



MEDICINSKA FAKULTETEN

Lunds universitet

Institutionen för logopedi, foniatri och audiologi

Reliabiliteten ur tre aspekter hos ett nytt svenskt grammatiktest

**Sara Clarén
Sara Hellsten**

**Logopedutbildningen, 2003
Vetenskapligt arbete, 20 poäng**

Handledare: Kristina Hansson och Karl Radeborg

SAMMANFATTNING

I Sverige råder stor brist på standardiserade logopediska instrument. Det är därför angeläget att ta fram sådana. Ett grammatiktest är under färdigställande av Kristina Hansson och Ulrika Nettelbladt vid institutionen för logopedi, foniatri och audiologi vid Lunds Universitet. Grammatiktestet skall standardiseras och beräkningarna av reliabiliteten är ett led i detta arbete. Det grundar sig på modern forskning kring grammatiska svårigheter hos svensktalande barn med språkstörning och bygger på processbarhetsteorin. Testet är avsett att användas för att diagnostisera barn med grammatisk språkstörning.

Vi har i denna uppsats undersökt Grammatiktestets reliabilitet ur tre aspekter; test-retest-reliabiliteten, interbedömarreliabiliteten samt Chronbachs alfa (homogeniteten).

Vår undersökning utfördes på 47 normalspråkiga svenska barn i förskoleåldern. Varje barn testades av samma testledare två gånger med en månads mellanrum. Barnens svar poängsattes av två oberoende bedömare. Båda bedömarna poängsatte alla barn. Transkriptionerna av barnens svar stämde i de flesta fall väl överens mellan de två bedömarna. I de fall bedömningarna av svaren var olika berodde det oftast på olikheter i bedömarnas transkriptioner.

Resultatet visar att reliabiliteten är hög ur alla tre aspekterna (test-retest-reliabilitet 0,70, interbedömarreliabilitet 0,99, Cronbachs alfa 0,84). Detta innebär att en förbättring av ett barns testresultat från en gång till en annan i hög grad beror på att barnet utvecklats, och inte på slumpen. Det visar också att man kan jämföra och dra slutsatser utifrån två olika personers bedömningar av samma barn, och att testet i hög grad mäter bara en faktor (troligtvis grammatisk förmåga).

Resultaten jämfördes med reliabiliteten hos Reynell developmental language scales (Reynell, 1987). Grammatiktestet klarar en sådan jämförelse bra.

TACK

till våra handledare, Kristina Hansson och Karl Radeborg för att ni alltid haft tid för oss

till Erik Hellsten för hjälp med de datorprogram vi använt

till Johan Pettersson för datorsupport och för att vi fått låna hans dator under arbetet

till barn, föräldrar och personal på berörda förskola i Kävlinge kommun för all hjälp.

INNEHÅLLSFÖRTECKNING

1. INLEDNING	4
2. BAKGRUND	5
2.1 Testning.....	5
2.1.1 Skäl att testa barn	5
2.1.2 Förutsättningar och svårigheter vid testning av barn	6
2.2 Språktestning.....	6
2.2.1 Språkbedömningar	6
2.2.2 Screeningtest och diagnostiserande test	7
2.2.3 Brist på standardiserade test.....	7
2.2.4 Elicitering	7
2.2.5 Svårigheter vid elicitering av grammatiska strukturer	8
2.3 Grammatiktestet	8
2.4 Psykometriska egenskaper	9
2.4.1 Standardisering.....	9
2.4.2 Tillförlitlighet—reliabilitet och validitet.....	10
2.4.3 Reliabilitet	10
3. SYFTE.....	13
4. METOD.....	14
4.1 Testmaterial	14
4.2 Försökspersoner	14
4.3 Tillvägagångssätt.....	15
5. RESULTAT	17
5.1 Test-retest-reliabilitet	17
5.2 Interbedömarreliabilitet.....	17
5.2.1 Transkriptionsöverensstämmelse	17
5.2.2 Bedömningsöverensstämmelse	18
5.3 Chronbachs alfa.....	18
6. DISKUSSION	19
6.1 Test-retest-reliabilitet	19
6.2 Interbedömarreliabilitet.....	19
6.3 Chronbachs alfa.....	20
6.4 Klinisk tillämpning.....	20
6.5 Slutsatser	21
7. REFERENSER.....	22
BILAGOR.....	24

1. INLEDNING

Vår uppgift i föreliggande studie är att undersöka reliabiliteten hos ett nykonstruerat grammatiktest för svenska. Reliabiliteten undersöks ur tre aspekter; test-retest-reliabilitet, interbedömarreliabilitet och Chronbachs alfa.

En viktig del av det logopediska arbetet är att bedöma barns språk. Det görs både i spontana observationer och med hjälp av test. I Sverige har det traditionellt funnits en förkärlek för kvalitativa bedömningsmetoder när det gäller kliniska logopediska bedömningar av barn. De kvantitativa standardiserade test som funnits har oftast inte varit baserade på det svenska språket, utan har varit översatta från andra språk. I USA däremot har man i stor utsträckning använt sig av kvantitativa test, ofta standardiserade, när man bedömt barns språk (Nettelblatt, 1995). För att kunna hålla en god kvalitet på den kliniska logopedverksamheten behövs kvantitativa språktest utformade utifrån det svenska språket. Risken är annars att de logopediska bedömningarna blir alltför subjektiva och ej jämförbara.

Många barn som kommer till logopedmottagningarna har grammatiska problem. Behovet av ett svenskt, standardiserat och reliabelt grammatiktest är därför stort. Kristina Hansson och Ulrika Nettelblatt vid institutionen för logopedi, foniatri och audiologi vid Lunds universitet har utformat ett svenskt grammatiktest där standardisering och normering håller på att sammanställas. Testet kallas än så länge Grammatiktestet, vilket också är det namn vi kommer att använda i vår uppsats.

2. BAKGRUND

2.1 Testning

Ett test är ett stickprov ur verkligheten, som används för att man på ett systematiskt sätt ska kunna registrera en viss aspekt av en människas beteende. Att studera hela det beteende som avses skulle vara alltför tidsödande. Istället tar man ut ett stickprov som ska variera i svårighetsgrad och som ska belysa olika sidor av det beteende man är intresserad av (Holmquist, 1983). Testsituationen är inte identisk med den verklighet där beteendet normalt förekommer, men måste vara tillräckligt lik den för att slutsatser ska kunna dras. Testet får inte vara godtyckligt, utan måste ha en viss grad av objektivitet. Med det menas att man eftersträvar att administration, poängsättning och tolkning av resultatet ska vara oberoende av testledarens subjektiva bedömningar (Anastasi & Urbina, 1997).

Att använda test för att bedöma mänskliga förmågor är en gammal företeelse. Redan i Kina på 200-talet f Kr testade man människor bland annat för att hitta den lämpligaste placeringen inom försvaret. De tidiga experimentella psykologerna i Europa under 1800-talet var intresserade av att formulera generella principer för hur människan fungerar. Inga avvikelser mellan testpersonerna sågs som normala. Man var till exempel intresserad av människans reaktionstid ur olika aspekter. Då testade man för att hitta likheter mellan människor. Skillnader sågs som fel i testet eller hos testpersonen (Anastasi & Urbina, 1997). Numera används test för att hitta skillnader mellan olika personer, eller olikheter i en persons beteende vid skilda tillfällen. Det är det faktum att människor är olika som är grunden för testning idag (Holmquist, 1983). Därför vill man konstruera test som ger en så stor interindividuell spridning av testresultaten som möjligt, så att individerna fördelar sig över i stort sett samtliga testpoängvärden (Smedler, 1993).

Det finns olika typer av test, och dessutom olika sätt att indela test i grupper. Ett sätt att dela in test är i direkta och indirekta test. Direkt testning är när man mäter precis det man är intresserad av, till exempel någons ordkunskap i ett glosförhör. Då man vill undersöka något som inte är direkt mätbart, kan man mäta något annat, och sedan dra slutsatser utifrån det. Det man mäter behöver inte nödvändigtvis ha likheter med det beteende man vill uttala sig om, men måste ha ett bevisat samband med det sökta. Detta är ett indirekt sätt att testa. Det kan handla om till exempel intelligens (Anastasi & Urbina, 1997). Gränsen mellan direkt och indirekt testning är inte alltid helt tydlig. Det handlar mer om ett kontinuum mellan dessa två.

