



SAHLGRENSKA AKADEMIN

Institutionen för neurovetenskap och fysiologi  
Sektionen för hälsa och rehabilitering  
Enheten för logopedi

**330**

## **Lyssnarförståelighet avseende tal från barn med avvikande tal – effekter av dialekt och uppspelningsmedium**

Emil Brynte

Examensarbete i logopedi  
30 högskolepoäng  
Vårterminen 2019

Handledare  
Tove Lagerberg  
Sofia Strömbergsson

## **Lyssnarförståelighet avseende tal från barn med avvikande tal – effekter av dialekt och uppspelningsmedium**

Emil Brynte

*Sammanfattning.* Denna studies syfte var att undersöka hur två olika faktorer vid bedömning av förståelighet kan påverka resultatet. En frågeställning var att undersöka om dialekt hos talare och lyssnare hade betydelse för lyssnarförståelighet. En ytterligare frågeställning undersökte om skillnad i lyssnarförståelighet fanns mellan uppspelningsmedium audio och video. Resultat från en lyssnargrupp Göteborgsbor (n=20) jämfördes med resultat från en lyssnargrupp Stockholmsbor (n=20). Stimuli omfattade inspelningar från barn med speech sound disorder (SSD) (n=14) från Göteborg respektive Stockholm. En lyssnargrupp (n=4) transkriberade baserat på videomedium och resultatet jämfördes med resultat från audiomedium-transkription. Resultatet visade signifikanta skillnader där Stockholmslyssnare hade en högre lyssnarförståelighet för samtliga stimuli. Skillnaden mellan audio och video var inte signifikant. Sammanfattningsvis visar föreliggande studie att dialekt kan spela roll vid bedömning av förståelighet men att överföringsmedium har en begränsad betydelse.

Nyckelord: Förståelighet, barn, tal- och språkstörning, dialekt, uppspelningsmedium.

## **Intelligibility assessment of kids with deviant speech – effects of dialect and presentation mode**

*Abstract.* The purpose of this study was to investigate how two different factors can affect the result on assessment of intelligibility. One research question examined whether dialect of speakers and listeners affected intelligibility. Another research question examined whether differences in listener intelligibility existed for audio or video medium. Results from a group (n=20) residing in Gothenburg were compared with results from a group (n=20) residing in Stockholm. Stimuli included recordings from children with speech sound disorder (SSD) (n=14) from Gothenburg and Stockholm. A group (n=4) transcribed from video-medium and results were compared with results from audio-medium transcription. The result showed that Stockholm listeners had a significantly higher intelligibility score for all stimuli. The difference between audio and video was not significant. To conclude the present study indicates that dialect can affect the result on assessment of intelligibility but that the presentation mode has a limited impact on the result.

Key words: Intelligibility, children, speech sound disorder (SSD), dialect, presentation mode.

Ett centralt ändamål för all mänsklig talad kommunikation är att förmedla ett förståeligt budskap till en annan individ (Flipsen, 2006). För att kommunikation ska vara framgångsrik och effektiv behöver lyssnaren kunna uppfatta och urskilja innebörden eller orden som talaren ämnar förmedla (Lagerberg, 2013). Begreppet förståelighet definierar Yorkston, Strand och Kennedy (1996) som till vilken grad lyssnaren förstår vad talaren har sagt. En behjälplig terminologi föreslås av Miller (2013) då han pratar om förståelighet som signalberoende, det vill säga baserad endast på information från audiosignalen, eller signaloberoende, baserad på information från samtliga intryck som syntax, semantik och ansiktsuttryck. För att lyssnaren ska kunna konvertera signalen från talaren till en tanke och kunna dra adekvata slutsatser krävs både signalberoende och signaloberoende faktorer (Mattys, Davis, Bradlow & Scott, 2012). Flera aspekter påverkar hur väl en talare blir förstådd, till exempel precision av talproduktion, talhastighet, intensitet av talsignalen samt yttrandens längd och komplexitet (Flipsen, 2006). Vidare menar Flipsen (2006) att följande faktorer kan ha betydelse för förståeligheten vid konversation: yttrandets lingvistiska innehåll; icke-verbala signaler (gester, hållning, ansiktsuttryck); kommunikativ miljö/innehåll (ämne som diskuteras, tidigare uttalanden).

En typiskt utvecklad fyraåring förväntas uppnå nära 100% förståelighet om hen talar till okända lyssnare (Flipsen, 2006). Talavvikelser hos barn kan dock i högre eller mindre grad påverka hur väl barnet blir förstått (Lagerberg, 2013). Olika typer av talavvikelser hos talaren kan resultera i varierande grad av nedsatt förståelighet (Lagerberg, 2013). SSD är ett övergripande begrepp för olika tal- och språkstörningar som kan påverka förståeligheten på grund av försenad och avvikande tal- och språkutveckling (Smith, Pennington, Boada & Shriberg, 2005), lägre grad av förståelighet är ofta kopplat till en högre grad av avvikelse (Gordon-Brannan, 1994). Kent et al. (1994) har dock visat att även en mildare grad av talavvikelse kan komma att leda till en lägre grad av förståelighet.

Eftersom framgångsrik kommunikation bygger både på talarens förmåga att uttrycka sig och på lyssnarens förmåga att förstå, kan logopediska insatser för barn med SSD ha olika fokus; å ena sidan att öka barnets uttrycksförmåga, och å andra sidan att rikta insatser till personer i barnets omgivning (Miller, 2013). Gällande talaren kan detta innebära intervention med målet att hjälpa hen uppnå eller bibehålla förståeligt tal och gällande lyssnaren kan det vara insatser med målet att kunna uppfatta talsignalen. För att mäta huruvida behandling varit framgångsrik eller ej kan man genomföra bedömningar av talarens förståelighet.

En faktor som vid bedömning av förståelighet antas påverka resultatet är vilken typ av uppgift talaren har. Vanligt förekommande uppgifter innefattar repetition, läsning, bedömning av ord, meningar, nonsensmeningar eller spontantal (Kent et al., 1994). En väl strukturerad metod, som ord- och meningslistor, innehåller tydliga målord vilket gör analysen enklare och öppnar upp för möjligheten att enkelt förklara till vilken grad stimuli är förståeligt. Vid bedömning av spontantal kompliceras en sådan bedömning av att facit som anger korrekt produktion saknas. Jämfört med högläsning av ord- och meningslistor har spontantalsbedömningar dock en högre ekologisk validitet, vilket innebär att det är en metod att föredra vid bedömning av förståelighet (Lagerberg, Åsberg, Hartelius & Persson, 2014). Två ytterligare faktorer som påverkar resultat vid bedömningen av förståelighet är *överföringsmedium* och *lyssnaregenskaper*. Dessa faktorer har

gemensamt att båda har visat ge en viss inverkan på bedömning av förståelighet (Miller, 2013; Hustad & Cahill, 2003; Klintö & Lohmander, 2017).

