltså en fråga om bedömning, särskilt i de tidiga inspelningarna. Det är därför viktigt att analyser som omfattar finita verb i den tidigaste perioden alltid åtföljs av en morfologisk analys.

3.5.3 V – verb

Alla verb som inte är definierade som finita enligt kriterierna ovan får taggen V, dvs. verb i infinitiv eller supinum. Vid tveksamma fall har jag hellre taggat med V än F. Verbmorfologin taggas som sagt separat.

Particip analyseras dock inte som V – inte ens verbala presens- eller perfektparticip som i *komma springande*, *bli sittande* och *bli målad*. Den perifrastiska passiven analyseras alltså som kopulaverb plus predikativ. SAG nämner visserligen att *komma* och *bli* närmast är att betrakta som hjälpverb i presensparticipkonstruktioner (Particip §37), men eftersom de är så sällsynta taggar jag dem som predikativ[10]. Presensparticipen kan man alltid söka ut genom att söka efter ord som slutar på -ande/-ende. En genomsökning av Hannakorpusen gav endast *spännande*, samt två particip ur en uppläst vers.

3.5.4 O – objekt

Både direkta och indirekta objekt får taggen O, så det kan finnas flera O i en sträng. Det förekommer också dubblerade objekt i satser, dels när det finns dubbla objekt i annex, dels i satsflätor (s. 57).

Gränsdragningen mellan objekt och verbpartikel kan vara svår i lexikaliserade verbfraser med substantiv (*äga rum*, *hålla vakt*, Mamban s. 67). Jag räknar substantivet som objekt om man kan ändra numerus eller species på substantivet utan att betydelsen förändras nämnvärt. Ett par svåranalyserade uttryck som är vanliga i korpusen är ”ha/få/göra ont” och ”få plats”. *Ha ont* och *få ont* tolkas enligt kriterierna ovan som verb + objekt (med viss tvekan, *Jag har/fick det onda i tån*), och *göra ont* i *det gör ont* tolkas också som verb + objekt. Eftersom hela *det gör ont* är ett lexikaliserat uttryck är det svårt att avgöra vilken av analyserna som är rimligast, men man kan notera att Hanna upprepade gånger använder uttrycket produktivt, t.ex. ”Jag gjorde ont” (2;2) i betydelsen ’jag gjorde illa mig”. *Ont* kan ju också utan vidare ta framförställda bestämningar (”det gör lite/mycket ont”). Om man vill undersöka hur just ordet *ont* används kan man alltid söka ut alla förekomster av det lexemet.

Få plats (i betydelsen *rymmas*) taggar jag däremot som verb + partikel, eftersom betydelsen blir en annan om man ändrar nominalet (se vidare under partiklar, s. 64).

[10] Visserligen var *komma v-ande* mycket vanligt i Hannas språk under en period när hon var runt 2 1/2 år. Då sa hon ofta ”Här kommer jag/pappa/xx springande/cyklande/åkande” och liknande. Hon övergeneraliserade också konstruktionen och sa ofta *Det kommer jag inte görande!*, vilket ju talar för att i alla fall hon betraktade participen som ett verb. Tyvärr har inget av detta kommit med i någon inspelning, men det finns i kompletterande dagboksanteckningar.

3.5.5 P – predikativ

I den här gruppen ingår subjektiv predikativ, objektiv predikativ och fria predikativ (t.ex. *själv*). Det utbrutna ledet i emfatiska omskrivningar utgör bundet subjektspredikativ (SAG IV, s. 515), och utbrytningsbisatsen analyseras som efterställt attribut till predikativet.

(74) MOT: e de pappan som spelar piano nu ?
%syn: $FSPQ M:ipre ¢UFOA M:basf 2;0

Skillnaden mellan subjekt och nominalt predikativ eller formellt och egentligt subjekt berör jag kort på s. 57. Det finns en hel del gränsfall i vuxenspråket och i barnspråket där det ofta fattas en del morfem kan det vara så gott som omöjligt att avgöra. När Hanna säger ”Där e (s)tor nalle” (1;9) och pekar på den stora nallen, vet man inte om hon menar ’där är den stora nallen’ eller ’det där är en stor nalle’. Om kontexten inte på något vis tyder på att ”där” är en äkta lokativ lutar jag åt att tolka ”stor nalle” som en predikativ (men jag analyserar ändå ”där” som A, dvs. $AFP). Det är dock en ganska liten andel som är så svåranalyserade, mest i de allra tidigaste inspelningarna. De flesta subjektiva predikativer är tydligare:

(75) CHI: oj , &bjum # e synnö [: sönder] !
%syn: $FP M:ipre 1;9

(76) CHI: de Hannas blong [: ballong] ?
%syn: $SPQ 1;11

Alla adjektivfraser i enledade yttranden får koden €P vilket betyder att det är ett adjektiviskt led, vars funktion jag inte kan uttala mig om. Till adjektiviska led räknas också adjektiviska pronomen och vissa adjektiviskt använda substantiv, t.ex. *slut* och *fel*.

3.5.6 A – verbfrasadverbial

I analysen görs ingen skillnad på bundna adverbial (prepositionsobjekt) och fria adverbial. Tidiga adverbial är *dä*, *borta* samt *lite* när det kombineras med ett verb som i ”sova lite” och ”(k)lappa lite voffen” (1;9). När *lite* däremot kombineras med ett substantiv räknar jag in det i nominalfrasen ”du åssså lite papper ?” (’vill du också ha lite papper’ 1;11).

Ett annat tidigt adverbial i barnspråket är *färdigt*, som i ”ätit färdigt”, ”spelat färdigt”. *Färdigt* är visserligen inte självklart ett adverbial. SAG tar bara upp ”färdigt” som predikativ till outtalat objekt (SAG III, s. 372), med exemplen *få*/*ha*/*göra (något) färdigt*. Det kan förekomma både före och efter ett objekt,

och kan alltså fungera som antingen partikel eller predikativ (jfr SAGs exempel *skriva färdigt artiklarna* vs *skriva artiklarna färdiga*, SAG III, s. 421). Jag väljer ändå att tagga *färdigt* som adverbial eftersom det i mitt material främst förekommer med verben *äta*, *läsa* och *spela*, och att anta att det finns ett outtalat objekt som är ”färdigt” är något långsökt i mina ögon. Man kan möjligtvis ”äta bullen färdigt”, men om man har gjort det är det inte *bullen* som är färdig, utan ätandet av den. Det betyder att *färdigt* mer liknar ett sättsadverbial i dessa fall. I fall där *färdigt* tydligt fungerar som predikativ sätter jag naturligtvis taggen P och om det placeras före ett utsatt objekt som partikel – L (se nedan).

I det prepositionslösa tidiga barnspråket kan det vara svårt att avgöra om en konstituent är adverbial eller objekt. Analysen utgår då ifrån min tolkning av vad som ligger närmast målspråket (se också diskussionen om utelämnade prepositioner på s. 72).

Sammansatta adverbial räknas som två konstituenter: ”där på golvet” blir alltså €AA. Om ett sammansatt adverbial skulle inleda en sats innebär det alltså att strängen blir en V3-sträng. I Hannas dokumenterade språk förekommer aldrig sammansatta adverbial satsinitialt, och bara en gång i det taggade vuxenspråket. Däremot förekommer några adjunktionella *så* i Hannas språk (och något enstaka dubblerat adverbial) som också ger V3-strängar (SAG IV, s. 694).

(77) CHI: så här högt så studsa den ti mej .
%syn: $AAFSA PP:prep 2;10

Ordet *så* taggas också som adverbial när det betyder ’på det här sättet’, ”den kan man ha så” (2;1), däremot som interjektion i betydelsen ’nu var det färdigt’,”pappa sitter där, så” (1;10). Intonationen skiljer. *Så* är ett mycket vanligt ord i satsbaser och motsvarar ofta snarare satsadverbial än verbfrasadverbial, men jag har inte fördjupat mig i analysen utan taggar alla adverbiella *så* som ser ut att befinna sig i satsbasen med A.

(78) CHI: å så ska pappa ha lammet .
%syn: $AFSA 2;8

3.5.7 I – satsadverbial

Denna grupp utgörs av alla slags adverbial som inte svarar på var, när, hur eller under vilka omständigheter. Här ryms *inte*, *åsså*, *ju* och liknande, men också frågeavslutaren *då*/*da* i ”Var är Hanna da?” (inte lika med *då* som tidsadverbial), som SAG kallar konklusivt satsadverbial (SAG IV, s. 146). Det förekommer mycket ofta i vuxenfrågor, och dyker upp hos Hanna i 2;8.

Fram till 1;10 förekommer ofta ordet ”mä” distinkt uttalat. Vid något tillfälle förekommer det i kombination med ”åsse” (”ha nä [=? Hanna] mä åsså” (1;7)) då det ser ut som ett *med* (additivt satsadverbial i postfokal ställning, SAG IV, s. 127). Men ordet är så svårtolkat i de andra fallen att jag har fått klassificera det som X (oanalyserbar konstituent).

3.5.8 L – verbpartikel

Det är långtifrån klarlagt vilka led som utgör verbpartiklar i svenskan. Vid transitiva partikelverb kan man använda placeringen som indikator; om ordet (adverbet eller prepositionen eller eventuellt andra ord) står före objektet är det en partikel. När det gäller små barn är dock placeringen ett något osäkert kriterium, eftersom det förekommer att barn placerar en tydlig verbpartikel efter det direkta objektet (Platzack, 2003). Betoningen och samhörigheten med verbet brukar nämnas som det säkraste tecknet på att något är en partikel. Men när det gäller intransitiva verb kan det ändå vara svårt att skilja mellan adverbial och partiklar, t.ex. för *bort* i ”ska du gå bort till mamma”. Trots alla svårigheter att definiera vad som är en partikel är det dock oftast lätt att känna igen dem och analysen har inte varit svår. I tvetydiga fall har jag valt partikelanalysen framför adverbialanalysen, eftersom en undersökning av verbpartiklar ingår i den vidare studien, och jag vill då ta med alla potentiella partiklar också. [11]

Gränsen mellan objekt och partikel som utgörs av nominal kan också vara oklar i några fall (jfr s. 61). Om betydelsen blir förändrad av att nominalet böjs i singular/plural eller av en bestämd slutartikel betecknar jag nominalet som partikel, t.ex. *få plats* = rymmas ≠ *få platsen*, *få platser* (jfr Mamban, s. 70).

Några andra knepiga fall tar jag upp här: SAG betecknar *ta vägen* som lexikaliserad verbfras i kvesitiva satser (”Vart ska du ta vägen?”, SAG VI, s. 742),

[11] Partikelförbindelsen har alltså ordgruppsbetoning, dvs. verbet är svagt betonat och partikeln starkt betonad. Prepositioner är oftast obetonade och därför brukar det vara lätt att känna igen en preposition som fungerar som partikeladverbial. I *’hoppa i ’vattnet* betraktas *i vattnet* som ett befintlighetsadverbial, medan *hoppa ’i ’vattnet* innebär riktning och gör *hoppa i* till partikelverb. Men detta verkar inte alltid stämma. I ett par vuxenyttranden (se nedan) är riktningsprepositionen obetonad. Pippis pappa ramlar (ner) ’i lådan när han dansar och pappadockan ramlar ’av stolen ner på golvet. Men båda exemplen uttalas med tydlig betoning på verbet respektive nominalet medan prepositionen är obetonad. Trots det har jag taggat prepositionen som partikel.

(1) MOT: han ramlar i lådan . 1;10

(2) MOT: han ramla av stolen .
%syn: $SFLO M:basf PP:prep 2;0

Observera att jag utgår ifrån att verbpartikel är en satsdelsbeteckning och inte en ordklass, jfr SAG III och IV, s. 417.

och det ligger närmast till hands att tagga *vägen* som partikel, även då det förekommer i andra kontexter, ”Min glass tog vägen!” (Ha 2;2) i betydelsen ’min glass försvann’. *Bli* i *låt bli lampan* klassas också som partikel. Verb + partikel kan mycket väl vara holofraser i barnspråket, ”g’in”= gå in, ”g’upp” = gå upp.

3.5.9 W – frågeord

V-frågeordet i en kvesitiv sats kan fungera som subjekt, objekt, predikativ eller adverbial. Eftersom mitt analyssystem bygger på att satsdelstaggarna bara är ett tecken långa har jag inget sätt att markera att satsdelen utgörs av ett v-frågeord, men eftersom kvesitiva frågor och frågeorden är intressanta vill jag kunna söka ut dem. Därför har jag valt att tagga v-frågeordets funktion som frågeled och inte dess funktion som satsled, så att ”vem såg du?” taggas $WFSQ. Om jag vill söka ut satser med topikaliserat objekt ($OFS*) får jag alltså inte med de satser som börjar med ett frågeordsobjekt. Å andra sidan kan jag söka ut alla kvesitiva satser på ett enkelt sätt. De hade annars varit svåra att hitta i och med att barnspråksformerna av frågeorden kan se ut på många sätt.

I den diskontinuerliga frasen”**Vad** sa du **för nåt**” (jfr *Du sa **vad för något*** $SFOQ) får v-frågeordet taggen W medan resten får satsdelsbeteckning för hela frasen (O i det här fallet). ”Vad sa du för nåt” får alltså strängen $WFSOQ. SAG betecknar ”Vad (. . .) för en/någon” som interrogativt flerordspronomen där frågeordet ensamt kan utgöra satsbas (SAG IV, s. 736f). ”Var (. . .) (för) någonstans” analyseras på samma sätt. Konstruktionen är mycket vanlig i korpusen.

Andra användningar av v-frågeord är:

ensamma	*Va?*	€WQ
avslutande	*Det e gott, va?*	$SFPWQ
i fragment	*Va gott*	€WP
bisatsinledare	få ja se *vem som filma nu* då?	¢WUVA

3.5.10 Q – fråga/utelämnat v-frågeord

Taggen Q står för frågeintonation och inte för en konstituent i satsen. Alla satser med frågeintonation får ett Q sist i strängen, även de med ett explicit frågeord eftersom ett v-frågeord inte behöver innebära att något blir en fråga. Bedömningen om ett yttrande har frågeintonation eller inte är egentligen redan gjord i transkriptionen, men det underlättar utsökningar att ha frågetecknet i

satskodningen också, eftersom interrogativa satser och deklarativa satser behöver analyseras separat.

Om Q står sist i satsen är det en fråga, t.ex. en rogativ (”Får han lite salt?” $FSOQ) som då kan skiljas från V1 deklarativer (”(Så) Får han lite salt.” $FSO). När yttrandena har prototypisk interrogativ respektive deklarativ intonation är det inte svårt att skilja dem åt, men i vissa yttranden i Hannakorpusen har intonationen varit svårbedömd, särskilt under en period (2;0 till 2;2) då Hanna verkade övergeneralisera frågeintonationen.

Om Q står först i satsen står det för att ett v-frågeord är utelämnat (från ett målspråksperspektiv). Det kan man oftast identifiera med hjälp av prosodin. Visserligen *kan* en kvesitiv sats med utelämnat frågeord ha samma intonation som en rogativ sats, men oftast är det lätt att avgöra vilken typ det handlar om. Jag frångår alltså här principen att inte tagga led som inte finns (hörs), men det skulle annars bli svårt att hitta de satser som saknar frågeord. En kvesitiv sats som saknar frågeord får alltså Q både först och sist: ”e de” $QFSQ.

3.5.11 U – bisatsinledare

I denna grupp ingår subjunktioner, adverb som inleder bisats och infinitivmärken. *Som* räknas som subjunktion.

Hanna har en period när många satser inleds med något som ser ut som subjunktionen *för att*, ”fatt”. Det skulle kunna vara ordet ”fast” också, och det är tänkbart att Hanna använder ”fatt” både adversativt och finalt eller kausalt. Den finala eller kausala tolkningen verkar ofta rimligare, så jag har valt att tagga alla förekomsterna av ”fatt” som subjunktioner. Det betyder att strängen också klassas som bisats, även om den har huvudsatsordföljd och står självständigt. Det betyder inte att jag tror att dessa satser är underordnade satser, men jag vill få med dem för att få en helhetsbild av bisatserna.

(79) CHI: xx fatt den e slut # den # .
%syn: ¢USFPS 2;4

(80) CHI: fatt den e så mycke sena [: senap] i den !
%syn: ¢USFPA PP:prep 2;4

3.5.12 C – konjunktion

Konjunktioner som samordnar hela huvudsatser ingår inte i analysen, men det gör däremot konjunktioner i pseudokoordinationer (se exempel 81) och i satssammandragningar (se t.ex. exempel 62 på sid. 55), där konjunktionen inte kan utelämnas utan att satsen blir ogrammatisk.

(81) MOT: dockan ligger å sover .
%syn: $SFCF M:spre M:spre 2;0

Jag märker alltså inte ut samordningar i andra fraser, även om det hade varit intressant. Så ett subjekt som består av två samordnade nominalfraser ”pappa å ja” får ett enkelt S. Om man är intresserad av konjunktioner bör man utan större problem kunna söka ut dem genom att söka på lexem. Det man tyvärr missar då är de tidiga samordningar som saknar utsatt konjunktion. Utelämnade konjunktioner i verbfrasen är däremot taggade, se nästa avsnitt.

3.5.13 + – utelämnat led

Plusteckenet står för ytterligare ett undantag från principen att bara tagga led som uttalas. Barnens tidigaste pseudosamordnade verb samordnas utan konjunktion.

(82) CHI: pappa måste ligga vila föt [: först] .
%syn: $SFV+VA M:ipre M:infv M:basf 2;1

(83) CHI: dä ligge Lina sove .
%syn: $AFS+F 2;3

Prosodin talar dock klart för att det är en pseudosamordning och inte en uppräkning av verb, så man kan säga att plustecknet i likhet med Q-tecknet står för något som ”finns där” genom prosodin. Eftersom samordningarna kan se så olika ut är det svår att söka ut dem som saknar konjunktion om de inte markeras på något sätt. Jag markerar dock inte utelämnade konjunktioner i andra fraser, så samordningar av nominal utan konjunktion kan man alltså inte söka ut.

Två plustecken, ++, används i satssammandragningar, dvs. för led som är utelämnade under identitet i samordnade satser. Det kan vara ett led eller flera led som är utelämnade, men verbet måste finnas där för att det som är kvar ska betraktas som en samordnad sats. Strängen inleds med ¢-tecknet.

(84) FAT: ska du gå bort till mamma å spela in ?
%syn: $FSVLAQ M:ipre M:basf PP:prep ¢++CVLQ M:basf 2;0

Jag använder således inte ++-taggen för yttranden som inte innehåller verb (eller borde ha innehållit ett verb) även om de skulle vara en fortsättning eller ett svar på något som har sagts i den föregående texten. De inleds också med meningsfragmentsinledaren (€).

(85) MOT: ja , här får du lite spagetti .
%syn: $AFSO M:basf
CHI: å lite sås åsse .
%syn: €NI 2;0

I barnspråkets ofullständiga satser kan man inte alltid avgöra om det är frågan om en satssammandragning eller en underordnad sats (eller något annat), allra minst innan barnet börjar använda konjunktioner och subjunktioner. Men om man misstänker att exempel 86 är mer komplext än exempel 87 måste de taggas olika.

(86) CHI: nu han lamla ont knana .
%syn: $ASV M:basf ¢++OA PP:noll
MOT: oj , ramla han å fick ont i knäna igen ? 2;0

(87) CHI: han ha ont i knana .
%syn: $SFOA M:basf PP:prep 2;0

Jag utgår även här ifrån den respons Hanna får som också i efterhand verkar vara den rimligaste tolkningen och taggar 86 som en satssammandragning (*nu ramlade han och fick ont i knäna*). Men man kan naturligtvis inte veta att det är så man ska tolka yttrandet. Det kan också vara en underordning (*så att han fick ont i knäna*). I avsnitt 4.4 får jag anledning att återkomma till detta yttrande.

3.5.14 N – obestämbart nominal

I strängar med få konstituenter kan det vara svårt att avgöra vad nominalfraserna har för satsdelsfunktion. Det gäller i hög grad i barnspråket. I ”ha den åsse” kan det vara *den* som ska ha något eller någon som ska ha *den*. I sådana fall får nominalet taggen N. Det gäller även i meningsfragmnet.