2.1.1 Skäl att testa barn

Barn är mycket olika som individer. I klinisk verksamhet arbetar man med att diagnostisera beteendevikelser för att avgöra behov av insatser. I detta arbete har man stor nytta av att kartlägga barns olikheter, vilket man bland annat gör med hjälp av test. Man testar för att man vill "skaffa sig en bild av barnets intrapsykiska fungerande i olika avseenden" (Smedler, 1993: 11). Anledningen är ofta ett behov av fördjupad individriktad diagnostik. Men att testa för diagnostikens egen skull är meningslöst. Resurser måste finnas för att följa upp testresultatet och barnets föräldrar måste från början vara införstådda med testningens syfte, bland annat för att kunna ta emot resultatet på ett bra sätt. I arbete med vuxna används ofta intervju för att kartlägga olika beteenden. Intervju fungerar dåligt med barn, då det ligger för långt ifrån deras spontana uttryckssätt att analysera sig själva. Varför kan man då inte bara vara med barnet och observera beteendet ifråga? Det finns flera skäl till detta. En anledning

att använda ett test är att det är ett kontaktmedel, som stimulerar det man vill observera. Man får alltså fram information av förutbestämd art. En annan anledning är att man använder ett material som ofta är nytt för barnet, vilket stimulerar nyfikenhet och spontanitet och får barnet att glömma eventuell blyghet. Eftersom testaren använder samma material med flera barn, och under så lika förhållanden som möjligt, får han/hon lättare att se nyanserna i barnens beteende (Smedler, 1993). Om testet dessutom är standardiserat hjälper detta testledaren till en mer objektiv bedömning av beteendet.

2.1.2 Förutsättningar och svårigheter vid testning av barn

Vid testning förutsätter man en viss stabilitet över tid och rum hos det testade beteendet. Man vill att testresultaten ska vara giltiga inte bara här och nu. Detta antagande om stabilitet stämmer inte riktigt på barn, då de befinner sig i ständig och dynamisk utveckling. Man måste därför ha en någorlunda bra bild av hur utvecklingen ser ut, och hur olika faktorer förhåller sig till varandra och ha förväntningar om hur den observerade egenskapen förhåller sig till andra aspekter av utvecklingen. Har man inte det kan man aldrig vara riktigt säker på vad det är man egentligen mäter när man gör en indirekt testning (Smedler, 1993).

Det finns fler svårigheter med att testa barn. Speciellt förskolebarn påverkas lättare än vuxna av vem testledaren är och av situationella influenser. Resultatet påverkas också av vad testpersonen gjort precis före testningen, speciellt om aktiviteten lett till trötthet eller emotionell instabilitet. Om barnet inte svarar på en uppgift behöver det inte betyda att han/hon inte klarar av den. Många andra slumpfaktorer kan spela in i barnets svarsbenägenhet. Dessutom påverkar testledaren resultatet både genom sitt beteende i testsituationen och genom sina egna förväntningar på barnets prestation (Anastasi & Urbina, 1997).

2.2 Språktestning

2.2.1 Språkbedömningar

Inom klinisk logopedisk verksamhet spelar språkbedömningar en stor roll. De ska utgöra underlag för såväl diagnostik som för beslut om intervention. Språkbedömningar kan göras på flera sätt. Som logoped arbetar man ofta problemorienterat. Utifrån remissen angående barnet och det första intrycket man får vid mötet med barnet skaffar man sig en hypotes om problematiken. Denna verifierar eller förkastar man sedan med hjälp av en språkbedömning. Ett annat sätt att utföra en språkbedömning, kan vara att använda ett bestämt test eller testbatteri för alla barn och utifrån de resultat man får fram besluta om vidare åtgärder.

En språkbedömning kan se ut på olika sätt. Den kan baseras på allt ifrån observationer av spontan kommunikation till standardiserade test. Både observationer och test är viktiga och kompletterar varandra. Ett annat sätt att se på språkbedömningar är ur aspekten kvalitativt kontra kvantitativt. Enkelt kan man säga att test mäter språket på ett kvantitativt sätt, medan observationer ger en kvalitativ bild av barnets språk (Nettelbladt & Hansson, ms.). I klinisk verksamhet är det inte alltid så enkelt uppdelat. Test kan också användas på ett kvalitativt sätt, för att ytterligare kartlägga ett barns specifika förmåga. Ett exempel på ett kvalitativt test är Nelli (Holmberg & Sahlén, 1986). Man frångår då standardiseringen, och det är viktigt att hålla i minnet att man inte längre gör en kvantitativ bedömning. Det finns också standardiserade metoder för språkobservationer som gör dem kvantitativa till sin natur, till exempel beräkandet av MLU.

2.2.2 Screeningtest och diagnostiserande test

När man ska testa ett barns språkliga förmåga finns det generellt två typer av test, som har två olika syften; screeningtest och diagnostiserande test. Med den första typen avgör man om det alls finns en språklig avvikelse hos barnet och om man ska gå vidare med en utredning. Screeningtest utförs vanligen av personer som inte har en specifikt språklig utbildning, till exempel personal på BVC. Barn som får ett avvikande resultat remitteras vidare till specialister för en mer utförlig utredning. För att specialisten sedan ska kunna ställa en diagnos behövs diagnostiserande test för den eller de språkliga domäner där avvikelse misstänks (Lavesson, 1999). Testning utgör som tidigare nämnts bara en del av diagnosunderlaget och resultatet måste sättas i relation till spontant beteende innan diagnos ställs. Test är trots allt bara stickprov på vad barnet presterar.

2.2.3 Brist på standardiserade test

Inom det logopediska området i Sverige är det ont om standardiserade test. Bristen på standardiserade test gör att det finns en risk att många logopediska bedömningar inte blir så objektiva (Lavesson, 1999). Många av de test som används idag är översatta från något annat språk, och om de alls är standardiserade är detta gjort på originalspråket. Att använda översatta test med tillhörande ursprungliga normer är inte alltid så lyckat. Språkliga strukturer som är vanliga i ett språk behöver inte vara det i ett annat språk. Hur pass svår en struktur är kan också variera från språk till språk. Därför kan översatta test vara inriktade på strukturer som inte nödvändigtvis är relevanta i en språkbedömning på svenska. Det krävs därför en grundlig bearbetning och anpassning och en ny standardisering om ett test översätts. Ett exempel på ett test som anpassats och standardiserats till svenska är TROG (Bishop, 1989; Holmberg & Lundälv, 1998).

Att det finns så få test som är standardiserade på svenska beror bland annat på att det krävs väldigt många testpersoner för att genomföra en standardisering. Standardiseringen är svår genomförbar då den svensktalande populationen inte är så stor, och dessutom är utspridd över ett stort geografiskt område. Det krävs också en stor arbetsinsats för att standardisera ett test (Nettelbladt & Hansson, ms.). Eftersom grammatik uppfattas som speciellt svårundersökt av många logopedier (Lavesson, 1999) kan det här ändå finnas ett extra stort behov av ett standardiserat test med noggranna instruktioner för att de logopediska bedömningarna ska vara välgrundade.

2.2.4 Elicitering

För att locka fram språkproduktion hos barnet man testar finns olika så kallade eliciteringsstrategier att tillgå, till exempel imitation, benämning, meningskomplettering, modellmening, frågeord och återberättande. Vilken strategi man väljer har visat sig kunna påverka barnets respons (Nettelbladt, 1995). Vissa strategier är mer styrande än andra. Den mest styrande är imitation, då barnet ombeds repetera vad testledaren säger. Eftersom barn med språkstörning kan vara talgenererade och fåordiga kan det vara nödvändigt att använda en styrande strategi (Nettelbladt & Hansson, ms.). Med en mindre styrande strategi riskerar man nämligen att ställa alltför höga krav på barnets språkproduktion, och därigenom inte få något svar alls. Man bör också tänka på att eliciteringen måste ske på ett pragmatiskt rimligt sätt, för att få bra svar från barnet. Barnet får inte tycka att det är en oberättigad eller konstig fråga (Gerken, 2000). Håkansson och Hansson (2000) har visat att meningskomplettering är en eliciteringsstrategi som fungerar bättre med förskolebarn än vad eliciterad imitation gör.

2.2.5 Svårigheter vid elicitering av grammatiska strukturer

Som tidigare nämnts uppfattas testning av grammatik av många logopedier som svårare än testning av andra språkliga områden. Detta beror på att det ställs stora krav på eliciteringen (Lavesson, 1999). Det kan vara svårt att locka fram grammatiska strukturer med en styrande eliciteringsstrategi, speciellt om det handlar om längre strukturer där det krävs att barnet producerar en fullständig mening. Med en mindre styrande strategi blir å andra sidan svaren mer varierande och därmed mer svårbedömda (Nettelbladt & Hansson, ms.). Man kan behöva variera eliciteringsstrategi då man vill locka fram produktion av grammatik, beroende på vilken struktur man är ute efter.

