



SAHLGRENSKA AKADEMIN

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

323

**Acceptabilitet avseende tal påverkat av speech sound
disorder, kopplat till talavvikelser och olika
lyssnargrupper**

Charlotte Carlsson
Matilda Jonsson

Examensarbete i logopedi
30 högskolepoäng
Vårterminen 2018

Handledare
Tove Lagerberg
Sofia Strömbergsson

Acceptabilitet avseende tal påverkat av speech sound disorder, kopplat till talavvikelser och olika lyssnargrupper

Charlotte Carlsson
Matilda Jonsson

Sammanfattning. Studiens syfte var att undersöka acceptabilitet hos olika lyssnare avseende barn med speech sound disorder (SSD) och att kartlägga vilka talavvikelser som är vanligt förekommande i tal med låg acceptabilitet. Validitet och reliabilitet prövades för studiens typ av audience response system (ARS) vid acceptabilitet. Talmaterial av 14 barn med SSD och två barn med typisk talutveckling användes. Lyssnare var åtta barn, 17 vuxna naiva lyssnare och nio logopeder. Signifikant skillnad visades mellan barn och vuxna naiva lyssnare ($p=0,023$) vad gällde grad av acceptabilitet. Dentalisering och avtoning/påtoning var vanligt förekommande i tal med låg grad av acceptabilitet. Validitet för ARS-systemet som mätinstrument för acceptabilitet för denna grupp av lyssnare och SSD prövades mot en skattningsskala och korrelationen var signifikant, ($r=0,91$), inter och intrabedömarreliabilitet var måttligt god.

Nyckelord: Acceptabilitet, speech sound disorder, audience response system, lyssnare

Acceptability regarding speech affected by speech sound disorder, connected to deviations of speech and different groups of listeners

Abstract. The aim of the study was to examine acceptability of different listeners regarding children with speech sound disorder (SSD) and to analyze which speech deviances that occurred most in speech with low degree of acceptability. Validity and reliability were tested regarding this study's type of audience response system (ARS) measuring acceptability. Speech material from 14 children with SSD and two children with typical speech, were used. The participants were eight children, 17 adult naive listeners and nine speech-language pathologists. A significant difference was shown between children and adult naive listeners ($p=0,023$) regarding degree of acceptability. Dentalization and voicing/devoicing occurred frequently in the speech with low degree of acceptability. Validity of the ARS system when measuring acceptability of this study's listeners and SSD was tested in relation to a rating scale and the correlation was statistically significant, ($r=0,91$), inter- and intrajudge reliability were moderately good.

Key words: Acceptability, speech sound disorder, audience response system, listener

Talet är för de allra flesta ett grundläggande kommunikationsmedel. Tal kan till stor del sägas handla om informationsöverföring i form av det språkliga innehållet. Hur talet hos en individ uppfattas av lyssnare är också en aspekt av kommunikationen, vilket undersöktes i föreliggande studie.

Acceptabilitet vid talperception kan definieras på olika sätt. Vid tidigare forskning kring ämnet har man bland annat menat att begreppet avser i vilken grad "speech deviates from normal", det vill säga hur avvikande från typiskt tal något uppfattas vara (Henningsson et al., 2008, s. 5). Utifrån detta resonemang handlar det om något ljudmässigt avvikande i talet, vilket drar lyssnarens uppmärksamhet till sig och uppfattas som annorlunda. I samma publikation påpekas att tal med låg grad av acceptabilitet inte behöver uppfattas som svårförståeligt, men den låga acceptabilitetsgraden riskerar att fjärma lyssnarens uppmärksamhet från talets innebörd (Henningsson et al., 2008). Vidare menar Henningsson et al. (2008) att acceptabilitet av tal bör betraktas som ett kontinuum, där graden av acceptabilitet varierar. En annan förekommande definition av acceptabilitet av tal baserar sig på talets behaglighet: "acceptability (i.e., the degree to which speech has a pleasant quality, is agreeable to the ear, and has a natural intonation and pace)" (Liu, Wan & Wang., 2005, s. 10). Gemensamt för de två presenterade synsätten är att acceptabilitet inom det logopediska fältet handlar om hur talet uppfattas subjektivt hos en enskild lyssnare, det är således hos lyssnaren som graden av acceptabilitet avgörs och därför kan graden av acceptabilitet av ett och samma talmaterial variera beroende på lyssnaren. I detta examensarbete har följande definition formulerats av författarna: *Talets grad av acceptabilitet anger hur avvikande talet uppfattas av lyssnaren.*

Barn med speech sound disorder (SSD) är en heterogen grupp och är den vanligast förekommande pediatrika kommunikationsstörningen. Vad SSD innebär finns det ingen enhetlig definition av och det finns i nuläget heller inte något universellt klassificeringssystem för SSD (Waring & Knight, 2013). I Diagnostic and statistical manual of mental disorders: DSM-5 (2013) definieras SSD som varaktiga svårigheter i talproduktionen vilka påverkar artikulationen och/eller fonologin som gör talet svårförståeligt. Dessa svårigheter grundar sig inte i något medfött eller förvärvat medicinskt eller neurologiskt tillstånd. Enligt DSM-5 finns fyra kriterier för SSD: 1. Varaktiga svårigheter att forma språkljud, vilket påverkar talets förståelighet eller hindrar personens möjligheter att förmedla sig verbalt: stör den verbala kommunikationen. 2. Störningen leder till begränsningar i kommunikationsförmågan, som påverkar ett eller flera av följande avseenden: social delaktighet, studier eller arbete. 3. Uppkomsten av symtomen visar sig tidigt under barndomen. 4. Svårigheterna med talproduktionen ska inte kunna härledas till några medfödda eller förvärvade tillstånd såsom cerebral pares, läpp- käk- gomspalt, dövhet eller hörselnedsättning, traumatisk hjärnskada eller andra medicinska eller neurologiska tillstånd (Diagnostic and statistical manual of mental disorders: DSM-5, 2013, s.44). McGrath et al. (2007) definierar SSD som en störning vilket kännetecknas av förseningar i talproduktionen, med artikulations- och/eller fonologipåverkan. Följande definition av SSD används i föreliggande studie: *SSD innebär talsvårigheter, vilka påverkar produktionen av specifika språkljud som resulterar i talavvikelser.*

Med mätinstrumentet ARS kan enskilda personer framföra sina åsikter om något genom att trycka på en mentometerknapp, vilket möjliggör analys av vad det är i talet som

lyssnaren reagerar på. Man kan få datan av ARS-resultaten normaliserad, vilket justerar för variationen mellan deltagarna, då den kompenserar för att vissa lyssnare klickar oftare än andra. Den normaliserade datan kan därefter bearbetas i en kernel density estimation-analys (KDE-analys), vilket är en utjämnande analys som tydliggör lyssnartoppar i en kurva som representerar klickmängd över tid (Johansson, Strömbergsson, Robieux & McAllister, 2017). ARS har som svarsmetod visat sig medföra ett högre deltagarantal än genom svarsmetoden pappersformulär (Turban, 2009). Då det gäller användning av ARS som mätinstrument för acceptabilitet av tal, har detta tidigare testats i en studie av Johansson et al. (2017), där lyssnares reaktioner på avvikande röstkvaliteter undersöktes. Även lyssnares reaktioner rörande förståelighet av tal har tidigare mätts med ARS (Contardo, McAllister & Strömbergsson, 2014).