Betydelsen av olika *lyssnaregenskaper* och hur dessa påverkar bedömningen av förståeligheten har undersökts i flera studier. Egenskaper hos lyssnaren som kan påverka bedömningen inbegriper förstaspråk, hörselnedsättning, vana vid talaren och vetskap om samtalsämnet eller stimuli (Kent et al., 1994). Flera studier har till exempel visat att föräldrar skattar sina egna barn som mer förståeliga än okända barn (Coplan & Gleason, 1988; Lagerberg, Hellström, Lundberg & Hartelius; Kwiatkowski & Shriberg, 1992). Lyssnaregenskaper som kön, ålder, utbildningsnivå och vana vid talarens dialekt har granskats i flera studier utan att påvisa några signifikanta korrelationer till förståelighet (McHenry, 2011; Pennington & Miller, 2007). När graden av förståelighet ska bedömas påverkas detta både av vem som gör bedömningen och hur mycket kontextuell information som finns tillgänglig (Miller, 2013). Individens som gör en bedömning har givetvis varierande vana av den dialekt som talaren har men hur detta kan påverka en bedömningssituation förefaller inte ha studerats i Sverige. Det förefaller därav finnas ett behov av att undersöka även hur dialekt påverkar en bedömning av förståelighet på barn med tal- och språkstörning.

*Överföringsmedium* har studerats av bland annat Hustad och Cahill (2003) som undersökte skillnad i förståelighet mellan video- och audiomedium, avseende förståelighet hos personer med mild och grav dysartri i anknytning till cerebral pares. Effekten av uppspelningsmedium avseende andra talare med andra talavvikelser har dock vad vi vet ej undersökts tidigare avseende förståelighet. Det finns dock en svensk studie av Klintö och Lohmander (2017) som undersöker effekt av överföringsmedium vid perceptuell bedömning. Klintö och Lohmander (2017) jämförde fonetiska transkriptioner baserade på audiomedium jämfört med videomedium avseende läpp-käk-gomspalt-tal och fann att fonetisk transkription till en viss grad påverkas av "visual cues" men att uppspelningsmedium inte verkar påverka resultatet. Denna studie påvisar likt Hustad och Cahill (2003) att uppspelningsmedium förefaller ha en viss inverkan som inte kunnat visa signifikanta resultat överlag utan endast inom vissa specifika fall. Det förefaller följaktligen inte finnas tidigare studier som har undersökt uppspelningsmediums betydelse eller betydelse av dialekt vid bedömning av förståelighet på barn med tal- och språkstörning. Eftersom detta är av betydelse vid exempelvis utvärdering av intervention vid reducerad förståelighet är det av stor betydelse att detta undersöks.

### *Syfte*

Syftet med föreliggande studie är att undersöka hur faktorerna dialekt respektive uppspelningsmedium av stimuli påverkar lyssnarförståelighet.

### *Frågeställningar*

1. Finns det en skillnad avseende lyssnarförståelighet beroende på om lyssnarna transkriberar stimuli innehållande tal från barn som kommer från samma område som lyssnaren själv eller inte?

2. Skiljer sig lyssnarförståelighet mellan transkriptionsbedömning baserat på olika överföringsmedium, audio- eller videostimuli?

Lyssnarförståeligheten förväntas vara högre för lyssnare som bor i samma stad som de barn som finns på inspelningarna. Det förväntade resultatet är vidare att videomedium är lättare att förstå och därmed att dessa transkriptioner kommer resultera i högre lyssnarförståelighet.

## Metod

Denna studie är en del av projektet *Funktionella konsekvenser av avvikelser i barns sammanhängande tal*.