Det finns fler hinder på vägen när man ska testa grammatik. De barn som oftast utsätts för sådan testning är barn med språkstörning. Det är en väldigt heterogen grupp. Vissa av dessa barn har problem både med språkförståelsen och språkproduktionen, medan andra bara har problem med produktionen. För att veta var problemet ligger krävs att såväl förståelse som produktion av grammatiska strukturer kartläggs. Barn med språkförståelseproblem är svårtestade, eftersom de kan ha svårigheter att förstå vad som förväntas av dem (Leonard, 1996).

Barn med språkstörning har också olika väl utvecklad fonologi. Stora fonologiska problem gör det svårt att bedöma den grammatiska förmågan, eftersom det kan vara svårt att förstå vad barnet säger. Det kan också vara svårt att avgöra om exempelvis ett utelämnande har sin grund i fonologiska eller grammatiska problem (Hansson, 1998; Leonard, 1996).

Barnen har också olika väl utvecklad lexikal förmåga. En god lexikal förmåga kan i vardagen i viss mån kompensera och dölja en bristande grammatisk förståelse. Genom att barnet förstår alla de ingående lexikala elementen i satsen det hör belastas dess förståelse av syntax och grammatiska morfem inte lika hårt för en adekvat tolkning av satsen. Dessa grammatiska problem kan synliggöras vid en testning. En bristande lexikal förmåga eller ordmobiliseringsproblem, å andra sidan, gör det svårt att bedöma grammatisk förståelse och produktion, eftersom grammatik förutsätter ordkunskap. Då man testar barn kan man heller inte vara säker på att barnets maximala språkliga förmåga kommer fram, då de sällan förstår att de förväntas göra sitt yttersta (Leonard, 1996).

2.3 Grammatiktestet

I vår studie har vi använt Grammatiktestet, som är ett nytt test utformat av Kristina Hansson och Ulrika Nettelbladt (under utarbetande). Testet består av 45 testuppgifter som testar 13 olika grammatiska strukturer. Testet grundar sig på aktuell forskning kring grammatiska svårigheter hos svensktalande barn med språkstörning. Många av de strukturer som testas har visat sig speciellt svåra att behärska för barn med språkstörning.

I en studie av Leonard, Salameh och Hansson (2001) fann man att svensktalande barn med språkstörning har svårt att använda genitiv (*kattens*), obestämd artikel (*en katt*) samt kongruens i nominalfras (*ett gult hus*) jämfört med både åldersmatchade och språkmatchade kontrollgrupper. Med språkmatchad grupp menas att barnen har samma MLU oavsett ålder. Barnen utelämnar eller ersätter dessa strukturer i högre grad än vad barnen i kontrollgrupperna gör. Även pluralis (*katter*) testades i denna studie, men skillnader kunde påvisas endast jämfört med den åldersmatchade kontrollgruppen.

Verbmorfologi och ordföljd hos svenska barn med språkstörning har undersökts i två olika studier (Hansson, Nettelbladt & Leonard 2000; Hansson & Leonard, 2003). Båda studierna visade att de största problemen för barnen med språkstörning, jämfört med både ålders- och språkmatchade kontrollgrupper, består i utelämnande av kopula i presens (*är*) samt i den regelbundna preteritumböjningen (*klippte*). Barnen med språkstörning hade även brister i ordföljd i negerade huvudsatser (*Han får inte måla*). De hade även problem med ordföljden i meningar som inte börjar med subjekt, det vill säga topikalisering (*Sen vilade de*). De hade signifikant fler ordföljdsfel än den språkmatchade gruppen. Däremot fann man i studierna bara små skillnader vad gäller oregelbunden preteritumböjning (*såg*) och i presensböjning (*klipper*), och endast mellan barnen med språkstörning och den åldersmatchade kontrollgruppen (Hansson, Nettelbladt & Leonard, 2000). Svenskspråkiga barn med språkstörning använder bisatser mer sällan än yngre barn som ligger på samma språkliga nivå. Bisatser dyker normalt upp före två års ålder. De första som kommer är vanligtvis relativbisatser (*... som dricker*) (Håkansson & Hansson, 2000).

Grammatiktestet utgår även från den så kallade processbarhetsteorin (Pienemann & Håkansson, 1999). Denna teori innebär att grammatiska strukturer kan delas in i fem olika nivåer, som ställer olika höga krav på språklig bearbetning. På första nivån bearbetar barnet ordet som en helhet. På nivå två omfattar bearbetningen ordets olika morfem, alltså bearbetning inom ordet, som till exempel tempus- eller numerusböjning. Barnet visar att det kan använda samma ord med olika form beroende på vad som menas (*klipper-klippte*). Den tredje nivån innebär att barnet kan utbyta information inom frasen, till exempel kongruera adjektivet med substantivet i en nominalfras (*ett gult hus*). På fjärde nivån kan barnet bearbeta information mellan olika fraser, och till exempel negera huvudsats eller kongruera adjektiv i predikativ (*huset är stort*). Den femte och högsta nivån innebär att barnet klarar av utbyte av information mellan olika satser, och kan skilja på huvudsats (*Han kan inte*) och bisats (*... att han inte kan*).

Processbarhetsteorin ligger till grund för testuppgifternas ordningsföljd i testet, så att nivåerna testas i stigande svårighetsgrad. Den har även påverkat urvalet av målord. Enligt processbarhetsteorin är det viktigt att få fram flera former av ett visst ord, för att se att barnet inte bara memorerat en viss form av ordet, och därmed befinner sig på nivå ett. Om flera former av ordet förekommer visar det att barnets bearbetande är på morfemnivå eller högre. Därför återkommer samma substantiv, adjektiv och verb i olika former i testet.

2.4 Psykometriska egenskaper

2.4.1 Standardisering

För att kunna jämföra flera barns resultat på samma test, eller samma barns resultat på olika test på ett objektivet sätt krävs att testen är standardiserade. Detta innebär att man på flera sätt försöker göra omständigheterna kring testningen så lika som möjligt från gång till gång:

- Testsituationen ska vara likadan vid alla testningar.
- Bedömningarna av de enskilda svaren ska göras på samma sätt av olika bedömare.
- Tolkningen av barnets resultat på testet ska ske på ett objektivet sätt.
- Testet ska vara normerat.

För att man skall kunna dra slutsatser utifrån resultaten krävs att testet är undersökt med avseende på reliabilitet och validitet och att dessa är goda (Holmquist, 1983). Det är dock viktigt att komma ihåg att inte ens en standardiserad testning någonsin kan bli helt objektiv.

För att kunna uttala sig om huruvida ett visst barns testresultat är normalt eller avvikande behöver man jämföra med hur andra barn presterar på samma test, det vill säga testet ska vara normerat. Det innebär att testet provats ut på ett stort antal individer. De som testet provats ut och normerats på skall som grupp ur vissa aspekter likna den grupp av individer man vill använda testet på, till exempel i fråga om ålder och modersmål när det gäller språktestning. Normeringsgruppen skall samtidigt vara ett representativt stickprov ur populationen (Smedler, 1993). Normerna man får fram visar inte bara gruppens genomsnittliga prestation utan också hur poängfördelningen ser ut över hela skalan av möjliga resultat. När det gäller barntest måste en sådan fördelningskurva göras för varje åldersgrupp. Om barnet inte presterar ett åldersadekvat resultat kan man i normerna se vilken åldersgrupp barnet liknar vad gäller det testade beteendet. Det gäller att hålla i huvudet att normering är socioekonomiskt och kulturellt specifik och att normerna dessutom kan förändras över tid. En normering gjord på 50-talet behöver därför inte vara giltig idag (Smedler, 1993). Beroende på vad testet mäter förändras normerna olika snabbt. Exempel på normer som är tämligen stabila skulle kunna vara svenska femtonåringars längd. Exempel på normer som förändrats väsentligt är svenska femtonåringars kunskaper i engelska språket.

2.4.2 Tillförlitlighet—reliabilitet och validitet

Även om det alltid finns faktorer i omgivningen som påverkar testningen vill man veta att man mäter det avsedda, och med stor precision. Man ska kunna kräva av ett test att det ger ett tillförlitligt resultat, det vill säga att det har en hög reliabilitet och en god validitet. En hög reliabilitet innebär att det inte är oidentifierade tillfälligheter som påverkar resultatet i alltför hög grad (Holmquist, 1983). Man bör dokumentera i vilken grad testresultatet är fritt från slumppåverkan, det vill säga hur tillförlitligt det är. För kvantitativa test kan tillförlitligheten uttryckas som en reliabilitetskoefficient (Nettelbladt & Hansson, ms.).