Att undersöka ett mätinstruments validitet och reliabilitet är av stor vikt, då de visar hur väl mätinstrumentet mäter det fenomen man avsett att mäta och även hur tillförlitligt mätinstrumentet är; att resultatet av ett fenomen är samma vid upprepade mätningar och oberoende av vem som utför testningen. Enligt Cicchetti (1994) kan utvärdering av mätinstrument vara av stor vikt i sökandet efter ett så passande mätinstrument som möjligt, vid undersökning av ett fenomen. Vidare menar Cicchetti (1994) att mätinstruments validitet såväl som reliabilitet kan undersökas på olika sätt. Validitet kan exempelvis undersökas genom att studera *face validity*, då man efter användning av mätinstrumentet i en testsituation utvärderar huruvida det tycks ha mätt fenomenet det egentligen var tänkt att mäta, exempelvis genom att jämföra med resultat från ett test som är väl beprövat och mäter samma sak. Exempel på ett sätt att undersöka reliabilitet är genom *test-retest*, då deltagare använder mätinstrumentet i samma test upprepade gånger, utifrån dessa resultat utvärderar man huruvida värdena mätinstrumentet ger tycks vara reliabla eller inte (Cicchetti, 1994). Utvärdering av ARS-systemet som mätmetod har bland annat gjorts i studien av Johansson et al. (2017) rörande reaktioner på avvikande röstkvalitet. Mätinstrumentet utifrån studien kan sägas vara något högkänsligt i jämförelse med en manuell mätning. Dock hade mätinstrumentet hög validitet, då det tydligt mätte samma fenomen som den manuella mätningen (Johansson et al., 2017). Validering av ARS som mätinstrument gjordes även i studien av Contardo et al. (2014) som undersökte förståelighet av barns avvikande tal. Antalet klick lyssnare gjort via ARS inom en ljudfil jämfördes med resultatet som noterats manuellt med hjälp av Percentage of consonants correct (PCC), vilket är ett mått på artikulationsavvikelse. Denna jämförelse visade på en stark korrelation, vilket tydde på god validitet då det var förväntat att hög grad av talavvikelse bedömt via PCC och låg förståelighet skulle korrelera (Contardo et al., 2014). I en undersökning av Strömbergsson och Tännander (2013) mättes på liknande sätt korrelationen mellan bedömning av tal från barn med SSD utifrån PCC och bedömning utifrån ARS, vilket även där visade på en stark korrelation. Edlund, Moubayed, Tännander och Gustafson (2013) undersökte i en studie den typ av ARS-system som används i föreliggande studie och resultatet tydde på att instrumentet hade god reliabilitet rörande att samtliga lyssnare klickade vid samma tidpunkter.

Bedömning av acceptabilitet genom gradering på skala är ett vanligt sätt att mäta acceptabilitet då det mäts på ett kontinuum, i övrigt finns det ännu inget vedertaget universellt sätt att bedöma acceptabilitet på (Bagnall & David, 1988). I studien av Dagenais, Brown och Moore (2006) och Trudeau (1987) användes just en sådan skattningsskala för att mäta acceptabilitet av tal. Eadie och Doyle (2004) uppger att en

skattningsskala är en valid mätmetod för bedömning av acceptabilitet av tracheoesophagealt tal.

Eftersom acceptabilitet är en ganska outforskad aspekt rörande hur lyssnare uppfattar tal, är det inte förvånande att det finns begränsad kunskap om hur acceptabilitet varierar mellan olika lyssnare. Studier som undersökt om det finns någon skillnad mellan olika lyssnare vad det gäller acceptabilitet av tal, visar inte alltid på enhetliga resultat. I en studie av Finizia, Lindström och Dotevall (1998) visades ingen skillnad mellan logopederna och vuxna naiva lyssnare vad gällde acceptabilitet av tal hos patienter som strålades för cancer. Dock fanns skillnad mellan lyssnarna, då de vuxna naiva lyssnarna hade lägre acceptabilitet till de laryngektomerade patienternas tal än vad logopederna hade. En studie, vilken undersökte acceptabilitet av inspelat dysartriskt tal producerat av personer med traumatisk hjärnskada, visade att det fanns en skillnad då äldre vuxna var mer accepterande gentemot det inspelade talet än yngre vuxna (Dagenais, Adlington & Evans, 2011). Andra studier har undersökt acceptabilitet hos olika lyssnare utifrån deras socioekonomiska förutsättningar och kön, där ingen skillnad fanns i acceptabilitet av tal varken mellan könstillhörighet (Ellis, Spiegel & Benjamin, 2002) eller socioekonomiska förutsättningar (Alcorn, Griffith & Miner, 1974).

Det tycks finnas få studier som undersökt vilka talavvikelser som uppfattas vara vanligt förekommande i tal med låg grad av acceptabilitet vid just SSD. En tidigare studie har dock undersökt vilka talavvikelser barn reagerade på då de lyssnade på tal från barn med LKG samt hur dessa reaktioner tedde sig i förhållande till logopeders bedömning (Nyberg & Havstam, 2016). Angående tecken på velofarynxinsufficiens (VPI) tycktes barnen inte reagera på milda sådana, barnen reagerade dock på kraftiga tecken på VPI medan logopederna reagerade även på milda sådana. Barnen var också känsliga för artikulationsavvikelser och reagerade på samma artikulatoriska svårigheter som logopederna. Barnen tycktes främst reagera på då dentaler producerades velart samt då /s/ producerades interdentalt/lateralt (Nyberg & Havstam, 2016). Flera tidigare studier tyder på att typiskt tal har en högre grad av acceptabilitet hos lyssnare än tal med talavvikelser. Bagnall och David (1988) bedömde acceptabilitet av tal hos barn med gomspalt. Resultatet visade att lägre acceptabilitetsgrad uppmättes för talet hos barnen med gomspalt än för typiskt tal. Liknande resultat framkom i studien av Eadie, Rajabzadeh, Isetti, Nevdahl och Baylor (2017), där lyssnarna uppfattade typiskt tal mer behagligt än talmaterial påverkat av grav adduktor spasmodisk dysfoni (ADSD). Acceptabilitet definierades i det fallet som uppfattad lyssnarbehaglighet (eng. listener comfort).

Man kan tänka sig att låg grad av acceptabilitet av tal i vissa fall kan leda till en negativ attityd gentemot talaren. I studien av Crowe (1987) undersöktes vilken attityd barn i fjärde och sjätte klass i USA, Ohio, hade gentemot barn som hade artikulatoriska talavvikelser (eng. articulatory errors). Resultatet visade att lyssnare som gick i sjätte klass var mer negativa till dessa artikulatoriska talavvikelser än vad elever i fjärde klass var. Varför dessa barn skiljde sig åt ansågs enligt studiens författare kunna bero på att utvecklingen av allt mer negativa attityder sker upp till en viss grad/ålder, där sedan den negativa attityden stabiliseras. Ingen specifik ålder för när denna stabilitet av negativa attityder sker angavs (Crowe, 1987). I studien av Nyberg och Havstam (2016) där det framgick att 10-åriga barn reagerade på små artikulatoriska talavvikelser, framgick det att flera barn

såg dessa talavvikelser som något negativt och menade att dessa talare lät barnsliga. Ett barn menade att hen upplevt att ett barn med talavvikelser blivit mobbat i skolan och andra barn i samma studie menade att de trodde att barn kunde bli mobbade på grund av sina talavvikelser (Nyberg & Havstam, 2016). En annan studie visade att vuxna naiva lyssnare i allmänhet tolkade fonologiska talavvikelser hos barn som ett tecken på lägre kognitiv förmåga (Burroughs & Tomblin, 1990). Ytterligare en studie där vuxna naiva lyssnare deltog visade att lärare i andra klass associerade svårförståeligt tal hos personer med SSD med beteendeproblematik (Overby, Carell & Bernthal, 2007). Dessa studier tyder alltså på att talavvikelser hos barn kan få allvarliga sociala konsekvenser.