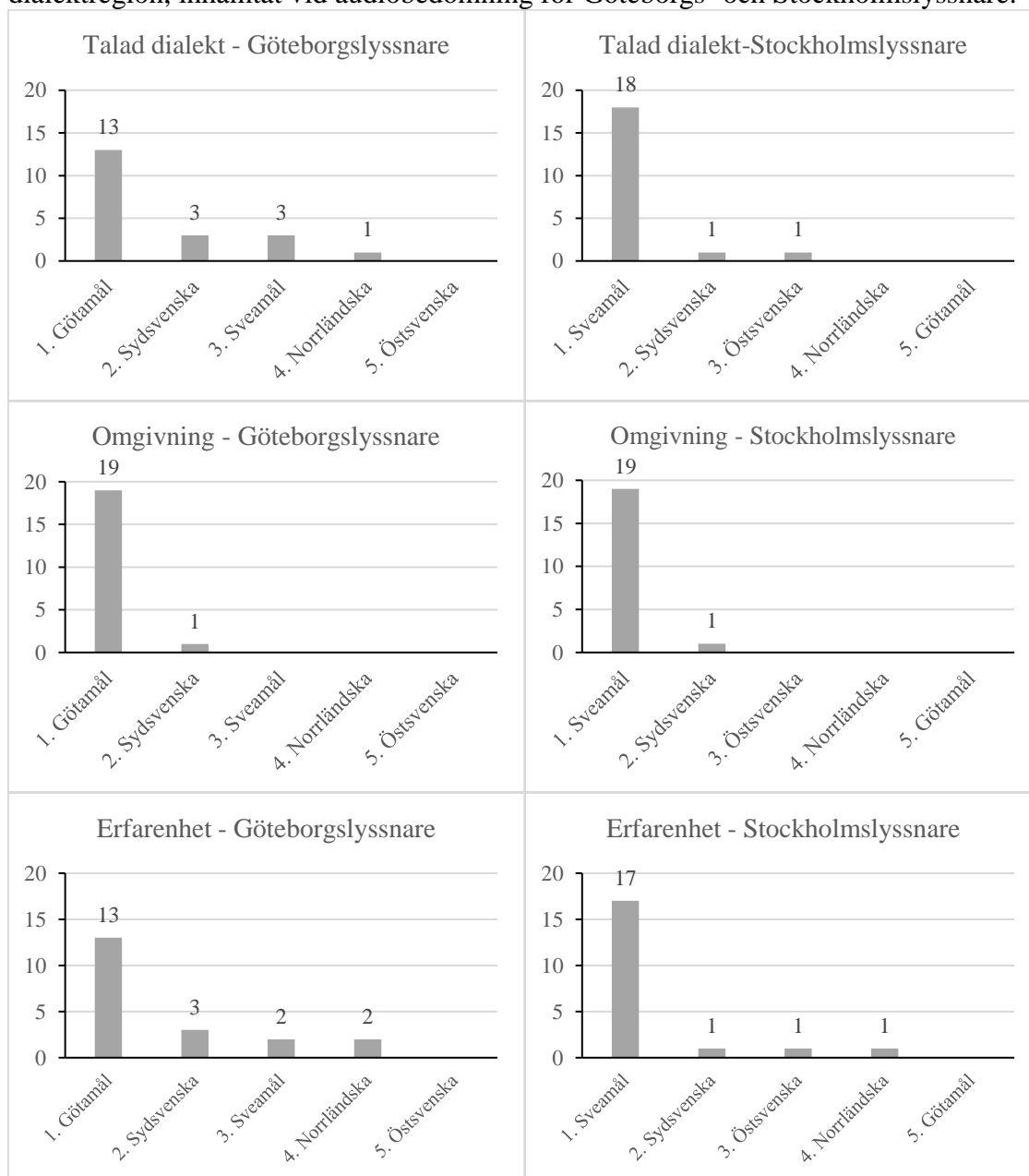
### Lyssnare

I föreliggande studie kommer alla deltagare benämnas som lyssnare, i enlighet med Hustad och Cahill (2003), trots att vissa deltagare har tillgång till videomaterial. I studien användes två olika lyssnargrupper som rekryterades från Göteborg, en för transkription från audioinspelning ( $n=20$ ) och den andra för transkription från videoinspelning ( $n=4$ ). Inklusionskriterier för lyssnare i både audio- och videogruppen var att de skulle ha studerat minst sex terminer på logopedprogrammet vid Göteborgs universitet och bodde i närområdet kring Göteborg. Inget särskilt urval genomfördes utan de som hade möjlighet att delta och uppfyllde kriterierna fick vara med i studien. Befintliga lyssnarbedömningar från audioinspelning ( $n=20$ ) från en lyssnargrupp från Stockholm användes. Dessa var genomförda av Emander och Kjerrman (2018) och inklusionskriterier var att de skulle ha studerat minst sex terminer på logopedprogrammet vid Karolinska institutet och de bodde alla i Stockholmsområdet. De lyssnare som rekryterades från Göteborg benämns hädanefter Göteborgslyssnare och lyssnare rekryterade från Stockholm benämns Stockholmslyssnare.

För transkriptionsbedömningen av audioinspelningar rekryterades nio studenter från termin sex och tio från termin åtta på logopedprogrammet vid Göteborgs universitet. Lyssnarna rekryterades både i samband med föreläsningar samt via elektronisk kommunikation. En nyexaminerad logoped med mindre än ett års yrkeserfarenhet rekryterades via sociala kontaktnät. Lyssnarna bestod av sjutton kvinnor och tre män och deras ålder var vid datum för bedömning 22–45 år ( $M = 27,9$  år,  $SD = 6,4$  år). Till transkriptionsbedömningen av videoinspelningar rekryterades fyra studenter från termin 8 på logopedprogrammet vid Göteborgs universitet. Dessa lyssnare rekryterades via sociala kontaktnät, alla var kvinnor, 23, 24, 27 och 30 år gamla. Lyssnargruppen från Stockholm bestod av 18 logopedstudenter från Karolinska Institutet och två nyexaminerade logopeder med mindre än ett års yrkeserfarenhet som rekryterades via sociala kontaktnät. Deras ålder var vid datum för bedömning 22–34 år ( $M = 26,7$  år,  $SD = 3,3$  år) och gruppen bestod av nitton kvinnor och en man. Samtliga lyssnare hade svenska som sitt starkaste språk och självrapporterad normal hörsel. Lyssnarna som genomförde audiobedömningen fick i samband med denna fylla i uppgifter om information relaterat till dialekt med hjälp av ett frågeformulär, vilket beskrivs i detalj

under rubriken ”frågeformulär”. Figur 1 visar lyssnares egenskattade svar angående dialektinformation för respektive lyssnargrupp.

*Figur 1.* Lyssnarnas svar från frågeformulär på följande frågor: självskattad talad dialekt, omgivande dialekt, erfarenhet av dialekt. Sorterat på antal svar inom respektive dialektregion, inhämtat vid audiobedömning för Göteborgs- och Stockholmslyssnare.

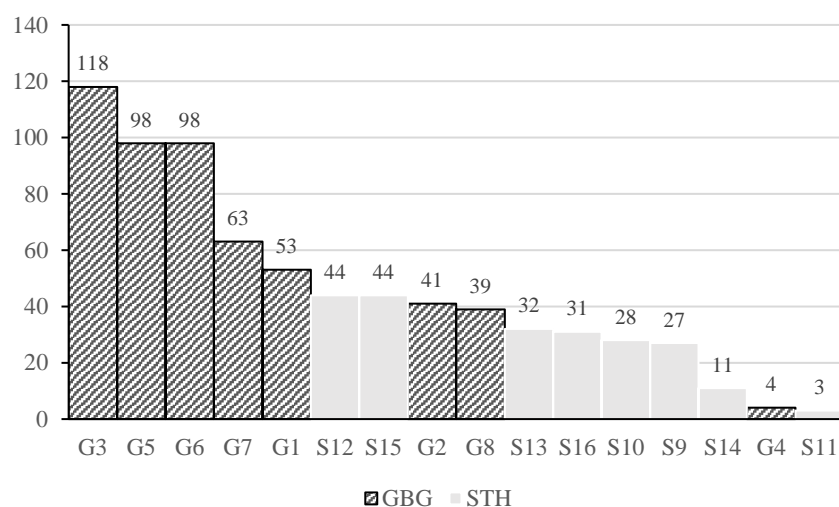


### Stimuli

Stimuli bestod av sexton cirka en minut långa videoinspelningar av åtta barn från Göteborg respektive Stockholm (se tabell 1) som samlats in under 2017 och använts inom projektet *Funktionella konsekvenser av avvikelser i barns sammanhängande tal* (Emander & Kjerрман, 2018). Stimuli som beskrivs som ”video” innehåller både video-