Validitet innebär att testet mäter det det är avsett att mäta. Hög validitet betyder alltså att testet i hög grad mäter det avsedda, låg validitet att det i praktiken till viss del också mäter något annat än det avsedda. Omständigheter specifika för testsituationen, jämfört med hur beteendet fungerar i verkligheten, påverkar validiteten negativt. Ett exempel på detta kan vara att testsituationen ofta kräver mer uthållighet än en "verklig" situation, och därmed mäter testet till viss del uthållighet, och inte enbart det eftersökta beteendet (Anastasi & Urbina, 1997). Om slumpfaktorer i hög utsträckning påverkar testresultatet, det vill säga om reliabiliteten är låg, är aldrig resultatet ett bra mått på det man vill mäta. Hög reliabilitet är alltså en förutsättning för att få en hög validitet. Reliabiliteten sätter så att säga den övre gränsen för möjlig validitet. Hög reliabilitet garanterar dock inte en hög validitet (Smedler, 1993). Om testet är tillräckligt reliabelt och dessutom har hög validitet kan man dra giltiga slutsatser utifrån resultatet.

2.4.3 Reliabilitet

Reliabilitet innebär att det finns en stabilitet i poängen en viss person får när han bedöms två gånger med samma test eller med likvärdiga test (Anastasi & Urbina, 1997). Stabiliteten kan också gälla de poäng två oberoende bedömare ger personen utifrån en och samma testning. Reliabiliteten är vanligtvis lägre för ett test ämnat för barn, än för ett vuxentest. Detta har att göra med barns föränderlighet och spontanitet. Reliabiliteten uttrycks i de flesta fall som en korrelationskoefficient, och kan anta värden mellan noll och ett. Ett värde över 0,70 kan sägas

vara bra för ett psykometriskt test och i synnerhet då man testar barn (Vogt, 1999). En annan sak som påverkar hur hög reliabiliteten blir är graden av samstämmighet mellan testuppgifterna, det vill säga i vilken mån de mäter samma sak. Denna samstämmighet kallas homogenitet. En hög homogenitet är en förutsättning, men ingen garanti, för en hög reliabilitet. Man kan öka reliabiliteten genom att öka antalet uppgifter i testet, förutsatt att testet är homogent. För att bli säkrare på resultatet ställer man då fler frågor som handlar om samma sak (Holmquist, 1983). Reliabiliteten påverkas också av spridningen av testpoängen. För att få en hög reliabilitet ska testet differentiera mellan olika grad av förmåga, det vill säga barnens testpoäng ska vara spridda över i stort sett hela poängskalan. Antalet testpersoner i beräkningarna skall också vara högt för att man ska få en hög reliabilitet.

För att beräkna reliabiliteten kan man gå tillväga på olika sätt. Man kan använda test-retest-metoden, interbedömmarmetoden, parallelltestmetoden, split-half-metoden eller homogenitetsmetoden (Chronbachs alfa). Det finns inget sant reliabilitetsvärde som man med hjälp av de olika metoderna försöker nå. De olika sätten att estimera reliabilitet kan sägas avspegla olika aspekter av reliabiliteten. Beroende på hur testet är tänkt att användas, väljer man den metod eller de metoder som passar bäst. Då man i logopedisk verksamhet ofta testar om barn och detta kan ske med olika testledare kommer vi i den här uppsatsen använda test-retest-metoden och interbedömmarmetoden. För test som är avsedda att endast mäta en sak är homogenitetsmetoden den idag mest använda metoden för att mäta reliabilitet. Därför kommer vi att använda även denna. Dessa tre metoder förklaras mer ingående nedan.

Test-retest-metoden innebär att man upprepar testningen och kontrollerar hur väl de två resultaten korrelerar med varandra. De två testsituationerna skall förutom vad gäller tidpunkten vara helt identiska (Smedler, 1993). Tiden man låter gå mellan de två testningarna är av stor betydelse när man testar barn. För kort tid innebär att en inlärningsfaktor tillkommer, för lång tid kan innebära att barnets beteende hunnit förändrats på grund av utveckling. Testet ska vara lika svårt för barnet båda gångerna (Holmquist, 1983).

Interbedömmarmetoden innebär att två eller flera av varandra oberoende bedömare gör bedömningar av en och samma testning, och sedan kontrolleras samstämmigheten mellan dessa bedömningar (Smedler, 1993). Praktiskt kan detta genomföras genom att flera bedömare sitter med i rummet under testningen, eller att de efteråt gör sin bedömning utifrån en inspelning av testningen. Låg interbedömmarreliabilitet visar att testet mäter olika saker beroende på vem som utvärderar svaren, och/eller att bedömare har olika bedömningskriterier. Det är lätt att underskatta subjektiviteten i människors bedömningar. Interbedömmarreliabiliteten för ett test bör därför rapporteras så snart en människa är inblandad i mätning eller registrering av testpersonernas svar. Endast då en dator registrerar, bedömer och tolkar testresultatet behövs ingen mätning av interbedömmarreliabilitet (McReynolds & Kearns, 1983). I praktiken är det mycket ovanligt att man räknar interbedömmarreliabilitet i de fall det finns mallar för rättning, till exempel på test som skall besvaras genom val av svarsalternativ. Interbedömmarreliabiliteten kan redovisas på olika sätt. Den kan redovisas som en korrelationskoefficient, vanligen Pearsons r , eller som en procentuell överensstämmelse mellan de två bedömningarna. I vår studie kommer vi att redovisa interbedömmarreliabiliteten på båda dessa sätt eftersom vart och ett har olika fördelar och de kompletterar varandra.

Fördelen med en procentuell överensstämmelse är att den är enkel att räkna ut och förstå (McReynolds & Kearns, 1983). Den tar dock inte hänsyn till exempelvis om två testledares bedömningar av ett barn är så gott som, men inte helt samstämmiga. Fördelen med en korrelationskoefficient är att den uttrycker en sådan samstämmighet. Nackdelen med en

korrelationskoefficient är att den är okänslig så länge testpersonernas inbördes rangordning är densamma och deras testresultat ligger lika långt ifrån varandra vid de båda bedömningarna. Till exempel korrelerar de parvisa värdena 1-2, 3-4, 5-6 således perfekt med varandra trots att det andra värdet genomgående är högre än det första. En procentuell överrensstämmelse skulle däremot visa att värdena inte stämde överens alls.

Homogenitetsmetoden innebär att man kontrollerar hur väl varje testuppgift stämmer överens med de övriga, det vill säga hur homogent testet är. Man vill att alla uppgifter ska testa samma faktor. Om testet i själva verket testar flera olika saker, som dessutom är oidentifierade, kan man inte dra några giltiga slutsatser utifrån poängsumman. Man vet inte vad testet har visat att testpersonen är bra respektive mindre bra på (Anastasi & Urbina, 1997). Den reliabilitet man får ut med homogenitetsmetoden kallas Chronbachs alfa. Detta värde visar i vilken mån testuppgifterna drar åt samma håll. Chronbachs alfa är idag det vanligaste måttet på reliabilitet vad det gäller test som är tänkta att mäta bara en faktor (Nunnally & Bernstein, 1994).

3. SYFTE

När man använder kvantitativa metoder i kliniskt arbete bör man ställa höga krav på psykometriska egenskaper. Man vill undvika att slumpen och andra ej kontrollerade faktorer skall påverka resultatet. Annars blir ett enskilt testresultat omöjligt att tolka på ett tillfredställande sätt. Svenska logopedier har hittills saknat ett standardiserat svenskt grammatiktest. Ett sådant test är nu konstruerat av Kristina Hansson och Ulrika Nettelblatt (under utarbetande).

Syftet med denna uppsats är att mäta test-retest-reliabiliteten, interbedömarreliabiliteten och Chronbachs alfa för Grammatiktestet. Vi vill alltså undersöka:

- I vilken grad får ett barn samma resultat då det testas två gånger inom en kort tidsrymd?
- I vilken grad får ett barn samma resultat då två oberoende testledare bedömer en och samma testsituation?
- Hur homogent är testet, det vill säga i vilken grad testar det bara en faktor?

