Avdelningen för logopedi, foniatri och audiologi
Institutionen för kliniska vetenskaper, Lund

Prosodi hos skolbarn med autismspektrumtillstånd

MarieHelene Dotevall & Louise Lendt

Logopedutbildningen, 2014
Vetenskapligt arbete, 30 högskolepoäng

Handledare: Ulrika Nettelbladt, Annika Dahlgren Sandberg
& Sofia Strömbergsson
SAMMANFATTNING


Nyckelord: autismspektrumtillstånd (AST), Aspergers syndrom (AS), prosodi, perceptuell analys, instrumentell analys, lyssnarbedömning
INNEHÅLLSFÖRTECKNING

1. INLEDNING .............................................................................................................. 5
   1.1 Studiens syfte ................................................................................................. 5

2. BAKGRUND ............................................................................................................ 5
   2.1 Autismspektrumtillstånd .................................................................................. 5
       2.1.1 Kommunikation vid AST ......................................................................... 6
   2.2 Prosodi ............................................................................................................. 6
       2.2.1 Svenskans prosodi .................................................................................. 7
       2.2.2 Prosodins domän .................................................................................. 8
   2.3 Prosodi och AST .............................................................................................. 8
       2.3.1 Taltonläge och intonation ..................................................................... 9
       2.3.2 Betoningsmönster ............................................................................... 10
       2.3.3 Frasering och yttrandelängd ................................................................. 10
       2.3.4 Taltempo ............................................................................................... 11
       2.3.5 Röstbetingade aspekter ....................................................................... 11
       2.3.6 Dialektala drag ...................................................................................... 12
       2.3.7 Form och funktion ................................................................................ 12
       2.3.8 Sammanfattande kommentarer .............................................................. 13
   2.4 Frågeställningar .............................................................................................. 14

3. METOD ................................................................................................................... 14
   3.1 Deltagare ........................................................................................................ 14
   3.2 Material .......................................................................................................... 15
   3.3 Databearbetning ............................................................................................ 15
   3.4 Kalibrering av skattningsprotokoll ................................................................. 16
   3.5 Genomförande ............................................................................................... 16
   3.6 Forskningsetiska överväganden ................................................................... 17
   3.7 Statistisk bearbetning .................................................................................... 17
       3.7.1 Modifiering av data inför statistisk analys ............................................. 18

4. RESULTAT ............................................................................................................... 18
   4.1 Instrumentell analys ....................................................................................... 18
       4.1.1 Jämförelser mellan grupper .................................................................. 18
   4.2 Perceptuell analys ......................................................................................... 20
       4.2.1 Jämförelser mellan grupper .................................................................. 20
       4.2.2 Prosodiska skillnaders påverkan på lyssnarintryck .............................. 21
   4.3 Korrelationer med språkliga mått ................................................................. 21

5. DISKUSSION .......................................................................................................... 22
   5.1 Resultatdiskussion ....................................................................................... 22
       5.1.1 Instrumentell analys ............................................................................. 22
       5.1.2 Relation till språkliga mått .................................................................... 24
       5.1.3 Perceptuell analys ............................................................................... 25
       5.1.4 Subtila variationers påverkan på lyssnarintrycket .................................. 26
       5.1.5 Problematisering av normalitetsbegreppet ........................................... 27
   5.2 Metoddiskussion ........................................................................................... 27
5.2.1 Perceptuell analys
5.2.2 Talmaterial
5.2.3 Instrumentell analys
5.3 Framtida forskning
5.4 Sammanfattande slutsatser och kliniska implikationer

6. TACK

7. REFERENSER

BILAGOR
Bilaga 1. Bedömningsprotokoll av prosodi
1. INLEDNING


Prosodisk förmåga förefaller variera kraftigt bland personer med AST (Peppé, McCann, Gibbon, O’Hare & Rutherford, 2007), och vissa individer uppvisar prosodiska mönster i enlighet med kontrollgrupper (Green & Tobin, 2009; Shriberg et al., 2001). Vidare har prosodisk förmåga hos personer med AST visat sig vara relaterad till språklig förmåga (DePape, Chen, Hall & Trainor, 2012; Peppé et al., 2007) såväl som till autismspecifika kommunikationsstörningar (Diehl et al., 2009; Paul, Augustyn, Klin & Volkmar, 2005). Detta kan tyda på att prosodi samvarierar med en eller flera andra förmågor hos personer med AST.

1.1 Studiens syfte

Syftet med föreliggande studie var att jämföra expressiv prosodi hos barn med AST och barn med TU. Ytterligare ett syfte var att undersöka på vilket sätt instrumentella data relaterar till hur lyssnare uppfattar prosodi, varför analyser genomfördes både instrumentellt och perceptuellt. Vidare undersöktes om prosodiska skillnader mellan grupperna är kopplade till språkliga mått.

2. BAKGRUND

2.1 Autismspektrumtillstånd


2.1.1 Kommunikation vid AST

I DSM-5 utgör kvalitativt nedsatt förmåga att interagera och kommunicera socialt ett av de två huvudsakliga symtomområdena vid AST. I de fall språkliga svårigheter förekommer kan de variera från avsaknad av tal, försenad språkutveckling, nedsatt språkförståelse till överdrivet formellt och bokstavligt språk. Även när språkliga förmågor såsom ordförråd och grammatik är typiskt utvecklade, finns svårigheter med användning av språk i social interaktion (APA, 2013). Bristande förmåga att inleda och upprätthålla samtal är vanligt förekommande, vilket försvärrar ett omfördelat kommunikativt utbyte. Svårigheter att ”läsa mellan raderna” leder ofta till att ytteransen uppfattas bokstavligt, något som kan orsaka språkliga missförstånd (Frith, 1989). Det är även vanligt att situationer tolkas på ett konkret sätt samt att personer med AST har svårigheter att förstå om ett budskap är riktat till dem (Dahlgren, 2007).


Den här typen av svårigheter med socialt samspel genomsyrar alla aspekter av språkutveckling hos barn med AST, och kan betecknas som pragmatiska svårigheter (Nettelbladt, 2013). De förekommer i någon mån hos alla personer med AST, oberoende av funktionsnivå, ålder och språklig förmåga, och har även visat sig vara bestående (Tager-Flusberg, Joseph & Folstein, 2001). Pragmatiska svårigheter kan emellertid förekomma även utanför autismområdet, och numera utgör social (pragmatic) communication disorder en fristående diagnos i diagnostmanualen DSM-5 (APA, 2013). En gemensam nämnare vid såväl AST som pragmatiska svårigheter anses vara bristande theory of mind (ToM). ToM kan beskrivas som förmågan att kunna ta andra människors perspektiv, och underlättar därför förståelsen för andra människors tankar och känslor. ToM är nödvändig för social interaktion, inte minst eftersom det möjliggör antaganden och förutsägelser om andras avsikter och beteenden (Nettelbladt, 2013; Perkins, 2007).

2.2 Prosodi

Prosodi kan beskrivas som talets rymiska, melodiska och dynamiska aspekter, och tillhör talets suprasegmentella egenskaper. Detta innebär att prosodin är en domän som spänner över
enskilda segment. Prosodisk information finns på stavelse-, ord-, fras- och ytrandenivå och behövs för att ge struktur åt det talade språket (Bruce, 2012).

Perceptuellt yttrar sig prosodiska drag i termer av tonhöjd, röststyrka, längd, intonation, betoning, pausering, tempo och rytm (McCann, Peppé, Gibbon, O’Hare & Rutherford, 2007). Dessa komponenter är relaterade till variationer i de akustiska dimensionerna duration, intensitet och grundtonsfrekvens (F0). Variationer i dessa dimensioner förmedlar tillsammans grammattisk, pragmatisk och affektiv information såväl som information om den individuella talaren. Det senare kan handla om vanemässiga sätt att variera exempelvis intonation och röststyrka samt signalering av regional härkomst genom dialektala drag. Således kan prosodi delas in i form; hur de prosodiska dimensionerna realiseras och varieras, och funktion; den kommunikativa funktion som realiseringen av prosodiska drag fyller (Peppé, 2009).