Alla nominalfraser i enledsyttrandena får koden N, för obestämbart nominal. Två nominal i följd taggas €NN. I dessa fall kan man rätt ofta misstänka att det rör sig om possessiva uttryck, så att de egentligen utgör ett led, men ibland kan det också vara ett exempel på det som Brown (1973) kallar *agent and object*, vilket skulle utgöra en sats i mitt system. I fallet ”Hanna bok” (1;7) vet man inte om Hanna påpekar att det är hennes bok, eller om det betyder att Hanna ska läsa bok.

3.5.15 R – rest/räkneord

I den här gruppen förekommer fristående räkneord och några fristående ordningstal i €-strängar.

3.5.16 X – oanalyserbart

X står för ohörbart, eller obegripligt. Ett X står för det som tycks vara ett ord och XX för det som verkar vara flera ord. Det innebär att de delar som är analyserbara taggas. Om satsen är användbar beror sedan på vad man vill ta reda på. Om man t.ex. är intresserad av V2 kan en sträng som $XFSO vara intressant. Jag analyserar också oavslutade yttranden (som avslutas med symbolen +... (trailing off) eller +/. (avbrutet av en annan talare). Att dessa yttranden inte är fullständiga markeras med ett X sist i strängen. Vill man bara ta med fullt ut analyserade yttranden kan man söka ut alla strängar som inte innehåller X.

3.6 Verbmorfologi

I satsanalysen markeras verben efter sin funktion som finit eller infinit verb. Som jag redovisade i avsnitt 3.5.2 kan man använda både syntaktiska och morfologiska kriterier för att bedöma om ett verb är finit eller infinit. Eftersom det kan vara svårt att avgöra syntaktisk finithet i det tidiga barnspråket har jag valt att göra en oberoende morfologisk analys av verbformerna. I denna analys är det därmed enbart verbets form som avgör klassificeringen (med några få undantag som redovisas nedan). Det gör att jag kan få en uppfattning om hur barnet tillägnar sig verbens morfologi och att jag kan undersöka finithet med två av varandra oberoende mått.

Lisa Christensen (under arbete) har gjort en omfattande undersökning av hur verbmorfologin lärs in av hennes två undersökta barn. Hennes analysmetod fångar många intressanta aspekter och jag har därför valt att basera min morfologianalys på hennes metod. Christensen konstaterar till att börja med att den gängse indelningen av verb i fyra konjugationer (se t.ex. SAG II, s. 557 ff) inte fungerar bra när man undersöker hur barn lär sig böjningsformerna. Hon menar att två mycket framträdande egenskaper, som barnen också mycket tidigt lägger märke till, är den akuta accenten i vissa presensformer och -it-ändelsen i vissa supinumformer. Hon föreslår en annan konjugationsindelning som baserar sig på verbens stammar.

(88) Verbböjning

Stamtyp	Imperativ	Infinitiv	Presens	Supinum	Preteritum
1 (svag)	gunga	gunga	gungar	gungat	gungade
2a (svag)	lek	leka	leker	lekt	lekte
2b (stark)	sitt	sitta	sitter	suttit	satt
3a (svag)	bo	bo	bor	bott	bodde
3b (stark)	gå	gå	går	gått	gick

Den stamtyp som är viktig för morfologiinlärningen är stamtyp 2, där skillnaden mellan infinitivens grava accent och presensformens akuta accent utgör en tydlig signal åt barnen att skilja presens från infinitiv. Hos de starka verben i stamtyp 2 är också supinumändelsen -it mycket framträdande.

Man måste också skilja på hjälpverb och huvudverb när man redogör för hur barnet lär sig finita verbformer. Jag utgår från Christensens klassificering av verbformer, och definitionerna är hämtade från Christensen (under arbete). Till skillnad från Christensen räknar jag particip som predikativer och inte verbformer. Hannamaterialet är taggat för verbmorfologi fram till och med 2;2. Efter det har endast ovanliga verbformer markerats, dvs. alla former utom presens, infinitiv och basformerna (se nedan).

M:basf – basform Basformerna består av alla verb av stamtyp 1 och 3, som inte har tydliga böjningssuffix, t.ex. *ramla* och *få*, och dessutom den korta infinitivformen eller imperativ av *vara*, *va*. Eftersom infinitiven och imperativen också sammanfaller i stamtyp 1 och 3 kan man med hjälp av den morfologiska analysen inte avgöra om ett verb används i imperativ. Den analysen har jag istället lagt i satsledsbeteckningen (F för imperativ och V för infinitiv, se avsnitt 3.5.2). Samma resonemang gäller för de många suffixlösa basformer som har tydlig preteritumanvändning: de får morfologitaggen M:basf, men satsdelsbeteckning F. Jag följer också Christensen i att betrakta hjälpverben *ha* och *få* som suffixlösa basformer, även i de sammanhang det följs av ett huvudverb (men i satsanalysen taggas de också som F).

M:infv – suffigerad infinitivform Infinitivformer som är bildade genom att ett -a-suffix läggs till stammen betecknas som suffigerade infinitiver, och utgörs av alla verb av stamtyp 2 som slutar på -a, t.ex. *bygga* och *sitta*.

M:impv – imperativ De enstaviga verbformer av stamtyp 2, t.ex. *lägg* och *bygg* utgör säkra imperativer. Här har jag dock måst avvika från Christensens

metod att bara titta på verbets morfologi utan hänsyn till kontext, eftersom jag har en liten men inte försumbar andel stamformer i mitt material, som tydligt inte är direktiva. Dessa har fått taggen M:stam (se nedan).

M:ipre – icke-suffigerad presens En stor del av verben i presens utgörs av de modala hjälpverben och presens av *vara*. Dessa verb bildas inte genom böjningssuffix utan har ordformer som måste lagras i lexikon. De taggas med M:ipre och består alltså av *ska*, *kan*, *vill*, *måste* och *är*. Några få deponensverb i presens räknas också in här, t.ex. *finns*.

M:spre – suffigerad presens I denna grupp ingår till största delen stamtyp-2-verb eftersom den akuta accenten är tydligt urskiljbar. Suffixet är transkriberat som -e(r), -å(r) eller -ö(r), t.ex. *sittå*, *bygge* och *sovå*. Det ingår dock också några verb av stamtyp 1, där slut-r:et är tydligt. (”går fötter ” (1;10)). Till den här gruppen räknas också verben *kör* och *gör*, eftersom den infinitiva formen har -a, även om de sammanfaller med imperativen (också här har kontexten fått spela roll, vilket även Christensen har tillåtit sig).

M:supm – supinum Supinumformens ändelse är ofta tydlig och alla verbformer som slutar på -t eller -it får taggen M:supm. I perfektkonstruktioner där huvudverbet saknar supinumsuffix klassificeras det som basform, även om det har ett ha-hjälpverb (”Mast ha slå sej” (2;1)).

M:iptm – icke-suffigerad preteritum Till denna grupp räknas alla starka preteritumformer, vilket för det mesta innebär enstaviga verbformer med vokalväxling, t.ex. *gick* och *tog*. De starka verben *komma* och *sova* visar inte upp vokalväxling i preteritum, vilket medför att preteritumformen sammanfaller med imperativen, *kom* och *sov*. I likhet med Christensen har jag dock låtit kontexten avgöra hur de ska klassificeras också här. Det modala hjälpverbet *skulle* räknas som icke-suffigerad preteritum, medan *kunde* räknas till de suffigerade, eftersom -de-suffixet är tydligt.

M:sptm – suffigerad preteritum När det dentala preteritumsuffixet sätts ut markeras verbformen som M:sptm. Även icke-målspråkliga förekomster av dentalsuffix räknas hit, t.ex. ”vilde” = *ville* (2;8).

M:stam – verbstam I barnspråk förekommer inte sällan bara verbstammen utan ändelse. Christensen har inte några sådana i sitt material och har därför inte behövt klassificera dem. Men i Hanna-materialet förekommer de ibland.

(89) CHI: dåt , lapp öjan .
mål: xx, klappa örat (Hanna klappar själv örat och kommenterar detta.)
%syn: €X $VO M:stam 1;9

M:oanb – oanalyserbar I denna grupp placeras verb med ohörbar eller svårtolkad ändelse, samt verb med svårtolkad accent.

M:ssss – s-form Alla s-former, dvs. passivformer, deponensverb och andra verbböjningar som slutar på -s får utöver koden för övrig verbmorfologi koden M:ssss, *kräktes* taggas alltså M:sptm M:ssss.

3.7 Prepositionsfraser

Riktigt små barn använder inte prepositioner, i alla fall inte i kombination med en rektion. Hur prepositionsbruket växer fram i svenskt barnspråk har visats av bl.a. Lange (1976) och Josefsson och Håkansson (2000). Jag markerar alla utsatta prepositioner samt utelämnade prepositioner i obligatoriska kontexter på alla nivåer i satsen.

Det är naturligtvis inte helt oproblematiskt att avgöra vad som är en obligatorisk kontext för prepositioner. Verben *rita* och *bygga* som förekommer flitigt i inspelningarna kan kombineras både med ett objekt ”bygga/rita en bil” eller med ett bundet adverbial ”bygga/rita med den (legobiten/kritan)” eller ”bygga/rita på den (legoplattan/papperet)”. Ett yttrande som ”bygga den” blir då svårtolkat. Trots det har jag försökt skilja objekt från adverbial när den icke-språkliga kontexten ger tydliga ledtrådar, men ofta har jag taggat nominalet med N för obestämbart nominal. I dessa fall är taggen PP inte tillämplig. Det betyder att PP-taggen inte förekommer särskilt ofta i de tidigaste inspelningarna, där de flesta nominal bara taggas som N.

Ett verb som *titta* har lite oklar status. I första hand tänker man sig nog att *titta* kombineras med ett bundet adverbial, *titta på ngt*. Det förekommer i vuxenspråket ofta i imperativ (får man förmoda) tillsammans med nominal i meningsfragment: ”titta, Pippi!”, ”titta, katten!”. Det förekommer dock också att uttryck som ”titta Pippi” utgör en prosodisk enhet, och då låter det som om verbet ”titta” tar ett nominalt komplement. När barnet säger ”titta x” är det därför inte rimligt att utgå ifrån att x är en prepositionsfras med utelämnad preposition. Jag har därför taggat ett sådant komplement som $VN, dvs. som ett nominal med oklar status. *Titta* förekommer också i kombination med bisatser, och då nominala bisatser i vuxenpråket:”ska vi titta vad Karin gör”. Man kan förmoda att den nominala bisatsen är en rektion i en prepositionsfras med

utelämnad preposition, särskilt som prepositionen inte så gärna kan utelämnas i absolut position (SAG IV, s. 533) ”Vad Karin gör ska vi titta (efter/*∅) om en stund”. Men eftersom konstruktionen ”titta vad xx” aldrig förekommer med preposition i vuxenspråket (däremot 5 gånger utan preposition) taggar jag också bisatsen som nominal med oklar status, och inte adverbial med utelämnad preposition.

(90) CHI: titta Molly gö .
%syn: $VN M:basf ¢SF M:spre 1;10

PP:noll – utelämnad preposition Jag märker upp prepositionsfraser på alla nivåer, oavsett satsdelsfunktion, där vuxenspråket skulle ha haft en preposition.

(91) CHI: dä golvet .
%syn: €AA PP:noll 2;0

(92) CHI: de vila min kna .
%syn: $SVA M:basf PP:noll 2;0

Det kan vara svårt att höra prepositionerna. Där utelämnade prepositioner (taggen PP:noll) förekommer i vuxenspråket kan det vara för att talhastigheten är så hög att man hade behövt bättre ljudkvalitet för att kunna höra dem. Och detta kan naturligtvis gälla även i barnspråket (även om det inte kommer upp i samma hastigheter).

I ettords- och tvåordsyttranden kan man sällan veta att något är ett adverbial som kräver preposition i målspråket. Men ibland framgår det faktiskt tydligt ur kontexten. I exemplet nedan pekar Hanna på lådan där papperet finns.

(93) MOT: då går vi å hämtar pennor och papper .
CHI: pappe , lådan xx pappe .
%syn: €N $AXN PP:noll
mål: i lådan (finns det/hämtar vi) papper 1;6

PP:prep – utsatt preposition Utsatta prepositioner i alla slags prepositionsfraser får koden PP:prep. Prepositionella verbpartiklar taggas dock bara som partiklar och inte som prepositioner.

3.8 Slutkommentar

Metoden och manualen har vuxit fram under arbetet med taggningen och jag har lagt till och dragit ifrån taggar allt eftersom de har visat sig nödvändiga

eller överflödiga. Vissa barn är utpräglade *fillers*-barn, dvs. de använder protomorfem som kan förekomma i olika kontexter (Pepinsky m.fl., 2001). För ett sådant barn kan det vara bra att lägga till en särskild tagg för detta. Hanna har inte använt många fillers, och jag har gått efter principen att ”tagga stort” och analyserat tidiga förekomster av sådana morfem fullt ut, även om det är långt kvar till de säkra förekomsterna.

För utsökningarna använder jag de CLAN-kommandon som finns att tillgå, mest *combo* och *freq*. Att skriva söksträngar som tar fram det man söker och inte något annat har visat sig vara den stora utmaningen, vilket det också varnas för i CLAN-manualen (MacWhinney, 2006b, s. 64). Beslutet att låta flera satser ingå i ett och samma yttrande visade sig ha negativa konsekvenser för utsökningarna. Ofta måste man söka ut ett fenomen i flera led, man *pipar* sökningen. Om man t.ex. vill söka ut alla strängar som har V (infinit verb) men saknar F (finit verb) kan man först söka ut alla strängar som har V och därefter skicka sökresultatet till en ny sökning som tar fram alla strängar som inte har F. Men om det finns flera satssträngar på samma rad kan den första sökningen hitta ett V i den ena strängen, och den andra ett F i den andra strängen. Det betyder att jag har måst gå igenom alla förekomster för hand, något som jag ändå avsåg att göra. För den statistik som är framtagen i nästa kapitel har dock de reguljära uttryck som finns att tillgå i combo inte alltid räckt till. Jag har då utfört kompletterande sökningar med *grep*, som är ett betydligt kraftfullare sökverktyg som bl.a. finns tillgängligt via sökfunktionen i textredigeringsprogrammet TextWrangler. Trots detta tekniska problem vidhåller jag att en yttrandeindelning som baserar sig på TKE:er är det enda rimliga. För att få en god uppfattning om satsens tolkning måste man lyssna på hela yttrandet, och eventuella tekniska problem som en sådan indelning för med sig kan lösas på andra sätt.

Jag hade kunnat lägga upp kodningssystemet på många andra sätt. I det avslutande kapitlet kommer jag att utvärdera systemet och nämna några uppenbara problem. Men taggningen har ändå fungerat bra, vilket jag hoppas kunna visa i nästa kapitel.

Kapitel 4

Analys av utdata

I detta kapitel ska jag redovisa utfallet av min analysmetod och visa att det ger en god bild av när Hanna har tillägnat sig olika egenskaper hos den finita satsen. Utfallet ger samtidigt en bred beskrivning av Hannas grammatiska språkutveckling mellan 1;6 och 2;10.

Det taggade Hannamaterialet gör det lätt att söka ut statistik över olika fenomen. Men eftersom jag har haft principen att ”tagga stort” måste statistiken undersökas kvalitativt för att vara användbar. Det är främst de första förekomsterna av varje nyhet som måste skärskådas, t.ex. om de första få exemplen av finita verbformer verkligen kan tolkas som finita, om de första icke-subjektsinledda deklarativerna ser ut att vara produktivt genererade eller om de består av ett fåtal mer eller mindre lexikaliserade fraser och så vidare.

Jag kommer att presentera en del tabeller över Hannas grammatiska språkutveckling och gå igenom hur språket bakom siffrorna ser ut. I det avslutande avsnittet kommer jag att sammanfatta framväxten av den finita satsen och diskutera vilka samband man kan se och vilka man inte kan se. Kapitlet handlar till största delen om kategorin $-strängar, dvs. i stort sett huvudsatssträngar, men jag redovisar också utvecklingen av andra grammatiska fenomen. Jag inleder med en översikt över Hannas genomsnittsliga yttrandelängd och fördelningen av de olika strängtyperna, och ger en kort redogörelse för Hannas upprepningar. Därefter går jag igenom verbmorfologin och huvudsatsen, för att till sist beskriva utvecklingen av de underordnade verbfraserna (¢-strängarna) och funktionsorden hos Hanna.

4.1 Språklig förmåga i stort

I min yttrandeanalys har jag alltså delat in satser och satsmotsvarigheter i olika strängtyper. Jag har skiljt ut meningsfragment (€-strängar) som inte innehåller något verb, och som jag bedömer inte heller behöver innehålla något verb. Förekomsterna i Hannamaterialet (hos Hanna och de vuxna) utgörs av vissa verblösa huvudsatsekvivalenter som också fungerar i målspråket, t.ex. uppmaningar ”(Du.) Inga fötter däruppe!” (V2;4[1]) eller utrop ”Va gott!”, ”Va bra!”. Andra är typiska ett- och tvåordsyttranden i barnspråket, mestadels så kallade benämningar med ett enkelt nominal ”Bil”, ibland med ett demonstrativt element före, ”Dä bil”.

Till huvudsatserna ($-strängarna) räknas förutom alla prototypiska huvudsatser också alla Hannas tidiga ett- och tvåordsyttranden med verb i infinit form, negativa imperativer (”Inte röra”), som ju också finns i målspråket, och senare rotinfinitiver. De satser som har fått inledaren ¢-strängar består i materialet av bisatsföregångare, bisatser, infinitivfraser, exklamativer och det samordnade ledet i satssammandragningar om det innehåller (eller borde ha innehållit) ett verb. För närmare definitioner av de olika strängtyperna, se avsnitt 3.4.

När Hannas språk blir mer avancerat blir hennes yttranden längre och fördelningen av de olika strängtyperna förändrar sig. Innan jag redovisar det behöver jag dock gå igenom Hannas spontana imitationer, eftersom de visade sig påverka siffrorna på olika sätt.

4.1.1 Spontana imitationer

I transkriberingen av materialet har jag angivit när Hanna spontant upprepar något som en annan talare säger i ett nära föregående yttrande. Jag har inte fördjupat mig teoretiskt i denna analys, men det visar sig att Hannas upprepningar måste tas med i beräkningen för viss utdata. Under arbetet med transkriptionerna skiljde jag mellan ordagranna imitationer och upprepningar i något förändrad form, *imit* och *rep* (se s. 27). När jag i den följande redovisningen talar om upprepningar eller spontana imitationer menar jag upprepningar av båda sorterna. Tabell 4.1 visar att upprepningar spelar en viss roll i Hannas språk fram till och med 1;10, men att de är mycket få från 1;11 och framåt.

I den typiska upprepningen fångar Hanna upp *ett* av orden i det föregående yttrandet, ofta men inte alltid det sista, och återger det, antingen exakt eller i något förändrat skick. De flesta upprepningsyttranden är därmed klassificerade som fragment, €-strängar, vilket blir viktigt att komma ihåg i avsnittet om

[1] V2;4 står för data från vuxenspråket i inspelning 2;4.

Tabell 4.1: Upprepningar

Insp.	Alla upprepn antal[2]	andel	Varav €	Varav $
1;6	18	6%	16	2
1;7	9	3%	7	1
1;8	21	6%	16	4
1;9	44	11%	34	8
1;10	76	20%	54	15
1;11	7	2%	4	2
2;0	3	1%	0	3
2;1	2	0%	0	2
2;2	1	0%	0	1
2;3	1	0%	0	0
2;4	1	0%	0	1
2;5	1	0%	1	0
2;6	0	0%	0	0
2;7	0	0%	0	0
2;8	0	0%	0	0
2;10	1	0%	1	0

förhållandet mellan strängtyper.

(94) MOT: ja , man ska akta sej lite för bin .
MOT: dom e lite farliga .
CHI: faelit . %com: rep %syn: €P
MOT: lite .
CHI: ite . %com: imit %syn: €A
MOT: man får <va för> [/] va försikti .
CHI: titit . %com: rep %syn: €A
CHI: jaˆa . 1;9

Denna typ av ettordsupprepningar upphör nästan helt efter 1;10. De få upprepningar som förekommer från och med 1;11 är alla (med något undantag) klassificerade som huvudsatssträngar.

(95) MOT: ska dom kuddarna va där borta ?
CHI: ka # kudden va dä bortˆta .
%com: rep %syn: $FSVA M:ipre M:basf 2;1

[2]I de tidiga inspelningarna förekommer också några upprepningar av interjektioner som inte är medtagna i satsanalysen, varför det totala antalet upprepningar inte nödvändigtvis är summan av de båda strängtyperna.