och audiosignal men när stimuli beskrivs som ”audio” består detta endast av audiosignal. Studiens författare genomförde en granskning av dialekt för stimuli från Göteborg (götamål) och Stockholm (sveamål) som båda bedömdes kunna representera respektive dialektregion. Materialet innehöll talarnas svar på frågan ”Vad skulle du göra om du fick en miljon?”. Ursprungligen fanns trettio inspelningar inom projektet, av dessa valde projektmedarbetare ut fjorton barn med tal- och/eller språkstörning (SSD) och två barn med typisk tal- och språkutveckling. Ytterligare en inspelning användes som övningsexempel vid transkriptionerna och inkluderades inte i studien för analys. Stimuli valdes utifrån kriteriet att talets avvikelsegrad skulle variera över hela urvalet. För varje stimulus fanns information tillgänglig angående antal talavvikelser. Denna information baserades på tidigare genomförd manuell uppmärkning av avvikelser från förväntat uttal samt av oförståeliga sekvenser, så kallad annotering. Exempel på sådana annoteringar kunde vara ”dentalisering”, ”förenkling av konsonantförbindelse” eller ”oförståeligt”. I föreliggande studie redovisas antal avvikelsetyper per stimulus, som en beskrivning av talavvikelsegrad; för mer detaljerad information om annoteringen hänvisas till Emander och Kjerrman (2018). En jämförelse mellan stimuli från Göteborg och Stockholm för antal annoterade talavvikelser kan ses i figur 2. Stimuli skiljer sig åt signifikant avseende talavvikelsegrad (antal annoterade avvikelser) vilket analyserades med t-test. Resultatet visade att Göteborgsstimuli ( $M = 64,3$ ,  $s = 38,0$ ) har en högre grad av talavvikelser än Stockholmsstimuli ( $M = 27,5$ ,  $s = 14,4$ ),  $p < ,05$ .

Figur 2. Antal annoterade talavvikelser för stimuli från Göteborg (G) och Stockholm (S).



### Utrustning

Vid både audio- och videobedömningar användes VLC media player för att spela upp stimuli som fanns i en spellista. Lyssnarna transkriberade med stiftpenna på olinjerat papper och hade tillgång till sudd vid behov. För lyssnarna som bedömde audio användes hörlurarna Logitech USB Headset H390 (on-ear) och de som bedömde videoklipp använde de mer bekväma hörlurarna Sony MDR-XD150 (over-ear). Video-gruppen använde mer bekväma hörlurarna eftersom deras bedömning innehöll cirka fem timmars lyssnande, mot audio-gruppens cirka en timme.

## *Frågeformulär*

För att samla in lyssnares dialektinformation användes samma frågeformulär som använts i Stockholm. Formuläret innehöll tre frågor som berörde dialekt. Lyssnarna fick svara på vilken dialekt de talar, vilken de främst omges av idag samt vilken de hade mest erfarenhet av att höra. Sveriges dialekter delades då in i sex huvudområden i enlighet med Kierkegaard, 2017, och lyssnarna fick välja den som stämde bäst in på respektive fråga. Dessa huvudområden var sydsvenska, göta-, svea-, gotländska, norrländska och östsvenska mål. Föreliggande studie ämnade jämföra dialekt för lyssnare från Stockholm och Göteborg. Enligt Kierkegaard (2017) faller regionen Stockholm under sveamål och Göteborg under götamål.

## *Tillvägagångssätt*

*Audiobedömning.* Då studien ämnade jämföra resultatet från Göteborgslyssnare med sedan tidigare insamlade data genomförd av Emander och Kjerrman (2018) för Stockholmslyssnare, replikerades deras tillvägagångssätt i möjligaste mån. Alla lyssnare genomförde transkriberingen på institutionen för logopedi på Göteborgs universitet, förutom en som skedde på Angereds närsjukhus. Endast testledaren och lyssnarna fanns med i rummet och samtliga lokaler var avskilda, tysta rum för att optimera bedömningsförhållanden. Av praktiska skäl genomförde sex av lyssnarna transkriptionerna med två personer i rummet, medan resten av lyssnarna var ensamma i rummet. Även de lyssnare som var två i rummet under bedömningen genomförde transkriptionen individuellt. Lyssnarna delades in i fem grupper (A-E) om fyra personer i varje. Grupperna slumpades inte utan de första fyra personerna blev grupp A och de nästkommande fyra blev grupp B och så vidare. Med tanke på rimlig kognitiv belastning och tidsåtgång för lyssnarna fick respektive grupp lyssna på fyra ljudfiler, varav minst en var ett barn med typisk språkutveckling. Varje stimulus transkriberades därmed av fyra lyssnare. Det tog cirka en timme per lyssnare. Samtliga lyssnare fick identiska muntliga och skriftliga instruktioner som gick i linje med fastställda anvisningar i Lagerberg et al. (2014). Lyssnarna instruerades att transkribera alla ord och delar av ord de uppfattade ortografiskt och markera stavelser de inte förstod med siffran noll ("0"). Transkriptionerna skulle även inkludera onomatopoetiska ord (ex. "pang"), omtagningar av hela ord eller meningar. Lyssnarna instruerades att inte gissa, i enlighet med Weiss (1982). Utfyllnadsord som "eh" eller omtagningar av stavelser skulle inte transkriberas. Samtliga lyssnare fick lyssna på ett övningsexempel som transkriberades individuellt, följt av diskussionsmöjligheter med testledaren. Lyssnarna fick lyssna i sitt eget tempo, valde själva vilken ljudfil de ville börja med och när de ville gå vidare till nästa audiofil samt de kunde även justera ljudvolymen fritt. Lyssnarna fick lyssna på varje audiofil maximalt två gånger, de fick inte spola tillbaka men de kunde pausa audio-filen utan begränsningar. Under hela transkriptionen fanns testledaren till hands för att reda ut frågetecken och vid behov förtydliga instruktioner. En lyssnare missuppfattade instruktionerna och skrev två separata transkriptioner per audiofil och för rättning användes den andra transkriptionen utifrån att lyssnaren då hade lyssnat två gånger.

För rättning av samtliga transkriptioner användes ett befintligt facit ("master transcript") som hade skapats i enlighet med Kwiatkowski och Shriberg (1992) inom ramen för det



större projektet. Den logoped som genomförde inspelningarna hade direkt efter varje inspelning gjort en ortografisk transkription med hjälp av videospelningen. Därefter hade en genomgång utförts samt en eventuell rättning av transkriptionerna där logopeden och barnens vårdnadshavare hade tittat på videospelningen för att komma så nära som möjligt det barnets avsåg att säga.

