



GÖTEBORGS UNIVERSITET  
SAHLGRENSKA AKADEMIN

Institutionen för neurovetenskap och fysiologi  
Enheten för logopedi

232

**Utprovning av SWITCH**  
**Ett svenskt förståelighetstest för barn**

Anna-Karin Ahlman  
Andrea Börjesson

Examensarbete i logopedi  
30 högskolepoäng  
Vårterminen 2011

Handledare  
Tove Johannisson  
Christina Persson



# Utprovning av SWITCH

## Ett svenskt förståelighetstest för barn

Anna-Karin Ahlman  
Andrea Börjesson

*Sammanfattning.* SWITCH- Swedish Intelligibility test for children, är ett test som mäter förståeligheten i barns tal genom repetition av enskilda ord. Testet är datorgenererat och består av 1000 ordlistor som slumpmässigt väljs ut för varje ny bedömning. Syftet med studien var att undersöka likvärdigheten, reliabiliteten och validiteten i ordlistorna. Ordlistor provades ut på tio barn med typisk språkutveckling (N-grupp) och tio barn med talavvikelse (T-grupp). Barnen audio-inspelades när de repeterade två listor. Tjugo logopedstudenter bedömde förståeligheten. Procent korrekta konsonanter (PCC) beräknades. Resultaten på barnens båda ordlistor jämfördes för att beräkna likvärdigheten. Inter och intrabedömarreliabilitet undersöktes. Validitet undersöktes genom att undersöka eventuella samband mellan PCC och förståelighet samt genom att jämföra förståelighet mellan N-gruppen och T-gruppen. Resultaten visade att ordlistorna som utgör SWITCH var likvärdiga samt att metoden var reliabel och valid.

Nyckelord: Förståelighetstest, barn, enskilda ord, reliabilitet, validitet

*Abstract.* SWITCH- Swedish Intelligibility test for children, is a test that measures intelligibility of children's speech through repetition of single words. The test is computer generated and consists of 1000 word lists randomly selected for each assessment. The purpose of this study was to examine the equivalence, reliability and validity in the word lists. Word lists were tested on ten children with typical language development (N group) and ten children with speech deviation (T group). The children were audio recorded when they repeated two lists. Twenty speech therapy students assessed the intelligibility. Percentage of consonants correct (PCC) were calculated. The results of each child's two word lists were compared to calculate the equivalence. Inter- and intraraterreliability were examined. Validity was examined by examining correlations between PCC and intelligibility, and by comparing the intelligibility of the N-group and T-group. The results showed that the word lists in SWITCH were equivalent and that the method was reliable and valid.

Key words: Intelligibility test, children, single-word, reliability, validity

Förståelighet är ett mått på hur mycket en lyssnare förstår av det en talare avser att säga (Hartelius & Lohmander, 2008). Det är ett mått som används för bedömning vid allt tal där förståeligheten på något sätt är nedsatt. Nedsatt förståelighet i talet kan uppstå som en följd av talstörning eller språkstörning (Klein & Flint, 2006). Talstörningar kan vara orsakade av strukturella avvikelser, som vid exempelvis läpp-käk- gomspalt (Whitehill & Chau, 2004). De kan vara neurologiskt betingade som vid dysartri, eller av ett mer okänt ursprung som vid stamning (Hartelius & Lohmander, 2008). Även fonologiska språkstörningar med talavvikelse till följd påverkar talets förståelighet i olika stor utsträckning (Klein & Flint, 2006). Hos barn i förskoleåldern är fonologiska problem till följd av språkstörning den vanligaste språkliga svårigheten (Nettelbladt & Salameh, 2007). Syftet med muntlig kommunikation är att bli förstådd (Kent, Miolo & Bloedel, 1994) och av den anledningen ses ökad förståelighet av både kliniker, patienter och närstående som det viktigaste målet vad gäller i princip all intervention vid talavvikelse och talstörningar (Hodge & Gotzke, 2007a; Klein & Flint, 2006; Yorkston & Beukelman, 1978). En bedömning av förståelighet kan visa på effekter av olika sorters intervention (Kent, Weismer, Kent & Rosenbek, 1989).

Trots att förståelighet är ett viktigt mått vid bedömning vad gäller grad av talavvikelse så finns det idag ingen etablerad metod för hur förståelighet ska bedömas (dos Santos Barreto & Zazo Ortiz, 2008; Whitehill & Chau, 2004). Interventioner saknar ofta förankring i en reliabel mätmetod för förståelighet (Kent et al., 1994; Klein, 1996; Shriberg & Kwiatkowski, 1992). Det finns flera olika metoder och olika sorters material som kan användas vid förståelighetsbedömning, och därmed ingen enighet kring metod och material (dos Santos Barreto & Zazo Ortiz, 2008).

Det finns olika metoder för att elicitera tal. Talaren kan läsa (Hartelius & Lohmander, 2008) eller repetera ord eller meningar (Hodge & Gotzke, 2007a; Morris, Wilcox & Schooling, 1995; Zajac, Plante, Lloyd & Haley, 2010;) Spontantal är också en vanlig eliciteringsmetod (Hodge & Gotzke, 2007a; Kent et al., 1994). Därefter får en lyssnare som inte känner till materialet eller vet vad talaren pratar om bedöma förståeligheten i det talaren säger (Hartelius & Lohmander, 2008).

Det finns olika metoder för att bedöma förståelighet och de metoder som används mest är skalskattning och ordidentifikation av enskilda ord eller meningar (Kent et al., 1989). Metoder som använder sig av skalor innebär att en bedömare får skatta en persons förståelighet på en skala från låg till hög förståelighet (Kent et al., 1989). Ordidentifikation innebär att en bedömare avgör vilket/vilka ord han uppfattar av det talaren säger, antingen i enskilda ord eller i meningar (Kent et al., 1989). Identifikation av enskilda ord kan göras med olika metoder. Bedömningen kan göras antingen genom open-set (fri transkription av det en lyssnare uppfattar att talaren säger) eller genom closed-set (där lyssnaren väljer ett ord utifrån olika alternativ). Vilken metod som används beror på syftet med bedömningen (Hodge & Gotzke, 2007a).

Det finns fördelar och nackdelar med såväl skattning av förståelighet som med ordidentifikation. Den senare metoden anses dock vara den metod som har högst reliabilitet (Kent, 1992). Skattning är visserligen en tidseffektiv metod, men den saknar den precision och reliabilitet som identifikation av ord eller meningar ger (Kent, 1992). Ordidentifikation är en metod som underlättar för lyssnaren att enbart fokusera på förståeligheten i den auditiva talsignalen, och inte förvillas av variabler som i första hand innebär en påverkan på röstkvalitén såsom nasala genomslag eller hyperfoni (Whitehill, 2000). Hon menar att lyssnare som använt skattning bedömt talare som mindre förståeliga än vad de är eftersom de har tagit med röstkvalité i sin bedömning.

Ordidentifikation är därför en metod som underlättar för lyssnaren att enbart fokusera på förståeligheten i talet och inte bedöma talet utifrån röstkvalité (Kent et al., 1994; Whitehill, 2000). Ordidentifikation är alltså en metod med god validitet, reliabilitet och användarvänlighet (Kent et al., 1989).

När förståelighet bedöms genom ordidentifikation finns det olika slutsatser angående vilket material som är bäst att använda. Weston och Shriberg (1992) menar att transkription av spontantal ger högst extern validitet, eftersom det återspeglar barnets tal i en naturlig kontext. Inom den kliniska verksamheten är transkription av spontantal dock en tidskrävande metod (Morris et al., 1995). Vidare menar de att en annan begränsning med spontantal är att det inte går att veta säkert vad barnet säger, vilket gör det svårt att veta om det som lyssnaren uppfattar är rätt. Spontantalet hos en och samma person kan också variera avsevärt från ett tillfälle till ett annat, vilket innebär att test före och efter en intervention aldrig blir desamma och därmed inte helt jämförbara (Morris et al., 1995). Zajac et al. (2010) tar upp att ord är ett bra material eftersom det är beständigt, vilket gör det möjligt att jämföra resultat mellan olika barn även över tid och mellan olika kliniska verksamheter. Vidare menar Kent et al. (1989) att enskilda ord är att föredra framför meningar. De menar att ord är tillräckligt långa för att en bedömning av talarens fonologiska processer ska kunna genomföras, men korta nog för att personer med svåra motoriska talavvikelser ska klara av att producera dem. Flera studier har använt enskilda ord för bedömning av barns förståelighet (Hodge & Gotzke, 2007a; Morris et al., 1995; Whitehill & Chau, 2004; Zajac et al., 2010).