2.2.1 Svenskans prosodi

Jämfört med många andra språk är svenskans prosodiska system relativt komplext. På ordnivå finns exempelvis en betydelseskiljande funktion hos betonings- och accentkontrasten liksom skillnader i kvantitet (Nettelbladt, 2007). Ordbetoning innebär att vissa stavelser framhävs på bekostnad av andra, med hjälp av starkare intensitet, ökad längd och högre tonhöjd. Vid uttal av ett isolerat ord ligger betoningarna på en av stavelserna (lämna 'Japan och ja 'pan), men på satsnivå förlorar däremot vissa av orden sin betoning (Lindblad, 2010). Svenskans två ordaccenter realiseras med hjälp av variationer i tonhöjd (Cruttenden, 1997). Dessa brukar kallas accent I och accent II; jämför Polen respektive pälen (Bruce, 2012). Med kvantitet åsyftas skillnader i långd hos vokaler och konsonanter, där kort vokal följs av lång konsonant och lång vokal följs av kort konsonant; jämför matt respektive mat (Bruce, 2012).

Den betydelseskiljande funktion som ordbetoning, ordaccent och kvantitet har är emellertid relativt marginell i förhållande till andra kommunikativa funktioner som omfattas av prosodin (Bruce, 2012). Intonationen har en mer övergripande funktion, som sträcker sig över både fraser och ytranden (Cruttenden, 1997). Intonation kan beskrivas som variationer i tonhöjd och har sin främsta akustiska motsvarighet i variationer i grundtonsfrekvensen (F0). Med hjälp av intonationen signaleras dialekt, sinnesstämma och attityd. En talare kan även i samtal använda sig av intonation för att kommunicera att denne har talat färdigt alternativt vill behålla ordet. Intonationen behövs också för att framhäva viktig information, skilja fråga från påstående och för att strukturerar talet genom frasering (Lindblad, 2010).

Prosodisk frasering syftar på hur en talare med prosodins hjälp kan gruppera ord inom fraser och fraser inom ytranden samt signalera gränser mellan olika enheter (Bruce, 2012). Denna funktion är viktig inte minst för att underlätta förståelsen av tvetydiga fraser (lämna chocolate, cake and buns och chocolate-cake and buns; McCann & Peppé, 2003). Prosodiska fraser innehåller ofta färre antal ord än syntaktiska fraser, som återfinns i skrivna texter. Den kortare fraslängden i talat språk föreligger troligen för att skapa gynnsammare förutsättningar för talaren, genom att underlätta för andningen, men även för lyssnaren, genom att talflödet lättare bearbetas i mindre enheter (Bruce, 2012).
2.2.2 Prosodins domäner

Distinktionen mellan prosodins grammatiska, pragmatiska och affektiva domäner är vanligt förekommande i litteraturen kring AST. Ytterligare en domän är socio-indexikala aspekter, som utgörs av bland annat regional dialekt och talarens vanemässiga sätt att variera prosodiska drag (Peppé, 2009).

**Grammatisk prosodi** avser prosodiska ledtrådar som används för att förmedla grammatisk information inom en sats. Hit hör exempelvis funktioner som ordbetonning, fråge- och påståendeintonation (Paul et al., 2005) samt frasering (Peppé et al., 2007).

**Pragmatisk prosodi** syftar på att med prosodins hjälp lägga särskild tonvikt på ny, oväntad eller särskilt viktig information i ett yttrande (Paul et al., 2005). Ibland används detta för att markera semantiska kontraster; exempelvis kommer ordet blå i en fras som en blå Ferrari att betonas med extra emfás om föregående talare precis nämnit en röd Ferrari. Denna prosodiska funktion brukar betecknas som **kontrastiv betoning** (Kaland et al., 2013). Kontrastiv betoning är en prosodisk komponent som är starkt beroende av förmågan att ta lyssnarens perspektiv i beaktning, och antas därför vara associerat med ToM (Kaland et al., 2013).

**Affektiv/emotionell prosodi** syftar på hur en talare kan använda variationer i tonhöjd, intensitet och tempo för att förmedla känslor och sinnestillstånd (Le Sourn-Bissaoui, Aguert, Girard, Chevreuil & Laval, 2013). Exempelvis uttrycks generellt positiv affekt med högre och mer varierad tonhöjd, medan negativ affekt snarare kännetecknas av lägre och mindre varierad tonhöjd (McCann & Peppé, 2003).

2.3 Prosodi och AST


Flera forskare har påtalat att atypisk röst och prosodi kan påverka lyssnarintrikyet av individer med AST negativt (Peppé, 2009; Shriberg et al., 2001). Tal med överdriven intensitet och/eller förhöjt taltonläge kan exempelvis framstå som överdrivet ihärdigt eller intensivt, medan ett för långsamt tempo kan ge ett överlägset intryk. Dessa komponenter hos talet kan vara oavsiktliga, men har ändå en inverkan på hur individen uppfattas av omgivningen (Shriberg et al., 2001). Studier har därutöver visat på signifikanta samband mellan avvikande prosodi hos personer med AST och autismrelaterade störningar i social och

Detta tyder sammantaget på att lyssnarintrycket av talet hos personer med AST kan ha en väsentlig betydelse för hur individen uppfattas av omgivningen, och därmed också för förutsättningarna att interagera och kommunicera effektivt med andra. Vidare har observationer visat att prosodiska svårigheter, i de fall där de förekommer, kvarstår upp i vuxen ålder (DePape et al., 2012; Shriberg et al., 2001) och inte förbätras i samma utsträckning som andra språkliga förmågor (DeMyer et al, 1973; Simmons & Baltaxe, 1975).

Mot bakgrund av att personer med AST anses ha specifika svårigheter med affekt och pragmatik, har det föreslagits att prosodiska svårigheter hos gruppen främst omfattar de affektiva och pragmatiska domänerna medan grammatisk prosodi är relativt välutvecklad (Shriberg et al., 2001). Peppé, McCann, Gibbon, O’Hare och Rutherford (2007) fann i enlighet med den hypotesen att en grupp barn med AST i störst utsträckning avvekar från kontroller ifråga om pragmatisk och affektiv prosodi, och i minst utsträckning ifråga om grammatiskt prosodi. Bristande ToM hos personer med AST har antagits ligga till grund för svårigheterna med prosodins affektiva (Rutherford, Baron-Cohen & Wheelwright, 2002) och pragmatiska (Kaland et al., 2013) domäner.

2.3.1 Taltonläge och intonation

Instrumentella analyser av F0 har visat på signifikant högre taltonläge hos barn med AST jämfört med åldersmatchade kontroller (Sharda et al., 2010), och ytterligare några studier har noterat samma trend om än inga signifikanta gruppsskillnader påvisats (Diehl et al., 2009; Nadig & Shaw, 2012). Ett relativt stort antal studier har även funnit överdrivna variationer i F0 hos personer med AST, vilket står i kontrast till de ofta stereotypa beskrivningarna av monotont och robotliknande tal som förekommit i litteraturen. Stora F0-variationer och/eller stort F0-omfång har observerats vid analys av varierande talmaterial, såsom isolerade ord (Filipe, Frota, Castro & Vicente, 2014), imitationsuppgifter (Hubbard & Trauner, 2007), narrativer (Diehl et al., 2009), samtal (Green & Tobin, 2009; Nadig & Shaw, 2012; Sharda et al., 2010), högläsning (Green & Tobin, 2009) och standardiserade fraser (Diehl & Paul, 2013). Enligt DePape, Chen, Hall och Trainor (2012) utgör atypiska variationer i tonhöjd en primär bidragande faktor till udda prosodi hos gruppen med AST. Det förefaller dock finnas en betydande individuell variation, med både extremt smala och överdrivet vida intonationsspann rapporterade (Baltaxe, Simmons & Zee, 1984; DePape et al., 2012; Green & Tobin, 2009). Några studier har funnit en skiljelinje mellan subgrupper inom autismspektrumet, där individer med AS uppvisat överdrivna (DePape et al., 2012) eller typiska (Kaland et al., 2013) F0-variationer jämfört med deltagare med högfungерande autism (HFA), vilka snarare uppvisat små F0-variationer i enlighet med definitionen monotoni. I likhet med Kaland et al. (2013) rapporterade Green och Tobin (2009) F0-variationer i enlighet med en kontrollgrupp hos en subgrupp av deltagarna med AST.
Vidare har stort F0-omfång hos personer med AST kopplats till mindre användning av intonation i kommunikativt syfte. Det visade DePape et al. (2012) i en studie av hur tonhöjdstoppar placeras för att markera relevant information i yttranden. De individer med AS som hade stora variationer i F0 uppvisade trots detta en mer atypisk placering av tonhöjdstoppar jämfört med de individer med HFA som hade små variationer i F0, vilka presterade mer likt kontroller. Artikelförfattarna tolkade detta som att gruppen med AS använder intonation mindre kommunikativt. Atypisk placering av tonhöjdstoppar trots stora variationer i F0 har noterats även av Hubbard och Trauner (2007).