Författaren jämförde varje transkription med facit och antal korrekt uppfattade stavelser beräknades, vilket gav en procentuell lyssnarförståelighet för var och en av transkriptionerna. För att göra det möjligt att undersöka interbedömarreliabilitet gjordes 25% av transkriptionerna om av en logoped aktiv inom projektet. En median beräknades utifrån procentvärden för de fyra lyssnarbedömningarna som fanns för respektive audio-fil, vilket gav en lyssnarförståelighetspoäng för varje audio-fil.

*Videobedömning.* Videobedömningen genomfördes på institutionen för logopedi på Göteborgs universitet under samma förhållanden som för audiobedömningen. Gruppen bestod av fyra lyssnare (dvs samma antal lyssnare per stimulus som för audiobedömningen) och varje lyssnare transkriberade samtliga videoklipp. Endast testledaren och lyssnarna fanns med i rummet och samtliga lokaler var avskilda, tysta rum för att optimera bedömningsförhållanden. Utanför rummet var det en relativt trafikerad väg samt en helikopterplatta som användes under sittningen. Det sistnämnda utgjorde ett större störningsmoment på grund av dess höga ljudnivå och alla lyssnare tog en gemensam paus tills ljudet minskade. Varje lyssnare satt vid en avskild plats i rummet för att minimera störningsmoment från andra lyssnare. Bedömningen genomfördes under en eftermiddag uppdelad på fem pass med två korta (10min) och två långa (20min) pauser. Lyssnarna fick identiska instruktioner som för audio-bedömningen och beräkning av korrekt antal stavelser samt påföljande förståelighetspoäng beräknades på samma sätt.

### *Statistisk analys*

Förståelighetsbedömningarna resulterade i sexton medianvärden från audiogruppen och sexton från videogruppen (8 Göteborgsstimuli och 8 Stockholmsstimuli) som sedan analyserades. IBM SPSS Statistics och Microsoft Excel användes för den statistiska analysen. Eftersom studiens data inte var normalfördelad användes det icke-parametriska testet Wilcoxon Signed Ranks för jämförelsen mellan lyssnargrupper beroende på uppspelningsmedium och lyssnaregenskaper. För statistisk analys användes en signifikansnivå på  $p < ,05$ .

Interbedömarreliabilitet undersöktes med ICC (Intra Class Correlation) för samtliga stimuli för lyssnarna som gjort transkription från audio och videobedömningen samt för 25% av stimuli avseende korrekt antal stavelser jämfört med facit. Enligt Fleiss (1986) innebär  $ICC > 0,75$  hög reliabilitet,  $ICC = 0,4-0,75$  medel-låg reliabilitet och  $ICC < 0,4$  låg reliabilitet. Interbedömarreliabilitet beräknades med hjälp av ICC (two-way mixed, average measure). För audiogruppen kombinerades flera lyssnares bedömningar för att få ett material att beräkna interbedömarreliabilitet och således kunde fyra bedömningar jämföras i enlighet med Lagerberg (2014). Bedömningsmetoderna visade sig ha mycket god reliabilitet  $ICC = 0,98$  för audiogruppen och  $0,91$  för videogruppen samt  $0,99$  för beräkning av antal korrekta stavelser.

## Etiska aspekter

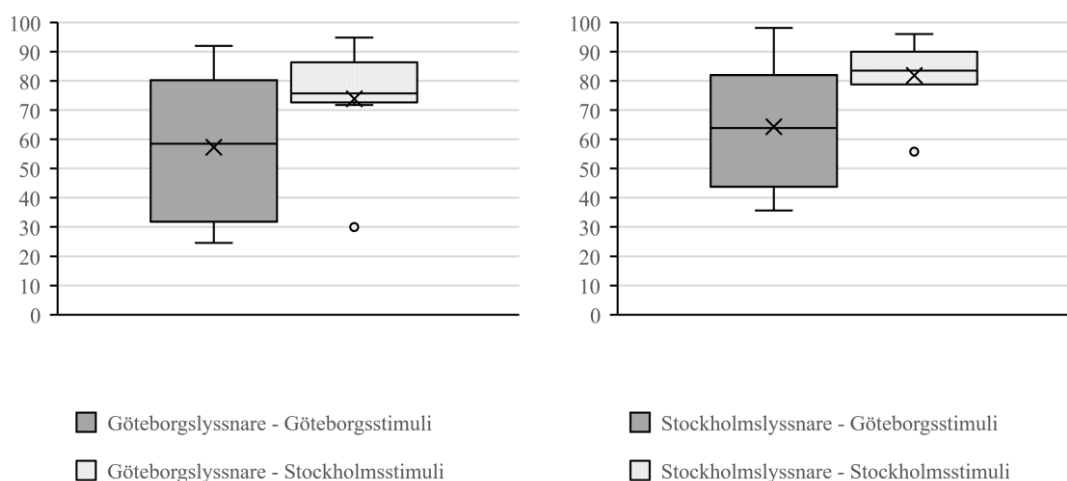
Innan varje tillfälle för transkription signerade lyssnarna informerat samtycke angående studien och de informerades om att medverkan var frivillig och att de utan förklaring eller anledning kunde avbryta sin medverkan. Alla transkriptioner gjordes anonyma och varje lyssnare tilldelades en kod. För lagring av avidentifierade audio- och videofiler användes USB-minnen med kodlås. Det fanns en godkänd etikansökan (2016/1628–31/1) för forskningsprojektet som föreliggande studie är en del av.

## Resultat

### Dialekt

För stimuli från Stockholm hade Stockholmslyssnare högre lyssnarförståelighet ( $M = 81,9$ ,  $s = 12,1$ ) än Göteborgslyssnarna ( $M = 73,8$ ,  $s = 19,4$ ); denna skillnad visade sig vara signifikant ( $p = ,05$ ). Avseende stimuli från Göteborg hade Stockholmslyssnare högre lyssnarförståelighet ( $M = 64,3$ ,  $s = 22,2$ ) än Göteborgslyssnare ( $M = 57,3$ ,  $s = 25,3$ ); också denna skillnad visade sig vara signifikant ( $p < ,05$ ). Medianvärden för respektive lyssnargrupp per inspelning redovisas i figur 3. Medelvärde på STH-stimuli var 81,9 för STH-lyssnare och 73,8 för GBG-lyssnare, en skillnad på 8,1. Medelvärde på GBG-stimuli var 64,3 för STH-lyssnare och 57,4 för GBG-lyssnare, en skillnad på 7,1.