Barns förmåga att imitera är en fascinerande egenskap som förmodligen alltid har intresserat såväl föräldrar som forskare. De språkliga upprepningarna inspirerade den behavioristiska synen på språkinlärning och t.ex. Freudenthal (1925), som nämndes i inledningen till denna avhandling redovisar ett eliciteringsexperiment av ettordsimitationer och beskriver fenomenet som *retningar* och *reaktioner* enligt den tidens terminologi. Även om man idag har avfärdat behaviorism som förklaring till barns språkinlärning liknar Hannas ettordsimitationer fram till 1;10 onekligen betingade reflexer. I exempel 94 ser det ut som om Hanna återger orden på ett kreativt sätt eftersom hon ändrar ordformen, men exempel som: MOT: ”den e rätt statisk.” CHI: ”tatisk.” (1;10) antyder att Hanna inte nödvändigtvis återger ord som finns i hennes lexikon. De få upprepningarna efter 1;10 ger inte samma intryck av reflexmässighet, utan i dessa kommer upprepningen med en liten fördröjning som ger intryck av att Hanna reflekerar över orden på ett annat sätt.

Barns förmåga eller oförmåga till verbala imitationer är förmodligen relaterat till deras verbala korttidsminne (Speidel, 1993) och därmed språkinlärning. Därför får jag säkert anledning att återkomma till Hannas spontana imitationer. Men för närvarande är jag bara intresserad av att redovisa upprepningarna, eftersom de många ettordsimitationerna fram till 1;10 påverkar både bilden av yttrandelängd och av strängtyper.

4.1.2 Yttrandelängd

I avsnitt 2.4.3 diskuterar jag MLU-måttet och hur en jämförelse av MLU mellan olika barn bygger på att man har definierat termerna *yttrande* och *morfem* eller *ord* på samma sätt. Vad jag kan utläsa av beskrivningar av både Stockholmskorpusen och Göteborgskorpusen ligger definitionerna och transkiptionerna i våra korpusar nära varandra. Innan jag reodvisar siffror över yttrandelängd vill jag påminna om vissa detaljer.

I Hannakorpusen är orden inte morfemindelade, så när jag beräknar MLU är det på grundval av antalet ord per yttrande, inte morfem, MLU(w(ord)). Vissa ljud med oklar status är transkriberade som ord (se sidan 21), och upprepade ord kan skrivas inom ett och samma yttrande (se exempel 16 på sidan 26). Yttrandet i exempel 96 ser alltså ut att vara 5 ord långt, men det är oklart vad yttrandet säger om Hannas språkförmåga.

(96) CHI: bi mang [<] ä , na:ngä , mangen .
mål: ? vagn ? , vagnen , vagnen 1;6

När jag nu använder MLU(w)-måttet är det inte för att relatera det till andra barn, utan för att kunna visa på progressionen av yttrandelängd i Hannas in-

spelade språk. Jag har redan gett siffror för MLU(w) i tabell 2.3. De siffrorna baserar sig på alla yttranden i alla inspelningar. Om man vill koppla yttrandelängd till språklig förmåga kan man dock behöva ta hänsyn till vad yttrandena består av. Återkopplingar t.ex. består i typfallet av ettordsyttranden och när barnets kommunikativa förmåga ökar, ökar också återkopplingarna vilket kan dra ner den genomsnittliga yttrandelängden (jfr MLU-profilen för Freja, Lange och Larsson 1977, s. 22, och Richthoff 2000, s. 25). Hannas återkopplingsmorfem var i vissa inspelningar så många att de påverkade MLU-värdet [3]. Man kan instruera MLU-kommandot att räkna bort alla återkopplingsord och de yttranden som bara består av återkopplingsord och på så sätt få en bättre uppfattning om hur långa yttranden barnet producerar. Eftersom mitt uttalsnära transkriptionssätt inte skiljer svarsordet *ja* från det personliga pronoment *jag* innebär det dock att alla subjekts-”ja” också räknas bort, och att det nya MLU-värdet därför ändå blir missvisande.

Eftersom korpusen är taggad för typ av yttranden valde jag istället att räkna ut MLU(w) på grundval av alla klassificerade yttranden, dvs. alla yttranden som har fått en %syn-kod av något slag [4] (se tabell 4.2). Det betyder att yttranden som bara består av interjektionsfraser eller vokativfraser samt yttranden som är helt otolkbara är borträknade. Interjektionsfraser, vokativfraser, konjunktioner och otolkbara ord som ingår i ett klassificerat yttrande räknas dock in i yttrandelängden.

MLU(w) progressionen är tydlig i materialet och man kan urskilja 6 tydliga steg i utvecklingen. I den de kommande avsnitten undersöker jag vad de olika MLU-stegen skulle kunna relateras till. Om man jämför med MLU-värdena i tabell 2.3 kan man se att värdena är avsevärt högre i den nya beräkningen. Det beror bl.a. på att yttranden som bara består av återkopplingsord är mycket vanliga i vissa inspelningar. I 2;0 är till exempel 113 av de totalt 382 yttrandena (dvs. . 30 %) återkopplingsyttranden. I de kursiverade värdena i sista kolumnen är de yttranden som består av upprepningar borträknade (jag anger bara värden i de inspelningar där de skiljer sig). Hannas spontana upprepningar består som jag har visat till stor del av ettordsyttranden och är som flest under 1;9 och 1;10. MLU(w)-värdet utan upprepningar blir 2,6 i de båda inspelningarna och MLU-steget i 1;9 blir tydligare.

[3] Till återkopplingar räknar jag både det som Norrby (1996) kallar *uppbackningar* och svarsord, i linje med Plunkett och Strömqvist (1992).

[4] Sökströngen är combo +t%syn +s(€*+$*+¢*) +t*chi +d HaXyYYm.cha | mlu -t%mor.

Tabell 4.2: Yttrandelängd

Insp.	Tot. antal yttranden	%syn-yttranden	Ord	MLU(w)	
1;6	317	241	425	1,8	
1;7	337	212	450	2,1	
1;8	338	232	521	2,2	*2,3*
1;9	411	329	818	2,5	*2,6*
1;10	385	298	710	2,4	*2,6*
1;11	289	178	498	2,8	
2;0	382	246	886	3,6	
2;1	418	256	847	3,3	
2;2	358	296	1240	4,2	
2;3	311	196	756	3,9	
2;4	275	190	742	3,9	
2;5	287	173	699	4,0	*4,1*
2;6	296	198	845	4,3	
2;7	260	191	796	4,2	
2;8	335	252	1277	5,1	
2;10	251	187	838	4,5	
V2;0	381	316	1559	4,9	
V2;4	304	251	1388	5,5	

MLU(w) beräknat på enbart klassificerade yttranden, dvs. yttranden med en %syn-tier

Anm. V2;0 och V2;4 betyder vuxenspråk i 2;0 resp. 2;4

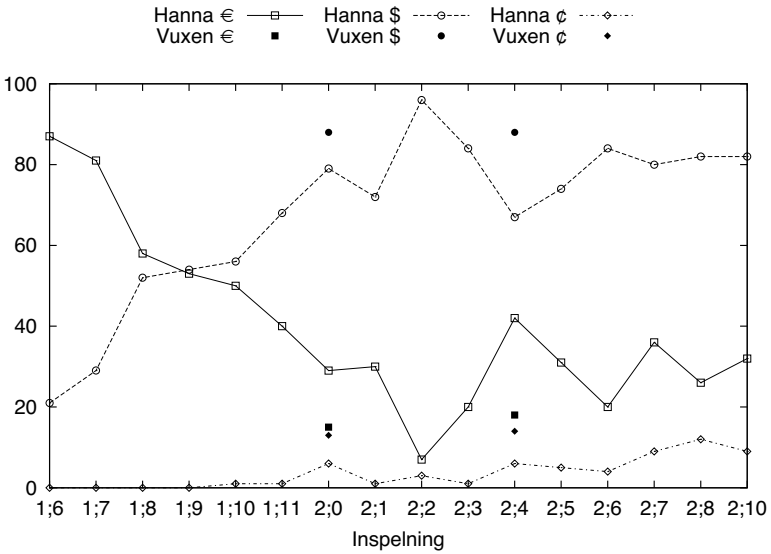
Tabell 4.3: Antal strängar av varje typ

Insp.	Antal €	Antal $	Antal ¢
1;6	211	48	0
1;7	172	61	0
1;8	139	115	0
1;9	186	164	0
1;10	169	144	2
1;11	73	119	1
2;0	71	195	15
2;1	77	184	3
2;2	22	284	8
2;3	39	164	2
2;4	79	128	11
2;5	55	128	9
2;6	39	167	7
2;7	69	152	18
2;8	66	207	30
2;10	61	153	17
V2;0	48	278	42
V2;4	45	222	34

4.1.3 Strängtyper

I tabell 4.3 redovisar jag hur många strängar av varje typ det finns i de olika inspelningarna. Ett yttrande kan innehålla flera strängar. Tabellen visar att antalet meningsfragment tydligt minskar mellan 1;10 och 1;11. De strängar som inleds med € består till allra största delen av enkla substantiv, nästan hälften av alla €-strängar utgörs av €N (665 av 1528). När meningsfragmenten minskar i 1;11 är det delvis en effekt av att Hanna slutar att upprepa ord från tidigare yttranden.

De spontana imitationerna fram till och med 1;10 skiljer sig som jag tidigare nämnde från Hannas andra yttranden ganska markant. När jag i figur 4.1 redovisar hur stor andel strängtyperna utgör av Hannas klassificerade yttranden, har jag därför valt att räkna bort alla imitaioner. I figuren visas hur många av varje strängtyp som finns per 100 yttranden. Siffrorna baserar sig alltså på alla klassificerade yttranden, dvs. de som redovisas i tabell 4.2 över yttrandelängd, men med alla yttranden som består av spontana imitationer borträknade. Varje yttrande kan innehålla en eller flera strängar av olika sorter, men de flesta yttranden består bara av antingen en €-sträng eller en $-sträng. Jämförelsesiffror för vuxenyttranden finns i figuren i punkterna 2;0 och 2;4 på x-axeln.



Figur 4.1: Antal av varje strängtyp per 100 klassificerade yttranden, med spontana imitationer borträknade

Figuren visar att förhållandet mellan meningsfragment och huvudsatser skiftar under perioden 1;8 till 1;11, om man räknar bort de spontana ettordsimitatonerna. Från och med 1;11 har jag alltså klassificerat fler av Hannas yttranden som huvudsatser. Linjerna som representerar meningsfragment och huvudsatser är i stort sett spegelbilder eftersom de flesta yttranden är klassificerade som antingen det ena eller det andra. ¢-strängarna (under- och samordnade verbfraser) ingår ofta, men långtifrån alltid, i $-strängar och påverkar inte den symmetrin så mycket.

Man kan se att Hannas andel meningsfragment under hela perioden ligger något högre än i vuxenspråket. Skillnaden mellan Hannas och de vuxnas meningsfragment måste undersökas grundligare. I 2;4 ser de många meningsfragmenten ut att delvis bero på att Hanna övar på eller leker med ord, t.ex. ”bäkon, beakon, bäkon [: bacon]” och ”domkyrkan, kyr(k)klockan, domklockan, dumklockan”. Men figur 4.1 visar framförallt att Hannas produktion av huvudsatser är större än hennes andel meningsfragment från och med inspelning 1;11.

Jag har i manualen beskrivit svårigheten med att klassificera de tidiga tvåordsyttranden som antingen $-strängar eller €-strängar. När jag nu i efterhand kan undersöka yttrandena samlat bekräftas min misstanke att klassificeringen inte passar för inspelningarna 1;6 till och med 1;8. Jag har klassificerat vissa verblösa yttranden som $-strängar. När jag undersöker dessa strängar kan jag se tre anledningar till detta: prosodi, referens och ordföljd.

Yttranden som *låter* som satser har fått inledaren $, trots att de är svårtolkade. I vissa finns ord som påminner om verb och som snarast verkar fungera som *filler*-element:

(97) CHI: ala kudden .
%com: jag vill ha kudden
%syn: $ XN
FAT: vill du också ha en kudde å sitta på ? 1;7

Dessa bedömningar kan jag också i efterhand acceptera. De skiljer sig från andra tvåordsyttranden i kraft av sin satsintonation.

De övriga två anledningarna är dock ytterst tveksamma. I avsnitt 3.2.2 försöker jag skilja mellan *benämningar* och *satser* genom att skilja ”dä” som som demonstrativt pronomen från ”dä” lokativt adverbial och huruvida nominalet pekar ut en predikationsbas eller fungerar som ett predikativ. När jag i efterhand jämför de yttranden som består av ”dä” + nominal (benämningar som har fått en €-inledare och satser som har fått en $-inledare) i de tre tidigaste inspelningarna kan jag inte acceptera uppdelningen. Den framstår som alltför godtycklig. På samma sätt har jag också klassificerat yttranden som har en ordföljd som passar in på predikativa satser med utelämnad kopula som $-strängar.

Men i efterhand kan jag konstatera att det finns en skillnad på ”den min” i 1;7 och ”mamma arg” i 2;1. I 2;1 *låter* det som en sats, men det gör det inte i 1;7.

När jag utarbetade manualen angav jag inte prosodiska egenskaper som ett kriterium för att skilja på satser och meningsfragment. Först i den samlade analysen i efterhand har jag insett att prosodin ofta har varit avgörande. De tveksamma fallen är dock inte så många. Det finns 3 tveksamma $-klassificeringar i 1;6, 15 i 1;7 och 2 i 1;8, vilket innebär att det främst är punkterna i 1;7 som ändrar sig i figur 4.1, så att ökningen av andel satssträngar blir i stort sett försumbar mellan 1;6 och 1;7 (och därmed också minskningen av andel meningsfragment).

4.2 Verb och verbmorfologi

Alla huvudsatser innehåller ett finit verb i svenska. För att ett barn ska kunna producera satser behöver hon lära sig verb, och för att satserna ska kunna bli målspråkligt finita måste hon lära sig de finita verbformerna. Att barn redan tidigt skiljer på finita och infinita former nämnde jag redan i kapitel 3 (s. 59). Framväxten av verbmorfologi har undersökts av bl.a. Plunkett och Strömqvist (1992), Hansson (1997) och Platzack (1992), och de kan visa att de första verbformerna liknar infinitivformen, att presens- och supinumformerna tenderar att komma sedan och att preteritumformerna kommer sist.

Christensen (under arbete) skiljer i sin undersökning av verbmorfologin mellan basformer, suffigerade former och icke-suffigerade former av verben (se avsnitt 3.6). Hennes två undersökta barn visar en kronologi där basformerna (t.ex. *bada*) kommer före de suffigerade infinitiverna (t.ex. *sitta*) och de icke-suffigerade preteritumformerna (t.ex. *sprang*) klart sist. Däremellan visar de båda barnen olika utvecklingsgångar. För båda barnen gäller dock att den icke-suffigerade presensformen (modala hjälpverb i presens och *är*) har ett mer plötsligt förlopp än de andra verbformerna som kommer mer gradvis. Christensens kronologier baserar sig på dagboksmaterial och utgör en bra jämförelse till de andra undersökningarna som baserar sig på sporadiska inspelningar där samtalet oftast handlar om den pågående situationen och man därför kan anta att verbformer som refererar till dåtid är underrepresenterade (jfr avsnitt 2.5).

Jag har taggat alla verb i inspelningarna fram till och med 2;2 för verbmorfologi enligt Christensens modell. Innan jag redovisar framväxten av verbmorfologi hos Hanna visar jag i tabell 4.4 att förhållandet mellan antal förekomster av verb och antalet verbtyper (lemman) varierar på ett oregelbundet sätt mellan inspelningarna. De olika verbtyperna är utsöka för hand utifrån ordlistor över strängar som innehåller taggarna F och V. I tabellen framgår också hur stor del

av det totala antalet ord verben utgör. I 1;8 kommer den första tydliga ökningen av andelen verb. Man kan notera att den inte innebär någon större skillnad i yttrandelängd, men däremot är det den ökningen som gör att andelen $-strängar ökar i 1;8 (i figur 4.1). Den andra verbökningen vid 2;0 sammanfaller dock med en ordentlig ökning av MLU(w) och andelen verb av det totala antalet ord ser från 2;0 ut att kunna jämföras med vuxenspråket. Man kan också se att skiftet mellan meningsfragment och huvudsatser mellan 1;10 och 1;11 som jag redovisade i förra avsnittet inte beror på att andelen verb blir större.

Tabell 4.4: Antal och andel verb fram till 2;2

Inspelning	1;6	1;7	1;8	1;9	1;10	1;11	2;0	2;1	2;2	V2;0
Antal verb (tokens)	46	43	101	145	129	111	265	241	416	438
Antal verb (types)	9	21	15	27	38	34	50	60	67	66
Token/type	5	2	7	5	3	3	5	4	6	7
Andel verb av alla ord	9%	7%	16%	16%	16%	18%	26%	23%	32%	27%

I tabell 4.5 redovisas Hannas verbformer. Det går inte att avgöra om basformerna föregår de suffigerade infinitiverna utan vid undersökningsperiodens början finns båda typerna med ungefär samma fördelning som i jämförelsesiffrorna i V2;0. Någon närmare kronologi över de andra verbformerna kan jag inte göra utan att analysera varje form med avseende på typer. Men i likhet med Christensens barn visar Hanna upp ett mer plötsligt förlopp när det gäller den icke-suffigerade presensformen. Det ser i tabellen ut som om de icke-suffigerade presensformerna kommer i två omgångar, först vid 1;9 (markerat med kursiv stil) och sedan vid 2;0 (markerat med kursiv fetstil). Övriga verbformer ser ut att etableras mer gradvis. I det följande ska jag ge en översiktlig beskrivning av de olika formerna, med tonvikt på det tidigaste förekomsterna i varje kategori.

Som framgår av tabell 4.5 dyker de första presensformerna upp vid 1;7, och de är av typen suffigerad presens (*bygger*, *läser*, *sitter* och *sätter*). Ytterligare ett par presensformer har identifierats i 1;8 (*bygger*). För dessa båda inspelningar gäller dock att både vokalkvaliten på ändelsevokalen och ordaccenten är mycket svårtolkade, och taggningen måste betraktas med viss skepsis. Ett par av presensformerna förekommer dessutom i verb-partikelkombination där accentskillnaden är upphävd eftersom betoningen ligger på partikeln (gäller ett par förekomster av *sitter fast*).

Tabell 4.5: Verbformer

Insp.	1 basf	2 infv	3 spre	4 ipre	5 Σpres	6 sptm	7 iptm	8 Σptum	9 impv	10 supm	11 stam	12 oanb	13 Alla
1;6	23	17	0	0	0	0	0	0	5	0	0	1	46
1;7	14	17	9	0	9	0	0	0	1	0	1	1	43
1;8	37	53	2	1	3	1	0	1	0	0	5	2	101
1;9	32	61	15	*21*	36	0	2	2	0	2	6	6	145
1;10	59	25	26	*6*	32	1	0	1	0	3	2	7	129
1;11	45	13	11	*28*	39	0	2	2	2	7	0	3	111
2:0	84	53	16	***99***	115	0	3	3	1	4	0	5	265
2;1	111	25	19	***69***	88	1	4	5	0	1	1	10	241
2;2	123	74	19	***169***	188	4	5	9	1	7	7	7	416
2;3	–	–	–	–	–	0	2	2	2	0	0	1	5
2;4	–	–	–	–	–	6	1	7	3	6	2	0	18
2;5	–	–	–	–	–	4	7	11	0	9	1	3	24
2;6	–	–	–	–	–	8	9	17	0	2	0	0	19
2;7	–	–	–	–	–	4	21	25	1	4	2	1	33
2;8	–	–	–	–	–	24	10	34	0	5	1	1	41
2;10	–	–	–	–	–	11	5	16	0	2	0	1	19
V2;0	148	66	31	153	184	2	19	21	8	7	0	4	438
V2;4	–	–	–	–	–	5	23	28	29	8	1	1	67

1= basform, 2=suffigerad infinitiv, 3=suffigerad presens, 4=icke-suffigerad presens, 6=suffigerad preteritum
7=icke-suffigerad preteritum, 9=imperativ, 10=supinum, 11=stamform, 12=oanalyserbar form

I 1;9 till och med 1;11 förekommer något fler presensformer. Från och med 1;9 blev det också gradvis lättare att identifiera presensformer med hjälp av vokalkvaliten på ändelsevokalen och den akuta accenten. I 1;9 kommer också de första icke-suffigerad presensformerna, dvs. några hjälpverb och presens av *vara*. Dessa tidiga yttranden är svåranalyserade och man bör komma ihåg min princip att ”tagga stort” när man tolkar dem. Förekomsterna består av ”(s)ka” (1 ggn), ”kan” (2ggr) och ”e” (18 ggr). Hjälpverben i 1;9 är inte så välartikulerade men förekommer i satser tillsammans med infinita verb[5]. Nio förekomster av ”e” är av slaget ”e de?” som Hanna använder omväxlande med ”dä?” som frågefras, och det är förstås mycket svårt att veta om frasen är segmenterad för Hanna. Det förekommer dock några andra fraser med ”e”, dels i kombinationen ”där e bil/röda/stor nalle”, dels i fraserna ”e sönder” och ”e trött” och ”månen e ute”.