Trots en grundläggande förmåga att producera mer varierad prosodi har en repetitiv användning av ett begränsat antal F0-konturer rapporterats hos barn med AST (Green & Tobin, 2009). DePape et al. (2012) föreslår att just denna begränsade variation av prosodiska konturer kan vara bidragande till det ibland monotona intrycket.

### 2.3.2 Betoningsmönster


### 2.3.3 Frasering och yttrandelängd

Studier som undersökt perception och produktion av *prosodisk frasering*, även kallat *chunking*, har funnit att personer med AST presterar nästan i enlighet (Peppé et al., 2007) eller helt i enlighet (Diehl & Paul, 2013; Paul et al., 2005) med kontroller. Thurber och Tager-


Sammantaget indikerar dessa studier att personer med AST inte uppsvisar avvikande frasering eller fraslängd. Resultaten beträffande icke-flyt och yttrandelängd förefaller peka i olika riktning, men kan möjligen förklaras av olika val av talmaterial och/eller nivå av fungerande hos deltagarna.

2.3.4 Taltempo


2.3.5 Röstbetingade aspekter

Röstanvändning är nära förknippad med prosodi. Rösten har inte samma språkliga funktion som vissa prosodiska komponenter men kan vara viktig för att kommunicera exempelvis

Beträffande röststyrka har instrumentella analyser visat på små (Diehl & Paul, 2013) eller inga (Filipe et al., 2014) skillnader i intensitet mellan grupper med AST och kontroller, medan perceptuella bedömningar har indikerat överdriven röststyrka hos talare med AST (Shriberg et al., 2001). Grupper med AST har även rapporterats använda variationer i röststyrka för att uttrycka olika typer av affekt i mindre utsträckning än kontroller (Hubbard & Trauner, 2007).

### 2.3.6 Dialektala drag


### 2.3.7 Form och funktion

Som tidigare nämnts har bristande ToM hos personer med AST förmodats ligga till grund för de prosodiska avvikelserna, vilka därför främst har antagits omfatta prosodiska funktioner inom de pragmatiska och affektiva domänerna. Som framgått har dock ett antal instrumentella analyser funnit atypiska mönster i realiseringen av prosodiska drag även när lyssnare bedömt den prosodiska kontrasten som tillräcklig. Detta har rapporterats både för prosodins grammatiska (Filipe et al., 2014; Grossman et al., 2010) och pragmatiska (Kaland et al., 2013)
domäner. På liknande sätt har studier av affektiv prosodi visat på en adekvat förmåga att uttrycka prosodiska kontraster mellan olika affekter, men trots detta har personer med AST uppfattats som mer avvikande och onaturliga (Grossman et al., 2013; Paul et al., 2005). I studien av Grossman et al. (2013) rapporterades att deltagarna med AST något ensidigt tycktes använda sig av kontraster i tempo vid produktion av de olika affekterna, vilket artikelförfattarna beskriver som att en del av den ”affektiva kvaliteten” i rösten saknades.


2.3.8 Sammanfattande kommentarer

Litteraturen kring prosodi och autism är bitvis motsägelsefull och svårtolkad, och det är oklart huruvida detta bottnar i metodologiska problem eller heterogenitet bland personer med AST (McCann & Peppé, 2003). Betydande skillnader mellan studier avseende metod och genomförande gör det svårt att jämföra olika resultat sinsemellan. Exempelvis varierar val av analysmetod och talmaterial kraftigt, liksom kognitiv och språklig nivå hos deltagarna med AST. Samtidigt finns indikationer på att de ibland motstridiga resultaten till viss del kan förklaras av den heterogena symtombild som kännetecknar autismområdet. Peppé et al. (2007) fann exempelvis att den prosodiska förmågan varierade kraftigt inom gruppen med AST, men att alla avvek ifråga om minst en prosodisk aspekt. Dessutom har studier av impressiv prosodi visat att förmågan att tolka affektiv prosodi varierar också inom individer beroende på affektarnas intensitet (Grossman & Tager-Flusberg, 2012) och i vilken kontext de presenteras (Le Sourn-Bissaoui et al., 2013). Utöver de samband som rapporterats med autismspeciﬁka kommunikationsstörningar (Diehl et al., 2009; Paul et al., 2005) har prosodi visat sig vara kopplat till språklig förmåga, framförallt språkförståelse (McCann et al., 2007), såväl som till passivt ordförstå (DePape et al., 2012). Detta kan tyda på att prosodi samvarierar med kommunikativ och/eller språklig förmåga hos personer med AST.
2.4 Frågeställningar

Mot bakgrund av genomgången litteratur formulerades följande frågor för föreliggande studie:

- Finns skillnader i expressiv prosodi mellan barn med AST och barn med TU som kan fångas upp med en instrumentell analys?
- Finns en perceptuell motsvarighet till de instrumentella fynden och/eller föreligger gruppskillnader som endast kan fångas upp med en perceptuell analys?
- Påverkas lyssnares prosodiska helhetsintryck av instrumentella och/eller perceptuella gruppskillnader?
- Är prosodiska komponenter där gruppen med AST avviker från gruppen med TU relaterade till språklig förmåga?

3. METOD

3.1 Deltagare


Deltagarna med AST rekryterades från barn- och ungdomspsykiatrisk klinik (BUP) med hjälp av Maria Råstam, professor i barn- och ungdomspsykiatri i Lund, som också är forskningshuvudman för studien. Deltagarna med TU rekryterades via skola, och för att undvika eventuell integritetskränkning rekryterades grupperna från olika skolor.
3.2 Material

De prosodiska analyser som genomfördes i föreliggande studie grundar sig på talmaterial i form av narrativer, som spelades in i samband med de tidigare uppsatsarbeten (Chouaiby et al., 2013; Rindhagen & Rudling, 2013). Barnen fick med stöd av bildmaterialet ERRNI (Expression, Reception and Recall of Narrative instrument; Bishop, 2004) generera en berättelse. Ljudupptagningen gjordes med en traditionell filmkamera och utan adekvat utrustning att optimera talupptagningen. Inspelningarna av deltagarna med TU genomfördes i barnens skolmiljö och inspelningarna av deltagarna med AST på BUP, med undantag av en deltagare som spelades in på Avdelningen för logopedi, foniatri och audiologi. Narrativerna var i genomsnitt 139 sekunder långa, och längden skilde sig inte signifikant åt mellan grupperna ($p > .05$).

Mot bakgrund av studiens frågeställningar och tidigare forskning inom området utarbetade uppsatsförfattarna ett skattningsprotokoll som omfattar ett antal prosodisk- och röstbetingade aspekter (se bilaga 1). Protokollet användes vid de perceptuella bedömningarna.

I samband med de två tidigare uppsatsarbeten genomgick de aktuella deltagarna en omfattande testning av språkliga förmågor. I föreliggande studie undersöktes i vilken grad prosodiska gruppstillader varierade med språklig förmåga. För ändamålet användes resultaten av följande testmaterial:

- TROG-2 (Test for Reception Of Grammar; Bishop, 2003), vilket gav ett mått på grammatisk språkförståelse.
- NAP (Narrative Assessment Profile; Bliss, McCabe & Miranda, 1998), vilket gav ett mått på berättarförmåga.

3.3 Databearbetning

Inför de instrumentella analyserna bearbetades inspelningarna genom att extrahera ljudet ur videofilmerna med hjälp av programmet VLC media player. Därefter klipptes de för studien irrelevanta sektionerna bort och det aktuella analysmaterialet konverterades från stereo- till monoformat i programmet Audacity. Alla yttranden transkriberades ortografiskt och visuellt stöd av programmet Wavesurfer och med visuellt stöd av programmets inbyggda spektogram markerades start- och sluttid för yttrandena. Således definierades i studien en akustisk paus som tystnad som både registrerades auditivt och visuellt av uppsatsförfattarna, och ett yttrande syftar därmed på talmängd (antal ord) mellan akustiska pauser. Genom att yttrandena skiltes ut från sådant som inte skulle beaktas i den instrumentella analysen, såsom tystnad, störningsljud eller tal från annan talare, möjliggjordes automatisk extraktion av deltagarnas yttranden samt beräkningar av antal ord och talhastighet.