Figur 3. Medianvärden, övre och nedre kvartil samt spridning för lyssnarförståelighet avseende Göteborgslyssnare (vänster figur) och Stockholmslyssnare (höger figur), fördelade på stimuli från Göteborg respektive Stockholm, krysset markerar respektive grupps medelvärde och pricken markerar starkt avvikande värde.



## Uppspelningsmedium

Skillnaden i lyssnarförståelighet mellan audio ( $M = 65,6$ ,  $s = 23,4$ ) och video ( $M = 68,3$ ,  $s = 19,8$ ) var inte signifikant ( $p = ,196$ ). Den största skillnaden, där video ger en högre lyssnarförståelighet, är för två av de stimuli med högst antal annoterade talavvikelser (GBG5 och GBG6), vilket kan observeras i tabell 1.

Tabell 1. Medianvärden av lyssnarförståelighet per stimuli sorterat på antal annoterade talavvikelser för audio- respektive videogrupp, skillnad mellan dessa grupper samt ålder på stimuli, sorterat på antalet annoteringar per stimulus.

Stimuli	Annoteringar	Audio-grupp	Video-grupp	Skillnad (A-V)	Ålder
GBG 3	118	50,6	55,4	4,9	7;2
GBG 5	98	24,5	42,0	17,5	6;3
GBG 6	98	30,0	45,0	15,0	5;2
GBG 7	63	37,4	46,8	9,4	5;9
GBG 1	53	66,4	68,7	2,3	9;6
STH 12	44	75,9	67,7	-8,2	7;7
STH 15	44	75,6	75,2	-0,4	7;2
GBG 2	41	81,8	83,9	2,1	9;2
GBG 8	39	75,5	69,5	-6,0	8;11
STH 13	32	78,3	74,3	-4,0	7;6
STH 16	31	30,0	28,4	-1,6	6;6
STH 10	28	89,2	78,8	-10,4	8;1
STH 9	27	71,8	81,7	9,9	9;2
STH 14	11	75,0	83,6	8,6	7;4
GBG 4	4	92,0	96,9	4,9	7;3
STH 11	3	94,8	95,5	0,7	7;8

## Diskussion

Huvudsyftet med föreliggande studie var att undersöka hur förståelighetsbedömningar påverkas av dialekt och uppspelningsmedium. Angående första frågeställningen, avseende dialekt mellan Göteborgs- respektive Stockholmsgruppen kunde en signifikant skillnad påvisas när det gällde lyssnarförståelighet. För den andra frågeställningen, om förståelighetspoängen skiljde sig signifikant mellan audio- och videogrupperna kunde ingen signifikant skillnad påvisas.

För Stockholmsstimuli var resultatet förväntat, alltså att lyssnare från Stockholm hade en högre lyssnarförståelighet för Stockholmsstimuli än lyssnare från Göteborg. Tidigare har forskare diskuterat att egenskaper som lyssnarens vana vid talaren har en viss inverkan avseende förståelighet (Kent et al., 1994) och tidigare studier har visat att dialekt inte har en huvudeffekt för förståelighetsbedömning (Pennington & Miller, 2007). Resultaten i föreliggande studie tyder dock på att dialekt avseende Stockholmsgruppen har betydelse för förståelighetsbedömning. Däremot var resultatet för Göteborgsstimuli motsatt det förväntade, även här hade Stockholmslyssnarna en högre lyssnarförståelighet jämfört med Göteborgslyssnare. En närmare analys av skillnad mellan lyssnarförståelighet visade att lyssnare från Stockholm jämfört med de från Göteborg hade en högre

lyssnarförståelighet för alla stimuli förutom två. Stockholmslyssnarna hade generellt ett högre medelvärde avseende lyssnarförståelighet och en mindre standardavvikelse för båda dialekterna. Denna genomgående skillnad mellan lyssnargrupperna kan indikera att något annat än dialektskillnaden i sig spelade in. Möjligen kan skillnaden vara relaterad till skillnader i genomförandet, eventuellt att de två grupperna tolkade instruktionerna olika.

Den stora skillnaden i lyssnarförståelighet mellan lyssnare från Stockholm och Göteborg skulle kunna påverkas av faktorer som olika nivåer av färdighet gällande koncentrationsförmåga, erfarenhet / vana av lyssnande på avvikande tal hos barn, transkriptionsfärdighet, dagsform samt hörsel som inte har kontrollerats kliniskt. Det finns dock ingen stark anledning att tro på att dessa nämnda faktorer skulle variera systematiskt mellan dessa lyssnargrupper, den variationen bör vara slumpmässigt fördelad i båda grupperna.

Det oväntade resultatet avseende dialektens betydelse kan ha att göra med lyssnarnas olika erfarenhet av dialekter. Vid en närmare analys angående lyssnarnas dialektinformation kan en större andel olika dialekter som avvek från inspelningsregionen observeras hos Göteborgslyssnarna än Stockholmslyssnarna när det gäller ”talad dialekt” och ”erfarenhet”. Denna skillnad mellan lyssnargrupperna kan ha påverkat resultatet och i framtida studier skulle grupperna kunna matchas angående dialektbakgrund. Gällande dialekter kan det nämnas att varken götamål (Göteborg) eller sveamål (Stockholm) anses vara svårförståeliga dialekter inom respektive stad, då båda är vanligt förekommande i mediala sammanhang. En större skillnad skulle kunna förväntas om man jämför lyssnargrupper från södra respektive norra Sverige då dessa grupper antagligen inte har lika stor erfarenhet av att höra varandras dialekter som Stockholm och Göteborg. Andra egenskaper hos lyssnaren har visat sig spela en roll i förståelighet, bland annat genom att föräldrar till svårförståeliga barn har visat sig förstå deras tal bättre än andra lyssnare (Coplan & Gleason, 1988; Lagerberg, 2019; Kwiatkowski & Shriberg, 1992). Gällande dialekt, har denna lyssnaregenskap som nämnts tidigare inte studerats i större utsträckning och denna studie ger en god utgångspunkt till vidare forskning.