Zajac et al. (2010) har utvecklat ett test baserat på 50 enskilda ord för att bedöma förståelighet i talet hos barn med opererad gomspalt. Dessa 50 ord slumpades fram ur en ordbank bestående av 510 ord indelade i 50 set. Varje set innehöll fonetiskt jämförbara och likvärdiga ord. Som eliciteringsstrategi kunde antingen repetition eller läsning användas. I studien användes repetition av en testledare, dels p.g.a. att barnens läsförmåga inte skulle påverka förståeligheten i orden, dels p.g.a. att en testledare bättre skulle kunna hålla uppe motivationen hos barnen. Testet visade på god reliabilitet och validitet. Hodge och Gotzke (2007a) undersökte validiteten och reliabiliteten i förståelighetstestet SIP-CCLP (Speech Intelligibility Probe for Children With Cleft Palate), som är baserat på enskilda ord. Testet innehåller både en open-set bedömning och en closed-set bedömning samt en del som bedömer förståelighet i spontantalet. Författarna kunde visa på en stark korrelation mellan förståelighetspoängen i barnens spontantal och förståelighetspoängen i open-set bedömningen. Det ger testet en god validitet eftersom det tyder på att en bedömning av enskilda ord återspeglar hur förståeligheten ser ut i en mer naturlig kontext. I en studie av Morris et al. (1995) fick barn med olika nivå vad gäller tal- och språkutveckling repetera ordlistor ur förståelighetstestet P-SIM (The preschool speech intelligibility measure). Listorna innehöll 50 enskilda ord som eliciterades genom repetition. Repetition visade sig fungera bra som eliciteringsmetod, eftersom barnen förstod uppgiften snabbt och deltog väl under hela repetitionsuppgiften. Förståelighet på ordnivå jämfördes med artikulation. En tydlig korrelation fanns emellan dessa två begrepp, vilket stärker testets validitet. Testet visade vidare på god reliabilitet.

Artikulation och fonologi är två viktiga faktorer för att en persons tal ska bli förståeligt för andra (Kent et al., 1994). PCC (Percentage of consonants correct) är en metod för att bedöma hur många konsonanter barnet klarar av att producera korrekt i ett talmaterial (Shriberg & Kwiatkowski, 1982). Varje konsonant i talmaterialet bedöms som antingen korrekt eller inkorrekt. PCC-poängen beräknas genom att antal rätt

uttalade konsonanter som talaren sagt delas med det totala antalet konsonanter som finns i ett talmaterial (Shriberg & Kwiatkowski, 1982).

Subtelny (refererad till i Kent, 1992) menar att det är viktigt att reliabilitet och validitet i ett förståelighetstest undersöks. Reliabiliteten undersöks genom att utreda om testet är konsekvent, dvs. ger samma resultat oavsett vem som gör bedömningen och om en och samma bedömare är konsekvent i sin bedömning över tid (Malterud, 1998). Validitet innebär att det som avses att mätas verkligen är det som mäts (Malterud, 1998). Vidare menar Malterud att för att bedöma validiteten i ett test kan begreppsvaliditeten undersökas. Begreppsvaliditet innebär att relaterade begrepp jämförs med varandra för att se om de korrelerar. Eftersom förståelighet och artikulation ofta korrelerar (Stimley & Hambrecht, 1999) kan dessa ses som två relaterade begrepp. För att ytterligare undersöka testets validitet kan skillnader i förståelighet mellan barn med talavvikelse respektive utan talavvikelse undersökas (Zajac et al., 2010). En skillnad i förståelighet mellan dessa grupper tyder på att testet verkligen mäter förståelighet och att det inte mäter något annat.

Hartelius och Lohmander (2008) tar upp att bedömning av förståelighet inte är helt enkelt, eftersom det förutsätter att den som lyssnar inte känner till materialet eller vet vad som kommer sägas av talaren. Familjaritet, d.v.s vana att använda ett visst talmaterial, gör en lyssnare olämplig som bedömare av just det materialet (Hartelius & Lohmander, 2008). I en klinisk verksamhet innebär det ofta att det krävs två personer för att bedöma förståelighet, dels en person som eliciterar talmaterialet, dels en person som lyssnar och transkriberar materialet (Hartelius & Lohmander, 2008). De senaste åren har intresset för hur förståelighet ska bedöms på bäst sätt ökat (Kent et al., 1994; Whitehill & Chau, 2004). Flera studier tar upp att det behövs en reliabel och valid metod, som kan skapa enighet kring hur förståelighet bör bedömas (dos Santos Barreto & Zazo Ortiz, 2008; Whitehill, 2000; Zajac et al., 2010). Det är viktigt att samma bedömningsmetod används för samma patient, eftersom även en liten ändring i metod kan förändra förståelighetspoängen signifikant (Hustad, 2006).

Det finns ett svenskt test som mäter förståelighet i talet hos vuxna personer med dysartri, SWINT-Swedish Intelligibility Test (Lillvik, Allemark, Karlström & Hartelius, 1999). Johannisson, Lohmander och Persson (2010) prövade att använda testet för att mäta förståelighet i talet hos barn i 10-års-åldern. Resultaten visade dock att testet var mer känsligt för barnens läsförmåga än för förståeligheten i talet, vilket innebär att testet därmed inte är optimalt för barn. Vidare visar studier på att förståeligheten i barns tal även påverkas av familjaritet. Barn som repeterar ord som de inte är familjära med får en förståelighetspåverkan (Morris et al., 1995). SWINT är främst utformat för att passa den vuxna populationens vokabulär. Barn behöver därför ett förståelighetstest där de slipper läsa orden och som innehåller ord som är familjära för dem. I Sverige saknas det idag en valid och reliabel metod för att mäta förståeligheten i barns tal. Förståelighet hos barn bedöms i nuläget främst genom skattning av hur väl en lyssnare förstår barnets spontantal eller genom transkription av spontantal, ord eller meningar ur artikulationstest (Hartelius & Lohmander, 2008).

SWITCH (Swedish intelligibility test for children) är ett svenskt förståelighetstest under utveckling som bygger på ordlistor där barn får repetera enskilda ord som anses vara familjära för dem (Johannisson, 2011).

Syftet med föreliggande studie var att undersöka ordlistornas likvärdighet samt reliabiliteten och validiteten i SWITCH. Ordlistorna prövades ut på barn med typisk talutveckling och på barn med någon form av talavvikelse. I studien besvarades följande

frågeställningar: a) är ordlistorna i SWITCH likvärdiga. d.v.s. fick deltagarna samma resultat på två slumpmässigt utvalda ordlistor och b) är SWITCH en reliabel och valid metod för att bedöma förståelighet i barns tal?

## Metod

### *Deltagare*

*Barn med tal/språkavvikelse (T-gruppen).* Tio barn med påverkad förståelighet till följd av språkstörning deltog i studien. Vårdnadshavarna hade rapporterat att deras barn hade god hörsel och svenska som starkaste språk. Åldersintervallet hos deltagarna var 4:6-8:3, år:månader ( $M = 6:0$ ,  $s = 1:0$ ). För närmare information om deltagarna se tabell 1.

Inklusionskriterier för deltagare i T-gruppen:

- förståelighetspåverkan i talet
- befinner sig inom åldersintervallet 4:6-10:6 år

Deltagarna i T-gruppen rekryterades genom att information skickades ut via mejl till logopederna i Göteborg som behandlade barn med tal- och språkstörningar. Även muntlig information till logopederna förekom. Logopederna delade i sin tur ut informationen till de föräldrar som hade barn med påverkad förståelighet. I enstaka fall skickades blanketterna direkt hem till föräldrarna. Detta skedde i de fall där det förekom behandlingsuppehåll. För att barnen skulle få medverka i den aktuella studien skrev föräldrarna på en samtyckesblankett vilken krävde båda föräldrarnas underskrift. Inspelningarna gjordes i anslutning till behandling, på barnens skola/förskola eller i hemmet.