Inför de perceptuella bedömningarna extraherades ur narrativerna en minut långa ljudsekvenser, vilka utgjorde bedömningsunderlaget. Då en kortare sekvens bedömdes ge en representativ bild av deltagarnas narrativa produktion togs detta beslut för att reducera tidsåtgången vid lyssnarbedömningarna. De extraherade sekvenserna sträckte sig 30 sekunder före och 30 sekunder efter respektive inspelnings mittpunkt, med vissa justeringar för att inte behöva klippa mitt i ett yttrande.
3.4 Kalibrering av skattningsprotokoll

Inledningsvis genomfördes en utvärdering av skattningsprotokollet med hjälp av två expertlyssnare, båda legitimerade logopeder och verksamma inom röstområdet. Ingen av dessa deltog sedan i huvudstudien. Expertlyssnarna fick med hjälp av protokollet skatta tre barn och bidra med sina synpunkter, vilket resulterade i förtydligande av vissa begrepp samt viss omarbetning av svarsalternativen. Inför detta tillfälle hade brus filtrerats bort ur ljudfilerna för att undersöka om detta underlättade de perceptuella bedömningarna. Vid en jämförelse mellan brusreducerade och icke-brusreducerade filer ansåg dock lyssnargruppen att brusreduceringen försvårade bedömning av vissa parametrar, varför beslut om att använda icke-brusreducerade filer togs. Den från början tänkta konsensusbedömningen med tre bedömningstillfällen per barn omformades i samråd med expertlyssnarna, vilka ansåg att likvärdig reliabilitet bör kunna uppnås vid ett bedömningstillfälle per barn (dock med möjlighet till flera uppspelningar).

3.5 Genomförande

De perceptuella analyserna genomfördes med hjälp av tre expertlyssnare, alla legitimerade logopeder och verksamma inom röstområdet. Valet av lyssnargrupp baserades på logopedernas erfarenheter av liknande bedömningsprotokoll och av att analysera röster perceptuellt. Lyssnarna informerades skriftligt och mundligt om studiens syfte, men gavs inte kännedom om deltagarnas grupptillhörighet. Analysen hade formen av en konsensusbedömning, där en minutsspelning av respektive barn först skattades individuellt av samtliga lyssnare, följd av diskussion och slutligen en gemensam och definitiv bedömning. Vid uppspelningen varierades ordningen på sekvenserna utifrån faktorerna inspelningskvalitet, barnens kön och grupptillhörighet. De instrumentella analyserna genomfördes i programmet YIN (De Cheveigné & Kawahara, 2002) med hjälp av Sofia Strömbergsson, legitimerad logoped och doktorand vid KTH, Avdelningen för Tal, Musik och Hörsel. Grundtonsfrekvens (F0) analyserades automatiskt på allt tonande material i deltagarnas yttranden, vilket gav ett akustiskt mått på taltonläge. För varje yttrande beräknades F0-variation i form av standardavvikelsen från grundtonsfrekvensen, angivet i halvtonssteg.

I tabell 1 redovisas samtliga instrumentellt och perceptuellt analyserade komponenter. Som framgår har en del instrumentella komponenter en direkt motsvarighet i den perceptuella analysen, medan vissa prosodiska mönster endast analyserades perceptuellt respektive instrumentellt. Detta beror på att samtliga prosodiska komponenter inte kunde fängas upp med båda metoderna.

<table>
<thead>
<tr>
<th>Instrumentella parametrar</th>
<th>Perceptuella parametrar</th>
</tr>
</thead>
<tbody>
<tr>
<td>Genomsnittlig F0</td>
<td>Taltonläge</td>
</tr>
<tr>
<td>F0-variation</td>
<td>Intonation</td>
</tr>
<tr>
<td>Taltempo (stavelser/sek)</td>
<td>Taltempo</td>
</tr>
<tr>
<td>Yttrandelängd (antal ord/yttrande)</td>
<td>Yttrandelängd</td>
</tr>
<tr>
<td>Pausfrekvens</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Intensitet</td>
</tr>
<tr>
<td></td>
<td>Röstklang</td>
</tr>
<tr>
<td></td>
<td>Nasalitet</td>
</tr>
<tr>
<td></td>
<td>Betoningsmönsrten</td>
</tr>
<tr>
<td></td>
<td>Icke-flyt</td>
</tr>
<tr>
<td></td>
<td>Dialektala drag</td>
</tr>
<tr>
<td></td>
<td>Prosodiskt helhetsintryck</td>
</tr>
</tbody>
</table>

3.6 Forskningsetiska överväganden

Föreliggande studie har genomgått granskning av den etiska kommittén på Avdelningen för logopedi, foniatri och audiologi vid Lunds universitet. De två tidigare studierna, i vilka de aktuella deltagarna ingick, granskades och godkändes av regionala etikprövningsnämnden (EPN). Deltagarna och deras vårdnadshavare informerades då skriftligt om forskningens syfte: att jämföra språkanvändning hos barn med AST med den hos barn med typisk utveckling. Eftersom prosodi är nära kopplat till språkanvändning kunde materialet användas inom ramen för det tidigare forskningsetiska beslutet. Således behövdes enligt forskningshuvudmannen ingen kompletterande ansökan göras till regionala EPN.


3.7 Statistisk bearbetning

De statistiska beräkningarna genomfördes i IBM SPSS Statistics (version 22). Då data inte var normalfördelad användes icke-parametrisk statistik för samtliga statistiska analyser. Signifikansnivån sattes till p ≤ .05. För de statistiska jämförelserna mellan grupperna användes Mann Whitney U-test. För beräkning effektstorlek används Cohen’s $d$, där ett värde på $\pm 0.2$ räknas som en liten, $\pm 0.5$ en medelstor och $\pm 0.8$ en stor effektstorlek. För korrelationsberäkningar användes Spearmans rho ($r$), där 1 respektive -1 motsvarar en perfekt korrelation.
3.7.1 Modifiering av data inför statistisk analys

Bedömningarna av parametern *intensitet* i skattningsprotokollet visade sig vara problematiska för lyssnarna till följd av inspelningskvaliteten. *Taltonläge* upplevdes även svårbedömt då uppsatsförfattarna vid tillfället för bedömningarna inte hade kännedom om barnens åldrar mer än på gruppnivå. Mot bakgrund av detta kunde resultaten för intensitet och taltonläge inte anses tillförlitliga, varför beslut togs att utesluta dessa vid den statistiska bearbetningen.

De parametrar i protokollet skattades utifrån en sjugradig skala där höga respektive låga värden motsvarar större avvikelse (se bilaga), kodades om inför vissa av de statistiska beräkningarna. Det resulterade i en stigande skala där ett högre värde konsekvent motsvarar en större avvikelse. Detta gjordes för att skattnings i olika riktningar inte skulle ”ta ut” varandra vid beräkningar av medelvärde samt för att underlätta beräkningen av totalt antal avvikande skattnings. Även tillvalsparametrarna kodades så att högre värden motsvarade större avvikelse.

4. RESULTAT

Inledningsvis presenteras resultaten från den instrumentella analysen, följt av resultaten från den perceptuella. Därefter redovisas korrelationsberäkningar med språkliga mått.

Total längd för de kompletta narrativerna skilde sig inte signifikant åt mellan grupperna (*p* > .05), varför beslut togs om att använda hela ljudsekvenserna vid de instrumentella jämförelserna mellan grupperna. Detta gjordes i syfte att stärka reliabiliteten i beräkningarna. Samtliga data som användes vid den perceptuella analysen bygger däremot på de extraherade 1 minutssekvenserna. Dessutom genomfördes de instrumentella analyserna på yttrande-nivå vilket gav upphov till ett värde per yttrande och deltagare, jämfört med den perceptuella analysen där varje deltagare erhöll ett värde per bedömd parameter.

4.1 Instrumentell analys

4.1.1 Jämförelser mellan grupper

I tabell 2 presenteras resultaten från den instrumentella analysen. *Taltempo* redovisas som antal stavelser/sek och *pausfrekvens* som genomsnittligt antal pauser per minut. *Yttrandelängd* avser antal ord per yttrande. Samtliga komponenter utom pausfrekvens är beräknade på yttrandelängd, vilket framgår av antalet värden (*N*) per analyserad komponent.

Av tabellen kan utläsas att gruppen med AST uppvisade signifikant längre yttrandet, färre pauser samt högre och mer varierat taltonläge jämfört med kontrollgruppen. Effektstorlekarna var mestadels små, något som kan förklaras av att endast en andel av deltagarna med AST avvek från gruppen med TU. Gällande pausfrekvens var skillnaderna mellan grupperna däremot mer distinkta. För att undersöka om skillnaderna i genomsnittlig F0 kunde tillskrivas åldersfaktorer genomfördes korrelationsberäkningar mellan ålder och F0, men inget samband kunde påvisas (*p* > .05).