I 1;10 finns också några förekomster av ”kan”, ”(s)ka” och ”e” som är något tydligare[6], och i 1;11 blir de både tydligare och fler. Dessutom tillkommer några ”måste” och ”vill”. De många förekomsterna av suffigerad presens i 1;10 beror inte på upprepad användning av något särskilt lexem, utan de 26 förekomsterna fördelar sig ganska jämnt över 12 typer. I 1;9 och 1;11 finns också några enstaka förekomster av preteritum av *vara*, t.ex.”dä va nalle” och ”här va sol”.

De icke-suffigerade presensformerna – presens av *vara* och hjälpverben *kan*, *ska*, *måste* och *vill* – förekommer alltså i 1;9 till 1;11. I 2;0 ökar de märkbart. Andelen presensformer i förhållande till totalt antal verb ökar redan något 1;11, men ökningen är ännu mer markerad vid 2;0, och det är den icke-suffigerade presensformen som bidrar till det. Man kan också anta att det är de icke-suffigerade presensformerna som bidrar till att andelen verb ökar i 2;0.

Supinumformerna är aldrig särskilt många i någon av inspelningarna, men de är inte heller så många i vuxenspråket och har alltså förmodligen med inspelningssituationen att göra. De allra första supinumformerna (två förekomster av ”bajsat”) finns i 1;9, dvs. den inspelning då också den suffigerade presensformen börjar dyka upp. I 1;10 finns ytterligare några supinumformer av två olika typer, och jag återger nedan passagen de förekommer i (supinumfromerna är fetstilta i exemplet).

(98) MOT: Pippi hitta en spunk .
CHI: punk . %com: imit %syn: €N
MOT: ja .

[5] Yttranden är ”Hanna kn nte söa den.”(Hanna kan inte köra den), ”kan du hjäpa lite ?” och ”ka sova ligga dä.”

[6] Det finns fyra icke-imiterade yttranden med modala hjälpverb: ”mamma kan bygga lejon” (2 ggr), ”de ka pappa bygga” och ”de kan (r)ita de”.

CHI:	golbacke [: skalbagge] dä .	%syn: €NA
MOT:	där e en skalbagge ja .	
CHI:	Hanna # hitta golp +...	%syn: $SVOX M:basf
CHI:	dä Hanna eˆute golbacke .	%syn: $ASXAN
CHI:	åssåå **fått** # golbacke ute .	%syn: $IVOA M:supm
MOT:	har du en skalbagge ute ?	
CHI:	ja .	
MOT:	har du fått en skalbagge ?	
CHI:	nn .	
CHI:	N ha osse # **köpit** kalbagge .	%syn: $SFIVO M:basf M:supm
MOT:	har du köpt en skalbagge ?	
CHI:	Karin ha osse **söpit** galbagge .	%syn: $SFIVO M:basf M:supm
FAT:	mm .	
MOT:	Karin ?	
FAT:	var köpte ni dom nånstans ?	
CHI:	Hanna åssåå äˆute galbagge .	%syn: $SIAN
FAT:	fick du betala pengar för den ?	
CHI:	nä , galbagge gˆute .	%syn: $NXA
MOT:	e de en skalbagge ute ?	
FAT:	mm .	
CHI:	ja .	1;10

Det Hanna verkar vilja berätta är att hon och hennes kompis Karin också har hittat (eller sett) en skalbagge ute på gården precis som Pippi i boken. Om det är en riktig tolkning är det intressant att notera att hon inte verkar kunna producera den supinumform hon behöver. Hon nämner verbet *hitta* i inledningen till passagen, men tar till supinumformerna ”fått” och ”köpit”. Passagen avslutas med att MOT frågar om det är en skalbagge ute. Hannas ja-svar låter här något tveksamt, men hon nöjer sig med den tolkningen. Formerna ”fått”, ”köpt”, ”sett” och ”hittat” betecknar vanliga avslutade handlingar med ett tydligt resultat, och alla förekommer förmodligen ganska ofta i barnets input (de finns i alla fall i korpusen). Frågan uppstår om Hannas supinumformer är kreativt producerade eller helhetsinlärda. Att hon inte verkar kunna producera rätt form av *hitta* tyder på helhetsinlärning, men formen ”köpit”, en övergeneralisering som inte förekommer i input, indikerar att supinumformen är kreativt producerad. Att supinumformen genereras kreativt tyder också formen ”flygat” i 1;11 på.

Inte heller preteritumformerna är så vanliga i Hannas språk, men till skillnad från supinum finns här en progression. Det ser ut som om preteritum blir ungefär lika vanligt som i vuxenspråket vid 2;6. De allra första sporadiska formerna av icke-suffigerad preteritum består av några ”va” (i 1;9 och 1;11, se ovan) och det förekommer ett par suffigerade preteritumformer i 1;8 och 1;10:

”tappade den” och ”hoppade ner”. Den första inspelning som innehåller flera olika typer är 2;2.

Stamformerna och de oanalyserbara formerna fördelar sig ganska ojämnt över tiden, men har en topp mellan 1;8 och 1;10 (räknat som andel av det totala antalet verb). Stamformerna förekommer ofta tillsammans med en annan form av samma verb, t.ex.”bygg bugga bil” eller ”bäj mamma bäjja [: bära]” (1;7) men också som huvudverb i en sats, ”du få sov fädit” (1;9). De oanalyserbara formerna består bland annat av verb som är tvetydiga mellan stamformer och imperativer eller preteritumformer, t.ex.”läs, fös vi läsa denna åssåˆ”, ”dä kom mas (?)” (1;10). Suffigerade preteritumformer kan också vara svåra att skilja från i det här fallet presens + ”de”-objekt som i ”ti kör den [=? körde] på din tom [: torn]” (2;1). Detta problem uppstår dock oftare för verben i första konjugation som då har taggats som basformer: ”tappa de å” (1;9). Hälften av de många förekomsterna av oanalyserbara former i 2;1 beror på ett och samma lexem ”gå upp”, men här finns också något som kan vara en supinumform i ”ja knytihop” = ”jag har knutit ihop”.

Imperativformer, till sist, finns det ganska få av i hela barnmaterialet. I 1;6 finns 5 förekomster av ”kom” (i tre olika episoder) som är tydligt imperativa. Hanna använder i 1;6 ”kom” omväxlande med ”komma” för att uppmana mamma eller pappa att komma. Också i 1;7 finns en förekomst av ”kom”, då som upprepning. De två imperativformerna i 1;11, ”kom” och ”lägg”, är något tveksamma. Den första kan vara en upprepning och den andra är inte så tydligt uppmanande och kan vara en stamform. Också imperativformen i 2;0 är en upprepning, så den första tydliga imperativen, förutom ”kom”, finns i 2;2. Imperativformerna därefter är otvetydiga och förekommer i direktiva satser. Imperativformerna i den första hälften av korpusen utgörs alltså av ”kom”, en form som jag, utan att räkna på det, kan konstatera är oerhört vanlig i barnriktat vuxenspråk.

Eftersom mitt huvudintresse är etableringen av den finita satsen är jag i detta sammanhang mest intresserad av när man kan säga att Hanna har förstått hur man uttrycker finithet genom finita verbformer. Översikten av verbformer visar att det inte finns så mycket som tyder på att morfologin är genomskådad före 1;9. Det finns visserligen former som har tolkats som presensformer, men dessa är så otydliga både i prosodi och vokalkvalitet så att den tolkningen kan ifrågasättas. Man kan inte heller veta om infinitivformerna verkligen är infinitiver. Det enda vi kan säga är att Hanna före 1;9 nästan alltid använder verbformer som slutar på -a.

I 1;6 och 1;7 finns inte heller några satsadverbial som kanske kunde säga något om verbens finithet utifrån deras placering. I 1;8 finns det två typer av satser med ”inte”. Verben utgörs båda av basformer, men den ena kan tolkas

som finit, medan den andra närmast bör betraktas som infinit.

(99) CHI: passa inte .
%syn: $VI M:basf 1;8
(7 yttranden under en längre passage där Hanna bygger med lego)

(100) CHI: nej , titta inte katten !
%com: betyder förmodligen "vi ska inte titta på katten"
%syn: $VIN M:basf 1;8

Det ser alltså ut som om Hanna placerar ett infinit verb före negationen i exempel 100 vilket stärker intrycket av att vi inte kan säga något om hennes kunskaper om vare sig morfologi eller finithet fram till och med 1;8.

Andelen stamformer och oanalyserbara former är något högre i 1;8 till och med 1;10, något som jag tolkar som att Hanna redan i 1;8 har lagt märke till, och kanske försöker reproducera, den differentierade morfologin. Från 1;9 börjar Hanna att visa mer pålitliga tecken på att hon gör skillnad på olika verbformer. I 1;9 är skillnaden mellan suffigerad infinitiv och presens tydligare, om än fortfarande inte helt otvetydig. I 1;9 kommer också något som ser ut som föregångare till hjälpverb, samt några supinumformer. Fram till och med 1;11 blir de böjda verbformerna fler och hjälpverben blir mer artikulerade. Från 1;9 finns det också fler satsadverbial som kan säga något om verbens placering i satsen, och från 1;9 och framåt är de finita verben alltid placerade före negationen och de infinita efter, och Hanna visar alltså upp samma mönster som andra undersökta svenska barn. Vid 2;0 ökar antalet icke-suffigerade presensformer ordentligt och andelen finita verbformer kan jämföras med vuxennivåer. Jag drar slutsatsen att Hanna vid 2;0 vet vilka verb som är finita, även om hon bara använder presens, och att hon från 1;6 till 1;8 varierar verbformer ganska slumpmässigt. Perioden från 1;9 till och med 1;11 är en övergångsperiod som kommer att behöva undersökas grundligare.

4.3 Huvudsatser

4.3.1 Subjekt och verb

En svensk huvudsats består minimalt av ett subjekt och ett finit verb, utom i direktiva satser med ett verb i imperativ form där subjektet kan saknas. I barnspråket har jag tolkat yttranden som satser trots att de saknar antingen verb (finit eller infinit) eller subjekt (för diskussion av hur jag använder taggen $ se avsnitt 3.4.2). I tabell 4.6 nedan ger jag en översikt över alla $-strängar med hänsyn till andelen subjekt och andelen finita och infinita verb.

Tabell 4.6: Huvudsatser

1	2	3	4	5	6	7	8	9
Insp.	Antal $	Andel	Andel finita verb		$-str. m. enbart infinit verb			Andel
		m. subj	synt. F	morf. fin.	andel	ant. infv	supm	u. verb
1;6	48	17%	(10%)	(10%)	83%	16	0	7%
1;7	61	33%	(16%)	(16%)	52%	16	0	32%
1;8	115	24%	(4%)	(3%)	80%	50	0	16%
1;9	164	32%	24%	23%	57%	54	2	19%
1;10	144	42%	25%	23%	52%	21	1	23%
1;11	119	51%	39%	36%	37%	3	5	24%
2;0	195	80%	77%	61%	10%	4	2	13%
2;1	184	73%	70%	51%	21%	3	0	9%
2;2	284	85%	81%	70%	10%	4	1	9%
2;3	164	79%	82%		5%	1	0	13%
2;4	128	84%	73%		7%	0	2	20%
2;5	128	89%	84%		7%	0	3	9%
2;6	167	90%	92%		5%	2	1	2%
2;7	152	89%	85%		8%	1	2	7%
2;8	207	91%	94%		2%	0	2	3%
2;10	153	84%	86%		10%	3	0	4%
V2;0	278	93%	98%	77%	2%			0%
V2;4	222	85%	100%		0%			0%

2=Antal satssträngar, 3=andel satssträngar som innehåller taggen S, 4=andel satssträngar som innehåller taggen F, 5=andel satssträngar som innehåller ett morfologiskt finit verb, 6=andel satssträngar som innehåller ett V med inget F, 7=antal morfologiskt säkra rotinfiniver, 8=antal säkra rotsupinum, 9=andel satssträngar som varken innehåller F eller V
Anm. Satssträngarna fördelar sig på kolumnerna 4, 6 och 9.

Andelen subjekt (kolumn 3) ökar i perioden från 1;6 till 1;11, och mellan 1;11 och 2;0 ser det ut som ett språng i utvecklingen som gör att andelen subjekt hamnar nära jämförelsesiffrorna i vuxenspråket. Andelen subjekt ser sedan ut att öka ytterligare något i 2;5. Att andelen subjekt inte uppgår till 100 % i Hannas språk beror inte främst på direktiva satser utan subjekt. I tabell 4.5 över verbformer framgår att Hannas imperativformer är få. De beror istället på annan målspråklig subjektslöshet som fundamentsellips och huvudsatsekvivalenter som svar på frågor. I 2;10 finns många elliptiska svar på frågor vilket delvis förklarar den förhållandevis låga andelen subjekt. I V2;4 är de direktiva satserna däremot många vilket är den största orsaken till att bara 85% av satserna innehåller subjekt.

Subjekt i talspråk består ofta av pronomen, och man kan misstänka att Hannas subjektsökning i 2;0 beror på att hon har lärt sig att använda pronomen. I tabell 4.7 redovisas en hastig sortering av alla subjekt fram till 2;0 och den visar att Hanna ser ut att ha skiftat till pronomen redan i 1;11. I 2;0 ökar så såväl andelen pronomensubjekt som andelen satser med subjekt till vuxenlika nivåer. De led som har klassificerats som subjekt fram till och med 1;9 måste analyseras närmare. Problemet med att skilja benämningar från satser spelar in här, samt principen ”tagga stort” som har lett till att ”de” i frasen ”e de?” har taggats som subjekt. Från 1;10 finns det dock få sådana tveksamma subjektsklassificeringar.

Tabell 4.7: Andel pronomen som subjekt

	1;6	1;7	1;8	1;9	1;10	1;11	2;0	V2;0
Antal subjekt	8	20	28	53	61	61	156	258
% pronomen	(0%)	(20%)	(4%)	(38%)	11%	70%	88%	84%

De verbformer som i de tre första inspelningarna har klassificerats som finita har visat sig vara ytterst osäkra och jag markerar detta med parenteser runt andelen finita verb i kolumn 4 i tabell 4.6. I 1;6 finns 5 förekomster av imperativformen ”kom”, och de är taggade som F-verb, men enligt resonemanget i avsnitt 4.2 är de förmodligen inte produktiva.

Det har diskuterats om den imperativa satsen är en finit sats. Traditionellt har man betraktat imperativen som en finit verbform, men Platzack och Rosengren (1998) föreslår att imperativer är icke-finita i bl.a. svenska, baserat på den obefintliga morfologin. Bohnacker (1999) diskuterar också imperativens finithetsstatus och hon landar i att den isländska imperativen bör betraktas som finit, eftersom den liksom andra finita verb måste flytta förbi negation och even-

tuella subjekt, medan den engelska imperativen varken kan betraktas som finit eller infinit. Enligt det resonemanget bör också den svenska imperativen betraktas som finit.

Det går en tydlig gräns mellan 1;11 och 2;0 också när det gäller andelen finita verb i $-strängar (kolumn 4 och 5, där siffrorna i kolumn 5 alltså utgör en delmängd av siffrorna i kolumn 4). Ökningen av finita verb ser ut att ske mer språngvis än ökningen av subjekt. De morfologiskt säkra finita formerna utgörs av de suffigerade och icke-suffigerade presensformerna i tabell 4.5 över verbformer, samt några enstaka preteritum- och imperativformer. I 2;0 är det endast den icke-suffigerade presensformer som ökar märkbart. Det innebär att ökningen av andelen $-strängar som innehåller ett morfologiskt finit verb (kolumn 5) från 36% till 61% beror på att hjälpverb och *är* blir vanliga. Lägg märke till att hjälpverben *få* och *ha* inte är medräknade här eftersom de faller in under den morfologiska kategorin M:basf (basform). De klassificeras dock som syntaktiskt finita om de följs av ett huvudverb. Från 2;0 klassificeras alla basformer i finita kontexter med F (se 3.5.2), vilket är anledningen till att andelen $-strängar med syntaktiskt finita verb (kolumn 4) ökar mer (från 39% till 77%). När Hannas finita satser i 2;0 ökar markant beror det alltså på att hon har etablerat användningen av hjälpverb och finit form av *vara*. Det beror inte på att hon har lärt sig att tempusböja verb; det kan hon redan tidigare.

$-strängar med enbart infinit verb redovisar jag i tre kolumner. I kolumn 6 visar jag hur stor andel av $-strängarna som har en V-tagg, men saknar en F-tagg. Detta är ett resultat av min bedömning av verbet under taggningen (se avsnitt 3.5.3). V-taggen kan märka ut både infinitiver och supiner. Strängarna kan också innehålla oklassificerbara element (X). I kolumn 7 finns antalet strängar som saknar finit verb och som innehåller en morfologiskt säker infinitivform (M:infv). Den visar att det finns säkra förekomster av rotinfinitiver och att de avsevärt minskar i mängd från 1;11. Kolumn 8 (M:supm) visar alla supinumverb i $-strängar som inte föregås av något hjälpverb. Jag har bara tittat på huvudsatser här och alltså inte räknat med de sammanlagt 4 perfektkonstruktioner utan hjälpverb som jag har klassificerat som under- eller samordnade (¢-strängar). Satser (eller huvudsatsekvivalenter) med enbart infinit verb förekommer som sagt också i målspråket, t.ex.i svar på frågor. Den typen av infinita satser förekommer i många av inspelningarna och måste undersökas närmare. Ett stickprov visar att 7 av de 15 infinita satserna i 2;10 är av den typen.

I den sista kolumnen redovisas $-strängar som varken har F- eller V-tagg. I de flesta satser ser det ut att saknas en finit form av *vara*. Strängarna kan också innehålla oklassificerbara element (X) som skulle kunna vara verb. Dessa satser gör att andelen finita verb i satssträngarna inte riktigt kommer upp i vuxennivåer hos Hanna.

Inspelning 1;11 utgör en tydlig övergång mellan Hannas tidiga språk och hennes finita system. I tabell 4.4 i avsnitt 4.2 ser vi att antalet och andelen verb inte visar på någon tydlig progression jämfört med föregående inspelning. Inte heller Hannas MLU(w) visar på någon progression från 1;10 till 1;11 (tabell 4.2). Men andelen finita verb har ökat märkbart i 1;11, och andelen satser med enbart infinit verb har sjunkit, både i den syntaktiska analysen och i den morfologiska. Andelen subjekt är märkbart högre och Hanna har börjat använda pronomen. Skiftet sammanfaller med att Hanna slutar att upprepa, och med skiftet från €-strängar till $-strängar. I 2;0 är den finita satsen etablerad (om än inte fullständigt vuxenlik).

4.3.2 Interrogativa huvudsatser

Jag har i min satsanalys inte delat upp yttranden efter språkhandling. Däremot har jag markerat de interrogativa satserna med ett Q sist i strängen där jag har avläst prosodin som frågande. I tabell 4.8 redovisar jag hur många av de klassificerade yttrandena låter som frågor, dvs. har fått ett Q sist i strängen. Observera dock att jag inte har undersökt språkhandlingar och om satserna verkligen fungerar som frågor. Mitt spontana intryck är att många av de frågeformade satserna med frågeprosodi inte har som främsta syfte att efterfråga information.