Tabell 1

*Ålder, Kön, PCC och Diagnos hos deltagarna i T-Gruppen*

Barn	Ålder (år:mån)	Kön	PCC	Diagnos
T01	4:6	P	59	GE
T02	5:6	F	60	F
T03	6:7	P	54	GE
T04	5:11	F	61	F
T05	5:11	P	70	F
T06	8:3	P	81	FG
T07	6:6	F	65	F
T08	6.6	F	49	F
T09	5:2	P	61	F
T010	5:4	F	61	GE

Not. P = pojke, F = flicka, F = fonologisk språkstörning, FG = fonologisk och grammatisk språkstörning, GE = generell språkstörning, PCC = procent korrekt uttalade konsonanter.

*Barn med typisk talutveckling (N-gruppen).* Tio barn som inte hade träffat logoped deltog som kontrollgrupp i studien. Vårdnadshavarna hade rapporterat att deras barn inte hade någon hörselpåverkan och att de hade svenska som starkaste språk. Åldersintervallet hos deltagarna var 4:8-7:4, år:månader ( $M = 5:9$ ,  $s = 1:1$ ). För närmare information om deltagarna se tabell 2.

Inklusionskriterier för deltagare i N-gruppen:

- ingen pågående eller tidigare kontakt med logoped
- befinner sig inom åldersintervallet 4:6-10:6 år
- går i kommunal skola/förskola i Göteborg

Tabell 2

<i>Ålder, Kön och PCC hos deltagarna i N-Gruppen</i>			
Barn	Ålder (år:mån)	Kön	PCC
N01	4:8	P	79
N02	4:11	F	98
N03	7:3	P	99
N04	5:0	F	72
N05	5:3	P	97
N06	7:4	P	100
N07	7:3	F	100
N08	6:6	F	100
N09	5:10	P	96
N010	4:10	F	100

*Not. P = pojke, F = flicka, PCC = korrekt uttalade konsonanter.*

Deltagare i N-gruppen rekryterades genom att information och förfrågningar skickades ut till rektorer/verksamhetschefer på skolor och förskolor i centrala Göteborg. Via skolans pedagoger skickades information sedan ut till vårdnadshavarna. I de fall där barnen bedömdes kunna ta till sig information skriftligt riktades ett särskilt brev till dem. För att barnen skulle få delta krävdes ett godkännande från båda vårdnadshavarna. Testledarna åkte därefter ut till skolorna och samlade in svarsblanketterna samt bokade tid för inspelning.

I båda grupperna (T-gruppen och N-gruppen) var könsfördelningen jämn, med fem pojkar och fem flickor i respektive grupp. Samtliga barn genomgick en artikulationscreening (PCC- procent korrekta konsonanter), se tabell 1 och 2. Hur bedömningen av PCC gick till beskrivs under "*Tillvägagångssätt*".

*Lyssnargruppen.* I studien deltog 20 logopedstudenter, alla kvinnor mellan 20:2–34:9 år ( $M = 24:0$  år,  $s = 3:5$ ), varav tre hade studerat på logopedprogrammet i sju terminer, en fem terminer, nio i tre terminer och sju i en termin. Lyssnarna var tvungna att ha rapporterat god hörsel och svenska som starkaste språk för att kunna delta i studien. Lyssnare rekryterades på logopedprogrammet i Göteborgs universitet kontaktades via mejl av huvudhandledaren. De fick även muntlig information av författarna i samband med undervisning.



## *Etiska hänsynstaganden*

Studien har prövats och godkänts av den regionala Etikprövningsnämnden i Göteborg. Alla medverkande, inklusive barnens vårdnadshavare, fick tydlig information kring studiens syfte och upplägg. Allt material som spelades in under studiens gång förvarades säkert och endast personer som var behöriga hade tillgång till detta material. Deltagande i studien var helt frivilligt och kunde när som helst avbrytas.

## *Material*

*SWITCH.* I den aktuella studien användes 40 slumpmässigt utvalda ordlistor ur SWITCH. Testet består av 1000 ordlistor som har skapats utifrån en ordbank med 1243 ord. Orden har hämtats ur ett examensarbete i logopedi där Case, Forsber och Uppmann (2009) undersökte vilka ord som förekommer i barns tal. De undersökte vokabulär hos barn i åldrarna 5 - 8 och 9 - 12 år. Orden i SWITCH bygger på de ord som förekom i åldrarna 5-8 år. Varje ordlista består av 60 ord med alla svenskans fonem representerade. En typisk ordlista utgörs av; 16 stycken ord med två-kluster, fyra stycken med tre-kluster, 17 stycken med s/r kluster samt 23 ord utan kluster. I en typisk ordlista finns det 16 enstaviga ord, 32 tvåstaviga ord, nio trestaviga ord och tre fyrstaviga ord.

*SVANTE.* För bedömning av procent korrekta konsonanter (PCC) användes delar ur SVANTE-materialet (Svenskt artikulations- och nasalitetstest) som är ett normerat nationellt test utformat av svenska logopedier specialiserade inom läpp-käk-gomspalt (Lohmander et al., 2005). Materialet består av ord som innehåller svenskans alla fonem. SVANTE utfördes enligt de strategier som togs upp i manualen.

*Inspelningsutrustning.* Materialet spelades in med bärbar inspelningsutrustning (TASCAM HD-P2 Portable High-Definition Stereo Audio Recorder). En bordsmikrofon (SONY, ECM-MS957) placerades ca 30 centimeter från deltagarna. Avståndet kunde variera beroende på deltagarnas röststyrka. Alla deltagare samt det inspelade materialet försågs med koder.

## *Tillvägagångssätt*

*Förstudie.* Innan inspelningarna av deltagarna inleddes genomförde testledarna en förstudie med syfte att testa den valda eliciteringsmetoden för orden i SWITCH, få erfarenhet och samträning av testmaterialet samt att pröva om det valda testförfarandet fungerade. Under förstudien repeterade barnen en SWITCH-lista, benämnde därefter orden i SVANTE och avslutade med ytterligare en SWITCH-lista. Författarna valde att växla mellan uppgifterna för att minska risken för uttrötningseffekt hos barnet. Denna testordning kom sedan att användas under hela datainsamlingen. Förstudien genomfördes på institutionen för neurovetenskap och fysiologi vid Göteborgs Universitet, och prövades på tre barn (en 5-åring och två 8-åringar) i testledarnas bekantskapskrets. SVANTE är utformat på ett sådant sätt att barnen kan läsa sig till vad varje bild föreställer. För att inte läsförmågan skulle påverka utfallet valde testledarna att under inspelningarna täcka över orden. Det upptäcktes tidigt att det behövdes en

visuell modell av SWITCH-förfarandet hos de yngre barnen som ett komplement till den muntliga informationen.

*Inspelningsförfarande.* Inspelningarna ägde rum under januari-mars 2011. Alla deltagare i N-gruppen spelades in i ett enskilt rum på respektive skola/förskola för att minimera risken för störande bakgrundsljud. Deltagarna i T-gruppen spelades in i enskilt rum på förskola, på logopedmottagningen i samband med besök eller i barnets hem. Under alla inspelningar var båda författarna närvarande, en och samma person agerade genomgående testledare varav den andra skötte inspelningsutrustningen. Målet var att ha en så likvärdig inspelningssituation som möjligt för samtliga barn. Innan testningen satte igång hölls ett kort kontaktskapande samtal med barnen samt en kort genomgång av vad som skulle ske. Det förklarades för barnen att de skulle repetera ord och titta i en bok med olika bilder i som de skulle benämna. Före varje inspelning med varje barn genomfördes en kort förtestning med ord som inte hörde till SWITCH-materialet för att försäkra sig om att de hade förstått uppgiften. Varje inspelning tog ca 25-30 minuter att genomföra. Inspelningarna tog något längre tid för de yngsta barnen.

Endast ljudstimuli från deltagaren spelades in. Detta kunde genomföras genom att testledaren pekade på barnet när det var dags att repetera ordet. Testledaren kunde på så sätt skapa en paus mellan det att testledaren sade ordet och barnet repeterade det. I pausen sattes inspelningsutrustningen igång, deltagaren repeterade sedan ordet och därefter stoppades inspelningen. En paus lades in för att minska databearbetningen i efterhand och för att minska testledarens inverkan som en artikulatorisk modell på barnens uttal. I efterhand bearbetades ljudfilerna i programmet PRAAT genom att störande ljud klipptes bort och cirka en sekunds paus klipptes in före varje ord.