Vidare framgår av tabell 2 att standardavvikelserna inom gruppen med AST var större för ett antal parametrar. En jämförelse av spridningarna inom grupperna visade att signifikanta
gruppskillnader ($p<.05$) förelåg beträffande parameter 2 och 3, vilket innebär att gruppen med AST uppförde en signifikant större spridning för yttrandeläng och pausfrekvens.

Tabell 2. Medelvärde, medianvärde och standardavvikelse för respektive analyserad komponent och grupp. För gruppskillnader anges u-värde, z-värde, signifikansnivå (p) och effektstorlek (d).

<table>
<thead>
<tr>
<th>Komponent</th>
<th>TU medelvärde (median) [standardavvikelse]</th>
<th>AST medelvärde (median) [standardavvikelse]</th>
<th>U-värde (Z-värde) [P-värde]</th>
<th>Cohen’s d</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Taltempo</td>
<td>N = 414 3.78 (3.80) [1.40]</td>
<td>N = 484 3.91 (3.87) [1.47]</td>
<td>97734 (-.63) [.53]</td>
<td>.09</td>
</tr>
<tr>
<td>2. Yttrandelängd</td>
<td>N = 414 5.61 (5.0) [3.98]</td>
<td>N = 484 4.75 (4.0) [3.65]</td>
<td>87289 (-3.35) [.001]</td>
<td>.23</td>
</tr>
<tr>
<td>4. Genomsnittlig F0 (Hz)</td>
<td>N = 413 203.36 (201.76) [28.60]</td>
<td>N = 484 192.76 (191.10) [28.55]</td>
<td>79187 (5.42) [&lt;.001]</td>
<td>.37</td>
</tr>
<tr>
<td>5. F0-variation (halvtoner)</td>
<td>N = 413 24.0 (24.04) [2.42]</td>
<td>N = 484 23.06 (23.09) [2.54]</td>
<td>79260 (-5.40) [&lt;.001]</td>
<td>.38</td>
</tr>
</tbody>
</table>

De instrumentella komponenter där gruppskillnader förelåg presenteras i figur 1-4 fördelningarna baserat på medelvärde per deltagare. Den skuggade boxen representerar 50% av deltagarna inom varje grupp och den markerade horisontella linjen representerar medianvärdet. Avståndet mellan boxen och respektive ändmarkering motsvarar 25% av deltagarna.
4.2 Perceptuell analys

4.2.1 Jämförelser mellan grupper

Tabell 3 visar deskriptiva data för lyssnarbedömningarna. Samtliga bedömda parametrar som är inkluderade redovisas med medelvärde, medianvärde och standardavvikelse. Samtliga deltagare kunde inte bedömas på alla parametrar, varför antal (N) anges för varje parameter. Parameter 1-3 bygger på den omarbetade skalan (se under rubrik 3.7.1) och anger således endast grad av avvikelse i någon riktning. För parameter 4-9 i tabellen (motsvarande 6-11 i protokollet) har respektive möjlig skattning i protokollet getts ett numeriskt värde där 1 representerar ingen avvikelse och högre värden konsekvent motsvarar större avvikelser. I figur 5-7 illustreras skattningarnas spridningar för parameter 1-3 enligt den ursprungliga sjugradiga skalan.


<table>
<thead>
<tr>
<th>Komponent</th>
<th>AST medelvärde (median) [standardavvikelse]</th>
<th>TU medelvärde (median) [standardavvikelse]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intonation (1-4)</td>
<td>N=9  2 (2) [1]</td>
<td>N=10  1.50 (2) [.63]</td>
</tr>
<tr>
<td>Taltempo (1-4)</td>
<td>N=10  2.10 (2) [1]</td>
<td>N=10  1.60 (1.50) [.70]</td>
</tr>
<tr>
<td>Yttredelängd (1-4)</td>
<td>N=10  1.9 (2) [1]</td>
<td>N=10  1.80 (2) [.63]</td>
</tr>
<tr>
<td>Röstklang (1-4)</td>
<td>N=10  1.70 (1.50) [.82]</td>
<td>N=10  1.70 (2) [.68]</td>
</tr>
<tr>
<td>Nasalitet (1-2)</td>
<td>N=10  1.1 (1) [.32]</td>
<td>N=10  1.40 (1) [.52]</td>
</tr>
<tr>
<td>Betoningsmönster (1-4)</td>
<td>N=9   1 (1) [0]</td>
<td>N=10  1 (1) [0]</td>
</tr>
<tr>
<td>Icke-flyt (1-4)</td>
<td>N=10  2.20 (2) [1.14]</td>
<td>N=10  2 (2) [.94]</td>
</tr>
<tr>
<td>Dialektala drag (1-2)</td>
<td>N=8   1.13 (1) [.35]</td>
<td>N=10  1.20 (1) [.42]</td>
</tr>
<tr>
<td>Prosodisk helhetsintryck (1-3)</td>
<td>N=9  1.33 (1) [.5]</td>
<td>N=10  1.10 (1) [.32]</td>
</tr>
<tr>
<td>Totalt antal avvikande skattningar</td>
<td>N=9  14.22 (14) [3.15]</td>
<td>N=10  13.50 (15) [2.95]</td>
</tr>
</tbody>
</table>

Figur 5.

Figur 6.
Enligt beräkningar med Mann Whitney U-test förelåg inga signifikanta skillnader \((p > .05)\) mellan grupperna med avseende på någon bedömd parameter, inte heller för det totala antalet avvikande skattningar. Ett mönster kunde dock urskiljas där gruppen med AST generellt uppvisade en något större spridning (standardavvikelse) på flertalet bedömda parametrar. Vid en granskning av skattningarna noterades även att gruppen med AST erhöll fler "extrema" skattningar, här definierat som mer än ett steg från det värde som i protokollet definierades som \textit{inget avvikelse}. Det totala antalet extrema skattningar var dubbelt så många i gruppen med AST (14 st) som i gruppen med TU (7 st). Detta berodde inte främst på att fler antal deltagare erhöll någon extrem skatning inom gruppen med AST, utan på att fler deltagare erhöll ett större antal extrema skattningar. Den enskilda parameter som resulterade i flest extrema skattningar var intonation, där samtliga fyra deltagare som erhöll värden 2 (och således skattades som monotona) ingick i gruppen med AST.

Vidare förelåg ett signifikant samband \((r(20)= .63, \ p=.003)\) mellan antalet extrema skattningar hos deltagarna och skattningar av grupptillhörighet. Detta tyder på att de tre individer med AST som skattades troliga att ingå i AST-gruppen också skattades som de mest prosodiskt avvikande. Detta verifieras även av en granskning av det totala antalet avvikande skattningar per deltagare.

\subsection*{4.2.2 Prosodiska skillnaders påverkan på lyssnarintryck}

För att undersöka om lyssnarintrycket påverkades av de prosodiska komponenter där gruppen med AST avvek från gruppen med TU, genomfördes korrelationsberäkningar mellan lyssnarnas skattningar av prosodiskt helhetsintryck och respektive prosodisk komponent där gruppskillnader förelåg. Resultatet visade att helhetsintrycket inte uppvisade ett signifikant samband med genomsnittlig \(F_0\), \(F_0\)-variation, pausfrekvens eller yttrandelängd. Då inga perceptuella gruppskillnader förelåg genomfördes inga korrelationer med perceptuella mått.

\subsection*{4.3 Korrelationer med språkliga mått}

Då studien även syftar till att undersöka om prosodiska gruppskillnader var relaterade till språklig förmåga hos deltagarna, genomfördes över alla individer korrelationsberäkningar
mellan respektive prosodisk komponent där gruppskillnader förelåg, och mätten på språkförståelse (TROG-2), berättarförmåga (NAP) och ordförråd (WISC-4). För TROG-2 användes åldersekvivalenser och för NAP och WISC-4 råpoäng.


Tabell 4. Korrelationer mellan prosodiska komponenter och språkliga mått angivna i Spearmans rho (r).