Eftersom det tycks finnas ytterst lite forskning rörande hur acceptabilitet varierar mellan olika lyssnare och vilka fonologiska talavvikelser som är vanligt förekommande i tal med låg grad av acceptabilitet, avseende barn med SSD, var syftet med föreliggande studie att studera detta. Dessutom var syftet att undersöka om användningen av ARS-systemet vid acceptabilitet av tal från barn med SSD, visade på god validitet och reliabilitet. I föreliggande studie undersöktes följande frågeställningar:

1. Finns det skillnader i bedömd grad av acceptabilitet gentemot tal hos barn med speech sound disorder mellan tre olika lyssnargrupper; barn, vuxna naiva lyssnare och logopedier?
2. Vilka fonologiska processer kännetecknar tal med låg acceptabilitet?
3. Är ARS-systemet en valid och reliabel metod för bedömning av acceptabilitet?

Metod

Forskningssammanhang

Föreliggande studie ingick i ett större forskningsprojekt vid namn *Funktionella konsekvenser av avvikelser i barns sammanhängande tal*. Projektet bedrivs vid Karolinska institutet (KI) i samarbete med Kungliga Tekniska Högskolan (KTH) och Göteborgs universitet (GU) och syftet är att studera förståelighet och acceptabilitet avseende barns tal samt utvärdera användning av mätinstrumentet ARS-system inom detta forskningsfält.

Deltagare

Tre olika lyssnargrupper; barn, vuxna naiva lyssnare samt logopedier använde ARS-systemet som bedömningsinstrument. Utöver dessa deltog ännu en grupp vuxna naiva lyssnare som använde skattningsskala som bedömningsinstrument. Inklusionskriterie för samtliga lyssnare var svenska som modersmål. Barnen skulle vara mellan 6-10 år och de vuxna naiva lyssnarna skulle vara i arbetsför ålder, det vill säga mellan 18 - ca 65 år. Logopederna skulle ha minst 2 års klinisk erfarenhet som logoped och ha jobbat med barn i minst 2 år. Försök till rekrytering av deltagande barn gjordes inledningsvis via kontakt med skolklasser och barnkör. Då detta inte gav resultat rekryterades de deltagande logopedernas barn. Vuxna naiva lyssnare rekryterades genom ett bekvämlighetsurval av bekanta. Rekrytering av logopedier skedde via kontakt med logopedmottagningar i Göteborg med omnejd, via bekanta samt via handledaren till föreliggande studies

författare. Korrespondens kring rekrytering av deltagare skedde via mail, sociala medier och telefonkontakt. Samtliga personer som anmälde intresse för att delta, fick mail med information om tid och plats för utförandet. De mottog även en samtyckesblankett som innehöll beskrivning av föreliggande studie samt det större projektet studien ingick i. Lyssnarna som använt sig av ARS-system som bedömningsinstrument ombads även i sina samtyckesblanketter att delta vid ytterligare ett tillfälle, utanför ramen för detta examensarbete. Rekryteringen resulterade i åtta barn, nio vuxna naiva lyssnare samt nio logopedier vilka använde ARS-system som bedömningsinstrument och åtta vuxna naiva lyssnare vilka använde skattningsskala som bedömningsinstrument. Totalt deltog 34 lyssnare och resultat från dessa användes för att besvara studiens frågeställningar (se tabell 1).

Tabell 1

Ålder, kön och bedömningsmetod för samtliga lyssnare i studien.

	Antal lyssnare	Ålder	Kön
Barn (ARS)	8	6-10 ($M=8:5$, $s=1,20$)	6 P, 2 F
Vuxna naiva lyssnare (ARS)	9	24-58 ($M=33:9$, $s=12,06$)	4 M, 5 K
Logopedier (ARS)	9	26-58 ($M=42:6$, $s=10,66$)	9 K
Vuxna naiva lyssnare (skattning)	8	24-29 ($M=27:3$, $s=1,67$)	4 M, 4 K

Not. ARS=Bedömning med ARS, skattning=Bedömning med skattningsskala, P=pojke, F=flicka, K=kvinn, M=man

Material

Talmaterial. Talmaterialet fanns insamlat sedan tidigare inom ramen för det större projektet. Inspelningar fanns av totalt 19 barn med SSD samt 11 barn med typisk talutveckling. Ljudfilerna till föreliggande studie valdes ut av tre projektmedarbetare, vilka alla är verksamma logopedier och varav två är specialiserade på barn med avvikande tal, med avsikten att uppfylla en variation i grad av talavvikelser. Talmaterialet bestod efter urvalet av 16 ljudfiler varav samtliga innehöll en minuts spontantal av ett barn som besvarade frågan *Vad skulle du göra om du fick en miljon kronor?* Talare i föreliggande studie var 16 barn inom åldersspannet 4:1-8:1 år ($M=5:9$, $s=1,12$). Dessa rekryterades till projektet via logopedmottagningar i Stockholm och Göteborg, åtta talare spelades in i vardera stad. Inklusionskriterier för 14 av talarna var att de behövde vara i åldern 4-10 år, svensktalande, ha SSD samt att detta enligt behandlande logoped påverkade förståeligheten. Inklusionskriterier för de resterande två talarna var att de skulle vara i åldern 4-10 år, svensktalande och ha typisk tal- och språkutveckling. Majoriteten av talare var pojkar ($n=13$) och övriga flickor ($n=3$), alla barn hade svenska som starkaste språk. Bland de logopediska diagnoserna var den vanligast förekommande fonologisk språkstörning ($n=7$) bland talarna. Talet från samtalspartnern, vilket var en logoped, klipptes bort och fyra av ljudfilerna dubbletades för att kunna undersöka intrabedömarreliabilitet, vilket resulterade i 20 stimuli sammantaget. Utöver de beskrivna

ljudfilerna fanns även ett övningsexempel innehållande tal från ett barn med talavvikelser, vilket spelades upp först för att säkerställa att lyssnarna vände sig vid sättet att använda ARS-system respektive skattningsskala. Samtliga grupper fick lyssna till samma ljudfiler, ordningen på ljudfilerna varierades dock. Vid bedömningar genom ARS-system användes fem olika uppspelningsordningar och vid bedömningar genom skattningsskala användes tre olika uppspelningsordningar.

Enkät för demografisk data. En enkät för demografisk data hade arbetats fram inom det större projektet och godkänts av Regionala etikprövningsnämnden i Stockholm, denna användes i föreliggande studie. Enkäten inkluderade frågor om deltagares ålder, kön samt språk. De vuxna naiva lyssnarna och logopederna fick även frågor om högsta genomförda utbildning, om hur ofta man möter barn i förskoleåldern, om man var logoped och i så fall hur länge man arbetat som logoped, vad man arbetade med i dagsläget och hur länge man arbetat med barn. Till samtliga deltagare ställdes avslutningsvis en allmän fråga om man trodde att det fanns något annat som kunde påverka ens förmåga att förstå avvikande tal.