4. METOD

4.1 Testmaterial

Grammatiktestet består av 45 uppgifter som testar 13 olika grammatiska strukturer. Fördelningen av uppgifter per struktur visas i tabell 1 tillsammans med de förväntade målorden. För vissa strukturer krävs exakt förväntat målord. Vad gäller Negation i huvudsats, Ordföljd vid topikaliserings och Bisatser krävs endast att barnets svar innehåller förväntad struktur. Som framgår av tabellen återkommer samma målord vid flera tillfällen, vilket är önskvärt enligt processbarhetsteorin för att garantera att barnet använder formerna produktivt.

Tabell 1. Testade grammatiska strukturer, antal uppgifter per struktur samt målord i grundform.

Struktur	Antal uppgifter	Målord
Presens	3	klippa, blåsa, köpa
Pluralis	3	katt, apa, pojke
Obestämd artikel	3	katt, apa, äpple
Bestämd form	3	katt, apa, äpple
Regelbunden preteritum	3	klippa, blåsa, köpa
Kongruens i nominalfras	4	gul, stor, katt, äpple, hus, pojke
Modal + infinitiv	3	klippa, blåsa, köpa
Har + supinum	3	klippa, blåsa, köpa
Genitiv	3	katt, apa, pojke
Kopula och adjektivkongruens i predikativ	3	gul, stor
Negation i huvudsats	3	<i>modal + negation + infinitiv</i>
Ordföljd vid topikaliserings	8	<i>X + verb + substantiv</i>
Bisatser	3	som + verb

Efter genomförda testningar plockades en uppgift under Ordföljd vid topikaliserings bort, då instruktionen för denna visade sig vara felaktig. Det totala antalet testuppgifter blev då 44.

Testet består av en pärm med tillhörande testblankett. Varje uppslag i pärmen visar en eller två bilder. Barnet ska titta på bilden/bilderna och testledaren eliciterar strukturen med en strategi som är en kombination av meningskomplettering och modellmening. Vad testledaren skall säga står angivet på bildens motsatta sida (För genitiv till exempel; *Det är hundens mat och det är...*, för kongruens i nominalfras till exempel; *Här är en blå bil och här är...*)

4.2 Försökspersoner

Kriterierna för deltagande i studien var att barnen skulle vara enspråkiga med svenska som modersmål. Barnen fick heller inte ha fått logopedhjälp tidigare för andra problem än stamning. De skulle vara över tre år men inte över sex år och sex månader. Barn som är yngre än tre år anses inte vara testbara (Håkansson & Hansson, 2000). Vi utförde studien på 47 slumpvis utvalda förskolebarn i åldern tre år och fyra månader till sex år och tre månader. Av dessa var 25 flickor och 22 pojkar. En översikt av försökspersonerna ses i tabell 2. Alla barnen gick på en och samma förskola på en ort i Kävlinge kommun. Orten ligger nära

riksgenomsnittet vad gäller inkomster och utbildningsnivå (Folk- och bostadsräkningen, 1990).

Tabell 2. Översikt över försökspersoner vad gäller ålder (i månader) och kön.

Ålder	Pojkar	Flickor	Totalt
36-41	0	2	2
42-47	4	3	7
48-53	1	1	2
54-59	6	4	10
60-65	2	5	7
66-71	7	6	13
72-78	2	4	6
Totalt	22	25	47

4.3 Tillvägagångssätt

Skriftligt tillstånd att rekrytera barn och genomföra testningar inhämtades från rektorn på den aktuella förskolan. Muntlig information gavs till personalen på de berörda avdelningarna innan skriftligt tillstånd inhämtades från de enskilda barnens föräldrar. Brev (se bilaga 1) lämnades ut till föräldrarna till 84 barn, vilket är det antal barn som finns på förskolans avdelningar för barn över tre år. Personalen lämnade in 54 svar till oss. Ett svar var negativt till medverkan. Bland de positiva svaren uteslöt vi ett barn på grund av att det inte uppfyllde kriterierna som beskrevs ovan. Av de återstående 52 barnen var det fyra som inte ville vara med och ett som inte kunde medverka på grund av sjukdom, vilket alltså resulterade i att 47 barn testades. Lämpliga datum för testning bestämdes i samråd med personalen på respektive avdelning.

De båda testledarna testade ungefär hälften av barnen var. Barnens kön och ålder var jämt fördelade i dessa två grupper. Varje barn testades två gånger med ca en månads mellanrum. Samma barn testades konsekvent av samma testledare. Ett av testtillfällena för varje barn spelades in på ljudband. För vissa barn skedde detta vid första testtillfället och för andra barn vid andra tillfället. Praktiska skäl avgjorde när inspelningen skedde. Två barn ville inte delta i den andra testomgången. Ett av dessa barn spelades in vid första tillfället, och finns således med i beräkningarna av interbedömarreliabilitet. Det andra barnet skulle spelats in andra gången, och finns inte med i några av beräkningarna, eftersom bara en testledare kunnat lyssna på det barnet.

Testningarna genomfördes i olika rum på förskolan. Vilken lokal som användes av respektive testledare varierade, och därför förekom det att samma barn testades i två olika lokaler. Ett barn testades vid ett tillfälle i sitt hem. Omständigheterna i det fallet var i övrigt desamma som under alla andra testningar. Rummen var avskilda från övriga rum och nivån av störande ljud från rummen omkring var relativt låg. Under testningen befann sig endast en testledare samt testpersonen i rummet. Barnens svar under testningen skrevs ner av testledaren på en särskild blankett. På testblanketten skrevs också barnets ålder, datum för testningen samt ett kodnamn som barnet tilldelades.

De ljudinspelningar som gjordes avlyssnades av den testledare som inte träffat barnet. Barnets svar nedtecknades då på en ny testblankett. Även på denna antecknades barnets kodnamn. Barnens svar bedömdes på alla blanketter som rätt respektive fel utifrån de bedömningskriterier som tillhandahölls av testförfattarna (se bilaga 2). Dessa kriterier innebär bland annat att inget annat ord än det eftersökta godkänns på många av testuppgifterna, även om svaret i och för sig är grammatiskt korrekt. Detta beror på att testet bygger på processbarhetsteorin, där man vill se att barnet kan använda orden i olika former. Att ett svar inte godkänns betyder därför inte automatiskt att det är grammatiskt felaktigt. Alla testuppgifter utom tre stycken ger vid rätt svar ett poäng. Dessa tre avvikande uppgifter kan ge noll till två poäng vardera. Detta beror på att två strukturer (kopula och adjektivkongruens i predikativ) testas i samma uppgift. I beräkningarna behandlades dessa tre testuppgifter som sex separata uppgifter. Maxpoäng är således 47 trots att antalet testuppgifter är 44.

Då vi började beräkningarna av testresultatet uppstod det problem i de fall då vi på ljudinspelningarna inte hört vad barnet svarat och därmed saknade data. I dessa fall har vi lagt in det som ett felsvar från barnets sida. Vi bestämde oss för detta förfaringssätt då det är vad man troligtvis skulle göra om samma situation uppstod i kliniskt arbete.

Utifrån dessa anteckningar av barnens svar räknades reliabiliteten för testet fram. Test-retest-reliabiliteten räknades ut som en korrelationskoefficient (Pearson r) mellan de två totalsummor barnen fått vid de två testtillfällena.

Med tvåsidigt t-test kontrollerade vi om barnens resultat signifikant förbättrats från första till andra testtillfället.

Vid uträkningen av interbedömarreliabiliteten jämförde vi först de två olika bedömarernas transkriptioner av barnens svar. Vi såg att vi inte alltid transkriberat exakt lika. En procentuell överensstämmelse räknades ut för varje testuppgift. Ett medelvärde av dessa procentsiffror, för varje struktur, och även för hela testet, räknades ut.

Utifrån bedömningarna av svaren räknades interbedömarreliabiliteten fram dels som en korrelationskoefficient mellan barnens totalsummor med Pearsons r, och även som en procentuell överensstämmelse av barnens totalsummor. En procentsiffra räknades också fram för hur väl testledarnas bedömningar för varje testuppgift och varje barn, det vill säga 2162 bedömningar, stämde överens. (46 barn x 47 bedömningar = 2162 bedömningar)

Med test-retest-materialet som grund räknades Chronbachs alfa ut för varje testtillfälle.

För att genomföra vår studie sökte vi och fick godkännande från Forskningsetikommittén vid Lunds Universitet (LU 325-03).

5. RESULTAT

Vi har räknat ut reliabiliteten ur tre aspekter. Resultatet visas i tabell 3.

Tabell 3. Tre aspekter på reliabilitet.

Test-retest	Interbedömar	Cronbachs alfa
0,70	0,99	0,84

5.1 Test-retest-reliabilitet

Test-retest-reliabiliteten (Pearson r) var 0,70, vilket ses i tabell 3.