<table>
<thead>
<tr>
<th>Yttrandelängd</th>
<th>Pausfrekvens</th>
<th>Genomsnittlig F0</th>
<th>F0-variation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Berättarförmåga (NAP)</td>
<td>N = 20</td>
<td>N = 20</td>
<td>N = 20</td>
</tr>
<tr>
<td>-.25</td>
<td>.43</td>
<td>.32</td>
<td>.32</td>
</tr>
<tr>
<td>Ordförråd (WISC-4)</td>
<td>N = 19</td>
<td>N = 19</td>
<td>N = 19</td>
</tr>
<tr>
<td>-.49*</td>
<td>.47*</td>
<td>-.24</td>
<td>-.31</td>
</tr>
<tr>
<td>Grammatisk språkförståelse (TROG-2)</td>
<td>N = 19</td>
<td>N = 19</td>
<td>N = 19</td>
</tr>
<tr>
<td>-.27</td>
<td>.34</td>
<td>-.08</td>
<td>-.13</td>
</tr>
</tbody>
</table>

* Signifikant korrelation (p<.05)

5. DISKUSSION

5.1 Resultatdiskussion

5.1.1 Instrumentell analys


Gruppen med AST upprivade alltså lägre frekvens av akustiska (tysta) pauser och fler ord per yttrande jämfört med kontrollgruppen. Då pausmönster och yttrandelängd undersökt i mycket liten utsträckning hos personer med AST är föreliggande resultat svåra att jämföra med tidigare studier. Resultaten gällande pausfrekvens kontrasterar exempelvis med en perceptuell studie som rapporterade kortare yttranden hos barn med AST jämfört med kontroller, trots liknande talmaterial (Thurber & Tager-Flusberg, 1993). I den studien förefaller dock begreppet yttrande betraktas som en mening oberoende av akustiska pauser, vilket skiljer sig från hur ett yttrande definierats i föreliggande studie. Det är därför tänkbart att ett yttrande enligt vår definition innehåller flera meningar, något som möjligens skulle kunna förklara de motstridiga resultaten. Ett annat intressant fynd i sammanhanget, som däremot tycks gå i linje med föreliggande studies resultat, är de längre yttranden som noterats hos deltagare med AS jämfört med andra individer med AST (Hubbard & Trauner, 2007). Trots att yttrande inte definieras närmare eller analyseras systematiskt i den nämnda studien, utgör det en intressant observation mot bakgrund av att majoriteten (70%) av deltagarna i
föreliggande studie var diagnosstiserade med AS. Det är tänkbart att högfungerande individer med välutvecklad språklig förmåga är mer medvetna om sina kommunikativa svårigheter och därför kompenserar genom att bygga ut och tydliggöra sina yttranden, i enlighet med artikelöf att fattarnas resonemang.

Om fynden gällande yttrandelängd och pausfrekvens är relatater till atypisk fraser eller inte, är svårt att avgöra då denna prosodiska funktion inte studerades specifikt. Eftersom intrrycket av "pauser" i samband med prosodisk fraser sällan utgörs av verkliga akustiska pauser (Bruce, 2012), kan man dock tänka sig att den lägre pausfrekvensen troligtvis speglade andra sorters pauser. Enligt Bruce är akustiska pauser i talet främst kopplade till andning (något som i hög grad synkroniseras med frasgränser) eller tankeverksamhet i form av sökande efter ord och formuleringar. Således är det tänkbart att gruppen med AST uppsvisade lägre frekvens av aningspauser och/eller tankepauser. Det är visserligen även troligt att icke-flyt i form av omtagningar och upprendningar skulle fångas upp som akustiska pauser vid en instrumentell analys, varför det inte kan uteslutas att skillnaderna beror på en högre mängd icke-flyt inom gruppen med TU. Om så vore fallet vore det att andra sidan troligt att detta fångats upp i lyssnarskattningarna av icke-flyt, inte minst eftersom två av expertlyssnarna arbetar eller har arbetat kliniskt med stamningspatienter.

En hypotes som möjligtvis skulle kunna appliceras på resultatet beträffande pausfrekvens har framförts av Thurber och Tager-Flusberg (1993). Enligt artikelöfattarna kan just tysta pauser (i studien refererat till som "icke-grammatiska pauser") spegla sökande efter ord och formuleringar, och således betraktas som ett mått på kognitiv belastning. En lägre frekvens icke-grammatiska pauser vid produktion av bildelicerade narrativer tolkades därför i deras studie i termen av mindre engagemang i berättande hos barn med AST. Vidare korrelerade även låg frekvens av dessa pauser med kort längd på hela narrativet, vilket av artikelöfattarna tolkades som ytterligare ett tecken på lägre grad av engagemang bland deltagarna med AST. Vi fann visserligen inget samband (p > .05) mellan pausfrekvens och narrativernas längd inom gruppen med AST, men däremot mellan narrativernas längd och yttrandelängd (r(10) = -.76, p = .01), något som eventuellt skulle kunna beträttas utifrån resonemanget om engagemang i berättande. Det skulle i så fall kunna innebära att individer som använder frekventa tankepauser producerar längre och därmed kanske mer utförliga och utmålande narrativer. Det tendens till samband som förelåg mellan pausfrekvens och berättarförmåga (NAP), kan möjligen också tolkas i ljuset av detta då det tyder på att få pauser är relatater till sämre berättarförmåga.

Hypotesen att långa yttrandet är relatater till lägre grad av engagemang och sämre berättarförmåga (Thurber & Tager-Flusberg, 1993) förefaller dock svår att förena med resonemanget om långa yttrandet som en kompensationsstrategi (Hubbard & Trauner, 2007). Det senare borde snarare hänga ihop med ett ökat engagemang och en strävan att berätta detaljerat och ingående, vilket i så fall borde resultera i ett bättre resultat på NAP. Vidare bör det påpekas att grad av engagemang i berättande inte kan bedömas enbart utifrån akustiska pauser. För det första kan fler former av icke-flyt och tvekfenomen i talet spegla kognitiv belastning, såsom en ökad förekomst också av fyllda pauser, meningsfragment, omtagningar och reparationer (Eklund, 2004). Eftersom dessa talbeteenden inte analyserades i föreliggande studie kan inga slutsatser dras rörande kognitiva processer hos deltagarna. För det andra kan de aktuella pauserna vara kopplade till fler faktorer än kognitiv belastning, inte minst till genren genom att deltagarna betraktar bilderna eller byter sida. Vi vet således inte med säkerhet vad som sker i pauserna. För en utförligare kartläggning av pausmönstren hade fler paustyper behövts analyseras och jämföras eftersom olika slags pauser kan återspeglas olika

Vidare var det genomsnittliga taltonläget högre inom gruppen med AST jämfört med inom gruppen med TU, trots den stora majoriteten pojkar inom gruppen med AST. Då pojkars F0 kan förväntas vara lägre än flickors redan i den ålder som barnen befinner sig i (Baken & Orlikoff, 2000) är resultaten särskilt intressanta eftersom detta snarare borde ha resulterat i ett lägre taltonläge inom gruppen med AST. Resultatet är dock samstämmigt med några studier som också noterat högre taltonläge hos personer med AST (Diehl et al., 2009; Nadig & Shaw, 2012; Sharda et al., 2010).


Stora F0-variationer har i litteraturen bland annat kopplats till nivå av språkligt fungerande inom autismspektrumet, då personer med AS uppsatt stora variationer jämfört med personer med HFA (DePape et al., 2012). Till följd av det låga deltagarantalet kunde dock inga jämföranden göras mellan subgrupper inom autismspektrumet. Om stor F0-variation är kopplad till hög språklig nivå inom autismspektrumet, skulle visserligen det faktum att majoriteten av de aktuella deltagarna var diagnosierade med AS kunna utgöra en bidragande förklaring till de större F0-variationerna hos gruppen. Det är tänkbart att högfungeraande individer med välutvecklat språk, i likhet med resonemang kring yttrandelängd, är mer medvetna om svårigheter med att uttrycka prosodi på ett naturligt sätt, och därför använder stora variationer i intonation som en kompensationsstrategi. Möjligen är intonation ett relativt uppennbart sätt att kommunicera engagemang och inlevelse, varför det skulle vara förhållandevis lätt att medvetet modifiera i det syftet. Hubbard och Trauner (2007) har dock belyst att fler akustiska aspekter än F0 påverkar intrycket av intonation, såsom intensitet, duration och placeringen av tonhöjdstoppar i relation till innehållsrika aspekter i yttrandena. Det är tänkbart att dessa komponenter inte i samma utsträckning går att modifiera ”mekaniskt”, vilket skulle kunna förklara varför individer med AST trots stora variationer i F0 visat sig uppfattas som mer onaturliga och udda jämfört med kontroller (Grossman et al., 2013).