Utrustning. Inspelningarna av talarna gjordes med närmikrofon från Sennheiser, MKE-P-C. ARS-systemet som användes vid lyssnarbedömningen var utvecklat vid Kungliga Tekniska Högskolan, Stockholm och hörlurarna var från Sony, MDR-ZX660AP. Hörselscreening utfördes med hjälp av en audiometer från Interacoustics, AS608 Screening Audiometer tillhandahållen av C-A TEGNÉR AB samt hörlurar från Sennheiser.

Tillvägagångssätt

ARS-bedömningar och skattningsbedömningar genomfördes av lyssnare vid 14 olika tillfällen under mitten av februari, 2018. Samtliga bedömningar genomfördes i konferensrummet vid Enheten för logopedi, Göteborg. Båda författarna till föreliggande studie var närvarade vid alla tillfällen. Lyssnarna fyllde i samtyckesblankett, besvarade blanketten med demografiska frågor, deltog i en lyssnarbedömning samt genomgick en hörselscreening utförd av författarna till föreliggande studie. Samtliga lyssnare visade sig ha normal hörsel. Skattningsbedömningarna genomfördes under tre dagar vid åtta tillfällen där lyssnarna deltog enskilt. ARS-bedömningarna skedde under tre dagar vid sex tillfällen, där grupper om 2-8 personer deltog samtidigt. Under bedömning med ARS-system medverkade en medarbetare för det större projektet, som är anställd vid KI.

En diskussion fördes sinsemellan författarna inför utformningen av instruktionen hur ordet oacceptabelt skulle kunna presenteras på ett mer lättförståeligt sätt. Det beslutades då att oacceptabelt skulle representeras av ordet *konstigt* eftersom detta antogs kunna förstås av samtliga deltagare. Lyssnarna som skattade acceptabilitet på intervallskalor fick sitta vid en dator, via hörlurar lyssna på stimulen och därefter fylla i en skattningsskala efter varje stimuli. Ljudfilerna på datorn kodades med 1, 2 osv. i enlighet med skattningsformuläret där det stod *talare 1*, *talare 2* osv. Lyssnarna fick muntlig information om att de skulle lyssna på inspelningar av barns tal, som var i åldern 4-8 år och att de sedan skulle fylla i på en skattningsskala hur konstigt de tyckte talet lät, att de skulle lyssna på varje stimuli en gång samt välja själva när de ville lyssna på nästa. Ett förtydligande lades till efter några bedömningar om att det inte fanns några rätt eller fel

utan att den subjektiva uppfattningen var i fokus. Bedömningen tog ungefär 45 minuter för en lyssnare att genomföra.

Lyssnarna som använde ARS-system fick sitta med ryggarna mot varandra för att inte påverkas av de andra lyssnarnas aktivitet. Lyssnarna fick muntlig information om att de skulle lyssna på stimuli i hörlurar, att de hade varsin handkontroll med en mentometerknapp de skulle trycka på när man tyckte att talet lät *konstigt*, att barnen som talade var 4-8 år, att en paus i lyssnandet skulle ske efter några stimuli, att det inte fanns några rätt eller fel utan att den subjektiva uppfattningen var i fokus, att man ibland kommer klicka mycket och ibland klicka lite. Anledningen till att denna instruktion skiljer sig till viss del från den som lyssnarna vid skattningen fick höra, är att en tidigare utarbetad instruktion användes vid tillfällena med skattning, som var tänkt att användas även vid tillfällena med ARS-system. Efter en pilottestning av ARS-systemet som utfördes inom det större projektet vid KI i Stockholm, omarbetades instruktionen, men då var skattningarna redan genomförda. Då tillfällena med ARS-system ingår i det större projektet och data från dessa kommer att användas i andra sammanhang än föreliggande studie, beslutades det att i fortsättningen använda den nya instruktionen. Vid frågor från deltagarna angående vad *konstigt* betydde, gavs som svar att det är subjektivt och att författarna inte kunde säga mer än så.

Mätmetoder

ARS. Bedömning av acceptabilitet hos lyssnare gjordes med hjälp av ett ARS-system. ARS-systemet registrerade antal gånger varje deltagare klickade i varje stimuli och vid vilken tidpunkt. Registrerad data från ARS-systemet från lyssnarbedömningarna, bearbetades av medarbetare vid KTH inom det större projektet.

Skattningsformulär. Åtta av de 16 vuxna naiva lyssnare som rekryterades till föreliggande studie, bedömde acceptabilitet genom ett skattningsformulär med en intervallskala mellan 1 och 7 (se figur 1). Formuläret som användes är taget från magisterarbetet av Moreno och Niklasson (2016), där det användes för bedömning av acceptabilitet av tal från personer som stammar. Skalan angav hur *normalt* till *mycket konstigt* talet lät och hur *behagligt* till *mycket obehagligt* talet lät. I föreliggande studie användes endast skalan *normalt* till *mycket konstigt*, eftersom det gick i linje med föreliggande studies frågeställning samt dess definition av acceptabilitet. Skalan användes som grund för att undersöka validiteten hos ARS-systemet, då en gradering genom ett sådant formulär kunde jämföras med uppmätta klickfrekvenser från ARS-systemet.

Talare								
1	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	
Normalt	1	2	3	4	5	6	7	Mycket konstigt

Figur 1. Skala för bedömning av acceptabilitet

Annoteringar. Inför analys av talmaterialet avseende vilka talavvikelser som var vanligast förekommande samt vilka talavvikelser som visade på lägst grad av acceptabilitet enligt lyssnare, gjordes manuella annoteringar. Detta gjordes i programmet ELAN 5.1 (Elan, 2018) med hjälp av manual framställd av Max Planck Institute for Psycholinguistics (2017) och principer för annoteringen hämtades ur Ode och Öster Cattu Alves (2017) magisterarbete. Som underlag för analysen fanns en ortografisk transkription av talmaterialet utförd av en logoped involverad i det större projektet. Varje ljudfil analyserades på så sätt att det vid varje talavvikelse markerades tidpunkt och typ av talavvikelse. Klassificeringen av talavvikelse typer skedde med utgångspunkt från de processer som beskrivits i Nettelblatt (2007) (se tabell 2). I de fall där det var möjligt att klassificera talavvikelsen som antingen syntagmatisk eller paradigmatiske, gavs de paradigmatiske företräde. Detta innebar exempelvis att när ett ord som *katt* producerades /tat/ beskrevs det som en dentalisering snarare än en assimilation. Detta motiverades av konsekvensskäl; om två tolkningar var möjliga skulle det finnas ett sätt att göra ett konsekvent val. Det betyder inte att syntagmatiska beskrivningar var uteslutna, då valet gällde endast där båda tolkningarna var möjliga. Parallellt med föreliggande studie pågick en studie vid KI där samma talmaterial användes för att undersöka förståelighet hos olika lyssnare. En uppdelning av talmaterialet gjordes mellan de två författarna till studien vid KI och de två författarna till föreliggande studie inför annoteringen, varje enskild person annoterade således fyra ljudfiler. Interbedömarreliabilitet kontrollerades genom att tre ljudfiler utöver dessa annoterades av samtliga författare.

Tabell 2

Redovisning av de talavvikelse typer som användes för att klassificera föreliggande studies talares talavvikelser i språkljudsproduktionen. Beskrivningarna av talavvikelse typer, exempel och kommentarer är, där inget annat anges baserade på fonologiska förenklingsprocesser i Nettelblatt (2007).