Som man kan utläsa av tabell 4 var medelvärdet av barnens testpoäng vid andra testtillfället 2,6 poäng högre än vid första tillfället. Barnens resultat är signifikant bättre vid andra testningen ($t_{(44)} = -3,66$, $p < 0,001$, tvåsidig prövning).

Tabell 4. Medelvärde (M) och standardavvikelse (S) för barnens totalsummor vid de båda testtillfällena.

	M	S
Första testtillfället	30,4	6,3
Andra testtillfället	33,0	5,9

5.2 Interbedömarreliabilitet

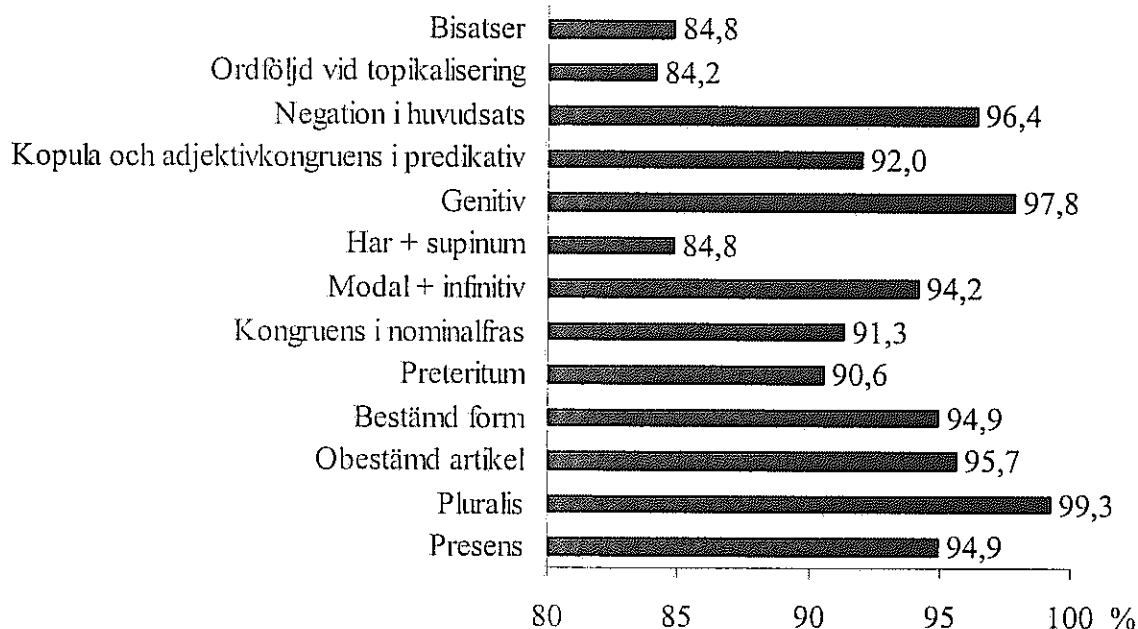
5.2.1 Transkriptionsöverensstämmelse

En översikt av den procentuella transkriptionsöverensstämmelsen för varje testuppgift visas i tabell 5. Som framgår av tabellen var variationen 26,1 procentenheter mellan de uppgifter som hade högsta respektive lägsta överensstämmelsen. Medianvärdet och standardavvikelsen visar dock att transkriptionerna i de flesta fall stämde väl överens.

Tabell 5. Översikt över testuppgifternas transkriptionsöverensstämmelse i procent, med avseende på minimivärde (Min), maximivärde (Max), medianvärde (Median), medelvärde (M) och standardavvikelse (S).

Min	Max	Median	M	S
73,9	100	93,5	91,6	7,1

Figur 1 visar hur väl transkriptionerna stämmer överens då man ser på de grammatiska strukturer testet är uppdelat i.



Figur 1. Medelvärden för transkriptionsöverensstämmelser per grammatisk struktur.

5.2.2 Bedömningsöverensstämmelse

Av de 2162 bedömningar vi gjorde stämde 2094 överens, i 68 fall var vi oense. Den procentuella överensstämmelsen var således 96,9 %. Barnens totalsummor var lika för båda bedömarna hos 41,3 % av barnen, det vill säga av 46 barn hade 19 fått samma testpoäng av båda bedömarna.

Korrelationen (Pearson r) mellan barnens testpoäng från de två olika bedömarna var 0,99, vilket ses i tabell 3.

5.3 Chronbachs alfa

De två värden vi fick fram med homogenitetsmetoden var identiska så när som på tredje decimalen, därför väljer vi att redovisa endast ett värde, vilket var 0,84 (se tabell 3).

6. DISKUSSION

Vi ville med den här studien undersöka Grammatiktestets reliabilitet i tre avseenden, test-retest-reliabilitet, interbedömarreliabilitet och Chronbachs alfa, det vill säga homogenitet. Som tabell 3 visar är alla tre värdena höga för att gälla ett psykometriskt test avsett för barn. Under arbetets gång har flera diskussionsfrågor väckts.

6.1 Test-retest-reliabilitet

Att test-retest-reliabiliteten är så pass hög innebär att man kan använda Grammatiktestet kliniskt för att utvärdera om ett barns grammatiska förmåga har förbättrats. Om testresultatet är tydligt bättre* vid en omtestning kan man dra slutsatsen att förbättringen av testresultatet inte beror på slumpen, utan att det är barnets förmåga som förbättrats.

Vi valde att testa med cirka en månads mellanrum. Detta trodde vi skulle vara ett lämpligt intervall då det borde vara för kort tid för att barnens grammatik skulle hinna utvecklas nämnvärt, men tillräckligt lång tid för att barnen inte skulle minnas de specifika svaren. Då vi senare jämförde barnens testpoäng från de två tillfällena såg vi en signifikant förbättring av resultaten. Man kan diskutera vad detta kan bero på, då det finns flera troliga orsaker. Det kan vara så att barnen faktiskt hunnit utvecklas språkligt på denna korta tid. De kan också kommit ihåg testet från förra tillfället, så pass mycket att de vid andra tillfället presterade bättre. Kanske var det så att barnen var lite mer osäkra eller blyga vid första testtillfället och därför presterade sämre då. Eftersom vi har sett en signifikant förbättring under en månad, behöver en observerad förbättring kliniskt alltså inte bero på terapin. Det är dock inte rimligt att barn med språkstörning utvecklas lika snabbt som de normalspråkiga barn som vi testat i denna undersökning.

6.2 Interbedömarreliabilitet

Att interbedömarreliabiliteten är hög visar att testet i mycket hög grad mäter samma sak oavsett vem som bedömer svaren. Detta innebär att man kan jämföra resultat som olika testledare fått fram av samma barn och sedan dra slutsatser om barnets eventuella förbättring, trots att olika testledare förstås påverkar barnet på olika sätt i testsituationen, vilket kan ge olika testresultat. Detta är något som troligen är svårt att komma ifrån.

En anledning till att interbedömarreliabiliteten blev så hög kan vara att de två bedömarna i denna studie är lika i många avseenden, till exempel vad gäller ålder, kön, utbildning och sociokulturell bakgrund. Detta kan ha påverkat våra bedömningar till att bli mer lika. I en vidare studie skulle man kunna ha fler bedömare, som skiljer sig åt i olika avseenden.

Då vi räknade ut transkriptionsöverensstämmelsen för varje testuppgift upptäckte vi att den varierade ganska kraftigt mellan uppgifterna. Överensstämmelsen är som lägst vid de

* För att beräkna sannolikheten att en förändring mellan två testningar av ett enskilt barn beror på tillfälligheter, bör man beräkna medelfelet för differenser (Anastasi & Urbina, 1997; Smedler, 1993).

uppgifter som testar har + supinum (*har klippt*), ordföljd vid topikalisering och bisatser. Då man jämför våra transkriptioner kan man se återkommande olikheter på dessa strukturer. Vid har + supinum har vi i vissa fall hört olika ändelser, till exempel *blåst* och *blåsit*. Vad gäller de övriga två ligger svårigheten troligen i att transkribera ett långt svar. Man kan tänka sig att längre strukturer ger lägre interbedömarreliabilitet.