*Artikulationsscreening- PCC (Percentage of consonants correct).* PCC bedömdes enligt Shriberg och Kwiatkowskis (1982) modell för bedömning av artikulation av konsonanter. Shriberg och Kwiatkowski beräknade PCC utifrån spontantal. Kent et al. (1994) menar dock att det kan vara svårt att beräkna PCC utifrån spontantal, eftersom det i mycket svårförståeligt tal kan vara svårt att veta vad barnet har sagt. Innehållet i spontantal kan variera vad gäller komplexitet och längd. Av den anledningen valdes SVANTE-materialet för bedömning av PCC. Enligt Shribergs och Kwiatkowskis (1982) modell ska enbart konsonanterna i varje ord bedömas, vokaler togs således ej med i bedömningen. Dialektala variationer bedömdes som korrekt, likaså snabbt vardagligt tal där det ansågs att fonemet skulle ha förekommit om ordet skulle ha sagts enskilt t ex.: ”hoppa hopprep” som i vardagligt tal kan bli ”hoppa hopprep”. För sje-ljudet accepterades [ç] (tje-ljudet) istället för [h] (sje-ljudet) i ordet/garage/ eftersom det ansågs vara en konventionell allofon.

Författarna transkriberade inte varje enskilt ord utan valde att utifrån en egen blankett skriva ner de konsonanter som uttalades korrekt i varje ord genom att bedöma och diskutera transkriptionen tillsammans.

I den aktuella studien användes 64 ord ur SVANTE. Orden med s-kuster utslöts eftersom inga kluster används enligt PCC-metoden (Shriberg & Kwiatkowski, 1992). PCC för varje enskilt barn redovisas i tabell 1 för T-gruppen och tabell 2 för N-gruppen. Utifrån de 157 konsonanter som fanns med i orden som användes ur SVANTE beräknades PCC enligt formeln:

$$PCC: \frac{\text{Antal korrekt uttalade konsonanter}}{\text{Det totala antalet förekommande konsonanter (157)}} \times 100$$

*Lyssnarbedömning.* För lyssnarbedömningen av SWITCH användes ett bedömningsformulär framtaget av författarna. Formuläret bestod av 13 sidor där den första sidan innehöll information om lyssnarbedömningen samt ett antal frågor som skulle besvaras av lyssnarna. Frågorna tog upp ålder (år:mån), kön och hörselstatus. I efterhand lades två frågor till som tog upp starkaste språk samt att lyssnarna fick skriva ned sina tankar kring lyssnarbedömningen. De lyssnare som inte hade möjlighet att besvara dessa frågor på plats besvarade dem i efterhand via mail. De återstående 12 sidorna i bedömningsformuläret bestod av tomma numrerade rader där lyssnarna, ortografiskt, skulle skriva ner de ord som de uppfattade att barnen sade. Innan lyssnartranskriptionen gjorde författarna en testbedömning för att pröva metoden och kunna förutse eventuella lyssnarfrågor och därmed bättre kunna planera informationen till lyssnarna.

Lyssnarbedömningarna ägde rum på Göteborgs universitet. Två lyssnare gjorde sina bedömningar i hemmet. De bedömningar som genomfördes vid Göteborgs universitet ägde rum vid fyra olika tillfällen. Lyssnarna satt i samma sal men genomförde bedömningarna enskilt med hjälp av hörlurar. När lyssnarbedömningen ägde rum fick de ingående information kring testförfarandet, både i form av skriftlig och muntlig information. De fick också möjlighet att ställa frågor om eventuella oklarheter förekom. En av författarna fanns alltid tillgänglig under bedömningen om frågor skulle uppkomma.

Lyssnarna delades slumpmässigt in i fem olika grupper (Grupp A-E), detta innebar att fyra lyssnare bedömde samma material. Materialet fördelades på detta sätt för att varje enskild lyssnare skulle få en hanterbar mängd ljudfiler att bedöma. Varje grupp av lyssnare bedömde material som uppskattades vara likvärdigt avseende PCC och kön. Att materialet fördelades utifrån PCC beror på att varje grupp skulle få lika stor ”auditiv börda”. Att lyssna på mycket svårförståeligt tal kan vara mer ansträngande är att lyssna på tal som är mycket lätt att förstå.

Tabell 3 presenterar medelvärdet, standardavvikelsen samt lägsta och högsta PCC-värdet för varje lyssnargrupp. För att säkerställa att det inte förekom några signifikanta grupp-differenser vad gäller PCC-poäng genomfördes analys med Kruskal Wallis. Denna analys indikerade inga signifikanta skillnader i PCC-poäng mellan de olika lyssnargrupperna,  $U = ,426, p = ,98$ .

Tabell 3  
*PCC inom de olika lyssnargrupperna, Grupp A-E*

<i>Grupp</i>	<i>M</i>	<i>SD</i>	<i>Lägsta</i>	<i>Högsta</i>
<i>A</i>	78	16	61	99
<i>B</i>	77	26	49	100
<i>C</i>	79	15	65	100
<i>D</i>	78	24	54	100
<i>E</i>	79	22	59	100

*Medelvärden (M), standardavvikelse (SD) samt högsta och lägsta PCC-värdet i varje grupp angivet i procent.*

Lyssnarna bedömde en hel SWITCH-lista från ett barn i taget. Ordningen på SWITCH-listorna var randomiserade så att inte listor från samma barn kom efter varandra. Denna metod valdes för att minska eventuell familjaritet med barnens tal. Varje lyssnargrupp lyssnade på sex listor från T-gruppen och sex listor från N-gruppen. Dessa numrerades från 1-12 inom varje lyssnargrupp. Bedömningen var utformad så att varje lyssnare kunde genomföra transkriptionen i sin egen takt. Varje lyssnare transkriberade ortografiskt enligt open-set principen. De ord lyssnarna inte uppfattade eller var osäkra på markerade de med ett "X" i formuläret. Weiss (1982) menar att förståelighetspoäng som inkluderar gissningar har lägre reliabilitet. Vidare menar han att osäkerhet tyder på att en lyssnare inte har förstått ordet. Bedömningen skedde via VLC mediaplayer. De uttalade orden låg i enskilda filer och lyssnarna kunde lyssna på varje ord två gånger.

Lyssnarna hade en obligatorisk paus på 15 minuter efter ca 45 minuter transkribering för att minimera transkriptionstiden till en rimlig tidsperiod och för att minska eventuell uttrötningseffekt. Sammanlagd tid för lyssnarbedömningen var cirka 3,5 timme per lyssnare, inklusive rast. För bedömning av intrabedömarreliabilitet fick deltagarna på nytt bedöma de två listorna från det barn som hade lägst PCC av de barn som de redan hade bedömt. Detta blev naturligt ett barn från T-gruppen. Att listor från barn med lägst PCC duplicerades beror på att det har visat sig svårare för lyssnare att vara konsekventa i sina bedömningar när de har bedömt tal med kraftigt nedsatt förståelighet (Hodge & Gotzke, 2007a). Lyssnarna fick därför på nytt bedöma de barn med lägst PCC vilket bör ha gett en säkrare beräkning av intrabedömarreliabiliteten. För att minska risken för att lyssnarna skulle känna igen materialet hade dessa listor placerats i början respektive i slutet av lyssnarbedömningen.

*Förståelighetsberäkning (F %).* Författarna räknade för hand ut det antal ord som lyssnarna hade uppfattat korrekt och överförde det till en procentsats, *F %*, genom formeln;

$$F \% = \frac{\text{Antal rätt transkriberade ord} \times 100}{\text{Det totala antalet ord}}$$

De ord som bedömdes som rätt, förutom de som var korrekt transkriberade, var homofoner (ord som uttalas likadant men stavas på olika sätt) och felstavningar. Exempel på homofoner var en/än, häller/heller och scen/sen. Oskar/åskar bedömdes som en homofon trots skillnad i accent eftersom den skillnaden ansågs vara mycket svår att särskilja. Vid analys fann författarna att enstaka lyssnare hade skrivit ner båda orden trots instruktioner om att enbart skriva ner ett ord. Det var därför svårt att veta om resterande lyssnarna kan ha funderat på om det kan ha varit det ena eller andra ordet.

## Dataanalyser

Icke-parametriska tester användes genomgående eftersom data inte var normalfördelad. Signifikansnivån bestämdes till  $p < ,01$ .