5.1.2 Relation till språkliga mått

Om stora F0-variationer föreligger i högre grad hos individer med AS jämfört med andra individer med AST, skulle man följa att kunna förvänta sig ett samband mellan F0-variation och språkliga mått, något som dock inte var fallet. Sambanden som yttrandelängd och pausfrekvensens uppsatt med ordförrådsdeklen ur WISC-4 tyder däremot på någon form av relation mellan språklig förmåga och prosodiska mått. En möjlig förklaring till varför ett gott ordförråd tycks hänga ihop med korta yttranden och många pauser, skulle kunna vara att barn...
med välutvecklat ordförråd funderar mer över formuleringar och ordval. Det skulle i så fall kunna ge upphov till en större mängd tysta pauser i deras narrativer. Vidare ligger det nära till hands att tänka sig att erfarenhet av böcker bidrar positivt till ordförrådsutvecklingen, och att barn med gott ordförråd som dessutom har en medvetenhet kring narrativ struktur i högre grad skulle vara motiverade att elaborera med språket i syfte att göra berättelsen mer intressant. Ett sådant resonemang skulle även kunna kopplas till tidigare diskussion kring engagemang i berättande, och stärks av det nästintill signifikanta sambandet mellan ett stort antal pauser och god berättarförmåga.

Inget av de använda testmaterialen kan dock anses utgöra ett mått på generell språklig förmåga, varför inga långtgående slutsatser kan dras av dessa resultat. Vidare är det möjligt att den prosodiska variationen inom gruppen med AST snarare är kopplad till autismspecifika kommunikationsstörningar, något som rapporterats av vissa forskare (Diehl et al., 2009; Paul et al., 2005). Då ingen tillgång fanns till det diagnostiska utredningsmaterialet kunde detta dock inte undersökas.

5.1.3 Perceptuell analys

Till skillnad från många studier på området som använt perceptuella skattningar i syfte att bedöma prosodiska funktioner som korrekt eller felaktiga, fick lyssnargruppen istället skatta grad av avvikelse för enskilda röst- och prosodibetingade drag. Det innebär att både den perceptuella och den instrumentella analysen i första hand rör prosodins formmässiga aspekter, varför inga slutsatser kan dras om den funktionella användningen av prosodi.

Trots att den instrumentella analysen visade på signifikanta gruppsskillnader fanns ingen motsvarighet till dessa i den perceptuella bedömningen, där gruppen med AST alltså inte skattades mer avvikande varken när det gäller intonation eller yttrandelängd. Ingen av de övriga skatade komponenterna gav heller upphov till statistiska gruppsskillnader. En bakomliggande faktor kan vara att de längre sekvenserna användes vid den instrumentella analysen samt att denna genomfördes på yttrependivå. Detta innebar ett stort antal värden per deltagare till skillnad från den perceptuella bedömningen, där varje deltagare endast erhöll ett värde per bedömd parameter. Det är även tänkbart att de korta sekvenserna som användes vid den perceptuella analysen inte var representativa för talarens hela narrativa sekvens. Vidare kan yttrandelängd enligt definitionen talmängd mellan akustiska pauser vara svårbedömt i löpande tal. Det är möjligt att lyssnare ibland tolkar frasgränser som akustiska pauser, vilket kan resultera i skatningar av kortare yttranden, eller missar akustiska pauser som placeras mitt i frasen, vilket tvärtom kan resultera i skatningar av längre yttranden.

Perceptuella skattningsar av variationer i intonation är ovanliga i litteraturen på området, men det är ändå värt inträffande att avsaknaden av perceptuella gruppsskillnader står i motsättning till de impressionistiska beskrivningarna av intonation hos personer med AST som ofta förekommit i litteraturen. Termer som robotliknande, monoton eller enformig antyder att intonation utgör ett framträdande område av atypisk prosodi bland individer med AST. Intonationsparametern var förvisso den komponent där gruppen med AST erhöll flest extrema skattningsar. Det är dock anmärkningsvärt att de extrema skattningsarna pekade i motsatt riktning jämfört med den instrumentella analysen. Flera deltagare med AST uppfattades alltså ha extremt små variationer i intonation, trots att ingen deltagare i gruppen med AST akustiskt sett hade mindre F0-variation än någon i gruppen med TU. Vid en närmare granskning av datamaterialet framgick vidare att de deltagare med AST som erhöll skattningsar av monoton
intonation istället utgjordes av den ”subgrupp” som akustiskt sett uppvisade F0-variationer i likhet med kontrollgruppen. Omvänt motsvarades inte de stora F0-variationerna inom gruppen med AST av perceptuella skattningar av varierad intonation, utan samtliga dessa deltagare erhöll genomsnittliga skattningar.

Likartade resultat har tidigare rapporterats av Nadig och Shaw (2012), och man kan möjlichen spekulera i artikelförfattarnas hypotes att en stor tonhöjdsvariation som placeras godtyckligt i frasen inte uppfattas som ett tecken på ökat engagemang eller uttrycksfullhet. Det skulle i så fall kunna förklara varför den större F0-variationen inte tolkades i termer av varierad intonation i den perceptuella bedömningen. Godtycklig användning av intonation kan i sin tur tyda på att stora F0-variationer inte nödvändigtvis används på ett kommunikativt meningsfullt sätt hos personer med AST. Samband som rapporterats mellan stora F0-variationer och atypisk placering av tonhöjdsstoppar hos personer med AS (DePape et al., 2012) förefaller ge stöd åt detta.


Samtidigt är det möjligt att de varierade beskrivningarna speglar en verklig heterogenitet hos personer med AST, något som den generellt sett större spridningen i skattningarna av gruppen med AST antyder. En stor variation i prosodin bland personer med AST har stöd i både perceptuella (Peppé et al., 2007; Shriberg et al., 2001) och instrumentella (Green & Tobin, 2009; Paul et al., 2008) studier. Dessutom ger den heterogenitet som finns vid själva symtombildens vid AST, beträffande såväl språklig som kognitiv nivå, skäl att tro att även den prosodiska förmågan varierar. Den signifikant större spridningen inom gruppen med AST för de instrumentella måtten på yttrandelängd och pausfrekvens tycks också peka i denna riktning.

5.1.4 Subtila variationers påverkan på lyssnarintrycket

Avsaknad av gruppskillnader även ifråga om prosodiskt helhetsintryck tyder på att lyssnarintrycket inte påverkades av de (instrumentella) prosodiska skillnaderna mellan grupperna. Också avsaknaden av statistiska samband mellan skattningar av helhetsintryck och de prosodiska komponenter där gruppskillnader föreligger pekar i denna riktning. Mot bakgrund av studier som belyst den påverkan på lyssnarintrycket som till och med subtila prosodiska avvikelse kan ha (Diehl & Paul, 2013; Filipe et al., 2014; Grossman et al., 2010; Paul et al., 2005), framstår resultatet som något oväsent. I studien av Nadig och Shaw (2012) skattades exempelvis gruppen med AST signifikant mer avvikande gällande det prosodiska helhetsintrycket, trots att inga gruppskillnader kunde påvisas avseende någon enskild prosodisk komponent.
Det är samtidigt tänkbart att skattningar av prosodiskt helhetsintryck i föreliggande studie till viss del baserades på andra aspekter än prosodiska. Denna hypotes grundas i att inget samband ($p > .05$) föreligger mellan helhetsintrycket och det totala antalet avvikande skattningar i protokollet. Vidare upplevdes exempelvis endast ett av de fyra barn som erhöll skattningar av monoton intonation ge ett avgivande helhetsintryck, trots att intonation anses vara en tämligen central del i prosodibegreppet. Kanske skulle resultatet sett annorlunda ut om lyssnarna fått bedöma kortare ljudsekvenser i likhet med studien av Nadig och Shaw (2012), där endast 2-3 sammanhängande yttranden analyserades. Det är tänkbart att detta skulle eliminera distraherande faktorer.

5.1.5 Problematisering av normalitetsbegreppet


5.2 Metoddiskussion

Föreliggande studie har en rad begränsningar. Även om många studier på området i likhet med vår bestått av små deltagargrupper, utgör detta en uppenbar svårighet att generalisera resultatet till att gälla andra högfungerande barn med AST. Därutöver var grupperna endast åldersmatchade, vilket gjorde att grupperna inte var jämförbara språkligt eller könsmässigt.