	Avvikelse typ	Exempel	Kommentarer
Paradigmatiska processer	Frikativisering och Affrikativisering	Bil -> [βi:l]	
	Klusilering	Fot -> [pu:t]	
	Nasalering	Lejon -> [ˈneʝɔn]	
	Lateralisering	Banan -> [baˈla:l]	
	Avtoning/påtoning	Sked -> [ʃe:ɖ]	
	Försvagning	Öra -> [œ:ja]	
	Labialisering	Skidor -> [fi:dɔr]	
	Dentalisering	Kaka -> [ˈta:ta]	
	Palatalisering	Sova -> [çɔ:va]	
	Velarisering	Teve -> [ke:vɛ]	

	H-sering	Röd -> [hø:d]	
	Vokalsubstitution	Hus -> [hɪ:s]	
Syntagmatiska processer	Utelämning av obetonad stavelse	Ambulans -> [ha:s]	
	Utelämning av final konsonant	Båt -> [bo:]	
	Tillägg av dummykonsonant	Vante -> [ha:na:n]	
	Förenkling av konsonantförbindelse	Klocka -> [ˈkɔka]	
	Vokalepentes	Blomma -> [bøˈlɔma]	
	Neutralisering	Hus -> [hø:s]	
Avvikelse typer som lagts till i detta arbete, utöver de som beskrivs i Nettelbladt (2007)	Multipla processer*	Skulle -> [døle]	Konstruerat exempel
	Utelämning av initial konsonant	Fin -> [i:n]	Konstruerat exempel
	Övriga avvikelser**		

**Multipla processer*: Användes i de fall där det gick att tillämpa tre eller fler befintliga vedertagna förenklingsprocesser vid beskrivning av en avvikande språkljudsproduktion, till exempel vid ordet *skulle* blev /døle/, /sk/ blev alltså ett /d/. Här sker dels en förenkling av konsonantförbindelse men även dentalisering samt påtoning.

***Övriga talavvikelser*: Användes i de fall där talavvikelsen inte gick att beskriva med någon av de ovan nämnda talavvikelse typerna, t.ex. vid läspning (interdentalisering).

Etiska hänsynstaganden

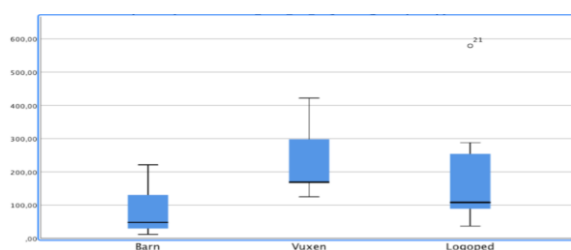
En godkänd etikansökan av Regionala etikprövningsnämnden i Stockholm fanns för det större projektet inom vilket föreliggande studie gjordes. Samtliga lyssnare och vårdnadshavare till de barn som deltog som lyssnare fick information om studien och gav skriftligt samtycke till att delta i studien. Samtyckesblanketterna förvarades i låst skåp och arkiverades därefter på KI. Lyssnarna i studien fick information om att de när som helst kunde avbryta bedömningen, utan att uppge någon anledning till detta. Lyssnarnas personuppgifter avidentifierades genom kodning inför bearbetning av data, det avidentifierade resultatet och demografisk data arkiverades efter databearbetning på KI. Talarna och deras föräldrar hade gett sitt medgivande till att ljudinspelningarna talarna medverkat i skulle få användas i lyssnarstudier. Lyssnarna hade ingen tystnadsplikt

gällande ljudfilerna, dock var talmaterialet redan avidentifierat när föreliggande studies författare fick ta del av det. Talmaterialet förvarades därefter i låst konto på den digitala förvaringstjänsten box.com samt under databearbetningen i de två författarnas privata datorer.

Statistisk analys

Resultat från ARS-bedömningen på samtliga ljudfiler användes för analys. Bland de dubblade ljudfilerna användes resultatet från det första lyssningstillfället förutom när intrabedömarreliabilitet hos lyssnarna analyserades. De statistiska beräkningarna gjordes i statistikprogrammet SPSS. Icke-parametriska test användes för att deltagarantalet var lågt samt att data inte var normalfördelad, om inget annat angetts.

Vid analysen av om de olika grupperna skiljde sig åt när det gällde grad av acceptabilitet, användes det icke-parametriska oberoende testet Kruskal-Wallis då det var tre olika gruppers resultat på samtliga ljudfiler som jämfördes. En tvåsidig prövning med beslutad signifikansnivå på 0,05 gjordes och därefter undersöktes resultaten i form av en parvis jämförelse. För att styrka resultatet i den parvisa jämförelsen användes även det icke-parametriska testet Mann-Whitney. Interbedömarreliabiliteten mellan samtliga lyssnarna som använde ARS-system undersöktes genom en analys av intraclass correlation (ICC). Vid beräkning av ICC användes *two-way model, absolute agreement* och *random effect* och konfidensintervallet sattes till 95%. Därefter analyserades skillnader i antal klick på varje ljudfil, mellan den lyssnare som i den enskilda ljudfilen hade högst antal klick och den lyssnare i samma ljudfil som hade lägst antal klick. Summan av skillnaderna i antal klick delades på antal ljudfiler, vilket resulterade i ett medelvärde som beräknades enskilt för varje grupp. Detta gjordes då studier av rådata visade stora skillnader mellan lyssnarna trots god ICC. Ett boxdiagram över lyssnarnas klick (se figur 2) visade på en stor spridning inom och mellan grupperna samt att det fanns en tydlig outlier i logopedgruppen.



Figur 2. Boxdiagram över spridning och median av klickfrekvens på gruppnivå.

Den manuella annoteringen kontrollerades genom en beräkning av interbedömarreliabiliteten mellan de fyra annoterarnas analys av antal annoterade talavvikelser. Antal annoteringar jämfördes i fyra ljudfiler med hjälp av punkt för punkt samstämmighet. En marginal på sex antal annoterade avvikelser tilläts för att räknas som överensstämmelse mellan annoterarna. Utifrån detta kunde andel (%) samstämmighet mellan de olika annoterarna presenteras. Inför beräkningen av vilka ljudfiler som fått mest klick av lyssnarna och därmed hade lägst grad av acceptabilitet, gjordes en normalisering av datan. Vid normaliseringen tilldelades alla lyssnare en lika stor "reaktions-pott" vilket var satt till värdet 100, denna pott fördelades över varje lyssnares totala antal klick över samtliga ljudfiler. Ett klick hos en lyssnare som klickade totalt 200 gånger sammantaget

på alla ljudfiler, blev på så sätt värt 0,5 “poäng”, medan ett klick hos en lyssnare som klickade totalt 50 gånger sammantaget på alla ljudfiler var värt 2 “poäng”. Eftersom normaliseringen resulterade i en sammanlagd siffra per ljudfil, behövdes lika många deltagare i varje grupp. Därför uteslöts en vuxen person vars yrke kan ha påverkat bedömningen och att denna lyssnare hade mycket hög klickfrekvens i jämförelse med resterande i gruppen samt en logoped vars resultat var en tydlig outlier (se figur 2). Endast de ljudfiler som spelats in i Göteborg användes till analys av vilka talavvikelser som var vanligast förekommande i tal med låg grad av acceptabilitet. Detta beslut togs efter undersökning av samstämmighet mellan annoterarna. De två ljudfiler som visade på lägst acceptabilitet studerades och mönster hos dessa ljudfiler i form av vilka talavvikelser som var vanligast enligt annoteringarna, analyserades. Därefter gjordes en jämförelse av hur dessa två ljudfiler förhöll sig till varandra samt till resterande ljudfiler angående talavvikelse typer och mängd talavvikelser.