En annan svårighet i att transkribera lika ligger i ljudinspelningens kvalitet, det vill säga förhållandena för förste och andre bedömare är inte lika vad gäller ljudkvalitet. Att lyssna på ett ljudband ger inte samma möjligheter att höra rätt och fråga om, som den testledare har som utför testningen. Detta kan man tänka på då man i framtiden använder testet kliniskt och kanske transkriberar från en ljudinspelning. Detta borde vara extra viktigt att tänka på då man testar barn med fonologisk språkstörning, då det är ännu svårare att uppfatta vad dessa barn säger. Interbedömarreliabiliteten borde därför bli lägre då man testar barn med fonologiska svårigheter.

I de testuppgifter där transkriptionsöverensstämmelsen är mycket hög består svaren vanligen av ett ord, vilket ytterligare visar att kortare strukturer är lättare att transkribera korrekt än längre.

Bedömningskillnaderna i sin tur beror nästan uteslutande på olikheter i transkriptionen. Men då testledarna transkriberat olika kan svaren ändå resultera i samma poäng, till exempel accepteras både *blåst* och *blåsit*. Detta förklarar att överensstämmelsen mellan bedömningarna är högre än mellan transkriptionerna.

Förfarandet att ge barnen noll poäng på de testuppgifter då vi utifrån ljudbandsinspelningen inte hört vad barnet svarat kan endast ha påverkat interbedömarreliabiliteten i negativ riktning. Då de flesta barn fått ett poäng på de flesta testuppgifter drar en nolla oftast ner korrelationen. Den mycket höga interbedömarreliabiliteten är med andra ord inte överestimerad.

Vad gäller Grammatiktestet håller vi med McReynolds och Kearns (1983) om att interbedömarreliabiliteten bör redovisas. Trots att det i detta fall finns klara riktlinjer om hur svaren ska bedömas blir inte bedömningarna exakt lika, eftersom det finns olikheter i transkriptionerna av barnens svar.

6.3 Chronbachs alfa

Att Chronbachs alfa är så pass hög innebär att testet är homogent till en hög grad, det vill säga det mäter i princip bara en faktor. Med viss försiktighet skulle man då kunna dra slutsatsen att den faktorn är barnets grammatiska förmåga.

6.4 Klinisk tillämpning

I de fall barnets svar bedömts som fel är det inte alltid så att svaret är grammatiskt felaktigt. Barnet måste också använda rätt målord för att svaret ska bedömas som korrekt. Då testet bygger på processbarhetsteorin måste barnet visa att det kan använda samma ord i olika former. Fantasifulla och kreativa barn med god grammatisk förmåga kan ibland hitta på svar där de tänkta målorden inte finns med men där grammatiken är korrekt. Det är sålunda möjligt

att barn med god grammatisk förmåga ändå får ett lågt resultat på Grammatiktestet. Kliniskt är det därför mycket viktigt att även använda testet på ett kvalitativt sätt, det vill säga att gå in på uppgiftsnivå och se vilken typ av fel barnet gör.

Det är intressant att diskutera hur dessa resultat står sig jämfört med reliabiliteten hos andra språktest som används kliniskt idag. Då det inte finns några språktest som är standardiserade på svenska, jämför vi med de siffror som finns för Reynell developmental language scales (Reynell, 1987) som är standardiserat på norska. Test-retest-reliabiliteten är där uträknad för varje del av testet. Den varierar mellan 0,41 och 0,89. Reynells interbedömarreliabilitet är uträknad på två av deltesten och dessutom för de olika åldersgrupperna som finns med i normeringen. Interbedömarreliabiliteten varierar mellan 0,93 och 0,99. Då man jämför dessa resultat med våra siffror (se tabell 2) ser man att reliabiliteten för Grammatiktestet är god. Grammatiktestets reliabilitet ligger inte över reliabiliteten för alla Reynells delresultat men klart över ett tänkt medel av dem.

Med tanke på den höga reliabiliteten tycker vi att Grammatiktestet är ett bra logopediskt verktyg då man vill mäta ett barns grammatiska förmåga. Om man ändå vill höja reliabiliteten ytterligare föreslår vi, grundat på vad Holmquist (1983) skriver, att Grammatiktestet utökas med fler testuppgifter. Detta är dock en avvägning då man inte får trötta ut testpersonen.

6.5 Slutsatser

Det frågor vi ställde angående Grammatiktestet inför detta arbete har nu fått sina svar:

- Till godtagbart hög grad ($r = 0,70$) får ett barn samma resultat när det testas två gånger inom en kort tidsrymd.
- Till mycket hög grad ($r = 0,99$) får ett barn samma resultat när två oberoende testledare bedömer en och samma testsituation.
- Testet har en bra homogenitet ($r = 0,84$) det vill säga testet mäter i hög grad bara en faktor.

7. REFERENSER

- Anastasi, A. & Urbina, S. (1997). *Psychological testing*. Upper Saddle River, New Jersey: Prentice-Hall.
- Bishop, D. (1989). *Test for the Reception Of Grammar (T.R.O.G.)*. University of Manchester: The Author, Age and Cognitive Performance Research Center.
- Folk- och Bostadsräkningen (1990), del 6: Inkomst och utbildning. Statistiska Centralbyrån, Stockholm.
- Gerken, L. (2000). Examining young children's morphosyntactic development through elicited production. L. Menn & N. Bernstein Ratner (red.) *Methods for studying language production (45-52)*. Mahwah, New Jersey: Erlbaum.
- Hansson, K. (1998). *Specific Language Impairment in Swedish - Grammar and Interaction*. Doktorsavhandling vid Institutionen för Logopedi och Foniatri. Lunds Universitet.
- Hansson, K. & Leonard, L. B. (2003). The use and productivity of verb morphology in specific language impairment: an examination of Swedish. *Linguistics*, 41-2: 351-379.
- Hansson, K. & Nettelbladt U. (under utarbetande). *Grammatiktest*. Malmö: Pedagogisk Design.
- Hansson, K., Nettelbladt, U. & Leonard, L. B. (2000). Specific language impairment in Swedish: the status of verb morphology and word order. *Journal of Speech, Language, and Hearing Research*, 43: 848-864.
- Holmberg, E. & Lundälv, E. (1998). *T.R.O.G.* Göteborg: SIH Läromedel.
- Holmberg, E. & Sahlén, B. (1986). *NELLI—Neurolingvistisk undersökningsmodell för språkstörda barn*. Malmö: Utbildningsproduktion AB.
- Holmquist, R. (1983). *Psykologiska test—fakta och fördomar*. Stockholm: Psykologiförlaget.
- Håkansson, G. & Hansson, K. (2000). Comprehension and production of relative clauses: a comparison between Swedish impaired and unimpaired children. *Journal of Child Language*, 27: 313-333.
- Lavesson, A. (1999). *Förskolebarn på förstagångsbesök. En kritisk granskning av Bruces och Wigforss "Språklig snabbscreening av barn i förskoleåldern, 2-6 år"*. Examensarbete i logopedi. Institutionen för logopedi och foniatri, Lunds universitet.
- Leonard, L. B. (1996). Assessing morphosyntax in clinical settings. D. McDaniel, C. McKee & H. Cairns (red.) *Methods of assessing children's syntax (2-38)*. West Lafayette, Indiana: MIT Press.

- Leonard, L. B., Salameh, E-K. & Hansson, K. (2001). Noun phrase morphology in Swedish-speaking children with specific language impairment. *Applied Psycholinguistics*, 22: 619-639.
- McReynolds, L. & Kearns, K. (1983). *Single-subject experimental designs in communicative disorders*. Austin, Texas: Pro. ed.
- Nettelblatt, U. (1995). Hur undersöker man språklig och kommunikativ förmåga hos barn? Erfarenheter från logopedisk forskning. I M. Linnarud (red), *Språk—utvärdering—test*. Rapport från ASLA:s höstsymposium, Karlstad, 10-12 november 1994. ASLA, Svenska föreningen för tillämpad språkvetenskap. Uppsala, s. 89-93.
- Nettelblatt, U. & Hansson, K. (ms.). *Språkstörning hos barn*.
- Nunnally, J. & Bernstein, I. (1994). *Psychometric theory*. New York: McGraw-Hill.
- Pienemann, M. & Håkansson G. (1999). A unified approach toward the development of Swedish as L2: A processability account. *Studies in Second Language Acquisition*, 3: 383-420.
- Reynell, J. (1987). *Reynell developmental language scales*. Oxford: NFER-Nelson.
- Smedler, A-C. (1993). *Att testa barn. Om testmetoder i barnpsykologiska utredningar*. Stockholm: Psykologiförlaget.
- Vogt, W. P. (1999). *Dictionary of statistics and methodology. A nontechnical guide for the social sciences*. Thousand Oaks, California: SAGE.