Den rogativa satsen (ja-/nej-frågan) har det finita verbet på första plats, men jag har inte krävt att det ska finnas ett utsatt subjekt. I rogativa frågor är subjekten svåra att urskilja, även i vuxenspråket. I de kvesitiva frågorna kan subjekt och frågeord sammanfalla (se 3.5.9), så i de strängarna är en S-tagg inte heller nödvändig. I själva verket är det mycket ovanligt med kvesitiva frågor som saknar S i materialet, dvs. det är ovanligt att v-frågeord utgör subjekt vilket också Santelmann (2003) och Waldmann (2008) kan konstatera. De kvesitiva frågorna innehåller typiskt en kopula som också ofta utelämnas (eller är svår att höra) i barnspråk, så de kvesitiva strängarna behöver inte innehålla ett finit verb (t.ex. ”Va svarta tåget?” $WSQ (2;0)). De övriga frågorna består mest av underställande frågor med verbfras i infinitiv, med eller utan subjekt: ”läsa den?”, ”pappa bygga den?”. För fullständighetens skull finns också €-strängarna med frågeprosodi med, meningsfragment som ”den?” och ”vilken då?”.

Alla kvesitiva frågor fram till och med 1;11 innehåller (eller borde ha innehållit) presens av *vara*. Alla utom två består av ”e de?”-frasen, och de två som saknar finit verb innehåller frågeord: ”vae lammen?” och ”ha katten tans?” (var är katten någonstans). Den enda rogativa satsen är visserligen syntaktiskt välformad, ”kan du hjälpa lite?”, men den har en intonation som avviker något från den typiska frågeintonationen. När hjälpverben blir vanligare i 2;0 kom-

Tabell 4.8: Frågor

Insp.	Rogativa	Kvesitiva		Övriga	Σ frågor	Andel	=€...Q
	$F...Q	$Q	$W			av $	
1;6	0	0	0	0	0	0%	0
1;7	0	1	0	1	2	3%	0
1;8	0	0	1	4	5	4%	3
1;9	1	9	0	2	12	7%	7
1;10	0	0	0	11	11	8%	4
1;11	0	4	1	7	12	10%	2
2;0	28	2	10	6	46	24%	4
2;1	7	1	7	2	17	9%	0
2;2	56	0	9	3	68	24%	1
2;3	12	1	9	4	26	16%	0
2;4	10	0	1	1	12	9%	0
2;5	4	0	4	2	10	8%	4
2;6	11	0	5	1	17	10%	0
2;7	2	0	6	2	10	7%	0
2;8	5	2	12	2	21	10%	2
2;10	8	0	2	2	12	8%	0
V2;0	98	0	36	3	137	49%	5
V2;4	56	0	17	3	76	34%	11

mer också de rogativa frågorna med besked. Alla finita verb i de rogativa frågorna i 2;0 utgörs av icke-suffigerad presens, och 25 av de 28 strängarna har subjekt som alla består av pronomen (de övriga saknar subjekt). En jämförelse med vuxeninput i samma inspelning visar att bara lite mer än hälften av de rogativa frågorna konstrueras med hjälpverb. Det ser ut som om Hanna lär sig den rogativa satstypen med hjälpverb först.

(101) CHI: ka den sitta äta da ?
%syn: $FSV+VAQ M:ipre M:infv M:infv 2;0

(102) CHI: kan du hjälpa den ?
%syn: $FSVOQ M:ipre M:infv 2;0

(103) CHI: kan du öppna denna ?
%syn: $FSVOQ M:ipre M:basf 2;0

(104) CHI: få han sätta den da ?
%syn: $FSVOAQ M:basf M:infv 2;0

Andelen frågor är som störst i 2;0 och 2;2 där de rogativa frågorna är överrepresenterade. Det finns inget i inspelningssituationerna i 2;0 och 2;2 som förklarar

varför Hanna frågar så mycket, utan inspelningarna äger rum i likadana leksituationer som alla inspelningar fram till 2;3. Många av frågorna i inspelningarna uttrycker inte heller språkhandlingen *fråga*, utan är snarare ett sätt för Hanna att beskriva vad hon gör ("(S)ka den sitta äta dä? (S)ka den sitta äta pallen? (S)ka den sitta dä?" säger Hanna medan hon försöker att sätta pappadockan på en pall vid dockornas matbord). Frågeprosodin överanvänds också i dessa inspelningar, så att också satser som verkar ha en deklarativ funktion kan få en något frågande prosodi.

4.3.3 Deklarativa huvudsatser

Eftersom jag inte har märkt ut direktiva satser kan de inte skiljas från deklarativa på något enkelt sätt (expressiva satser och andra som har bisatsform är klassificerade som ¢-strängar). Men imperativer är, som jag tidigare har redovisat, få i Hannas språk och jag har därför inte lagt ner tid på att skilja ut dem från de deklarativa satserna. Under denna rubrik redovisar jag alltså alla $-strängar som innehåller ett finit verb, men med rogativa och kvesitiva frågor borträknade. Frågor som har vanlig påståendeform är dock medräknade samt alltså direktiva satser med verb i imperativ.

4.3.3.1 V1-deklarativer

De flesta svenska deklarativa huvudsatser har ett fundament, där i stort sett valfritt satsled kan placeras, och ett finit verb på andra plats. Ganska många satser har dock det finita verbet på första plats, antingen genom fundamentellips av ett obligatoriskt satsled (främst subjekt eller objekt, men också ellips av bundna adverbial) eller genom så kallad narrativ inversion, som kan betraktas som fundamentsellips av ett konjunktionellt adverbial (Mörnsjö, 2002). Dessa satser bedömer Loman och Jörgensen (1971, s. 54) som "felkonstruerade meningar", även om de markerar dem för att kunna beräkna dem separat. V1-deklarativerna har sedan dess uppmärksammats i en rad undersökningar och i SAG anges att deklarativa huvudsatser utan fundament förekommer i olika kontexter, dock sällan i neutralt eller formellt skriftspråk (SAG IV, s. 692). Håkansson (1994) undersöker V1-deklarativer med subjekt i små L1-barns språk. Hon konstaterar att de är mycket vanligare än i vuxenspråk och att de är nästan helt frånvarande i L2-barns språk. Josefsson (2003a) konstaterar också att satser med finit verb–subjekt–(XP) (inklusive rogativa frågor) är vanliga i språket hos de tre L1-barn hon undersöker. Josefsson undersöker också V1-deklarativer som beror på fundamentsellips av subjekt och kan konstatera att de är mycket vanliga i det tidiga barnspråket och minskar med tiden.

I tabell 4.9 redovisar jag den deklarativa satstypen med hänsyn till om det finita verbet är på första eller andra plats. Jag skiljer inte på olika sorter av V1-deklarativer och subjekt behöver inte finnas i strängarna. De 5 V1-satserna i 1;6 har verbet i imperativ form, men i övrigt är de direktiva satserna försumbara i Hannas inspelade språk. I de tidigaste inspelningarna finns många ettordssatser med finit verb (1;6 till 1;8) som också betraktas som V1-satser. De finita deklarativa satserna är så få i dessa tre inspelningar att jag inte har satt ut några procentsiffror. I V2;0 och V2;4 är de direktiva satserna så många att resultaten blir missvisande. Jag har därför räknat bort de direktiva satserna för att kunna ge ett rimligt jämförelsevärde över V1-deklarativer i vuxenspråket (de kursiverade siffrorna). Att döma av det analyserade vuxenspråket i Hannakorpusen är 10–15% av de deklarativa satserna fundamentslösa i målspråket (11 resp. 13% i tabellen).

Tabell 4.9: Deklarativa satser

Insp.	Dekl med F	V1	Andel	V2	Andel
1;6	5	5		0	
1;7	10	5		5	
1;8	5	3		2	
1;9	29	10	34%	17	59%
1;10	36	13	36%	22	61%
1;11	42	5	12%	33	79%
2;0	113	26	23%	83	73%
2;1	113	21	19%	91	81%
2;2	172	20	12%	150	87%
2;3	115	20	17%	93	81%
2;4	84	11	13%	73	87%
2;5	102	6	6%	94	92%
2;6	139	7	5%	128	92%
2;7	122	10	8%	111	91%
2;8	175	19	11%	154	88%
2;10	123	8	7%	113	92%
V2;0	138	22	16%	107	78%
V2;0	*130*	*14*	*11%*	107	*82%*
V2;4	148	44	30%	103	70%
V2;4	*119*	*15*	*13%*	103	*86%*

Tabellen visar att Hanna vid 2;1 kommer ner i en andel V1-deklarativer som ligger nära målspråksnivån. I 2;0 där den finita satsen blir vanlig är andelen V1-deklarativer fortfarande förhållandevis hög (23%). Anledningen till att summan av andelen V1-satser plus V2-satser inte alltid blir 100% är att det finns några satser med det finita verbet på senare plats i satsen, både målspråk-

liga och icke-målspråkliga. Dessa är dock få och jag redovisar dem inte i detta sammanhang. I 1;11 där de deklarativa satserna med finit verb inte är så många utgör dock de fyra V3-satserna 10% av alla satser.

Många av Hannas V1-deklarativer består av satser som saknar subjekt. De kan jämföras med den målspråkliga processen att stryka ett subjekt från fundamentet, men med tanke på att Hanna fram till 2;0 inte sätter ut så många subjekt över huvud taget kan man inte veta om Hannas tidiga satser kan betraktas på samma sätt. Några av dem återges nedan:

(105) CHI: sitte vast [: fast] .
%syn: $FL M:spre 1;7

(106) CHI: bygge husen .
%syn: $FO M:spre 1;8

(107) CHI: oj , &bjum # e synnö [: sönder]!
%syn: $FP M:ipre 1;9

(108) CHI: äte bulle .
%syn: $FO M:spre 1;10

(109) CHI: måste laga den .
%syn: $FVO M:ipre M:basf 1;11

Bara tre V1-deklarativer fram till 2;0 innehåller subjekt och de redovisas nedan. Notera att exempel 110 innehåller en tveksam verbform och exempel 112 ett ganska otydligt pronomensubjekt.

(110) CHI: läse [=? läsa] mamma den .
%syn: $FSO M:oanb 1;8

(111) CHI: titta den -_ # sitte pappe lada [: låda] .
%syn: $VN M:basf $FSA M:spre PP:noll 1;10

(112) CHI: kan -, # a hämta fam elefanten .
%syn: $FSVLO M:ipre M:basf 1;11

4.3.3.2 V2-deklarativer

I den prototypiska satsen står det finita verbet på andra plats och (nästan) valfritt satsled i fundamentet. Enligt Jörgensen (1976) är fördelningen av satsled i fundamentet ganska stabil över de olika materialgrupper han undersöker (med undantag av nyhets- och telegramläsning), med ca 60 % subjekt i fundamentet, runt 10 % objekt och runt 20 % adverbial. Santelmann (1995) och Josefsson (2003a) har visat att andra led än subjektet förekommer i små barns satsfundament tidigt, även om Josefsson visar att icke-subjektsinledda deklarativer

kommer lite efter de subjektsinledda hos barnen. Waldmann (2008) kan också i en detaljerad genomgång visa att hans undersökta barn tidigt har fundament som består av olika sorters satsled. Håkansson och Nettelbladt (1993) konstaterar att inverterade deklarativa huvudsatser finns tidigt i små barns första språk, men att de är sena i andraspråket och hos SLI-barn. [7]

Hannas fundament redovisas i tabell 4.10. Satsdelarna i fundamentet kan bestå av de analyserade satsleden A–adverbial, I–satsadverbial, O–objekt och P–predikativ, där I och P förekommer marginellt, samt oklassificerbar satsdel, X, och obestämbart nominal, N. I tabellen redovisar jag siffror för subjekts-, adverbials- och objektsfundament. Inte heller här ger jag några procenttal för de tre första inspelningarna eftersom antalet satser är så litet.

Tabell 4.10: Satsled i fundamentet

Insp.	Dekl V2	$SF	Andel	$AF	Andel	$OF	Andel
1;6	0	0		0		0	
1;7	5	2		2		0	
1;8	2	1		1		0	
1;9	18	6	33%	6	33%	0	0%
1;10	22	14	64%	7	32%	0	0%
1;11	33	16	48%	7	21%	1	3%
2;0	84	57	68%	18	21%	3	4%
2;1	91	54	59%	16	18%	11	12%
2;2	150	89	59%	42	28%	17	11%
2;3	93	59	63%	23	25%	6	6%
2;4	73	56	77%	7	10%	8	11%
2;5	94	63	67%	10	11%	19	20%
2;6	128	92	72%	16	13%	16	13%
2;7	111	76	68%	31	28%	3	3%
2;8	154	84	55%	41	27%	26	17%
2;10	113	72	64%	26	23%	8	7%
V2;0	107	59	55%	42	39%	5	5%
V2;4	103	65	63%	27	26%	8	8%

Som framgår av tabell 4.10 finns det inte så många deklarativa huvudsatser med fundament och finit verb i de tidiga inspelningarna, men de få som finns fördelar sig redan från början på subjektsinledda och icke-subjektsinledda satser. Hannas genomsnittsliga värden i undersökningsperioden skiljer sig inte mycket från Jörgensens siffror, även om det finns en lite större andel subjektsfundament överlag. I vissa inspelningar kan man dock se avvikande värden.

[7] De visar dock att den första inspelningen av det ena L1-barnet (Karin 1;11) inte innehöll några inverterade deklarativer med fundament, utan de inverterade satserna (35% av alla) var verbinitiala.

Objekt i fundamentet är helt frånvarande fram till 1;11 (och det enda exemplet i 1;11 är något tveksamt). I 2;8 har hon bara 55 % subjekt i fundamentet, och 17 % objekt av, som vi såg, varierat slag. Frågan är hur produktiva de olika ordföljdsmönstren är. Den kanoniska ordföljden subjekt–finit verb är kanske inte så ifrågasatt när det gäller produktivitet, men när det gäller den omvända ordföljden finns misstanken att barnen från början bara använder en särskild sorts konstruktion som inte generaliseras (jfr Waldmann, 2008). Nedan går jag igenom Hannas deklarativer för att försöka svara på om de är kreativa yttranden eller en uppsättning fasta fraser.

Fram till 1;8 är som sagt verbmorfologin skakig och man kan inte vara säker på att verben verkligen är finita, men de satser som har taggats som V2-deklarativer redovisas nedan.

(113) CHI: den sitte dä .
%syn: $SFA M:spre 1;7

(114) CHI: da buke [: bygger] -, <hämta den>[>] .
%syn: $AF M:spre $VO M:basf 1;7

(115) CHI: ja buggö dä .
%syn: $SFA M:spre 1;7

(116) CHI: tumme läsä .
%syn: $XF M:spre 1;7

(117) CHI: nu buggö den .
%syn: $AFO M:spre 1;7

(118) CHI: Hanna bygge # den .
%syn: $SFN M:spre 1;8

(119) CHI: däj [!] e biten -_ mä .
%syn: $AFS M:ipre €X 1;8

Från 1;9 finns det något fler V2-deklarativer och jag går först igenom de inverterade satserna, och ger sedan exempel på några subjektsinledda.

Inverterade deklarativa huvudsatser I talspråk består fundamentet i inverterade deklarativa huvudsatser mest av lätta led vilket illustreras av de 79 inverterade satser som finns i vuxeninputen. De innehåller i fallande ordning *nu*, *så*, *där*, *här*, *de*, *då*, *den*, *sen*, *dom*, *såna* och *annars*. Endast ett fundament innehåller en utbyggd fras, en konditionalsats (ytterligare en konditionalsats förkommer i ett annex följt av ett adjunktionellt *så*, men FV3-strängar är inte medräknade här). Satserna fördelar sig ganska jämnt på satser med finit form av *vara*, finit hjälpverb och finit huvudverb.

Hannas inverterade deklarativer i 1;9 återges nedan. Alla, utom exempel 121 som är en imitation, inleds med *här* eller *där*. Fram till 1;9 har de flesta inverterade satser finit form av *vara*, men det förekommer också ett par andra. Ordet ”se” i 126 har jag tolkat som en hjälpverbsfiller, dvs. en hjälpverbsföregångare. Det vanligaste mönstret är alltså ”där är...”, vilket kan jämföras med Tea (Waldmann, 2008) vars inverterade deklarativer under en längre tid uteslutande består av satser som inleds ”där/här är...”, medan övriga icke-subjektsinledda deklarativer bildar v3-satser.

(120) CHI: hä sitˆsitˆtä [= accent 1] .
%syn: $AF M:spre 1;9

(121) CHI: van [!] e de .
%com: imit
%syn: $PFS M:ipre 1;9

(122) CHI: däj e bil .
%syn: $AFN M:ipre 1;9

(123) CHI: dä e jöta , tack .
%syn: $AFP M:ipre 1;9

(124) CHI: där e to [: stor] nalle .
%syn: $AFP M:ipre 1;9

(125) CHI: dä va nalle .
%syn: $AFS M:iptm 1;9

(126) CHI: dä , dä se Lelli sova .
%syn: €A $AFSV M:oanb M:infv 1;9

I 1;10 visar Hanna att hon har lärt sig mönstret ”här/där kommer *subj*”, men det finns också en annan inverterad deklarativ med suffigerad presens (exempel 130), som är säskilt intressant eftersom den inte har någon given förebild i input. Yttrandet refererar till en bild i en bok, på Annika som gråter. Fortfarande inleds nästan alla inverterade satser med *här* eller *där*, utom exempel 131. Förutom finit form av *vara* förekommer hjälpverben *ska* och *kan*. I 1;11 finns det förutom *här* och *där* ett par *nu* som inledare. Här dyker också det första, något tveksamma, objektet i spetsställning upp.

(127) CHI: däj komme Lisa .
%syn: $AFS M:spre 1;10

(128) CHI: ja hä komme Hannas # bok .
%syn: $AFS M:spre 1;10

(129) CHI: tit där komme pa [/] pappa Pippi .
%syn: $AFS M:spre 1;10

(130) CHI: dä gåde [: gråter] Annika .
%syn: $AFS M:spre 1;10

(131) CHI: den ka pappa bygga .
%com: target med den/på den
%syn: $AFSV M:ipre M:infv PP:noll 1;10

(132) CHI: de där kan man göra .
%syn: $OFSV M:ipre M:infv 1;11
('de där' syftar på att rita på papper, vilket man får göra till skillnad från att rita på väggen.)

(133) CHI: hä kan ja rita .
%syn: $AFSV M:ipre M:infv 1;11

(134) CHI: nu e slut .
%syn: $AFP M:ipre 1;11

(135) CHI: u ka ja <gå ut>[>] jobba .
%syn: $AFSVL+V M:ipre M:basf M:infv 1;11

Vid 1;11 finns alltså olika sorts led i fundamenten, kombinerade med olika sorters verb, men antalet V2-deklarativer är inte så stort (33 st enligt tabell 4.10). Från 2;0 finns det ett bättre underlag för att undersöka den finita satsen. I den inspelningen har Hanna 84 finita deklarativa satser med fundament och 21 av dessa är inverterade. De börjar på (i fallande ordning) *där*, *här*, *de*, *den*, *då* och *nu*, och kombineras med verbformerna *kommer*, *passar*, *är*, *var*, *får*, *kan*, *måste* och *har*. Det är en begränsad mängd kombinationer (9 typer), men variationen är ändå ganska stor[8]. Några exempel

(136) CHI: nu kan ja titta me Karin .
%syn: $AFSVA M:ipre M:basf PP:prep 2;0

(137) CHI: hä komme planet .
%syn: $AFS M:spre 2;0

(138) CHI: dä [!] passa de:n !
%syn: $AFS M:basf 2;0

I de två följande inspelningarna bygger Hanna på med nya kombinationer. I spetsställning tillkommer *annars*, *sen* och *snart* och de kombineras med en

[8]Kombinationerna består av *där var* (2 ggr), *där är*, *där passar* (4 ggr), *här kommer* (3 ggr), *här får* (3 ggr), *nu kan* (2 ggr), *det kan* (2 ggr), *då måste* och *den har*.

mängd nya verbformer, både i presens och preteritum. I 2;1 blir andelen objekt i spetsställning något större (11 st, alla utgörs av *de* eller *den*). De första exemplen på nominalfraser och andra lite tyngre fundament finns i korpusen i 2;3, med ”gussgussguss finns det” (Hanna svarar på vad det finns för sånger i en bok, och ”gussgussguss” är namnet på en ramsa), ”nästa gång ska vi tafera [: fotografera]”. I 2;8 finns det tydliga tecken på att hon har generaliserat regeln riktigt. Här finns det utbyggda nominalfraser (”dom här böckerna”, ”sån här blomma”) och adjektiviska adverbial (”högt, högt har ja klättrat”). I den inspelningen har hon också den högsta andelen XP–FV–subj–YY-strukturer, ca 45 %. Utbyggda fundament är som sagt lågfrekventa i målspråket, så det faktum att de inte förekommer i de tidigare inspelningarna behöver inte betyda att Hanna inte behärskar dem tidigare.