*Likvärdighet.* Alla barn repeterade två slumpmässigt utvalda listor, lista 1 och lista 2. Likvärdigheten mellan listorna beräknades med Spearmans Rho ( $r_s$ ) utifrån medelvärden för varje barn. Medelvärdet för varje lista beräknades utifrån de fyra lyssnare som hade bedömt samma barn. För att ytterligare undersöka korrelationen genomfördes samma beräkning utifrån rådata, d.v.s. alla lyssnarnas individuella bedömningar. Beräkningar utifrån medelvärden kan ha jämnat ut skillnader mellan olika lyssnare, därför var det av intresse att göra beräkningar där alla lyssnarnas bedömningar inkluderas i beräkningen. Medelskillnaden, d.v.s. skillnaden mätt i procentenheter mellan lista 1 och lista 2 hos varje enskilt barn beräknades utifrån rådata.

*Reliabilitet.* Interbedömarreliabilitet undersöktes genom att beräkna samstämmigheten mellan de 4 lyssnarna som bedömde samma material (Grupp A-E). För denna analys användes ICC (intraclass correlation coefficient). Analysen undersöker hur olika värden följs åt men inkluderar också skillnaden mellan resultat i beräkningen. Intrabedömarreliabilitet beräknades med Spearmans Rho ( $r_s$ ) utifrån de listorna som varje lyssnare bedömde på nytt.

*Validitet.* Detta undersöktes genom att det kontrollerades om testet kunde visa skillnader i förståelighet mellan barn med och utan talavvikelser. Validiteten undersöktes också genom att beräkna korrelationen mellan förståelighet och artikulation (PCC). Med Mann Whitney U undersöktes skillnaden i förståelighet mellan gruppen med talavvikelser och gruppen utan talavvikelser, T- respektive N-gruppen. Med Spearmans rho undersöktes korrelationen mellan PCC och förståelighet. Medelvärdet för förståelighet hos varje enskilt barn ( $M_{tot}$ ) beräknades genom att medelvärdet för lista 1 ( $M_1$ ) och lista 2 ( $M_2$ ) adderades och delades med två.

$$M_{tot} = \frac{M_1 + M_2}{2}$$

## Resultat

### Likvärdighet

Korrelationen utifrån medelvärden var  $r_s = ,96$  ( $p < ,01$ ) och utifrån rådata  $r_s = ,95$  ( $p < ,01$ ). Hos 13 lyssnare varierade förståeligheten med  $\leq$  fem procentenheter mellan lista 1 och lista 2. Hos fyra lyssnare varierade förståeligheten med  $\leq$  10 procentenheter mellan lista 1 och lista 2. Hos två lyssnare varierade förståeligheten  $\leq$  15 procentenheter mellan lista 1 och lista 2. Hos en lyssnare var skillnaden som störst 23 procentenheter mellan lista 1 och 2. Medelskillnaden mellan lista 1 och 2 beräknat utifrån rådata, angett i procentenheter, var  $M = 5,1$  ( $s = 4,6$ , min = 0, max = 23). Tabell 4 listar alla barn samt de resultat de erhöll av varje enskild lyssnare, medelvärdet för varje lista samt deras totala förståelighet ( $M_{tot}$ ).

Tabell 4

Förståelighet angett i procent hos deltagarna så som det har bedömts av lyssnare L1-L20.

Deltagare	Lista1					Lista2					DFM	$M_{tot}$
	L1	L2	L3	L4	$M_1$	L1	L2	L3	L4	$M_2$		
T04	35	33	52	55	44	35	32	37	40	36	8	40
T06	78	68	75	78	75	85	72	82	88	82	7	79
N03	92	93	95	92	93	98	92	98	97	96	3	95
N04	72	57	83	78	73	85	80	82	92	85	12	79
Deltagare	Lista1					Lista2					DFM	$M_{tot}$
	L5	L6	L7	L8	$M_1$	L5	L6	L7	L8	$M_2$		
T08	15	10	15	22	16	10	10	20	12	13	3	15
T09	28	28	38	33	32	33	25	38	33	32	0	32
N02	88	93	98	95	94	83	88	95	88	89	5	92
N06	95	97	100	100	98	95	97	98	98	97	1	98
Deltagare	Lista1					Lista2					DFM	$M_{tot}$
	L9	L10	L11	L12	$M_1$	L9	L10	L11	L12	$M_2$		
T05	48	53	57	60	55	58	48	45	62	53	2	54
T07	28	27	28	27	28	27	23	30	25	26	2	27
N01	38	48	43	55	46	42	47	45	53	47	1	47
N07	85	78	83	83	82	78	82	73	78	78	4	80
Deltagare	Lista1					Lista2					DFM	$M_{tot}$
	L13	L14	L15	L16	$M_1$	L13	L14	L15	L16	$M_2$		
T02	17	12	10	20	15	17	17	22	20	19	4	17
T03	20	28	28	25	25	22	17	22	23	21	4	23
N05	90	88	80	80	85	85	73	77	78	78	7	82
N08	92	97	90	90	92	90	95	95	87	92	0	92
Deltagare	Lista1					Lista2					DFM	$M_{tot}$
	L17	L18	L19	L20	$M_1$	L17	L18	L19	L20	$M_2$		
T01	17	22	23	18	20	28	27	27	32	29	9	25
T010	30	23	18	37	27	30	22	27	30	27	0	27
N09	83	90	75	82	83	80	83	83	82	82	1	83
N010	90	92	92	88	91	83	92	87	90	88	3	90

Not.  $M_{1-2}$  = medel för varje barns förståelighet som har beräknats utifrån de fyra lyssnarna.  
DFM = differensen mellan medelvärdena från lista 1 och lista 2.  $M_{tot}$  = den totala förståeligheten beräknat utifrån medelvärdena på lista 1 och lista 2

## Reliabilitet

*Interbedömarreliabilitet.* ICC hos de lyssnarna som bedömt samma material låg mellan ,91 - ,99.

*Intrabedömarreliabilitet.* Korrelationen var  $r_s = ,80$  hos samtliga lyssnare ( $p < ,01$ ) Medelskillnaden mellan de två bedömningarna var  $M = 4,4$  ( $s = 3,7$ , min = 0, max = 13) procentenheter hos lyssnarna.

## Validitet

Medelvärde vad gäller förståelighet, beräknat utifrån rådata, var  $M = 83\%$  ( $s = 14,4$ , min = 38 och max = 100) för N-gruppen och  $M = 34\%$  ( $s = 19$ , min = 10 och max = 88) för T-gruppen. Mann Whitney U, där varje barns medelvärde ( $M_{tot}$ ) användes, visade på signifikant skillnad i F % mellan de båda grupperna ( $U = 2,0$ ,  $p < ,01$ ). Korrelationen mellan F % och PCC beräknat utifrån medelvärdena var  $r_s = ,94$ ,  $p < ,01$ .

## Diskussion

Syftet med den aktuella studien var att undersöka likvärdighet, reliabilitet och validitet i SWITCH- Swedish intelligibility test for children. Detta undersöktes genom att 40 ordlistor ur SWITCH provades ut på barn med talavvikelse och på barn med typisk talutveckling. I Sverige finns det idag inget test som mäter förståeligheten i barns tal, vilket kan tyckas märkligt med tanke på att förståelighet är en av de viktigaste faktorerna när framgång med en intervention undersöks (Hodge & Gotzke, 2007a; Yorkston & Beukelman, 1978).

Resultaten i den aktuella studien visade att listorna i SWITCH är likvärdiga. Medelskillnaden mellan de olika listorna låg på fem procentenheter och reliabiliteten var  $r_s = ,96$ . Resultaten stämmer överens med tidigare studier där ordlistor och repetition som eliciteringsmetod användes. Zajac et al. (2010) undersökte likvärdighet, reliabilitet och validitet hos ordlistorna i ett förståelighetstest för barn med gomspalt. Medelskillnaden mellan de olika listorna som ett och samma barn repeterade var en procentenhet och reliabiliteten var  $r = ,97$  ( $p < ,01$ ) mellan första och andra listan. I föreliggande studie har medelskillnaden beräknats på ett annat sätt varför resultaten inte är helt jämförbara. I den aktuella studien beräknades medelskillnaden mellan de olika listorna utifrån rådata, dvs. varje lyssnares bedömning av barnens förståelighet. Zajac et al. (2010) beräknade medelskillnaden genom att beräkna ett medelvärde för alla deltagare för lista 1 och lista 2. Skillnaden mellan dessa medelvärden i procentenheter beräknades sedan. En beräkning som följer den formel som Zajac et al. (2010) använde visar att skillnaden mellan lista 1 och lista 2 var två procentenheter i föreliggande studie. I den aktuella studien valdes inte denna metod för beräkning av skillnaden mellan lista 1 och 2 då denna typ av beräkning kan ha dolt eventuella skillnader i bedömningen som kan förekomma mellan olika lyssnare i bedömningen av förståelighet.