De stimuligrupper (Stockholm respektive Göteborg) som användes i föreliggande studie skiljde sig signifikant avseende antal annoterade talavvikelser. Detta är en faktor som kan ha påverkat resultatet. Göteborgsstimuli hade ett avsevärt högre medelvärde avseende antal annoterade talavvikelser jämfört med Stockholmsstimuli. Den högre standardavvikelsen för Göteborgsstimuli för båda lyssnargrupperna stärker tesen av Gordon-Brannan (1994) att en lägre grad av förståelighet ofta är kopplat till en högre grad av variation, vilket följaktligen stämmer med föreliggande studies resultat. Stimuli som samlats in av logopedier i Göteborg och Stockholm har antagits också kunna representera respektive dialektområden, detta bekräftas också vid författarens granskning av stimuli gällande dialekt. Urvalet av dessa stimuli valdes för att representera en variation av avvikelsegrad över hela gruppen, dock inte inom respektive dialekt. För detta arbetes syfte skulle det ha varit fördelaktigt om talargrupperna hade matchats gällande antal talavvikelser.

För den andra frågeställningen, att undersöka om det fanns en skillnad i förståelighet beroende på om stimuli som bedömdes hade baserats på audio eller video, kunde ingen

signifikant skillnad observeras. Studiens resultat korresponderar med Klintö och Lohmander (2017) som fastslog att uppspelningsmedium inte verkar påverka transkriptionsresultat. Genom videoinspelningar får lyssnaren tillgång till signaloberoende information som inte hade kunnat brukas om endast audio fanns tillgängligt. Inverkan av uppspelningsmedium på förståelighetspoäng har studerats i begränsad grad och till exempel Hustad och Cahill (2003) kunde inte påvisa en signifikant säkrare bedömning för videomedium jämfört med audiomedium för talare med dysartri. Videomedium förbättrade inte förståeligheten för fyra av fem talare, jämfört med endast audio. Talaren vars förståelighet förbättrades vid närvaron av video hade också högst grad av dysartri och samtidig påverkan av motorik. Föreliggande studie visade liknande resultat, att närvaro av visuell information hjälpte mest när stimuli innehöll hög förekomst av annoterade talavvikelser. I övrigt verkar inte den visuella informationen vara till större hjälp vid bedömning av förståelighet för barn med tal- och språkstörning. Dock kvarstår att i klinisk och framför allt i social miljö är de signaloberoende faktorerna en grund för att vår kommunikation ska kunna optimeras (Miller, 2013). Att video inte förbättrar förståeligheten för allt stimuli kan bero på att videogruppen i föreliggande studie inte uppmanades i att utnyttja tillgänglig visuell information, eftersom instruktioner och metod bestämdes att hållas lika som för audiobedömning. Lyssnare förväntades dock använda den visuella informationen eftersom man gör det i vanliga konversationer. En testsituation som denna kan eventuellt förväntas ändra detta då det inte blir en naturlig kommunikationsmiljö.

Föreliggande studie har ett antal begränsningar. Gällande instruktionerna som gavs inför transkriptionerna kommenterade flera lyssnare att de var osäkra angående betydelsen av ordet ”stavelser”. Detta hade kunnat ersättas med en tydligare term, eller demonstrera detta. Instruktionen innehöll också termen ”ortografisk”, många lyssnare var osäkra och frågade om en förklaring och det hade för tydlighetens skull vara tillräckligt med till exempel ”skriv ner det du hör”. Även termen ”gissa inte” är ett uttryck som olika lyssnare kan ha tolkat olika och blir till en subjektiv bedömning av individen som olika lyssnare kan ha gett olika innebörd, med grund i att vissa lyssnare tydligt hade gissat mer än andra. Instruktionen att inte gissa används dock i andra förståelighetstest, som till exempel WEISS Intelligibility test (Weiss, 1982). Lyssnares olika tolkning av denna instruktion har ingen tydlig orsak utan kan troligen härledas till personliga faktorer hos lyssnarna. En studie av Johannisson, Lohmander och Persson (2014) undersökte olika typer av förståelighetsbedömningar och kunde visa på att reliabiliteten i deras studie blev bättre om gissningar exkluderades. I föreliggande studie hade det med denna bakgrund varit intressant att undersöka hur resultatet hade förändrats om gissningar exkluderades. Ytterligare tydliggörande hade behövts genom att till exempel säga att man inte ska ”rätta” exempelvis grammatiken eller ordföljd. Liknande osäkerheter som har presenterats och att lyssnarförståelighet varierar mellan lyssnare förväntas dock finnas i studier som involverar perceptuella bedömningar och det visat sig vara fördelaktigt om förståelighetsbedömningar baseras på fler än en lyssnares bedömning (Hustad, Oakes & Allison, 2015). Flera av lyssnarna frågade om vilket perspektiv man skulle transkribera från. Försöksledaren förklarade då med ett exempel att transkriptionen skulle skrivas utifrån lyssnarens perspektiv och att fokus ligger på betydelsen. I fortsatt användande av lyssnarförståelighet enligt metoden i denna studie föreslås därför instruktionerna ”skriv ner det du förstår”. Om talaren säger ”tatt” (vilket lyssnaren hör) men samtidigt förstår att barnet menar ”katt” (vilket man förstår) är meningen att man ska skriva ”katt”. Vissa

lyssnare i studien undrade om påhittade ord eller felstavade ord skulle transkriberas, vilket besvarades nekande. Denna instruktion visade sig vara problematisk då flera felstavade och påhittade ord förekom i facit. Det finns ingen absolut gräns för vad ett felstavat eller påhittat ord innebär. Exempelvis valde projektet som spelade in stimuli att inkludera ord som är vanligt förekommande hos barn men inte hos vuxna, som till exempel ”gådde”.

För transkriberingen användes olinjerat papper, ett alternativ hade varit att använda en dator med VLC media player och textredigeringsprogram på samma skärm. Detta hade möjliggjort att lyssnarna i speciellt videogruppen hade kunnat hålla ögonen på talaren medan man skrev på datorn. Dessa förändringar avser att hjälpa lyssnaren att fokusera helt på transkriberingen och inte behöva lägga energi på andra saker. Textredigeringsprogram hade även underlättat rättningsprocessen då avläsning av svåra handstilar inte behöver göras.

Många faktorer påverkar en klinisk bedömning av förståelighet. Föreliggande studie har visat att bedömarnas dialekt kan ha en viss betydelse för förståelighetsbedömning. Vidare kan noteras att för förståelighetsbedömning av individer med stort antal talavvikelser kan användning av videomedium eventuellt resultera i att lyssnaren förstår talaren bättre. Av den anledningen finns det goda skäl att använda videomedium för dessa individer, förutsatt att syftet är att få god förståelighet samt likna en vardagssituation. Föreliggande studien är begränsad avseende lyssnargruppenas storlek för både audio- och videobedömningen och för vidare forskning bör grupperna vara större. Vidare forskning behövs således inom detta område för att kunna veta specifikt hur faktorer som dialekt och val av uppspelningsmedium påverkar en lyssnarbedömning av förståelighet.