Man kan alltså konstatera att inverterade deklarativer förekommer i Hannas språk från de första inspelningarna som innehåller finita satser. De allra tidigaste har adverbial i spetsställning medan spetsställda objekt verkar komma lite senare.

Subjektsinledda deklarativa huvudsatser De subjektsinledda deklarativa satser som finns i 1;9 återges nedan.

(139) CHI: solan -, # lysen [= accent 1] solen .
%syn: $SFS M:spre 1;9

(140) CHI: manen [=? månen] e ute .
%syn: $SFA M:ipre 1;9

(141) CHI: Hanna kn nte söa den .
%syn: $SFIVO M:ipre M:infv 1;9

(142) CHI: osj , du få sov fäˆdiˆt .
%syn: $SFVA M:basf M:stam 1;9

(143) CHI: de va juvi Nelly .
%syn: $SFXS M:iptm 1;9

Hanna ger i 1;8 (se exempel 113 och 118 ovan) och 1;9 prov på fulla nominalfraser, egenamn, och både personliga och opersonliga pronomen som subjekt i kombination med tematiska verb, finit form av *vara* och hjälpverb. Jag kan inte se att det finns någon anledning att tro att kombinationerna inte skulle vara kreativa.

I exempel 143 ser vi ett tidigt och något tveksamt exempel på ”det” som formellt subjekt. I Hannas språk finns många ”de”, de flesta taggade som N, obestämbart nominal, men många är också tolkade som subjekt eller objekt.

I 2;0 finns det exempel på ”de” som grammatiskt subjekt, expletivt subjekt och objekt, men också i icke-målspråklig svårdefinierad funktion. I enstaka yttranden ser ”e de...” ut att vara en oanalyserad konstruktion som Hanna bygger ut med en verbfras: ”den e de logö [: laga] dä.” (1;9) och ”sen e de # ligga påsen.” (1;11). Plunkett och Strömqvist 1992 undersökte användningen av ”det” och ”den” närmare och menar att ”the extensive usage of *det* and *den* in the combinatorial speech of all the children described here, together with their associated positional constraints, suggests that these forms may be fulfilling an important anchoring function in the emergence of syntactic form” (s. 511). En utförligare undersökning av hur Hanna använder *det* och *den* i både €- och $-strängar ger kanske fler ledtrådar till hur den deklarativa V2-satsen växer fram.

4.3.3.3 Sammanfattning: den deklarativa satsen

Jag slog tidigare fast att Hanna har lärt sig den den finita satsen med subjekt och finit verb vid 2;0 i och med att andelen subjekt och finita verb då ligger nära de vuxenpråkliga värdena. Men den fullständiga deklarativa satstypen med fundament ser ut att bli mer vuxenlik en månad senare när V1-deklarativerna blir färre. Å andra sidan förekommer fundament av olika slag och i kombination med olika verb på ett sätt som ser ganska kreativt ut en månad tidigare, i 1;11. Inspelning 2;1 som alltså är den första av en sammanhängande period med vuxenspråklig andel V2-satser, visar i många andra avseenden en tillbakagång i språkutvecklingen[9]. Yttrandena är kortare (men inte färre), andelen meningsfragment är större och andelen satser mindre. Satserna innehåller en större del infinitva satser, och överlag färre subjekt än inspelningen innan. Men det visar sig alltså att de deklarativa satserna samtidigt går ett steg framåt i och med att en större andel av de finita satserna har fundament.

4.4 Underordnade verbfraser

I ¢-strängtypen samsas bisatser (och exklamativer), infinitivfraser, omkategoriserade anföringssatser och satssammandragningar som innehåller ett verb. I detta avsnitt gör jag en noggrannare analys av dessa strängtyper. Jag tittar på vilka typer som finns och diskuterar de tidigaste förekomsterna. Omkategoriserade anföringssatser kommer jag dock inte att ta upp här.

[9]Hanna har i inspelning 2;1 nyss fått en lillasyster, och det kan vara en av anledningarna till den tillfälliga tillbakagången.

Tabell 4.11: Underordnade verbfraser

1 Insp	2 Antal ¢	3 antal b.s.	4 antal IF	5 omkat anf	6 ¢++
1;6	0	0	0	0	0
1;7	0	0	0	0	0
1;8	0	0	0	0	0
1;9	0	0	0	0	0
1;10	2	2	0	0	0
1;11	1	1	0	0	0
2;0	15	4	10	0	1
2;1	3	1	2	0	0
2;2	8	4	1	1	0
2;3	2	2	0	0	0
2;4	11	10	0	1	0
2;5	9	8	0	0	1
2;6	7	5	1	0	1
2;7	18	13	0	0	2
2;8	30	27	0	2	1
2;10	17	15	0	1	1
V2;0	42	25	10	1	5
V2;4	34	27	3	1	1

2=Antal ¢-strängar totalt, 3= antal bisatser, inklusive bisatsföregångare, 4=antal infinitivfraser, 5=omkategoriserade anföringssatser, 6=satssammandragningar

4.4.0.4 Underordnade satser

Lundin (1987) har undersökt tidiga bisatser i fem barns språk och hon kan konstatera att det finns föregångare till fullständiga bisatser som tycks sakna bisatsinledare. Dessa är dock mycket få. Lundin har sammanlagt exerperat 1815 fullständiga bisatser i dessa fem barns inspelningar från ca 1;10 till 3;6, och bisatsföregångarna är inte fler än kanske 30. Lundin skiljer ut de tidiga *protosatserna*, som kännetecknas av att de innehåller två verbenheter inom ramen för ett yttrande: EMBLA: ”Han är sönder ramla.” (2;0) dvs. ’Han är sönder för att han ramlade’, eller ’Han gick sönder när han ramlade’ (Lundin, 1987, s. 52). Den andra verbfrasen tolkar hon som ett bisatsfragment som saknar både inledare och subjekt. Det finns ca 10 exempel av dessa protosatser i hennes material. Det finns också *sambandssatser* som är två (eller flera) prosodiskt fristående satser där det finns en relation mellan dem som man i vuxenspråket skulle föredra att uttryckas med underordning: TOR: ”Knuffa! Trilla ner!” (2;3), dvs. ’Nallen knuffade henne så att hon trillade ner’ (ibid, s. 46). Det är oklart hur många sådana Lundin räknar med. Samtidigt med dessa sambandssatser finns också 5 subjunktionslösa[10] tidiga bisatser, yttrade av Ask och Freja. Gemensamt för dessa är att de yttras några månader innan Ask och Freja blir produktiva med fullständiga bisatser. Det finns också subjunktionslösa bisatser senare under barnens produktion, men de räknar Lundin till *defekta* bisatser som utesluts från undersökningen. De defekta bisatserna uppgår till 14 sammanlagt, så när barnen väl har börjat att producera fullständiga bisatser förekommer subjunktionslösa satser sällan. Till sist hittar Lundin också 11 subjektslösa relativer. Gemensamt för dessa är att de förekommer i de olika barnens tidiga produktion och de flesta förekommer i presenterings- eller utbrytningskonstruktioner (”Det/Där är NP (som) VP”). Lundin menar att de tidiga protosatserna både är resultat av att barnen har begränsad förmåga att producera längre yttranden och att de inte ”har de underordnande konjunktionerna i sitt aktiva ordförråd” (s. 55), medan de få subjunktionslösa tidiga bisatserna bara skulle bero på det sistnämnda.

I Hannakorpusen finns det några få yttranden som innehåller ¢-strängar i den period som föregår den etablerade finita satsen, dvs. före 2;0. I 1;10 och 1;11 finns tre förekomster (två olika typer) av kombinationer med ”titta” (exempel 144–145).

(144) CHI: titta hä tå # Molly .
MOT: mm .
CHI: ta Molly jö: .

[10]Lundin räknar med både subjunktionsinledda och satsbasinledda bisatser här.

MOT: va sa du ?
CHI: titta Molly gö .
CHI: da ## dä Molly gö dä . 1;10

(145) CHI: titta Hanna ritat . 1;11

De två yttrandena i 1;10 tolkade jag som ”Titta vad Molly gör”. I vuxeninput är konstruktionen vanlig. I samma inspelning finns t.ex. ”Ska du gå å titta va Karin gör?” och ”Ska vi se va Pippi gör för nånting?”. Men eftersom Hanna därefter använder *gör* utan objekt (”Dä Molly gör där”) kan man misstänka att hon inte har argumentstrukturen för verbet *göra* helt klar för sig, och eftersom intonationen inte ger någon klar ledning i dessa fall kan yttrandena lika gärna klassificeras som två separata verbfraser: ”Titta. Molly gör.”. Inte heller i exempel 145 ger intonationen någon klar ledning till om det ska betraktas som två seprata verbfraser (”Titta. (Den har) Hanna ritat”) eller en underordning (”Titta (vad) Hanna (har) ritat”).

Från 2;0 är ¢-strängarna klassificerade på säkrare grunder. När det gäller bisatser ska man dock komma ihåg min princip att ”tagga stort”. Den har inneburit att jag har klassificerat verbfraser som bisatser när de har låtit som bisatser, dvs. när verbfrasen i ¢-strängen har ingått i samma intonationsenhet som en matrissats, men också när de har fungerat som självständiga satser men varit inledda av en bisatsinledare. Många av bisatserna står självständigt som egna yttranden, något som också Lundin (1987) kunde konstatera.

Från 2;0 till 2;4 finns 9 bisatser som ingår i en matrissats, och alla utom 2 saknar bisatsinledare. De två som har bisatsinledare avviker på andra sätt (exempel 147 och 148). Några exempel ges nedan:

(146) CHI: ja måste titta denna flyge himle . 2;0

(147) CHI: xx vet inte sängen va e ? 2;0

(148) CHI: inte komma in de hon ka ligga ! 2;2

(149) CHI: da n gul glass ja ka äta . 2;2

(150) CHI: han dra ner byssorna han kan kissa . 2;3

Fram till 2;4 finns bara två fristående bisatser (”om vi har fyrkant” i 2;1 som inte har någon begriplig relation till kontexten, och ”mm # som komme nog sna(r)t ge(no)m mi (tt) fönstör” i 2;3 som har fa-ordföljd). I inspelning 2;4 finns däremot 10 bisatser som alla utom en är fristående. De flesta av dem (7 st) inleds med ”fatt” (för att) och verkar fungera som vanliga huvudsatser[11].

[11] Hur man ska skilja på konjunktioner och subjunktioner är inte helt självklart. *Så* och *för* kan i talspråk fungera som både konjunktioner och subjunktioner med mycket liten betydelseskillnad.

De övriga inleds med ”som”. Två av ”fatt”-satserna innehåller ett satsadverbial som visar att de har fa-ordföljd. Den enda bisatsen som inte är fristående utgör det första exemplet på en helt målspråklig bisats, den första utbrytningskonstruktionen:

(151) CHI: de e sena [: senap] <som vi> [/] som vi hanlat i fären me papp +... 2;4

I 2;5 finns 8 bisatser. Alla är i stort sett målspråkliga och de är av olika typer, *att*, *så (att)*, vanlig relativsats, utbrytningsbisats och en *när*-sats:

(152) CHI: eh att hon kan inte sova . 2;5

(153) CHI: kan du vända den hit så # ja kan se ? 2;5

(154) CHI: all sån däe # som ligge -_ den [/] de ha ja köpt . 2;5

(155) CHI: e bara lillebarn (s)om få ha den munnen . 2;5

(156) CHI: mm # de ska vi nä(r) vi a fikat färdit . 2;5

I den ena att-satsen i 2;5 finns också det första tecknet på af-ordföljd i bisats (*att* är realiserat som [t]):

(157) CHI: hon e tåt fyra .
MOT: en gång till # va sa du för nåt ?
CHI: t Yrsa [//] t # Klara snat e fyra . 2;5

Från 2;0 till 2;4 producerar alltså Hanna sporadiskt bisatser av olika typer, men i stort sett utan bisatsinledare. I inspelning 2;4 visar hon att hon har lärt sig ett par subjunktioner och i 2;5 ser bisatserna ganska fullständiga ut, med ett ganska tydligt exempel på af-ordföljd. Efter 2;5 ser man en gradvis större repertoar av bisatstyper; en satsfläta i 2;6, en expressiv sats och många utbrytningar i 2;7, v-frågebisatser och många när-satser i 2;8. Att produktionen inte är helt problemfri kan man se i ett exempel från 2;10 där flera underordnande satser i en och samma huvudsats blir för mycket. Observera dock att hon trots denna övermäktiga komplexitet klarar bisatsordföljden i den sista bisatsen (”när jag inte hade...”).

(158) CHI: ja har , kolla , förklä så # när ja kulˆle göra sånt där # ös # xxx att ja spillde ## efte när ja nte hade nåt förläd # ovanför min tröja .
(oklar tolkning) 2;10

Man kan också notera att alla att-satser före 2;4 som ingår i en matris består av konstruktionen ”Det är bäst/bättre (att)...”. Konstruktionen har alltid utelämnat *att*, också då den förekommer i den period då *att* finns i andra bisatser. Lägg märke till att ordföljden i det tidigaste exemplet är avvikande.

(159) CHI: den e bätte sitte den da . 2;0

(160) CHI: de bätte ja byggö . 2;2

(161) CHI: oj , de bätt ja gö själv . 2;2

(162) CHI: e de bäst man knäppör . 2;5

Lite senare använder Hanna också konstruktionen ”Det är för att...” ofta, och i en av de bisatserna förekommer af-ordföljd i en *för att*-sats (”de e fatt ma [/] mamma inte hitta den.” (2;7)). I samma inspelning använder hon dock ”fatt” för att inleda något som ser ut som en huvudsats (”fatt ja måste åssså ha en filt”). Konstruktioner av detta slag verkar alltså spela en viss roll för bisatsproduktionen av *att*-satser, men det är svårare att se att de spelar roll för bisatsproduktionen i stort. Men underlaget är i minsta laget för att man ska kunna dra några slutsatser.

De finns fyra bisatsföregångare som jag har tolkat som oinledda satsbasinledda bisatser.

(163) CHI: ja måste titta denna flyge himle . 2;0

(164) CHI: de ka leta upp # passa . 2;0

(165) CHI: xx vet inte sängen va e ? 2;0

(166) CHI: inte komma in de hon ka ligga ! 2;2

De två första saknar enligt min tolkning v-frågeord. I det första fallet skulle den målspråkliga varianten snarast ha ett *när* som inte finns i Hannas inspelade produktion förrän 2;5 (och då som bisatsinledare). I exempel 164 saknas ett *var*, och i tabell 4.8 framgår att lexemet finns i Hanna produktion vid 2;0. Men ett utsatt v-frågeord som bisatsinledare dröjer som sagt till 2;5 då första belägget på *när* kommer, och därefter kommer ett *vad* i en expressiv sats i 2;7 (”Va du snäll”). I exempel 165 finns visserligen ett utsatt *var*. Målet för yttrandet har jag tolkat som ”Jag vet inte var sängen är”, men eftersom hela satsen har en prosodi som är frågande, och en ordföljd som gör att det är svårt att segmentera yttrandet, är det svårt att hävda att ”va” verkligen fungerar som en bisatsinledare. I det sista exempel är den bisatsinledare som ligger närmast till hands *där*, men man kan också tänka sig att det finns ett kausalt förhållande mellan verbfraserna: ”Du får inte komma in där eftersom hon ligger och sover där”. I så fall skulle man kunna se bisatsföregångarna som logiskt knutna till en matrissats, utan att bisatstypen egentligen är specificerad.

Samordningar av satser förekommer sporadiskt (också i det taggade vuxenspråket). Den sats som är taggad som samordning vid 2;0 var svårklassificerad.

(167) CHI: nu han lamla ont knana .
MOT: oj , ramla han å fick ont i knäna igen ? 2;0

Taggningen baserar sig snarast på MOTs tolkning av yttrandet, men när man ser helheten är det svårt att motivera att det skulle finnas en skillnad mellan det här yttrandet och t.ex. protosatsen i exempel 164 ovan. Samordningarna från 2;5 ser dock ut som målspråkliga samordningar, även om de flesta saknar samordnande konjunktion.

(168) CHI: va [/] vatt på stajkpark plasen . (Hon menar stadsbiblioteket.)
CHI: <me Inga>[>] .
...
MOT: jaha !
CHI: å Pia # å Inga å lyssna på sage . 2;5

(169) CHI: man kan ta ut # nn # knuffa in den . 2;7

Om man antar att satssammandragningar innebär samordning av två huvudsatser och därefter ellips av de satsdelar som är gemensamma bör satssammandragning vara en mer komplicerad process än underordning av verb. De satssammandragningar man kan få syn på i Hannakorpusen dyker också upp något efter att bisatserna har blivit fullständiga. Skillnaden mellan satssammandragningar och pseudosamordningar av verb är inte alltid så lätt att avgöra, men som vi kommer att se i avsnitt 4.5 blir pseudosamordningar av verb synliga tidigare i Hannas produktion.

4.4.0.5 Infinitivfraser

Infinivfraser med infinitivmärke fungerar på samma sätt som ett nominalt led i satsen, antingen som ett argument till ett huvudverb (t.ex.*hjälpa någon att göra ngt*), ofta som egentligt subjekt (*det går inte att öppna*) eller som infinitivattribut i en nominalfras (*konsten att angöra en brygga*). Gränsen mellan hjälpverb och huvudverb som tar en infinitivfras som komplement är vag (SAG II, s. 536). Det märks bland annat på om infinitivmärket måste sättas ut (t.ex. *hålla på *(att)*), om det kan sättas ut optionellt (t.ex.*orka (att)*) eller om det inte kan sättas ut (t.ex. *kan *att*).

I Hannamaterialet finns sammanlagt 14 förekomster av infinitivfraser, alla utan infinitivmärke. Förekomsterna fördelar sig på fem olika typer, och ett par av dem kan tolkas som hjälpverb + huvudverb (exemplen i 170).

(170) CHI: ja ka fösöka leta upp den . 2;0

(171) CHI: xx orka vi leka sen . 2;0

(172) CHI: hä # måste gå öppna lite till . 2;0
CHI: de gå inte tända där . 2;1

(173) CHI: kan du hjälpa den dä sitta ? 2;0
CHI: kan du hjälpa mej bygga ? 2;1
CHI: nu få du hjäpa mej kasta # såna bollar . 2;6

(174) CHI: nu flickans tur # sitta . 2;0
CHI: nu ä (...) mammas tur stå ph [/] på tomet . 2;2

Jag känner inte till någon forskning om infinitivfraser i svenska barns produktion. Att infinitivmärket saknas är förmodligen en följd av att det inte är så framträdande i prosodin, och att Hanna inte har lagt märke till det. Observera att det inte är samma lexem som subjunktionen *att*. Infinitivmärket uttalas vanligtvis [ɔ] i spontant talspråk, medan subjunktionen alltid uttalas [at].

4.5 Funktionsord

Funktionsord som prepositioner, subjunktioner, konjunktioner och infinitivmärke är inte centrala för den finita satsens framväxt. Dessa ord är för det mesta umbärliga; vi tolkar utan större svårigheter yttranden som saknar dessa ord, och det uppstår mycket sällan missförstånd. Inte desto mindre lär sig barnen dem, så småningom. I tabell 4.12 har jag listat utsatta och icke-utsatta förekomster av prepositioner, konjunktioner och infinitivmärken i obligatoriska kontexter. För konjunktioner har jag bara tittat på samordningar av verbfraser, dvs. psuedosamordnade verb (”sitta och äta”) och samordnade satser med olika verbfraser (”då får han gå in köket och diska tallriken”), där den samordnande konjunktionen alltid är obligatorisk. För subjunktioner har jag listat förekomster av utsatta subjunktioner och icke-utsatta, både obligatoriska och icke-obligatoriska. Underlaget är för litet för att man ska kunna dra några generella slutsatser av det, men man kan ändå se några intressanta detaljer.

Prepositionerna kommer ganska plötsligt i inspelning 2;1. Subjunktionerna börjar att sättas ut i 2;4, även om nästan alla utgörs av ”fatt”. I 2;5 finns olika sorters subjunktioner. Alla nollsubjunktioner efter 2;4 är målspråkliga. Waldmann (2008) ser en liknande utveckling hos de barn han undersöker. Infinitivmärket är frånvarande i den tidiga produktionen. Den senare perioden kan jag inte säga något om eftersom infinitivfraserna kuriöst nog är nästan helt frånvarande då. Konjunktioner däremot, som är homofona med infinitivmärken, sätter Hanna ut första gången i 2;3 (”ja ka gå upp å hämta den”).