BILAGA 1

Lund, dd mm åå

Till föräldrar

Vi är två studenter på logopedprogrammet i Lund, som under året ska skriva vår magistersuppsats. Logopeder är en yrkesgrupp som bl. a. arbetar med utredning och behandling av barn men språkförsening/språkstörning. Vårt uppsatsarbete kommer att handla om ett nyligen utformat grammatiktest, som så småningom ska komma att användas vid testning av förskolebarn, för att se vilka av dem som behöver logopedhjälp. I nuläget är det ont om standardiserade och normerade test för detta ändamål, så detta test behövs för att vi ska kunna hjälpa kommande generationer bättre.

Det vi ska göra är att prova testets tillförlitlighet; dels om två testningar av samma barn med ca en månads mellanrum ger samma resultat, dels om två bedömare gör samma bedömning av barnet utifrån exakt samma testsituation. För att göra detta behöver vi testa ca 50 förskolebarn. Vi skulle vara mycket tacksamma om Ni vill ge Ert tillstånd att testa Ert barn för detta.

Testningen tar ca 15 minuter, och innebär att barnet tittar på bilder i en bok och svarar på några korta frågor. Testningen sker under lekfulla och avspända former och är helt ofarlig. Barnet deltar i undersökningen bara om han eller hon verkligen vill det. Ni kan när som helst avbryta barnets deltagande i undersökningen. Ca en månad senare görs testningen om på precis samma sätt. Vid ett av tillfällena krävs det att vi gör en ljudinspelning av testningen för att en annan bedömare senare ska kunna lyssna på exakt samma testsituation och göra sin bedömning. Inspelningarna kommer därefter inte att sparas. Ingen annan än vi som testat kommer att känna till barnens identitet och barnen kommer att vara helt anonyma i redovisningen.

Vi hoppas att Ni ger oss Ert tillstånd att testa Ert barn genom att fylla i medföljande blankett och lämna den till personalen på barnets avdelning på XX förskola senast Ring gärna om ni har några frågor!

Med vänliga hälsningar

Sara Clarén (046- 130756) och Sara Ringdahl (046- 148332)

BILAGA 1

Ja, mitt barn får vara med i undersökningen.

Har barnet haft logopedkontakt, och i så fall av vilken anledning? _____

Talar barnet något mer språk än svenska? _____

Nej, mitt barn får inte vara med i undersökningen.

Barnets namn: _____ Födelsedatum (ååmmdd): _____

Förälders underskrift: _____

GRAMMATIKTEST

K. Hansson & U. Nettelbladt

Korrekt svar

rutan)

(Sätt kryss i

A. PRESENS

1. klipper _____ Inga andra svar än målorden accepteras, även om det är korrekta presensformer, som t.ex. *äter, tar*

B. PLURALIS

4. pojkar _____ Det tänkta substantivet måste användas. Alternativ som kan accepteras är *kattor, aper*, som på skånska är uttalsvarianter, däremot inte tydlig användning av fel deklination som t.ex. *pojkor*. Inte heller korrekta pluralisformer av andra substantiv, t.ex. *killar, kissar*. Svar som innehåller korrekt pluralisform, men som är mera komplexa än vad som förväntas, t.ex. bestämd form pluralis (*pojkmarna, aporna*) accepteras.

Best

Obest

C. OBESTÄMD ARTIKEL / BESTÄMD FORM

7. en katt _____ art form8. katten _____ Obestämd artikel: tänkt substantiv måste användas med korrekt artikel. *En räv* eller *ett troll* accepteras t.ex. inte. Alternativ som (i sydsvenska dialekter) kan godkännas är *ett äppel*, som räknas som uttalslättning ("undvikande" av konsonantkombination). Däremot accepteras inte fel genus, t.ex. *ett katt, en äppel/äpple*. Svaret accepteras om barnet lägger till något, t.ex. *en liten apa*.Bestämd form: endast målorden accepteras, inte andra substantiv och inte pronomen (*den, han*)

D. PRETERITUM

13. klippte _____ Tänkt verb måste användas med korrekt preteritumform. Fel konjugation accepteras inte som t.ex. *klippade*E. KONGRUENS I NOMINALFRAS
formObest Best
art

BILAGA 2

16. ett gult hus _____

18. det gula äpplet _____

Alla tre delarna måste vara med, d.v.s. artikel + adjektiv + substantiv. Svar som *gult hus, stora katten, ett hus, en pojke* accepteras inte. Endast de exakta målorden accepteras, alltså accepteras inte t.ex. *ett stort hus, ett stort troll, det gröna äpplet, en stor människa, en stor katt*. Däremot accepteras om barnet lägger till ett adjektiv, som t.ex. *den stora gula katten* eller *den jättestora katten*.

svar

F. MODAL + INFINITIV
rutan)

Korrekt

(Sätt kryss i

20. ska köpa _____

Ett modalt hjälpverb och en infinitivform krävs. Infinitivformen måste vara av rätt verb *ska ta, ska plocka, ska klippas* accepteras inte. Annat modalt hjälpverb accepteras däremot, d.v.s. *vill (köpa), kan (klippa), måste (blåsa), får (köpa)* eller preteritumform på hjälpverbet, t.ex. *skulle köpa*. Man accepterar även om barnet lägger till en negation, t.ex. *ska inte blåsa*. Enbart hjälpverb plus negation accepteras inte, t.ex. *kan inte*.

G. HAR + SUPINUM

23. har klippt _____

Hjälpverbet *har* måste vara med och supinumformen måste vara av rätt verb. Övergeneraliseringar accepteras, t.ex. *har klippit, har blåsit, har köpit* som är mycket vanligt förekommande och som kan betraktas som uttalslättnad. Svar som innehåller hela måstrukturen men som är mera komplexa accepteras också, t.ex. *har blivit klippt, har blivit färdigklippt*. Däremot accepteras inte andra variationer, t.ex. *har blåsat*

H. GENITIV

26. kattens _____

Endast målformerna accepteras. Enda undantaget är om barnet svarar med bestämd pluralisform, t.ex. *pojkarernas, kattarnas*. Då finns hela målformen med. Obestämd form med genitiv *-s* accepteras inte, t.ex. *apas*. Ersättning med pronomen, t.ex. *hans, dens* eller annat substantiv t.ex. *killens*, accepteras inte.

I. KOPULA OCH ADJEKTIVKONGRUENS I PREDIKATIV

Adjektiv

Kopula

29. x är gul _____

BILAGA 2

För att barnet ska få poäng för kopula måste både ett subjekt, kopula och en predikatsfyllnad vara med. Svar som *katten gul*, *katten e* eller *e gul* accepteras alltså inte, däremot *huset e stor*.

För att få adjektivpoäng krävs korrekt form av tänkt adjektiv (det exakta målordet) *katten e vit*, *huset e gult*, *huset e stor* eller *katten e blå* accepteras inte. Om subjektet är ett annat gör inget – bara det har samma genus och numerus som adjektivet. Även om subjektet är ett pronomen accepteras svaret, t.ex. *han är gul*, *människorna är stora*, *dom är stora*, *trollen e stora*, *det är stort*. Om kopula inte är med, så att det blir en nominalfras, t.ex. *ett stort hus*, *dom stora* accepteras **inte**, eftersom målstrukturen är adjektiv i predikativställning. Inte heller *katten gul* eller *gul katt* accepteras.

J. NEGATION I HUVUDSATS

32. kan inte måla _____

Andra modaler än *kan* accepteras (*vill*, *ska*, *måste*, *får* t.ex.). En infinitivform måste finnas, men behöver inte vara det exakta målverbet. Negationen måste vara i rätt position. Bara *kan inte (de)* accepteras inte.

K. ORDFÖLJD VID TOPIKALISERING

36. sen gungade dom _____

Det måste finnas ett topikaliserat ord, ett verb och ett subjekt (i nämnd ordning) för att svaren ska kunna accepteras: *sen gick dom till en restaurang*, *sen går han till dagis*, *sen sover dom*, *sen går han*, *då ska dom sjunga*, *så läser mamma* är exempel på bra svar. Exempel på svar som inte accepteras är: *läser mamma saga*, *sen gick till en restaurang*, *sen dom sover*

L. BISATSER

43. som plockar blommor _____

Svaret måste innehålla *som* och ett (finit) verb, men det kvittar vilket. Det ska alltså vara en bisats med minimalt subjekt (*som*) och verb. Svar som *som ätit kaka* (inget finit verb), *när han äter kakor* (inte relativbisats) accepteras inte, inte heller om barnet upprepar samma verb i mer än ett svar, d.v.s. två *som ätit*, då accepteras bara det ena.