5.2.1 Perceptuell analys

Gällande lyssnarbedömningarna kan konsensusformen problematiseras. Då konsensusbedömningar innebär en kompromiss för varje bedömd komponent är det rimligt att detta ger färre avvikande skattningar än vad som kanske skulle ha varit fallet vid individuella bedömningar. Inte minst vid skalar med ett normalvärde i mitten (parameter 1-5, se bilaga 1) är det troligt att ett genomsnittsvärde väljs om det finns uppfattningar om avvikelse i båda riktningar. I de fall där man inte kan enas gäller dessutom majoritetsprincipen, vilket alltså innebär att endast majoritetsuppfattningen speglas i de fallen.

Under inspelningarna av narrativerna förekom ibland prompting från testledarens sida, något som kan tänkas ha påverkat lyssnarnas bedömningar. Detta kan ha gett en fingervisning om
vilka barn som behövde hjälp med uppgifterna - något som kan ha styrt skattningar av inte minst grupptillhörighet.

Det bör även poängteras att lyssnarna främst hade erfarenhet av att på liknande sätt skatta röstproblem och icke-flyt, aspekter som i högre utsträckning än prosodi är föremål för kartläggning och intervention i kliniken. Det är därför möjligt att dessa komponenter lätta kunde skattas reliabelt och samstämmigt. Vidare kan 1-7-skalan med extrema värden i båda riktningar ha skapat en osäkerhet hos bedömarna eftersom man traditionellt skattar röststörningar på en stigande skala från 0-10.

5.2.2 Talmaterial

Talmaterialet i form av narrativer valdes i syfte att använda en längre sekvens som kunde ge en helhetsbild av barnets expressiva prosodi. Ett stort talmaterial ger en representativ bild inte minst av en talares intonation, yttrandelong, pausfrekvens, icke-flyt och taltempo. Emellertid kan narrativer genererade med bildstöd påverka vissa prosodiska drag, exempelvis i form av en "uppräknande" intonation eller genom att yttrandelong styrs av bildernas innehåll. Pausmönster kan också antas påverkas av den här typen av talmaterial eftersom en talare rimligtvis producerar en något längre paus inför varje ny bild. Det är därför möjligt att narrativer med fria teman skulle ha resulterat i en mer naturlig prosodi. Standardiserade fraser hade samtidigt varit önskvärt för att lättare kunna jämföra resultaten sinsemellan, eftersom graden av engagemang och/eller emotionellt innehåll kan variera beroende på berättelsernas karaktär och därmed påverka såväl temporalas aspekter som intonation. Inom ramen för de tidigare studierna forskningsetiska beslut fanns dock inga möjligheter att samla in nytt talmaterial av de aktuella deltagarna, varför detta inte var genomförbart.


Med tanke på att kännedom om hemort såväl som föräldrarnas härkomst saknades var även dialektala drag svårbedömt. Resultatet hade troligtvis blivit mer tillförlitligt om deltagarna kunnat rekryteras på basis av likartad geografisk och social bakgrund. Dessutom hade det kanske underlättat att inkludera fler skattningsalternativ än ja och nej, i enlighet med det ursprungliga protokollet som efter utvärderingen reviderades. Ett nej förutsätter en avsaknad av dialektala drag, vilket troligen är mycket sällsynt.
5.2.3 Instrumentell analys

Vid inspelnningarna av narrativen var filmkameran placerad långt ifrån barnet och vissa ord blev därför ohörbara när barnen mumlade eller störningsljud förekom. Till följd av detta kan ord ha missats i samband med transkriptionerna. Dessutom genomfördes markeringar av yttrandens början och slut manuellt av uppsatsförfattarna utifrån var tystnad registrerades visuellt och auditivt. Det innebär att ingen hänsyn togs till pausernas längd. Avsaknaden av ett nedre gränsvärde för hur korta tystnader som skulle registreras kan ha försvårat gränssnäglingen kring vad som var att betrakta som akustisk paus eller inte. Vidare kan det faktum att endast tydiga pauser analyserades i studien utgöra en begränsning, eftersom pausfrekvens därmed inte speglar deltagarnas kompletta användning av pauser. En analys och jämförelse av fler paustyper, som exempelvis fyllda pauser, hade behövts för att bättre kartlägga deltagarnas pausbeteende.

När det gäller den instrumentella analysen måste resultatet betraktas i ljuset av att inspelningskvaliteten inte var optimal för ändamålet och innehöll varierade grad av brus som på vissa håll kan ha stört registreringen av F0. I två av sektionerna förekom dessutom ett klockljud relativt nära mikrofonen. Vid ett antal stickprov av bruset i de olika filerna visade sig dock endast klockljuden, och alltså inte övrigt brus, vid ett fåtal tillfällen registreras som en ton inom programmens förinställda gränsvärden för tal (100-500 hz). Det kan inte uteslutas att övrigt brus inte registrerades och registrerades som en ton någonstans i talmaterialet. Men stickproven kunde alltså inte påvisa att dessa låg inom gränsvärdena för tal. Mot bakgrund av detta beslöts att inkludera samtliga F0-data, och för att säkerställa att ljudsekvenserna med klockljuden inte hade en för stor inverkan på resultatet genomfördes samtliga statistiska beräkningar ytterligare en gång utan dessa, vilket inte förändrade resultaten.

5.3 Framtida forskning


Det är vidare tänkbart att logopeder med klinisk erfarenhet av AST i högre utsträckning kan känna igen och upptäcka prosodiska avvikelser i talet, varför en framtida studie skulle kunna använda en lyssnargrupp bestående av exempelvis logopeder i autismteam. Mot bakgrund av att flera studier har lyft fram det udda intryck som atypisk prosodi kan medföra i mötet med andra människor finns det även anledning att låta naiva lyssnare utgöra ett komplement till expertlyssnare, eftersom naiva lyssnare kan anses representera människor som personer med AST möter i sin vardag. Istället för att lyssna efter prosodiska delkomponenter kunde dessa
bedöma exempelvis hur naturlig en talare låter, något som sedan skulle kunna relateras till instrumentella analyser för att få en uppfattning om vilken inverkan prosodiska avvikelser kan tänkas ha på social interaktion och kommunikation.

5.4 Sammanfattande slutsatser och kliniska implikationer


Eftersom avvikande prosodi hos personer med AST i många fall utgör ett påtagligt socialt hinder, och dessutom visat sig vara bestående även när andra språkliga förmågor förbättrats, är det också av stor vikt att detta uppmärksammars mer i interventionen än vad som är fallet idag. Bristen på normer för typisk och avvikande prosodi gör emellertid bedömning och behandling av prosodiska svårigheter till en stor utmaning för logopeder. Vidare är prosodi svårt att styra viljemässigt varför en medveten kontroll av prosodi riskerar resultera i ett mekaniskt och onaturligt tal (Peppé, 2009). Med tanke på att samband rapporterats mellan expressiv och impressiv prosodi hos barn med AST (Peppé et al., 2007), är det dock tänkbart att intervention med fokus på att identifiera och tolka andras sätt att variera prosodiska drag skulle kunna bidra till en ökad medvetenhet om det egna sättet att använda prosodi.

Då generaliseringseffekten visat sig vara störst vid intervention som fokuserar på redan befintliga förmågor hos barn med AST (Lord, 2000), är det tänkbart att prosodisk intervention framförallt skulle vara effektivt för barn som uppvisar en grundläggande medvetenhet om prosodins roll och betydelse. Prosodisk intervention skulle i de fallen kunna utgöra en naturlig del i språkliga och pragmatiska interventionsprogram för barn med AST.
6. TACK

Ett stort tack till följande personer som bidragit med hjälp och stöd i uppsatsarbetet:

- **Ulrika Nettelbladt** för all optimism, handledning och stöd under skrivprocessen.
- **Annika Dahlgren Sandberg** för värdefulla kunskaper om AST samt all hjälp med metod och statistik.
- **Sofia Strömbergsson** för uppmuntran, handledning samt ovärderlig hjälp (och stort tålamod) med de instrumentella analyserna.
- **Rebecca Rindhagen** för ytterligare utredningar av barn och svar på våra frågor kring materialet och deltagarna.
- **Lotta Browall, Christina Askman och Susanna Whitling** för eminent utförda perceptuella bedömningar.
- **Viveka Lyberg Åhlander och Cecilia Lundström** för bra synpunkter och hjälp med utformningen av skattningsprotokollet.
- **Maria Råstam** för initiativ till projektet.
7. REFERENSER