Indirekt ger tabell 4.11 och 4.12 över ¢-strängar och funktionsord också en inblick i hur Hanna förmår att kombinera verbfraser inom ramen för en proso-

Tabell 4.12: Funktionsord

Insp	prep	nollprep	subj	nollsubj	konj	nollkonj	IM	nollIM
1;6	0	1	0	0	0	0	0	0
1;7	0	2	0	0	0	0	0	0
1;8	0	1	0	0	0	1	0	0
1;9	4	7	0	0	0	0	0	0
1;10	3	18	0	(2)	0	5	0	0
1;11	2	8	0	(1)	0	1	0	0
2;0	7	24	(1)	3	0	3	0	10
2;1	23	7	1	0	0	4	0	2
2;2	36	8	1	3	0	2	0	1
2;3	19	1	1	1	1	1	0	0
2;4	44	1	10	0	0	0	0	0
2;5	20	2	6	2	1	0	0	0
2;6	24	2	5	0	1	1	0	1
2;7	27	4	12	1	0	2	0	0
2;8	32	1	19	8	1	0	0	0
2;10	33	2	15	0	2	0	0	0
V2;0	76	2	19	6	14	0	7	3
V2;4	57	0	26	1	7	0	3	0

prep=prepostion, subj=subjunktion, konj=konjunktion, IM=infinitivmärke

disk fras. Jag har redovisat att Hanna kan kombinera hjälpverb och huvudverb från i alla fall 1;0. Nästan alla nollkonjunktioner i tabell 4.12 fram till 2;5 finns i pseudosamordnade verb. Hanna använder alltså pseudosamordning av verb lika tidigt som hjälpverb och huvudverb [12].

(175) CHI: Hanna sitt läsa den . 1;8

(176) CHI: gå ut titta Karin ! 1;10

(177) CHI: u ka ja <gå ut>[>] jobba . 1;11

I 2;0 kommer sedan verbfraskombinationer i form av infinitivfraser och bisatsföregångare samtidigt. I 2;5 blir bisatserna fullständiga. Satssammandragningar är däremot sällsynta.

[12] Det tidiga exemplet i 175 är analyserat som pseudosamordning eftersom *sitta och...* är ett vanligt förled i pseudosamordningar och för att betydelsen passar i kontexten. Man kan inte utesluta att det istället är ett exempel på en omtagning, dvs. ”Hanna sitt [/] läsa den”.

4.6 Den finita satsen hos Hanna

I detta avslutande avsnitt ska jag försöka sammanfatta vad man kan se i all den utdata som taggningen har genererat. Jag håller mig här till sådant som verkar vara av intresse när man ska undersöka hur den finita satsen etableras.

En förutsättning för att kunna producera satser i någon större omfattning är att man kan tillräckligt många verb, och en förutsättning för finita satser är att man också kan producera finita verbformer. I de två första inspelningarna utgörs mindre än vart tionde ord av ett verb hos Hanna. I 1;8 ökar andelen verb i förhållande till det totala antalet ord, men andelen verb är mindre än i målspråket och verben är fortfarande inte differentierade morfologiskt. Verbökningen i 1;8 innebär inte någon förändring av yttrandelängden, men däremot visar den sig i att jag har klassificerat fler yttranden som $-strängar än i 1;6 och 1;7. Perioden 1;6 till 1;8 är ett ett- till två-ordsstadium, där satsstrukturer ännu är ovanliga och verbmorfologin inte genomskådad.

Från och med 1;8 till och med 1;11 förändras inte andelen verb, men verbböjningen börjar att bli mer systematisk från 1;9. Skillnaden mellan suffigerad presens och infinitiv blir lite lättare att identifiera och supinum och några icke-suffigerade presensformer börjar dyka upp. I 1;9 har jag också bedömt att prosodin på yttrandenivå har förändrats, så att yttrandena i större utsträckning *låter* som satser. De första sporadiska tecknen på förmåga att kombinera fler verb inom samma prosodiska fras börjar att synas, dels genom att Hanna kombinerar hjälpverb och huvudverb, dels genom att hon pseudosamordnar verb.

Den finita satsen ser sedan ut att växa fram i tre steg från 1;11 till 2;1. I 1;11 har Hanna upphört med sina spontana imitationer på ordnivå, meningsfragmenten (€-strängarna) minskar och satsstrukturerna ($-strängarna) blir fler. I 1;11 visar Hanna också att hon kan konstruera deklarativa V2-satser med finit verb, subjekt, som övervägande består av pronomen, och olika slags fundament på ett sätt som ser ganska målspråkligt ut. De finita satserna är ännu inte så många. Om man räknar bort ettordsimitationerna i 1;9 och 1;10 innebär förändringen inte att yttrandena blir längre. Andelen satser med enbart infinit verb sjunker. I 1;11 visar Hanna alltså att hon har identifierat var olika led ska placeras i satsstrukturen. Hon *kan* producera finita satser, men de flesta satser är fortfarande inte finita.

I 2;0 är de finita satserna däremot många. $-strängarna är nästan uppe i vuxennivå. De icke-suffigerade presensformerna (hjälpverb och *är*) blir märkbart fler, vilket innebär att andelen satser med ett finit verb ökar drastiskt. Detta sammanfaller med en klar ökning av andelen subjekt i satsströngarna. Andelen pronomensubjekt ökar ytterligare något och når vuxennivåer. Satser med hjälpverb och huvudverb är vanliga och Hanna har också börjat att kombinera

verbfraser med hjälp av bisatsföregångare och infinitivfraser. Yttrandena blir därmed i genomsnitt nästan ett helt ord längre. I 2;0 har hon också hittat den rogativa satstypen.

I 2;1 kommer så det tredje steget när satsfundamenten faller på plats och andelen V2-deklarativer kommer upp till en målspråklig nivå. Detta borde också öka yttrandelängden, men i 2;1 är MLU(w)-värdet snarare något lägre. Ökningen av yttrandelängd vid 2;2 kan dock eventuellt hänga samman med att Hanna i den inspelningen har såväl fundament som finita verb och subjekt i satserna. I 2;1 har de finita satserna visserligen fundament, men den inspelningen visar också en liten tillbakagång när det gäller andel finita verb (främst icke-suffigerad presens) och subjekt.

Hannas finita satser ser i vissa avseenden ganska målspråkliga ut från 2;1, men förstås inte helt. Fram till och med 2;2 ser det ut som ett *optional infinitive stage* (Wexler, 1994). Även om satser med enbart infinit verb är förhållandevis ovanliga från 1;11 blir de ännu ovanligare efter 2;2. Rotinfinitver och rotsupiner finns dock med många funktioner varav en del är målspråkliga, och de bör undersökas närmare. Under hela undersökningsperioden saknar Hannas satssträngar finit verb oftare än i jämförelsematerialet vilket till största delen beror på utelämnade kopulaverb. Andelen subjekt ser ut att öka ytterligare lite från 2;5 till vuxennivå. Från 2;4, 2;5 kommer ordentliga bisatser med bisatsinledare. Andelen bisatser är enligt figur 4.1 i hela undersökningsperioden mindre än i vuxenspråket, men i 2;7 ser den ut att öka något. Nästa tydliga ökning av yttrandelängden kommer först vid 2;8, så det är tveksamt om det är andelen bisatser som bidrar till den ökade yttrandelängden.

Sammanställningen visar att satsen etableras i perioden 1;11 till 2;2. Under perioden 1;9 till 1;11 börjar den finita satsen ta form och den perioden blir därför intressant att undersöka mer i detalj. Bland annat behöver subjekten och deras relation till det finita verbet utredas noggrannare. Analysen av utdata indikerar också att de icke-suffigerade presensformerna kan vara intressanta att studera vidare. Många av de inverterade tidiga satserna tycks ha en finit form av *vara* eller ett finit hjälpverb. Att de modala hjälpverben har en särskild ställning i språkinlärning har tidigare föreslagits av bl.a. Håkansson (1989); Clahsen (1990); Håkansson och Collberg (1994); Eide (under arbete).

Kapitel 5

Avslutning

I detta avslutande kapitel ska jag helt kort sammanfatta resultaten och utvärdera analysmetoden.

Till att börja med kan jag konstatera att arbetet med CHAT- och CLAN-verktygen och med det länkade materialet har varit mycket givande. Jag har kunnat analysera materialet i flera omgångar och inte behövt förlita mig på att jag i det första analysskedet gjort en riktig bedömning av ett yttrande. Den slutgiltiga tolkningen har jag kunnat göra utifrån en samlad bedömning av flera yttranden av samma typ. Genom att alltid ha nära till barnets talade språk i undersökningarna har risken för över- och undertolkningar minskat. Man kan inte nog underskatta det skrivna språkets dominans. Ett yttrande i skrift ser ofta mer vuxenlikt ut än det är, och jag har många gånger fått revidera min uppfattning av vad Hanna säger efter att ha lyssnat till yttrandet igen. Beslutet att transkribera språket ljudnära istället för med standardiserad ortografi har lett till vissa problem i analysskedet. Men jag tror ändå att det var ett riktigt beslut. Det är svårt att frigöra sig från skriftspråket, men den avvikande stavningen påminner hela tiden om att talspråket är annorlunda.

Mitt första syfte med föreliggande avhandling var att undersöka om jag utifrån spontandata kunde avgöra när den finita satsen var etablerad i barnets språk. I kaptiel 4 har jag redovisat att det i alla fall utifrån Hannamaterialet har varit möjligt. Genom att jag har taggat barnyttrandena för yttrandetyp (strängtyp), ledföljd och verbmorfologi oberoende av varandra har jag fått en god bild av när Hanna kan sägas ha kunskap om hur en full satsstruktur uttrycks i svenska. Kartläggningen förde också med sig en detaljerad beskrivning av Hannas språkutveckling mellan 1;6 och 2;10 som relateras till tidigare undersökningar av svenskt barnspråk. Kartläggningen visar att Hannas finita satser tar form i

perioden 1;9 till 1;11. I 1;11 ser hon ut att veta hur en sats ska formas, men först i 2;0 blir subjekt och finita verb vanliga och finta satser därmed dominerande i produktionen. Det betyder att perioden fram till och med 1;11 är intressant att undersöka när det gäller hur den finita satsen växer fram, och inspelningarna från och med 2;0 (och även kanske 1;11) kan utgöra underlag för att testa Platzacks (2007) EPP-hypotes.

Det andra syftet var att beskriva, tillämpa och utvärdera den analysmetod som jag har utarbetat. Beskrivningen av satsanalysen finns i kapitel 3. I kapitel 4 redovisar jag en del av utfallet av analysen och jag fick då anledning att diskutera en del problem med analysverktyget. I det följande ska jag på ett mer systematiskt sätt gå igenom om analysen kan ge mig det jag behöver i de fortsatta studierna.

Analysen skulle ge mig möjlighet att söka ut och undersöka de fenomen som Platzack undersöker i sin pilotstudie av EPP-relaterade avvikelser i små barns språk (2007), vilket den gör. Jag kan inte bara söka ut de yttranden som exemplifierar fenomenen utan också se under vilken period de förekommer. Jag kan utan vidare söka fram och kartlägga satser som saknar fundament, dvs. med finit verb på första plats som inte är rogativa frågor. Jag kan också söka fram kvesitiva frågor som saknar v-frågeord och subjektslösa satser utan svårigheter. Jag kan enkelt söka fram satser som har det finita verbet på tredje plats, även om jag inte har visat det i utfallskapitlet. Vid sökning efter $-strängar med ett finit verb på tredje plats ($_ _F*) kommer också målspråkliga V3-satser med, så sökningen kräver alltid en genomgång av exemplen. Satser som enbart innehåller infinit verb kan sökas ut, men med vissa svårigheter (som jag kort beskriver på s. 73). Det hade varit önskvärt med mer avancerade reguljära uttryck i CLAN-programmet *combo*, alternativt mer avancerade grepkunskaper hos mig. En annan lösning är att sådana satser (rotinfinitiver) utgör en egen strängtyp. Jag kan också hitta utelämnade prepositioner i obligatoriska kontexter och utelämnade subjunktioner även om jag inte direkt kan söka ut de subjunktioner som är utelämnde i obligatoriska kontexter (Josefsson, 2003b).

Analysen skulle också göra det möjligt att hitta andra systematiska avvikelser när det gäller obligatoriska satsled. Vissa sådana är redan undersökta om än inte publicerade. Det rör sig bl.a. om avvikande placeringar av subjekt i förhållande till infinit verb och adverbial i förhållande till objekt. Under perioden 1;6–1;8, när den finita satsen ännu inte har börjat visa sig i Hannas produktion, varierar Hanna ordningen mellan subjektet och ett morfologiskt säkert infinit verb. Totalt finns 14 (6 olika verb) yttranden med ordföljden S-V(infv) och 9 (3 olika verb) med V(infv)-S (se exempel 178). Den senare ordföljden finns inte vuxenspråket och det är inte självklart hur man ska beskriva den. Från 1;9 verkar ordföljden ha stabiliserat sig och man finner bara undantagsvis subjekt efter

ett infinit verb. Under perioden 1;6–1;10 kan Hanna placera ett adverbial före ett direkt objekt, en ordföljd som inte heller den förekommer i vuxenspråket (se exempel 179). Det finns lika många exempel på ordningarna adverbial–objekt som objekt–adverbial fram till 1;10, men från 1;11 finns bara den förväntade ordföljden objekt–adverbial.

(178) CHI: Hanna , läsa Hanna den Mak [: Maxboken] . 1;8

(179) CHI: lappa lite voffen .
mål: klappa vovven lite 1;9

Dessa satser bör naturligtvis studeras närmare och knytas till en teori om språkinlärning. Ordföljdsmönstren utgör dock exempel på avvikelser som kan hittas i och med ledföljstaggningen. Jag undersöker också möjligheten att sortera alla $-strängar automatiskt så att de som inte passar in i det prototypiska satsschemat faller ut.

Det finns vissa problem med analysens indelning i strängtyper. Min klassificering av satser och satsmotsvarigheter utgår från vuxenspråket och jag försöker att använda så få strängtyper som möjligt. Jag skiljer på meningsfragment utan verb, huvudsatser och underordnade satser. Det finns två huvudsakliga problem med den indelningen.

Det första problemet är att de tidiga ett- och tvåordsyttrandena är svåra att ordna in i strängklasserna. De kriterier jag använde visade sig inte hålla (se s. 83). Ett alternativ är att dessa yttranden utgör en egen strängtyp, som då eventuellt också omfattar tidiga ett- och tvåordsyttranden med verb. Fördelen är att man då behandlar dessa yttranden som en egen typ av satsmotsvarigheter, vilket de kanske är (Bloom, 1973; Brown, 1973; Roeper, 2007). Men då får man istället andra avgränsningssvårigheter. Hur kan man avgöra om ett yttrande är ett typiskt barnspråkligt tvåordsyttrande eller en huvudsats som saknar ett eller flera led? Satser kan sakna kopulaverb under en längre tid i barnspråket, utan att de för den delen självklart utgör den tidiga sortens tvåordsyttranden. Hur gör man med ett tidigt ettordsyttrande som består av ett tempusböjt verb? I efterhand kunde jag se att jag hade använt mig av prosodiska kriterier för att identifiera en sats. Om detta är en framkomlig väg när det gäller att skilja på tvåordsyttranden och satser återstår att undersöka. Klart är i alla fall att analys av tvåordsyttranden kräver ett annat perspektiv än det jag har haft här, där intresset främst har berört den finita satsen. Om jag på allvar ska kunna undersöka hur den finita satsen växer fram måste jag dock gå djupare in på hur tvåordsyttranden ska betraktas.

Det andra problemet är att huvudsatskategorin ($-strängar) samlar en grupp olika satstyper som har olika förutsättningar. Den prototypiska huvudsatsen har

fundament, finit verb och subjekt i en förutsägbar följd. Men de direktiva satserna med verb i imperativ kan inte ha fundament, och de måste inte innehålla subjekt. Huvudsatsekvivalenter som består av verbfraser i infinitiv och t.ex. utgör svar på frågor (”Vad ska vi göra idag?” ”Åka till farmor och farfar.”) kan varken ha fundament eller subjekt. När man ska undersöka ordföljd i huvudsatsen utgör dessa andra satstyper en felkälla. Och när man i tabell 4.6 på s. 91 ser en nedgång i andelen finita verb i huvudsatser hos Hanna beror det att hon svarar på frågor med hjälp av ofullständiga satser, som i vuxenspråket.

Min samlade bedömning är ändå att analysmetoden har fungerat bra och fyllt sitt syfte, även om det finns viss anledning att fundera över antalet strängtyper och vilka typer av satser och satsmotsvarigheter man ska särskilja.

Det tredje syftet med avhandlingen var att presentera ett nytt barnspråksmaterial av en svensk förstaspråksinlärare. Syftet var vidare att diskutera metodfrågor kopplade till transkriptioner av talat språk, och då särskilt små barns språk, samt redogöra för de ställningstaganden som jag har gjort. Presentationen av Hannakorpusen och metoddiskussionen finns i kapitel 2.

Denna licentiatavhandling är tänkt som ett första led i mina undersökningar av den finita satsen hos små svenska barn. Mitt fortsatta arbete kommer att ske inom den generativa grammatikens ramverk, och jag kommer naturligtvis att inkludera fler barn i undersökningarna. I det fortsatta arbetet finns två huvudintressen. För det första vill jag undersöka om de syntaktiska avvikelserna man finner i finita satser hos små förstaspråksinlärare kan beskrivas som ett problem med att automatisera kunskapen om de språkspecifika EPP-egenskaperna på det sätt Platzack (2007) föreslår. För det andra är jag intresserad av den allra tidigaste språkinlärningen och framväxten av den finita satsen. Jag är dessutom intresserad av begreppet *full kompetens* och hur man kan vänta sig att det tidiga språket ser ut om barnet har samma kompetens, dvs. samma I-språk, som en vuxen.

Litteraturförteckning

Barton, Ellen. Toward a nonsentential analysis in generative grammar. I: Progovac, Ljiljana, Paesani, Kate, Casielles, Eugenia, och Barton, Ellen, redaktörer, *The Syntax of Nonsententials. Multidisciplinary perspectives*, band 93 av *Linguistik Aktuell/Linguistics Today*, kapitel 1, ss 11–32. John Benjamins, Amsterdam/Philadelphia, 2006.

Bloom, Lois. *One word at a time. The use of single word utterances before syntax*. Mouton, 1973.

Bohnacker, Ute. *Icelandic plus English: Language differentiation and functional categories in a successively bilingual child*. Doktorsavhandling, University of Durham, 1999.

Bohnacker, Ute. Nominal phrases. I: Josefsson, Gunlög, Platzack, Christer, och Håkansson, Gisela, redaktörer, *The Acquisition of Swedish Grammar*, band 33 av *Language Acquisition and Language Disorders*, kapitel 8, ss 195–260. John Benjamins, 2003.

Brown, Roger. *A First Language*. George Allen & Unwin Ltd., 1973.

Bruce, Gösta. Allmän och svensk prosodi. Lund, 1998.

Byrman, Gunilla, Lindquist, Hans, och Levin, Magnus, redaktörer. *Korpusar i forskning och undervisning. Corpora in Research and Teaching. Rapport från ASLA:s höstsymposium, Växjö, 11–12 november 1999.*, 1999. Växjö universitet.

Christensen, Lisa. The acqusition of tense. I: Josefsson, Gunlög, Platzack, Christer, och Håkansson, Gisela, redaktörer, *The Acquisition of Swedish Grammar*, band 33 av *Language Acquisition and Language Disorders*, kapitel 3, ss 31–74. John Benjamins, 2003.

Christensen, Lisa. The semantics of early verbs, under arbete.

Clahsen, Harald. Constraints on parameter setting: A grammatical analysis of some acquisition in stages in german child language. *Language Acquisition*, (1):361–391, 1990.

Demuth, Katherine. Collecting spontaneous production data. I: McDaniel, Dana, McKee, Cecile, och Cairns, Helen Smith, redaktörer, *Methods for Assessing Children's Syntax*, Language, Speech, and Communication, kapitel 1, ss 3–22. The MIT Press, Cambridge, Massachusetts, 1996.