## Referenser

- Coplan, J. & Gleason, J. R. (1988). Unclear speech: Recognition and significance of unintelligible speech in preschool children. *Pediatrics*, 82(3), 447–452.
- Emander, A., Kjerrman, N. (2018) *Validitets- och reliabilitetsprövning av förståelighetsbedömningar inhämtade med ett panelsvarssystem* (Magisteruppsats). Stockholm: Institutionen för klinisk vetenskap, intervention och teknik, CLINTEC, Karolinska Institutet.
- Fleiss, J. L. (1986). *The design and analysis of clinical experiments* (New York, NY: Wiley).
- Flipsen, P. Jr. (1995). Speaker–listener familiarity: Parents as judges of delayed speech intelligibility. *Journal of Communication Disorders*, 28, 3–19.
- Flipsen, P. Jr. (2006). Measuring the intelligibility of conversational speech in children. *Clinical Linguistics & Phonetics*. 20(4), 202-312.
- Gordon-Brannan, M. (1994). Assessment of speech intelligibility: a prerequisite for oral communicative competence. *Topics in Language Disorders*, 14(2). 18–27.
- Hustad, K., & Cahill, M. (2003). Effects of presentation mode and repeated familiarization on intelligibility of dysarthric speech. *American Journal of Speech - Language Pathology*, 12(2), 198-208.

- Hustad, K. C., Oakes, A., & Allison, K. (2015). Variability and Diagnostic Accuracy of Speech Intelligibility Scores in Children. *Journal of speech, language, and hearing research: JSLHR*, 58(6), 1695–1707. doi:10.1044/2015\_JSLHR-S-14-0365.
- Johannisson, T. B., Lohmander, A. & Persson, C. (2014). Assessing intelligibility by single words, sentences and spontaneous speech: A methodological study of the speech production of 10-year-olds. *Logopedics, Phoniatrics, Vocology*, 2014, Vol. 39, Iss. 14, Pp. 159-168, 39(14), 159-168.
- Kent, D. R., Miolo, G., & Bloedel, S. (1994). The intelligibility of children's speech. *American Journal Speech-Language Pathology*, 3, 81–95. doi: 10.1044/1058-0360.0302.81.
- Kierkegaard, L. (2017). *Hur många dialekter finns det i Sverige? Var går gränsen mellan olika dialekter? - Institutet för språk och folkminnen*. [online] Sprakochfolkminnen.se. Available at: <http://www.sprakochfolkminnen.se/sprak/dialekter/fragor-och-svar-om-dialekter/faq/2013-10-21-hur-manga-dialekter-finns-det-i-sverige-var-gar-gransen-mellan-olika-dialekter.html> [Hämtad 23 mars, 2019].
- Klintö, K., & Lohmander, A. (2017). Does the recording medium influence phonetic transcription of cleft palate speech? *International Journal of Language & Communication Disorders*, 52(4), 440-449. <https://doi.org/10.1111/1460-6984.12282>.
- Kwiatkowski, J. & Shriberg, L. D. (1992). Intelligibility assessment in developmental phonological disorders: Accuracy of caregiver gloss. *Journal of Speech and Hearing Research*, 35(5), 1095–1104.
- Lagerberg, T. (2013). Assessment of intelligibility in children. Gothenburg: Division of Speech and Language Pathology, Institute of Neuroscience and Physiology, Sahlgrenska Academy at University of Gothenburg. Tillgänglig: <https://gupea.ub.gu.se/handle/2077/34074>.
- Lagerberg, T., Åsberg, J., Hartelius, L., & Persson, C. (2014). Assessment of intelligibility using children's spontaneous speech: Methodological aspects. *International Journal of Language & Communication Disorders*, 49(2), 228–239.
- Lagerberg, T., Hellström, A., Lundberg, E., & Hartelius, L. (2019). An Investigation of the Clinical Use of a Single-Word Procedure to Assess Intelligibility (Swedish Test of Intelligibility for Children) and an Evaluation of the Validity and Reliability of the Intelligibility in Context Scale. *Journal of Speech, Language and Hearing Research (Online)*, 62(3), 668–681.
- Lam, J., Olsson, R. (2017) *Förståelighet hos barn med avvikande tal bedömda av lyssnare med svenska som andraspråk* (Magisteruppsats). Göteborg: Institutionen för neurovetenskap och fysiologi, Sektionen för hälsa och rehabilitering, Enheten för logopedi, Göteborgs universitet. Tillgänglig: <http://hdl.handle.net/2077/53353>.
- Mattys, S. L., Davis, M. H., Bradlow, A. R., & Scott, S. K. (2012). Speech recognition in adverse conditions: A review. *Language and Cognitive Processes*, 27(7-8), 953-978.
- McHenry, M. (2011). An exploration of listener variability in intelligibility judgments. *American Journal of Speech-Language Pathology / American Speech-Language-Hearing Association*, 20(2), 119. doi:10.1044/1058-0360(2010/10-0059).
- Miller, N. (2013). Measuring up to speech intelligibility. *International Journal of Language and Communication Disorders*, 48(6), 601-612.
- Pennington, L., & Miller, N. (2007). Influence of listening conditions and listener characteristics on intelligibility of dysarthric speech. *Clinical Linguistics & Phonetics*, 21(5), 393-403. doi:10.1080/02699200701276675.

- Schiavetti, N. (1992) Scaling procedures for the measurement of speech intelligibility. In Kent, R. D. (Ed.), *Intelligibility in speech disorders*. Philadelphia, PA: John Benjamins.
- Shriberg, L. D., Kwiatkowski, J. (1982). Phonological Disorders III: A Procedure for Assessing Severity of Involvement. *Journal of Speech and Hearing Disorders*, 47(3), 256 - 270.
- Smith, S. D., Pennington, B. F., Boada, R., & Shriberg, L. D. (2005). Linkage of speech sound disorder to reading disability loci. *Journal of Child Psychology and Psychiatry*, 46(10), 1057–1066. doi: <https://doi-org.proxy.kib.ki.se/10.1111/j.1469-7610.2005.01534.x>.
- Weiss, C. E. (1982) *Weiss intelligibility test*. Tigard, OR: CC Publications.
- Yorkston, K. M., Strand, E. A., & Kennedy, M. R. T. (1996). Comprehensibility of Dysarthric Speech: Implications for Assessment and Treatment Planning. *American Journal of Speech-Language Pathology*, 5(1), 55-65.