För att undersöka validiteten hos ARS-systemet gjordes en korrelationsanalys mellan summorna av de vuxna naiva lyssnarnas klick och skattningspoäng på varje stimuli. Korrelationstestet utfördes med hjälp av det icke-parametriska testet Spearman's rangkorrelation. Samstämmighet mellan vuxna naiva lyssnare undersöktes genom en analys av ICC, separat för de två olika mätsätten. Vid beräkning av ICC användes *two-way model*, *absolute agreement* och *random effect*, konfidensintervallet sattes till 95 %.

Intrabedömarreliabiliteten mellan samtliga lyssnare i ARS-bedömningen, undersöktes med punkt för punkt samstämmighet beräknat på de 25% (n=4) av ljudfilerna som dubblerats. Procentsatsen valdes i linje med tidigare studier om acceptabilitet (Eadie et al., 2015). En marginal på ett klick tilläts för att räknas som överensstämmelse mellan första och andra lyssningen. Även analyser med marginaler på två respektive tre klick utfördes.

Resultat

Jämförelse av acceptabilitet mellan lyssnargrupper, vilket analyserades med ett Kruskal-Wallis-test, visade att de olika lyssnargruppernas bedömningar av acceptabilitet skiljde sig åt ($H(2)=7.14$, $p=0.028$). Vid en parvis jämförelse framkom att skillnaden var signifikant endast mellan de vuxna naiva lyssnarna ($M=243$, $s=110,46$) och barns ($M=81$, $s=80,52$) bedömningar ($p=0,023$), där de vuxna naiva lyssnarna hade en högre sammanlagd klickfrekvens, alltså lägre grad av acceptabilitet än barnen. Variationen hos de vuxna naiva lyssnarna låg mellan 125-422 klick per lyssnare för samtliga stimulin och för barnen låg variationen mellan 12-221 klick per lyssnare för samtliga stimulin. Vid en parvis jämförelse mellan barn och logopeder ($M=186$, $s=172,27$) ($p=0,43$) samt de vuxna naiva lyssnarna och logopeder ($p=0,64$), framkom ingen signifikant skillnad. Logopederna varierade mellan 37-579 klick per lyssnare för samtliga stimulin. Resultaten ovan av den parvisa jämförelsen blev därefter bekräftade genom ett Mann-Whitney-test, alltså endast signifikant skillnad mellan barn och vuxna naiva lyssnare påvisades ($U=10$, $n_1=8$, $n_2=9$, $p=0,012$ two-tailed). En god interbedömarreliabilitet visades mellan samtliga lyssnare (ICC (r)=0,91, $n=16$) enligt gränsvärden från Cicchetti (1994). Vid jämförelse av skillnader i antal klick på varje ljudfil, mellan den lyssnare som i den enskilda ljudfilen hade högst antal klick och den lyssnare i samma ljudfil som hade lägst antal klick, blev $M=15$ och $s=7,27$ för barn, för vuxna naiva lyssnare blev $M=22$ och $s=9,67$ samt för

logopederna blev $M=35$ och $s=19,70$, på samtliga ljudfiler. Skillnaden i antal klick på en ljudfil, var i barnens fall som minst 0 och som mest 26 klick, bland vuxna naiva lyssnare som minst 7 och som mest 37 klick och bland logopederna som minst 1 och som mest 64 klick.

Analys av vilka talavvikelse typer som var vanligast förekommande i tal med låg grad av acceptabilitet baserades på annoteringar. Interbedömarreliabiliteten för dessa bedömningar beräknades med punkt för punkt samstämmighet och procentsatsen analyserades i förhållande till gränsvärdet på 80% (Cordes, 1994). Samstämmigheten mellan samtliga annoterare visade på 42% överensstämmelse, vilket tydde på låg samstämmighet. Jämförelser av annoterarna (se tabell 3) sinsemellan hade annoterarna från Göteborg 75% överensstämmelse, vilket var det par som var närmast gränsvärdet för god samstämmighet av alla par. Resterande jämförelser av par tydde på låg samstämmighet.

Tabell 3

Interbedömarreliabilitet mellan de olika bedömarna.

Bedömare	GBG1	GBG2	STHLM1	STHLM2
GBG1	100%	75%	25%	25%
GBG2	75%	100%	50%	25%
STHLM1	25%	50%	100%	50%
STHLM2	25%	25%	50%	100%

Vid manuell kvalitativ analys av vilka olika talavvikelse som fanns representerade i de olika ljudfilerna framkom följande resultat (se tabell 4). De ljudfiler som visade sig ha lägst acceptabilitet genom flest antal klick, kännetecknades av att *dentalisering* förekom flertalet gånger i de fyra ljudfiler med lägst acceptabilitetsgrad, *avtoning/påtoning* förekom flertalet gånger i de tre ljudfiler med lägst acceptabilitetsgrad, *tillägg av dummykonsonant* förekom endast i den ljudfil som hade näst lägst acceptabilitetsgrad. *Klusilering* samt *utelämnning av initial konsonant* förekom flest gånger i den ljudfil som hade allra lägst acceptabilitetsgrad. *Förenkling av konsonantförbindelse* var vanligt förekommande i flera ljudfiler och förekom ofta i de ljudfiler med allra lägst och näst lägst acceptabilitetsgrad. Även *utelämnning av final konsonant* förekom i flera ljudfiler, men flest gånger i den fil som hade allra lägst acceptabilitetsgrad. Där förekom även utelämnning av initial konsonant, förenkling av konsonantförbindelse och utelämnning av final konsonant ofta i kombination med varandra. I detta talmaterial fanns därför en mycket begränsad mängd konsonanter. Den ljudfil som hade näst lägst acceptabilitetsgrad hade flest annoterade övriga talavvikelse i jämförelse med de andra ljudfilerna. Vid analys av dessa talavvikelse framkom att talaren hade flera talavvikelse på ett och samma ord vilket liknade multipla processer, men i dessa fall kunde talavvikelse inte placeras under någon kategori. Resultatet blev att dessa ord var mycket avvikande från målorden. De fyra ljudfiler där författarna noterade flest talavvikelse, är också de fyra ljudfiler där lyssnarna gjort mest klick och de två ljudfiler som innehöll typiskt tal är de ljudfiler där lyssnarna gjort minst klick.

Tabell 4

Översikt av antal talavvikelsestyper per stimulus samt grad av acceptabilitet hos stimuli utifrån normaliserad data.

Stimulus	Grad av acceptabilitet (1=lägst grad av acceptabilitet)	Totalt antal annoterade talavvikelser	Avvikelsestyp						
			D	A/P	Tillägg av DM	K	Utelämning av IK	Förenkling av KF	Utelämning av FK
G5	1	98	19	6	0	10	12	12	17
G4	2	98	16	19	8	5	1	10	6
G3	3	63	14	8	0	2	0	7	1
G6	4	118	17	2	0	0	0	18	11
G2	5	39	6	0	0	3	2	5	3
G7	6	41	4	1	0	0	0	6	0
G8	7	53	10	1	0	0	2	11	5
G1	8	4	0	0	0	0	0	0	2

Not. D=Dentalisering, A/P=avtoning/påtoning, DM=Dummykonsonant, K=Klusilering, IK= Initial konsonant, KF=Konsonantförbindelse, FK=Final konsonant.