I den aktuella studien var inter- och intrabedömarreliabilitet hög. ICC gav ett värde mellan ,91 - ,99. som enligt Portney och Watkins (2000) indikerar på nästan perfekt

samstämmighet. Resultaten kan jämföras med Zajac et al. (2010) som fick ett ICC mellan ,96 - ,99 ( $p < ,01$ ). Intrabedömarreliabiliteten i föreliggande studie låg på  $r_s = ,801$ . Zajac et al (2010) hade en intrabedömarreliabilitet på  $r = ,92$ ,  $r = ,94$  och  $r = ,95$  ( $p < ,001$ ). Att Zajac et al. (2010) har tre värden och den aktuella studien endast ett värde beror på metodskillnader i beräkningen av intrabedömarreliabilitet. I den aktuella studien bedömde alla lyssnarna på nytt om ett barns båda listor ungefär tre timmar efter första bedömningen av samma filer. Värdet för intrabedömarreliabiliteten i den aktuella studien är beräknad utifrån samtliga lyssnare eftersom korrelationsanalysen kräver fler än två värden. I studien av Zajac et al. (2010) bedömde tre lyssnare på nytt sitt material efter tre veckor.

Validiteten i den aktuella studien var hög eftersom den kunde visa på en signifikant skillnad mellan gruppen med och utan talavvikelse. Vidare påvisades en stark korrelation mellan förståelighet och artikulation,  $r_s = ,94$  ( $p < ,01$ ). Zajac et al. (2010) visade på en signifikant skillnad vad gäller förståelighetspoäng mellan gruppen med och utan läpp-käk och gomspalt, vilket tyder på god validitet. Vidare undersöktes validiteten genom att jämföra artikulation (PCC) och förståelighet med varandra. Dessa visade på en måttligt hög korrelation  $r = ,79$  ( $p < ,001$ ).

Det bör dock hållas i åtanke att i föreliggande studie användes icke-parametriska tester genomgående på grund av att data inte var normalfördelad. Zajac et al. (2010) använde genomgående parametriska tester.

Morris et al. (1995) och Hodge och Gotzke (2007a) undersökte validiteten och reliabiliteten i förståelighetstest som bygger på enskilda ord som barnen repeterade. Interbedömarreliabiliteten i P-SIM, the Preschool Speech Intelligibility Measure (Morris et al, 1995) beräknades till  $r = ,91 - ,97$ . Validiteten beräknades till  $r = ,73$  genom att P-SIM värdet jämfördes med artikulationen hos barnen. Interbedömarreliabiliteten i SIP-CCLP, Speech Intelligibility Probe for Children with Cleft Palate, beräknades till  $r = ,96$  (Hodge & Gotzke, 2007a). Validiteten beräknades till  $r = ,88$  genom att förståeligheten i enskilda ord jämfördes med förståeligheten i spontantalet hos barnen. Deras resultat stämmer överens med föreliggande studies resultat där validiteten beräknades till  $r_s = ,94$  och stärker ytterligare att metoden att använda ordlistor för att mäta förståelighet i barns tal är en valid och reliabel metod.

I den aktuella studien var det ett relativt stort spann i lyssnarnas bedömningar vad gäller likvärdigheten hos varje barns två listor, från mycket små skillnader till relativt stora skillnader mellan bedömningarna av de två listorna hos en och samma lyssnare. Hos majoriteten av lyssnarna var variationen mellan lista 1 och lista 2  $\leq 5$  procentenheter. En lyssnare hade som mest 23 procentenheters skillnad mellan lista 1 och 2. Det resultatet får dock ses som en outlier eftersom det strider mot de övriga lyssnarnas bedömningar av just det barnet (N04). Denna studie och andra studier visar på att det förekommer variationer mellan lyssnare och inom lyssnare (Zajac et al., 2010). Detta beror på att testet bygger på perceptuella bedömningar. Det är svårt att nå 100 % samstämmighet mellan olika bedömare och att få en och samma lyssnare att vara konsekvent i sin bedömning (Kent, 1996). Detta ger indikationer på att det bör finnas ett spann inom vilket en lyssnare ska kunna befinna sig utan att det påverkar den förståelighetspoäng som barnet får. Det skulle kunna innebära att ett barn som befann sig inom exempelvis spannet 10 % - 20 % i förståelighet mätt med SWITCH skulle bli tilldelad en viss förståelighetspoäng. För att undersöka om det behövs ett spann och hur stort det i sådana fall bör vara behöver studier med fler lyssnare genomföras.



I den aktuella studien fanns det två barn i N-gruppen (N01, N04) med låga PCC-poäng, 79 % respektive 72 %, vilket är betydligt lägre än vad övriga barn i N- gruppen erhöll. Zajac et al. (2010) hade en PCC-gräns på 80 % satts för de barn som skulle agera kontroll. Schriberg och Kwiatkowski (1982) menar att barn med ett PCC mellan 60 % - 80 % har en lätt till måttlig artikulationspåverkan. I föreliggande studie fanns ingen PCC-gräns eftersom det är viktigt att normalvariationen inte påverkas. Det främsta syftet var ändå att undersöka om PCC och förståelighet korrelerade vilket de gjorde i föreliggande studie. Enligt Schriberg och Kwiatkowski (1982) som har utvecklat PCC bör PCC beräknas på barns spontantal. I föreliggande studie var talet ibland mycket svårförståeligt varför PCC beräknades utifrån ordnivå.

Det finns olika studier och olika resultat som beskriver korrelationen mellan artikulation och förståelighet. Det finns resultat som visar på att det finns ett samband (Morris et al., 1995; Zajac et al., 2010) och resultat som visar på ett mer begränsat samband (Ertmer, 2010). Föreliggande studie visar dock på en tydlig korrelation mellan PCC och förståelighet. Att korrelationen är hög i föreliggande studie och i studien gjord av Zajac et al. (2010) kan bero på att det var korrelationen mellan artikulation och förståelighet på ordnivå som undersöktes. Materialet var likvärdigt vad gäller både längd och komplexitet. När Ertmer (2010) beräknade korrelationen mellan förståelighet och artikulation hade förståeligheten beräknats utifrån meningar och artikulationen beräknats utifrån ord. Detta menar Ertmer (2010) kan ha påverkat resultaten. Påståendet styrks av Stimley och Hambrecht (1999) som fann att korrelationen mellan artikulation mätt utifrån enskilda ord och förståelighet mätt utifrån enskilda ord korrelerade mer än när artikulationen i enskilda ord jämfördes med förståeligheten beräknat utifrån spontantal. Hur korrelationen ser ut kan därmed bero på vilket talmaterial beräkningarna görs.

I föreliggande studie fick inte alla barnen i N-gruppen en förståelighet på 100 %. Studier indikerar att barn som är över 4 år bör ha en förståelighet på 100 % (Gordon-Brannan & Hodson, 2000; Weiss, 1982). Samtidigt har den aktuella studien och tidigare studier visat att barn utan talavvikelse sällan når upp till den nivån (Hodge & Gotzke, 2007a; Zajac et al., 2010). Hustad (2006) säger dock att en liten ändring av metod kan förändra resultatet i förståelighetsbedömningen signifikant. Således tyder det på att nya värden bör tas fram för varje nytt test som mäter förståelighet och att det kan vara svårt att dra paralleller mellan studier på grund av att de har olika metoder.