Demuth, Katherine. Morphological development: Prosodic constraints. I: Weissenborn, J. och Höhle, B., redaktörer, *Approches to Bootstrapping. Phonological, Syntactic and Neurophysical Aspects of Early Language Acquisition.*, band 24 av *Language Acquisition and Language Disorders*, ss 3–22. John Benjamins, 2001.

Dye, Cristina. The status of ostensibly nonfinite matrix verbs in child french: Results from a new corpus. I: Brugos, Alejna, Clark-Cotton, Manuella R., och Ha, Seungwan, redaktörer, *Proceedings of the 29th Boston University Conference on Language Development*, Boston, 2005. Cascadilla Press.

Eide, Kristin. Finiteness and inflection: the syntax your morphology can afford, under arbete.

Ekerot, Lars-Johan. Syntax och informationsstruktur. I: Hyltenstam, Kenneth, redaktör, *Svenska i invandrarperspektiv*, ss 79–108. Gleerups förlag, 1979.

Engstrand, Olle och Kadin, Germund. F_0 contours produced by Swedish and American 24-moth-olds: implications for the acquistion of tonal word accents. I: *Papers from the 17th Swedish Phonetics Conference*, ss 68–71. Department of Linguistics, Stockholm University, 2004.

Freudenthal, Gunnar. *Om typiska skiljaktigheter mellan modersmålsutveckling och språkkurs. Ett bidrag till språkundervisningens psykologi*. Gleerupska universitetsbokhandeln, 1925.

Håkansson, David. *Syntaktisk variation och förändring*. Doktorsavhandling, Lunds universitet, 2008.

Håkansson, Gisela. The acquisition og negative placement in Swedish. *Studia Lingustica*, (43):47–58, 1989.

Håkansson, Gisela. Verb-initial sentences in the development of Swedish. *Working Papers*, (42):49–65, 1994.

Håkansson, Gisela. Swedish language learner corpora. I: Josefsson, Gunlög, Platzack, Christer, och Håkansson, Gisela, redaktörer, *The Acquisition of Swedish Grammar*, band 33 av *Language Acquisition and Language Disorders*, kapitel 2, ss 23–30. John Benjamins, 2003.

Håkansson, Gisela och Collberg, Sheila Dooley. The preference for modal+neg: an l2 perspective applied to Swedish L1 children. *Second Language Research*, 2(10):95–124, 1994.

Håkansson, Gisela och Hansson, Kristina. Grammatiska problem hos barn med språkstörning. I: Nettelbladt, Ulrika och Salameh, Eva-Kristina, redaktörer, *Språkutveckling och språkstörning hos barn*, kapitel 6, ss 171–198. Studentlitteratur, 2007.

Håkansson, Gisela och Nettelbladt, Ulrika. Developmental sequences in L1 (normal and impaired) and L2 acquisition of Swedish syntax. *International Journal of Applied Linguistics*, 3:131–157, 1993.

Hansson, Kristina. Early swedish verb morphology. I: Dressler, W., redaktör, *Studies in Pre- and Protomorphology*, ss 73–82. Austrian Academy of Sciences, 1997.

Hansson, Kristina, Håkansson, Gisela, och Nettelbladt, Ulrika. Carpe corpus. Om att bygga upp en talspråkskorpus. I: Byrman, Gunilla, Lindquist, Hans, och Levin, Magnus, redaktörer, *Korpusar i forskning och undervisning. Corpora in Research and Teaching. Rapport från ASLA:s höstsymposium, Växjö, 11–12 november 1999.*, ss 109–119, 1999.

Jörgensen, Nils. *Meningsbyggnaden i talad svenska*. Nummer C7 i Lundastudier i nordisk språkvetenskap. Studentlitteratur, 1976.

Josefsson, Gunlög. The use and function of non-finite root clauses in Swedish child language. *Language Acquisition*, (10):273–320, 2002.

Josefsson, Gunlög. Input and output: Sentence patterns in child and adult Swedish. I: Josefsson, Gunlög, Platzack, Christer, och Håkansson, Gisela, redaktörer, *The Acquisition of Swedish Grammar*, nummer 33 i Language Acquisition and Language Disorders, kapitel 5, ss 95–134. John Benjamins, 2003a.

Josefsson, Gunlög. Non-target structures and non-target uses in child and adult Swedish. I: Josefsson, Gunlög, Platzack, Christer, och Håkansson, Gisela, redaktörer, *The Acquisition of Swedish Grammar*, nummer 33 i Language

Acquisition and Language Disorders, kapitel 7, ss 155–194. John Benjamins, 2003b.

Josefsson, Gunlög och Håkansson, Gisela. The PP–CP parallelism hypothesis and language acquisition: Evidence from Swedish. I: Powers, Susan M. och Hamann, Cornelia, redaktörer, *The Acquisition of Scrambling and Cliticization*, band 26 av *Studies in theoretical psycholinguistics*, ss 397–422. Kluwer, 2000.

Kadin, Germund och Engstrand, Olle. Tonal word accents produced by Swedish 18- and 24-month-olds. I: *Papers from the 18th Swedish Phonetics Conference*, ss 67–70. Department of Linguistics, Göteborg University, 2005.

Karlsson, Susanna. *Positioneringsfraser i interaktion. Syntax, prosodi och funktion*. Doktorsavhandling, Göteborgs universitet, 2006.

Lange, Sven och Larsson, Kenneth. Syntactical development of a Swedish girl Embla, between 20 and 24 months of age, 1973.

Lange, Sven och Larsson, Kenneth. *Studier i det tidiga barnspråkets grammatik*. Doktorsavhandling, Stockholms universitet, 1977.

Loman, Bengt och Jörgensen, Nils. *Manual för analys och beskrivning av makrosyntagmer*. Nummer C1 i Lundastudier i nordisk språkvetenskap. Studentlitteratur, 1971.

Lundin, Barbro. *Bisatser i små barns språk. En analys av fem barns första bistaser*. Doktorsavhandling, Lunds universitet, 1987.

Lust, B., Flynn, S., och Foley, C. Elicited imitation. I: McDaniel, Dana, McKee, Cecile, och Cairns, Helen Smith, redaktörer, *Methods for Assessing Children's Syntax*, Language, Speech, and Communication. The MIT Press, 1996.

MacWhinney, Brian. *The CHILDES project: Tools for analyzing talk*. Lawrence Erlbaum Associates, Publishers, 3 utgåvan, 2000. URL `http://childes.psy.cmu.edu/`.

MacWhinney, Brian. *The CHILDES Project, Tools for Analyzing Talk – Electronic Edition Volume 1: Transcription Format and Programs, Part 1: The CHAT Transcription Format*. Carnegie Mellon University, 2006a. URL `http://childes.psy.cmu.edu/`.

MacWhinney, Brian. *The CHILDES Project, Tools for Analyzing Talk – Electronic Edition Volume 2: Transcription Format and Programs, Part 2: The CLAN Programs*. Carnegie Mellon University, 2006b. URL `http://childes.psy.cmu.edu/`.

McDaniel, Dana, McKee, Cecile, och Cairns, Helen Smith, redaktörer. *Methods for Assessing Children's Syntax*. The MIT Press, Cambridge, Massachusetts, 1996.

Mörnsjö, Maria. *V1 Declaratives in Spoken Swedish. Syntax, Information Structure, and Prosodic Pattern*. Doktorsavhandling, Lunds universitet, 2002.

Nelson, Katherine. Individual differences in language development: Implications for development and language. *Developmental Psychology*, 17(2): 170–187, 1981.

Nettelbladt, Ulrika. Some reflections on transcribing. I: Allwood, Jens och Gregersen, Frank, redaktörer, *Proceedings of the XIVth Scandinavian Conference of Linguistics and The VIIIth Conference of Nordic and General Linguistics, August 16–21 1993. Four special Sessions. Spoken Language*, ss 47–62, 1994.

Nettelbladt, Ulrika och Salameh, Eva-Kristina. Språkstörning hos barn. I: Nettelbladt, Ulrika och Salameh, Eva-Kristina, redaktörer, *Språkutveckling och språkstörning hos barn*, nummer 1, kapitel 1, ss 13–33. Studentlitteratur, 2007.

Norrby, Catrin. *Samtalsanalys. Så gör vi när vi pratar med varandra*. Studentlitteratur, Lund, 1996.

Pepinsky, Thomas, Demuth, Katherine, och Roark, Brian. The status of 'filler syllables' in children's early speech. I: *Proceedings of the 25th Annual Boston University Conference on Language Development*, ss 575–586, 2001.

Pesetsky, David och Torrego, Esther. T-to-C movement: Causes and consequences. I: Kenstowicz, Michael, redaktör, *Ken Hale. A Life in Language*, kapitel 11, ss 355–426. The MIT Press, 2001.

Peters, Ann. Language learning strategies: Does the whole equal the sum of the parts? *Language*, 53(3):560–573, 1977.

Peters, Ann och Strömqvist, Sven. The role of prosody in the acquisition of grammatical morphemes. I: Morgan, J. L. och Demuth, K., redaktörer, *Signal to Syntax:bootstrapping from speech to grammar in early acquisition*. Lawrence Erlbaum Associates, Publishers, Hillsdale, N.J., 1996.

Platzack, Christer. Functional categories and early Swedish. I: Meisel, J., redaktör, *The acquisition of verb placement: functional categories and V2 phenomena in language development*, ss 63–82. Kluwer, Dordrecht, 1992.

Platzack, Christer. The verb phrase. I: Josefsson, Gunlög, Platzack, Christer, och Håkansson, Gisela, redaktörer, *The Acquisition of Swedish Grammar*, band 33 av *Language Acquisition and Language Disorders*, kapitel 6, ss 135–154. John Benjamins, 2003.

Platzack, Christer. Uninterpretable features and EPP: a minimalist account of language buildup and breakdown. *Working Papers in Scandinavian Syntax*, (75):97–128, June 2005.

Platzack, Christer. Uninterpretable features and EPP: a minimalist account of language buildup and breakdown. I: Liceras, Juana M., Zobl, Helmut, och Goodluck, Helen, redaktörer, *Role of Formal Features in Second Language Acquisition*, ss 97–128. Routledge, 2007.

Platzack, Christer och Rosengren, Inger. On the subject of imperatives: a minimalist account of the imperative clause. *Journal of Comparative Germanic Linguistics*, (1):177–224, 1998.

Plunkett, Kim och Strömqvist, Sven. The aquisition of scandinavian languages. I: *The Crosslinguistic Study of Language Aquistion, Volume 3*, kapitel 7, ss 457–549. Lawrence Erlbaum Associates, Publishers, 1992.

Potts, Christopher och Roeper, Tom. The narrowing acquisition path: From expressive small clauses to declaratives. I: Progovac, Ljiljana, Paesani, Kate, Casielles, Eugenia, och Barton, Ellen, redaktörer, *The Syntax of Nonsententials. Multidisciplinary perspectives*, band 93 av *Linguistik Aktuell/Linguistics Today*, kapitel 7, ss 183–201. John Benjamins, Amsterdam/Philadelphia, 2006.

Progovac, Ljiljana. The syntax of nonsententials: Small clauses and phrases at the root. I: Progovac, Ljiljana, Paesani, Kate, Casielles, Eugenia, och Barton, Ellen, redaktörer, *The Syntax of Nonsententials. Multidisciplinary perspectives*, band 93 av *Linguistik Aktuell/Linguistics Today*, kapitel 2, ss 33–72. John Benjamins, Amsterdam/Philadelphia, 2006.

Progovac, Ljiljana, Paesani, Kate, Casielles, Eugenia, och Barton, Ellen, redaktörer. *The Syntax of Nonsententials. Multidisciplinary perspectives*, band 93 av *Linguistik Aktuell/Linguistics Today*. John Benjamins, Amsterdam/Philadelphia, 2006.

Richthoff, Ulla. En svensk barnspråkskorpus: Uppbyggnad och analyser, Oktober 2000.

Roeper, Tom. *The Prism of Grammar. How Child Language Illuminates Humanism*. A Bradfoord Book. The MIT Press, 2007.

Rowland, Caroline F. och Fletcher, Sarah L. The effect of sampling on estimates of lexical specificity and error rates. *Journal of Child Language*, (33): 859–877, 2006.

Santelmann, Lynn. *The aquisition of verb second grammar in child Swedish: Continuity of universal grammar in wh-questions, topicalization and verb raising*. Doktorsavhandling, Cornell University, August 1995.

Santelmann, Lynn. The acqusition of Swedish *wh*-questions. I: Josefsson, Gunlög, Platzack, Christer, och Håkansson, Gisela, redaktörer, *The Acquisition of Swedish Grammar*, band 33 av *Language Acquisition and Language Disorders*, kapitel 9, ss 261–307. John Benjamins, 2003.

Söderbergh, Ragnhild. Project child language syntax and project early reading, 1975.

Speidel, Gisela E. Phonological short-term memory and individual differences in learning to speak: a bilingual case study. *First language*, (13):69–91, 1993.

Strömqvist, Sven. *Make-belief through words. A linguistic study in children's play with a doll's house*. Doktorsavhandling, Göteborgs universitet, 1984.

Strömqvist, Sven och Ragnarsdóttir, Hrafnhildur. On the acquisition of verb argument structure. *Linguistics*, 38(3):523–543, 2000.

Strömqvist, Sven, Ragnarsdóttir, Hrafnhildur, och Richthoff, Ulla. Input and production in the early development of function words. I: B. Höhle, J. Weissenborn, redaktör, *Approaches to Bootstrapping: Phonological, Syntactic and Neurophysical Aspects of Early Language Acquisition.*, band 2 av *Language Acquisition and Language Disorders*, ss 157–177. John Benjamins, 2001.

Strömqvist, Sven, Richthoff, Ulla, och Andersson, Anders-Börje. Strömqvist's and Richthoff's corpora: A guide to longitudinal data from four swedish children. *Gothenburg Papers in Theoretical Linguistics*, 66, 1993.

Stromswold, Karin. Analyzing children's spontaneous speech. I: McDaniel, Dana, McKee, Cecile, och Cairns, Helen Smith, redaktörer, *Methods for Assessing Children's Syntax*, Language, Speech, and Communication, kapitel 2, ss 23–53. The MIT Press, Cambridge, Massachusetts, 1996.

Teleman, Ulf. *Manual för grammatisk beskrivning av talad och skriven svenska*. Nummer C6 i Lundastudier i nordisk språkvetenskap. Studentlitteratur, 1974.

Teleman, Ulf, Hellberg, Staffan, och Andersson, Erik. *Svenska akademiens grammatik*. Norstedts Ordbok, 1999.

Tomasello, Michael och Stahl, Daniel. Sampling children's spontaneous speech: how much is enough? *Journal of Child Language*, (31):101–121, 2004.

Waldmann, Christian. *Input och output: Ordföljd i svenska barns huvudsatser och bisatser*. Doktorsavhandling, Lunds universitet, 2008.

Westergaard, Marit. *The Development of Word Order in Norwegian Child Language: The Interaction of Input and Economy Principles in the Acquisition of V2*. Doktorsavhandling, University of Tromsø, 2005.

Wexler, Ken. Optional infinitives, head movement and the economy of derivations. I: Hornstein, Norbert och Lightfoot, David, redaktörer, *Verb movement*, ss 305–350. Cambridge University Press, Cambridge, 1994.

NORDLUND
Småskrifter från Nordiska språk i Lund
ISSN 0281-5427

1. *Tre uppsatser* (1982)
a. Lena Ekberg: Vara och – en markör för subjektets 'icke-plats'.
b. Ulf Teleman: What Swedish is like. A mini-grammar for those who need it.
c. Börje Tjäder (utg.): Rapport från ett gränsland.

2. Ulf Teleman (red.) (1983): *Skrivundervisning på universitetsnivå. Rapport från det första ämnesdidaktiska seminariet i nordiska språk, Lund 1983*.

3. *Blandade studier* (1984)
a. Ulf Teleman: Har tal- och skriftspråk olika grammatiker?
b. Jan Svensson: Recension av C Platzack, Modern grammatisk teori.
c. Åke Pettersson: Frekvens och kompetens. Om elevers språkförmåga.
d. Ulf Teleman: Om att förstå lagen. Recension av Britt-Louise Gunnarsson, Lagtexters begriplighet.
e. Christer Platzack: Kommentar till Jan Svenssons recension.

4. Lars Wollin (red.) (1984): *Tiällmannstudier*.

5. Ulf Teleman (red.) (1984): *Språkhistoriens mening. Rapport från det andra ämnesdidaktiska seminariet i nordiska språk, Lund 1984*.

6. *Berätta och argumentera. Tre analyser av skoluppsatser*. (1984)

7. Ulf Teleman (red.) (1985): *Samhällsutveckling och språkbrukshistoria under 1700-talet*.

8. Lars Svensson (1985): *Förteckning över bibliografier rörande 1700-talets bokproduktion. Ett urval*.

9. Jan Svensson (red.) (1985): *Vad ska vi med grammatiken till? Rapport från det tredje ämnesdidaktiska seminariet i nordiska språk, Lund 1985*.

10. *Fyra uppsatser* (1986)
a. Bengt Audén: Studerandecentrerad språkhistoria på ämneslärarlinjen.
b. Eva Mårtensson: Det nya niandet.
c. Carin Sandqvist: Recension av Lars Melin och Sven Lange, Att analysera text.
d. Ulf Teleman: On visual feedback in writing.

11. Carin Sandqvist (red.) (1986): *Stilistik: Vad, varför och vartill? Rapport från det fjärde ämnesdidaktiska seminariet i nordiska språk, Lund 1986*.

12:1. Lars Wollin (1987): *Nysvensk bibliografi. Del 1. Sorterad efter språklig beskrivningsnivå, med författarregister. Editionsförteckning. Del 2. Sorterad efter ämnesinriktning, med författarregister*.

13. (1987) *Två uppsatser om språkbrukets formalisering*.
a. Cecilia Falk: Kommunikationsmönster inom familjen under efterkrigstiden.

b. Eva Mårtensson: När myndigheterna blev du med oss.

14. Eva Mårtensson & Jan Svensson (1988): *Offentlighetsstruktur och språkförändring*.

15. *Två uppsatser* (1989)
a. Sven Lange: Om bisatser i barnspråk.
b. Bo-A Wendt: Psalmspråket i nya kläder. Den språkliga bearbetningen av äldre psalmer i 1986 års psalmbok.

16. Tor G. Hultman (1990): *Barnet – han eller den? Om genuskongruensen vid ordet barn i gymnasisters språkbruk.*

17. Miranda Landen (1991): *Den fiktiva dialogen. Tal- och tankegestaltning i Hjalmar Söderbergs "Den allvarsamma leken"*.

18. Per Lagerholm (1991): *Språk i massmedierna – en forskningsöversikt*.

19. Jan Svensson (red.) (1994): *Journalistik i förvandling. Om språk och texter i Expressen*.

20. Lisa Christensen (1995): *Svenskans aktionsarter. En analys med särskild inriktning på förhållandet mellan aktionsarten och presensformens temporala referens*.

21. Ulrika Jakobsson (1996): *Familjelika betydelser hos* stå, sitta *och* ligga. *En analys ur den kognitiva semantikens perspektiv*.

22:1. Katarina Lundin & Robert Zola Christensen (2001): *Grannspråksförståelse i Öresundsregionen år 2000 – gymnasisters läsförståelse*.

22:2. Nils Jörgensen & Eva Kärrlander (2001): *Grannspråksförståelse i Öresundsregionen år 2000 – gymnasisters hörförståelse*.

23. Bo Rutström (2002): *Land och fæ. Strukturellt-rättsfilologiska studier i fornnordiskt lagspråk över beteckningar för egendom i allmänhet med underkategorier*.

24. Lena Ekberg (2004): *Tre uppsatser om semantisk förändring hos relationella lexem*.

25. Ulf Teleman (2005): *Om r-fonemets historia i svenskan*.

26. Dominika Skrzypek (2005): *The Decline of Nominal Inflection in Old Swedish. The Loss of Dative Case*.

27. Lena Ekberg (red.) (2007): *Språket hos ungdomar i en flerspråkig miljö i Malmö*.

28. Åsa Wikström (2008): *Den finita satsen i små barns språk*.

Skrifterna kan, med reservation för att vissa äldre skrifter är slutsålda, beställas från institutionen för självkostnadspris. Beställningen sker per brev eller e-brev till:

Nordiska språk, Språk- och litteraturcentrum, Box 201, 221 00 Lund
Exp_lk1@sol.lu.se