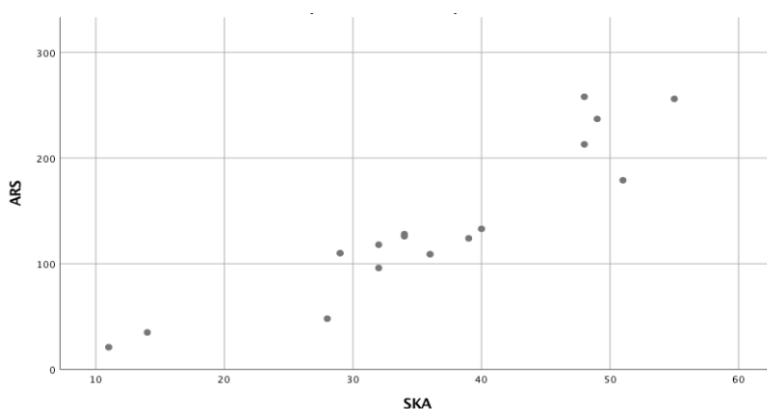
En översikt med antal talavvikelsestyper, vilket räknades utifrån talavvikelsens förekomst i samtliga ljudfiler från Göteborg beräknade av annoterare, sammanställdes (se tabell 5).

Tabell 5

Talavvikelsestyper samt summa av talavvikelser.

Avvikelsestyper	Summa av avvikelser
Dentalisering	86
Förenkling av konsonantförbindelse	69
Försvagning	58
Utelämning av final konsonant	45
Palatalisering	40
Avtoning/påtoning	37
Övriga avvikelser	28
Multipla processer	20
Vokalsubstitution	20
Klusilering	20
Utelämning av initial konsonant	17
Lateralisering	15
Nasalering	14
Labialisering	11
Tillägg av dummykonsonant	8
Velarisering	7
Utelämning av obetonad stavelse	6
Frikativisering och affrikativisering	6
Neutralisering och substitution av vokaler	4
H-sering	2
Vokalepentes	1
Totalsumma	514

För att undersöka validiteten i bedömningen med ARS-systemet vid acceptabilitet beräknades en korrelation mellan resultaten från ARS-systemet och skattningsformuläret. En stark positiv signifikant korrelation påvisades mellan resultaten ($r=0,91$, $p=0,00$, $n=16$) enligt gränsvärden från Cohen (1988). Sammanlagda poäng på ett stimuli för samtliga lyssnare vid skattningen varierade mellan 11-55 poäng ($M=36$, $s=12,46$) och antal sammanlagda klick på ett stimuli för samtliga lyssnare vid ARS-systemet varierade mellan 21-258 klick ($M=137$, $s=73,97$). Vid kontroll av interbedömarreliabilitet uppvisades god sådan i både gruppen av vuxna naiva lyssnare som använt ARS-system ($ICC(r)=0,89$, $n=16$) samt i skattningsgruppen ($ICC(r)=0,95$, $n=16$), enligt gränsvärden från Cicchetti (1994). De två ljudfiler som innehöll minst anoterade avvikelser och var typiskt tal visas som de två punkter längst ner i vänstra hörnet med minst klick enligt ARS, samt de fyra ljudfiler som innehöll flest antal anoteringar visas i högra hörnet med flest klick enligt ARS i scatter plot (se figur 3).



Figur 3. Scatter plot av korrelationen mellan resultaten från ARS-systemet och skattningarna.

Intrabedömarreliabiliteten mellan lyssnarnas klick vid första och andra tillfället för de fyra ljudfiler som dubblerats presenteras i tabell 6. Vid en marginal på ett klicks skillnad mellan de två bedömningarna hade barnen 53,1% punkt för punkt samstämmighet, de vuxna naiva lyssnarna 44,4 % punkt för punkt samstämmighet och logopederna 44,4 % punkt för punkt samstämmighet. Intrabedömarreliabiliteten sammantaget för alla grupper var 47,1 % punkt för punkt samstämmighet. Enligt Cordes (1994) visar 80% samstämmighet på en god reliabilitet vilket är betydligt högre än vad som framkom här. Ingen tydlig trend i förändring av klickfrekvens kunde ses mellan första och andra lyssningstillfället.

Tabell 6

Punkt för punkt samstämmighet vid en marginal på ett, två och tre klick.

Marginal	Barn	Vuxna naiva lyssnare	Logoped	Hela gruppen
1 klick	53,1 %	44,4 %	44,4 %	47,1 %
2 klick	71,9 %	58,3 %	72,2 %	67,3 %
3 klick	75 %	75 %	80,6 %	76,9 %

Diskussion

Syftet med föreliggande studie var huvudsakligen att undersöka hur barn mellan 6-10 år, vuxna naiva lyssnare och logopederna skiljde sig åt i grad av acceptabilitet avseende barn med SSD samt att undersöka vilka typer av fonologiska talavvikelser som är vanligast förekommande i tal med låg grad av acceptabilitet. Ett ytterligare syfte var att undersöka om ARS-systemet är en valid och reliabel metod för att mäta acceptabilitet hos barn med SSD. Resultatet visade en signifikant skillnad mellan barn och de vuxna naiva lyssnarnas bedömningar samt att dentalisering, avtoning/påtoning, tillägg av dummykonsonant, klusilering samt tal där det fanns en mycket begränsad mängd konsonanter var de talavvikelser som var vanligast förekommande i tal med låg grad av acceptabilitet. Resultatet visade att ARS-systemet förefaller ha en ganska god validitet och måttligt god reliabilitet vid bedömning av acceptabilitet hos dessa lyssnare och tal hos barn med SSD. Metodens lämplighet i logopedisk forskning och klinisk verksamhet diskuteras nedan.

Resultatet från jämförelsen av de olika gruppernas klickfrekvens visade att barnen överlag hade betydligt färre klick än vad de vuxna naiva lyssnarna och logopederna hade. Barnen tyckte möjligen inte att talet lät särskilt *konstigt* eller att de hade svårt att bibehålla koncentrationen. Vad gäller skillnad mellan logopederna och vuxna naiva lyssnarnas grad av acceptabilitet visade föreliggande studie att ingen signifikant skillnad förekom. De vuxna naiva lyssnarna hade enligt analys av rådatan högre klickfrekvens än logopederna, vilket kan tyda på att de vuxna naiva lyssnarna ändå hade något lägre grad av acceptabilitet, men det går inte utesluta att resultatet är slumpmässigt. Logopederna var troligen mer vana vid talavvikelser än de vuxna naiva lyssnarna, möjligen tyckte logopederna därför inte tal med SSD lät lika *konstigt* som vuxna naiva lyssnare tyckte. Resultat som både kan bekräfta och ifrågasätta denna hypotes har presenterats av Finizia et al. (1998) där man jämförde grad av acceptabilitet hos vuxna naiva lyssnare och logopederna. Resultatet visade att vuxna naiva lyssnare hade signifikant lägre acceptans till talet hos laryngektomerade personer än logopederna, men ingen skillnad uppkom mellan grupperna vad gällde tal hos strålade patienter. Det är inte helt lämpligt att jämföra föreliggande studie med studien av Finizia et al. (1998) då talarna där inte var barn med SSD och acceptabilitet mättes genom användning av en Visuellt analog skala (VAS). En intressant aspekt är att lyssnare i allmänhet möjligen har högre tolerans gentemot talavvikelser i barns tal, eftersom man vet att barn genomgår en utvecklingsfas, än gentemot tal av en vuxen där man möjligen förväntar sig ett mer typiskt tal. I föreliggande studie fick lyssnarna ingen information om att talarna hade talavvikelser, däremot visste lyssnarna i Finizia et al. (1998) att det handlade om tal från laryngektomerade och strålade patienter. Att som lyssnare ha vetskap om att en talare har avvikande tal skulle möjligen kunna bidra till högre tolerans gentemot talet. Föreliggande studies resultat är liknande det som visats enligt Contardo et al. (2014) där logopederna och vuxna naiva lyssnare fick markera svårförståeligt tal hos barn med SSD med hjälp av webbaserad ARS. Det påvisades ingen skillnad mellan gruppernas förståelighet. Påpekas bör att variablerna förståelighet och acceptabilitet skiljer sig åt, då förståelighet är ett mer välkänt och tydligt begrepp än vad acceptabilitet och begreppet *konstigt* är. Tal som anses låta *konstigt* kan vara förståeligt, men det är troligen så att man även upplever oförståeligt tal som *konstigt*.