Resultaten i föreliggande studie visar på att det kan diskuteras om alla deltagare var lämpliga att ingå i kontrollgruppen. Medelvärdet i förståelighet för N01 beräknades till  $M_{tot} = 47$  %. Detta är mycket lågt i relation till de övriga deltagarna och minskar medelvärdet för N-Gruppen totalt. N01 var den yngsta av alla deltagarna i N-gruppen, 4:8 år och hade låg PCC, vilket kan ha påverkat förståeligheten. Orden i SWITCH är framtagna för barn mellan 5 - 8 år och förståeligheten hos N01 kan därmed ha påverkats av att N01 inte kände till vissa ord. William och Stackhouse (2000) menar att det vid alla typer av ljudmässig repetition krävs auditiv förmåga och adekvata perceptuella tolkningssystem. Trots rapportering av god hörsel och normal språkutveckling går det inte att veta med säkerhet om alla barn i N- gruppen låg på samma nivå vad gäller dessa parametrar.

Lyssnarna i den aktuella studien hade olika mycket erfarenhet att lyssna på tal där förståeligheten är påverkad. Kent et al. (1994) menar att en talare som är van vid en viss talavvikelse har lättare att förstå den typen av tal. Detta stärks av Hustad och Cahill (2003) som i en studie fann att förståeligheten ökade hos personer med dysartri när

lyssnare fick höra på deras tal mer än en gång och blev vana vid det. Att lyssnarna inte gick i samma årskurs och hade varierande erfarenhet vad gäller att lyssna på tal med nedsatt förståelighet kan ha påverkat lyssnarbedömningarna och det kan förklara varför variationen var stor vid bedömningen av vissa barn. De som har studerat i en termin kan nästan ses som naiva lyssnare. Trots dessa olikheter hos lyssnarna visar analyserna att lyssnarna har hög samstämmighet. Inom den logopediska verksamheten kommer samma logoped göra bedömning före och efter intervention varför dessa skillnaden mellan olika lyssnare inte har betydelse för den kliniska verksamheten. Snarare talar resultaten om att variationer förekommer mellan olika lyssnare och att konsensus är svårt att nå. Detta är ett problem som förekommer inom metoden perceptuella bedömningar (Kent, 1996). Vid all perceptuell bedömning är det svårt att hitta 100 % enighet mellan olika lyssnare. För att denna enighet ska nås kan ett steg vara att tydliga instruktioner ges i hur bedömningen ska gå till (Kent, 1996).

Hur resultaten i föreliggande studie har påverkats av att det inte var verksamma logopedier som bedömde går bara att spekulera kring. En hypotes är att förståeligheten bör bli högre när verksamma logopedier gör bedömningarna eftersom de har mer vana av att lyssna på tal med nedsatt förståelighet.

Ett par deltagare hade mycket låg röststyrka. Reaktionen från lyssnarna gällande detta var att det kunde vara svårt att höra vilka ord barnen sade. Shriberg och Kwiatkowski (1982) har sett att kvaliteten på inspelningarna är en faktor som kan påverka förståelighetsbedömningen. I efterhand visade det sig också att VLC- mediaplayer klippte bort slutet på ett par av orden. Detta kan ha inneburit att en final konsonant inte spelades upp för lyssnarna. Författarna vet inte i vilken grad detta kan ha påverkat resultaten.

I föreliggande studie bedömde lyssnarna om samma material på nytt inom loppet av tre timmar för beräkning av intrabedömrareliabilitet. Detta kan ha påverkat utfallet eftersom förståeligheten hos barnen i de flesta fall ökade med  $\geq 8$  procentenheter mellan första och andra bedömningen. Detta kan tyda på att lyssnarna hade vant sig det material som de bedömde. Analysen av intrabedömrareliabiliteten kunde ha blivit säkrare om materialet inte hade bedömts på nytt vid ett och samma bedömningstillfälle.

Grupperna i föreliggande studie var relativt stora med totalt 20 barn i jämförelse med tidigare studier. Deltagarantalet i tidigare studier varierade mellan 15 - 38 barn, majoriteten av dessa studier hade under 20 barn (Hodge & Gotzke, 2007a; Morris et al., 1995; Whitehill & Chau, 2004; Zajac et al., 2010). I föreliggande studie var det 4 lyssnare som lyssnade på samma material. Det kan jämföras med studien gjord av Zajac et al. (2010) som hade fem lyssnare och Hodge och Gotzke (2007a) som hade sex lyssnare per barn. Morris et al. (1995) hade tre lyssnare som bedömde samma material. I förhållande till tidigare studier är antalet lyssnare per barn samt antalet medverkande barn i föreliggande studie relativt högt.

För att inspelningssituationen skulle bli så likvärdig som möjligt för samtliga barn var det alltid en och samma testledare som agerade artikulatorisk modell. Zajac et al. (2010) menar att reliabilitetsanalyserna kan påverkas om det förekommer inkonsekvens vad gäller artikulatorisk modell vid repetition.

SWITCH bygger på enskilda ord. Fördelen med att använda ett fast material, så som enskilda ord, är att förståeligheten går att jämföra över tid hos ett och samma barn (Morris et al., 1995) vilket innebär att effekterna av en intervention kan undersökas. Weston och Shriberg (1992) menar att spontantal bör användas som material eftersom det representerar hur barnets kommunikation ser ut i en naturlig kontext. Morris et al.

(1995) menar dock att spontantal varierar mycket från tillfälle till tillfälle, vilket gör att bedömning före och efter en intervention aldrig blir desamma, och därmed inte jämförbara. Kent et al. (1989) tar upp att en persons tal aldrig kan tillskrivas en viss "förståelighetspoäng", eftersom det kan variera så pass mycket beroende på vem det är som lyssnar och i vilken situation talare och lyssnare befinner sig. Det är snarare så att en viss talare har en variation av förståelighetspoäng som varierar beroende på hur van lyssnaren är vid den som talar, hur den fysiska miljön som omger talare och lyssnare ser ut samt kommunikativ motivation hos talare och lyssnare. Senare förespråkar Kent (1992) dock enskilda ord när förståelighet ska bedömas. Han menar att de är tillräckligt korta för att alla ska kunna uttala dem och att längre yttranden, såsom meningar, kan innebära alltför stor ansträngning för personer med stora motoriska svårigheter.

SWITCH är lätt att administrera och tidseffektivt. Repetition av en lista tog mellan 2-5 minuter för barnen att repetera och de hade inga större svårigheter att förstå repetitionsuppgiften. Varje lista i SWITCH består av 60 ord. I en studie av Hodge och Gotzke (2007b) där de undersökte förståeligheten hos barn med och utan språkstörning fann de att förståeligheten inte förändrades signifikant oavsett om det beräknades på 100 ord eller 225 ord. Detta undersöktes dock i spontantal men det kan ge indikationer på att talmaterialets storlek inte bör påverka utfallet. Det som därmed kan påverka utfallet är de ord vilket ett talmaterial består av. Orden i SWITCH är därför utvalda ur ett material med ord som förekommer i barns vokabulär och därmed anses vara familjära för barnet. Detta är en stor styrka i materialet. Tidigare när förståelighet i tal har bedömts har artikulationstest främst använts (Hartelius & Lohmander, 2008). Artikulationstest är främst framtagna för att undersöka olika fonologiska processer och är därmed i första hand inte utvalda att passa barns vokabulär. Som nämnts tidigare påverkar detta förståeligheten (Morris et al., 1995). I en studie gjord av Kent et al. (1994) tas det upp att komplexiteten i talmaterialet påverkar förståeligheten. Orden i SWITCH innehåller olika stavelselängder vilket innebär att komplexiteten i testet varierar. Detta stärker validiteten för SWITCH. I studien av Zajac et al. (2010) hade gruppen med talavvikelser som följd av läpp-käk-gomspalt högre förståelighet än i föreliggande studie, 70 % respektive 34 %. Det kan bero skillnader i typ av talavvikelse men det kan också bero på skillnader i talmaterial. I studien av Zajac et al. (2010) bestod talmaterialet av enstaviga ord.

När förståelighet bedöms ska lyssnaren inte veta vad talaren säger (Hartelius & Lohmander, 2008). Dock är det så att det inte kan förekomma helt unika ord för varje ny testning eftersom antalet ord inte är oändligt. Samtidigt är inte spontantal ett bra material att använda eftersom det varierar för mycket från gång till gång och därmed inte blir en reliabel och valid metod för att mäta effekten av en intervention (Morris et al., 1995). Ordbanken i SWITCH består av 1243 ord. Detta kan jämföras med studien av Zajac et al. (2010) där ordbanken bestod av 510 ord och P-SIM, The Preschool Speech Intelligibility Measure, (Morris et al., 1995) som har 600 ord i ordbanken. Jämförelsevis innehåller SWITCH mer än dubbelt så många ord vilket kan minska eventuell familjaritet med talmaterialet.

Inom den kliniska verksamheten är det av stort intresse att undersöka vilka processer som påverkar förståeligheten mest eftersom det är en faktor som bör styra val av behandling. Den fonologiska interventionen baseras idag på den normala fonologiska utvecklingen, alltså hur barn normalt lär sig olika språkljud (Bowen, 1998). Klein och Flint (2006) fann att de fonologiska processer som påverkar förståelighet mest inte alltid är den process som behandlas först. Detta strider mot det viktigaste målet med en

intervention alltså ökad förståelighet i talet (Hodge & Gotzke, 2007b; Klein & Flint, 2006; Yorkston & Beukelman, 1978). Genom att jämföra resultaten i ett förståelighetstest med resultaten ur ett artikulationstest kan det i framtiden dras slutsatser kring hur olika talavvikelser påverkar förståeligheten mest. För denna analys behövs en valid och reliabel metod för att mäta förståeligheten i talet.

### *Slutsats*

Att utveckla ett förståelighetstest som är anpassat för barn, både vad gäller metod och vokabulär, är av stor betydelse. Resultaten i föreliggande studie anses vara ett led på vägen till en mer enhetlig metod vad gäller bedömning av förståelighet av barns tal i Sverige. Ett reliabelt förståelighetstest leder till att effekten av olika interventioner kan bedömas och att det kan undersöka vilka artikulatoriska processer som påverkar förståeligheten mest.

I föreliggande studie testades SWITCH enbart på barn med talavvikelser till följd av språkstörning. Testet är tänkt att kunna användas även på talstörningar, såsom stamning och dysartri och bör därmed testas på dessa populationer för att säkerställa att SWITCH kan visa på grad av förståelighet även vid de diagnoserna.

## Referenslista

- Bowen, C. (1998). PACT: A broad-based approach to phonological therapy. Hämtad 19 maj, 2011, från <http://www.speech-language-therapy.com/clinphonology.html>.
- Case, C., Forsber, G., & Uppman, A. (2009). Ordfrekvens och ordformer: Analys av samtal mellan barn i åldersgrupperna 5-8 år respektive 9-12 år. *Examensarbete Linköpings universitet VT 09*.
- Ertmer, D. J. (2010). Relationships between speech intelligibility and word articulation scores in children with hearing loss. *Journal of speech, language and hearing research*, 53, 1075-1086
- Gordon-Brannan, M., & Hodson Williams, B. (2000). Intelligibility/severity measurements of prekindergarten children's speech. *American journal of speech-language pathology*, 9, 141-150
- Hartelius, L., & Lohmander, A. (2008). Talstörningar allmän del. I L. Hartelius, U. Nettelbladt & B. Hammarberg (red:er.), *Logopedi* (ss. 357-375) Lund: Studentlitteratur.
- Hodge, M., & Gotze, C. L. (2007a). Preliminary results of an intelligibility measure for English-speaking children with cleft palate. *Cleft Palate and Craniofacial Journal*, 44(2), 163-174
- Hodge, M., & Gotzke, C. L. (2007b). Effect of Sample Length on Children's Speech Intelligibility Scores. American Speech-Language-Hearing Association Annual Convention, Boston, MA. Convention, Boston, MA.
- Hustad, K. C. (2006). A closer look at Transcription Intelligibility for speakers with dysarthria: Evaluation of scoring paradigms and linguistic errors made by listeners. *American Journal of Speech-Language Pathology*, 15, 268-277.

- Hustad, K. C., & Cahill, M. A. (2003). Effects of presentation mode and repeated familiarization on intelligibility of dysarthric speech. *American Journal of Speech-Language Pathology*, 12, 198-208.
- Johannisson, T. B., Lohmander, A., & Persson, C. (2010). Assessing intelligibility in ten-year-olds: a methodological study. Manuskript
- Johannisson, T. (2011). SWITCH- swedish intelligibility test for children. Opublicerat.
- Kent, R. D., Weismer, G., Kent, J. F., & Rosenbek, J. C. (1989). Toward phonetic intelligibility testing in dysarthria. *Journal of speech and hearing disorders*. 54, 482-499.
- Kent, R. D. (Ed). (1992). *Intelligibility in speech disorders: Theory, measurement and management*. Philadelphia, PA: Benjamins.
- Kent, R. D., Miolo, G., & Bloedel, S. (1994). The intelligibility of children's speech: a review of evaluation procedures. *American Journal of Speech-Language Pathology*, 3(2), 81-95.
- Kent, R., D. (1996). Hearing and believing: Some limits to the auditory-perceptual assessment of speech and voice disorders. *American journal of speech-language pathology*, 5, 7-23
- Klein, E. S. (1996). *Clinical phonology: Assessment and treatment of articulation disorders in children and adults*. San Diego, CA: Singular
- Klein, E. S., & Flint, C. B. (2006). Measurement of intelligibility in disordered speech. *Language, speech and hearing services in schools*. 37, 191-199
- Lillvik, M., Allemark, E., Karlström, P., & Hartelius, L., (1999). Intelligibility of dysarthric speech in words and sentences: Development of a computerised assessment procedure in Sweden. *Logopedics phoniatrics vocology*, 24, 107-117
- Lohmander, A., Borell, E., Henningsson, G., Havstam, C., Lundeborg, I., & Persson, C. (2005). *SVANTE- Svenskt artikulations- och nasalitetstest*. Pedagogisk design.
- Malterud, K. (1998). *Kvalitativa metoder i medicinsk forskning*. Lund; Studentlitteratur.
- Morris, S. R., Wilcox, K. A., & Schooling, T. L. (1995). The preschool speech intelligibility measure. *American journal of speech-language pathology*, 4, 22-28.
- Nettelbladt, U., & Salameh, E-K. (2007). *Språkutveckling och språkstörning hos barn*. Stockholm; Studentlitteratur.
- Portney, L. G., & Watkins M. P. (2000). *Foundations of clinical research: Applications to practice*. New Jersey: Prentice Hall Inc.
- dos Santos Barreto, S., & Zazo Ortiz, S. (2008). Intelligibility measurements in speech disorders: a critical review of the literature. *Pro-Fono Revista de Atualizacao Cientifica*, 20 (3), 201-206.
- Shriberg, L. D., & Kwiatkowski, J. (1982). Phonological disorders III: A procedure for assessing severity of involvement. *Journal of speech and hearing disorders*, 47, 242-256.
- Shriberg, L. D., & Kwiatkowski, J. (1992). Intelligibility assessment in developmental phonological disorders: accuracy of caregiver gloss. *Journal of speech and hearing research*, 32, 1095-1104.
- Stimley, M.A., & Hambrecht, G. (1999). Comparisons of children's single-word articulation proficiency, single-word speech intelligibility, and conversational speech intelligibility. *Revue d'orthophonie et d'audiologie*, 23, 19-23.
- Weiss, C. E. (1982). *Weiss Intelligibility Test*. Oregon: C.C. Publications.
- Weston, A. D., & Shriberg, L. D. (1992). Contextual and linguistic correlates of

- intelligibility in children with development phonological disorders. *Journal of speech and hearing research*, 35, 1316-1332.
- Whitehill, T. L. (2000). Assessing intelligibility in speakers with cleft palate: A critical review of the literature. *Cleft Palate- Craniofacial Journal*, 39(1), 50-58.
- Whitehill, T. L., & Chau, C. H-F. (2004). Single-word intelligibility in speakers with repaired cleft palate. *Clinical linguistics & Phonetics*, 18(4-5) 341-355.
- Williams, P., & Stackhouse, J. (2000). Rate, accuracy and consistency: diadochokinetic performance of young, normally developing children. *Clinical linguistics & phonetics*, 14(4), 267-293.
- Yorkston, K. M., & Beukelman, D. R. (1978). A comparison of techniques for measuring intelligibility of dysarthric speech. *Journal of communication disorders*, 11, 499-512.
- Zajac, D. J., Plante, C., Lloyd, A., & Haley, K. L. (2010). Reliability and validity of a computer mediated single- word intelligibility test: preliminary findings for children with repaired cleft and palate. *Ej publicerat*

Ett varmt tack till rektorer, pedagoger, logopedier och handledare som har hjälpt oss att genomföra denna studie.