Vad gäller attityd i förhållande till acceptabilitet visade resultatet i studien av Crowe (1987) att lärarstudenterna hade mer positiv attityd gentemot barn med artikulatoriska

talavvikelser (eng. articulatory errors) än vad både fjärde- och sjätteklassare hade. Intressant är att resultatet skiljer sig från vad föreliggande studies resultat visade, där de vuxna naiva lyssnarna hade signifikant lägre grad av acceptabilitet än barnen. Utifrån detta kan man ifrågasätta om låg grad av acceptabilitet av tal leder till en negativ attityd gentemot talaren. Viktigt att påpeka är att lärarstudenterna gick en kurs rörande talets funktioner och lärares ansvar, de kan således inte jämföras med föreliggande studies vuxna naiva lyssnare, vilka hade blandade erfarenheter av barns tal. Resultatet i Crowe (1987) visade även att lyssnare som gick i sjätte klass var mer negativa till artikulatoriska talavvikelser än vad elever i fjärde klass var. Enligt Crowe (1987) verkar det finnas en övergångsålder där äldre barn successivt får en mer negativ attityd gentemot avvikande tal tills detta senare stabiliseras. De äldre barnen förväntade sig möjligen mer utav talaren och var mer medvetna om hur talet skulle låta. Till viss del kan detta även ses i föreliggande studie, där det yngsta barnet hade lägst klickfrekvens samt där två av barnen som var nio respektive tio år hade märkbart högre klickfrekvens än övriga barn som var i åldern 6-9 år. Dock var fyra av samtliga barn som deltog i föreliggande studie nio år, därav bör man inte dra några större slutsatser kring detta fenomen. Utifrån att resultatet i Crowes (1987) studie och föreliggande studie ger ett liknande resultat vad gäller acceptabilitet kontra attityd hos barnen, kan man ändå tänka sig att låg grad av acceptabilitet av tal kan leda till en negativ attityd gentemot talaren.

Analysen med ICC, som visade en god samstämmighet mellan lyssnarna som använt ARS-systemet, kan ifrågasättas då det är ett parametriskt test. Ifrågasättandet bekräftades av att samstämmigheten visades vara ganska låg baserat på jämförelser av skillnader i antal klick på varje ljudfil, mellan den lyssnare som i den enskilda ljudfilen hade högst antal klick och den lyssnare i samma ljudfil som hade lägst antal klick. En positiv aspekt rörande den låga samstämmigheten är att mätmetoderna tycks lämna utrymme för individuell spridning, vilket är en god grund för att mäta skillnader på individnivå. Negativt med den låga samstämmigheten är att det är en stor skillnad inom grupperna, vilket gör det svårt att dra slutsatser på gruppnivå. Det kan även ifrågasättas om lyssnarna förstod instruktionerna på samma sätt kring huruvida man skulle klicka flera gånger i ett långt yttrande eller inte.

De talavvikelser som enligt föreliggande studie var vanligast förekommande i tal med låg grad av acceptabilitet var andra talavvikelser än vad Nyberg et al. (2016) presenterade utifrån sin studie där talavvikelserna velarisering och lateralisering eller interdentalisering av /s/ var de talavvikelser barn reagerade mest på. Dessa talavvikelser fanns representerade i föreliggande studies talmaterial men var ej de mest framträdande i talmaterialet med lägst acceptabilitet. Det kan dock skilja en hel del mellan vilka typer av talavvikelser som uppstår vid LKG och SSD. Då barnens åsikter i studien av Nyberg et al. (2016) framkom via intervjuer skiljer sig även metoden från föreliggande studies metod. Vidare bestod lyssnarna i föreliggande studie förutom av barn, även av vuxna naiva lyssnare och logoped. I föreliggande studie användes inte ARS-systemets KDE-analys som kan visa tidpunkter i talmaterialet där klick gjordes. Detta då det enligt S. Strömbergsson (personlig kommunikation, 13 mars 2018), medförfattare till studien av Contardo et al. (2014), har visat sig uppstå slumpmässiga resultat då KDE-analys använts i tidigare studier. Istället utfördes i föreliggande studie en manuell kvantitativ analys av data, likt analysen i Hodson och Padens (1981) studie. Metoden kan dock problematiseras då man inte kan dra slutsatsen att de vanligast förekommande talavvikelserna i ljudfilerna

är just de som påverkat talet mest, då det finns fler tillfällen att producera vissa språkljud än andra och därav fler tillfällen att producera vissa talavvikelser (Hodson & Paden, 1983). Eftersom föreliggande studies talmaterial bestod av spontantal, var effekten svår att undvika. Resultatet i undersökningen av vilka talavvikelser lyssnarna uppfattade som mest avvikande blev även beroende av vilka talavvikelser som förekom i talmaterialet, även om talmaterialet noggrant valts ut med hänsyn till studiens ändamål. Något som trots dessa brister kan göra analysen godtagbar är att de talavvikelser som till antalet är flest sammantaget, inte är samma talavvikelser som tycks vara vanligast förekommande i de ljudfiler där klickfrekvensen var högst. Inför analys av vilka talavvikelser som förekom i ljudfilerna med lägst acceptabilitetsgrad utfördes en normalisering av datan. Positivt med detta var att den individuella variationen kring hur generös man var med sitt klickande jämnades ut. Negativt var att om lyssnarna skiftade i sitt sätt att klicka på grund av bristande koncentration eller förändrad referensram inför acceptabilitet, verkade normaliseringen ändå på samma sätt genomgående under dessa förändringar.

Interbedömarreliabiliteten mellan annoterarna i Göteborg visades vara ganska god enligt analys med punkt för punkt samstämmighet, låg samstämmighet visades mellan resterande annoterare. Därför beslutades att analysen av vilka talavvikelser som fanns representerade i tal med låg grad av acceptabilitet, enbart skulle baseras på de ljudfiler annoterarna från Göteborg annoterat. Resultatet kan ha påverkats av att riktlinjerna för hur annoteringarna skulle utföras ändrades under arbetets gång, vilket var nödvändigt inför mätning av samstämmigheten. Troligt är även att instruktionerna kring hur talavvikelser skulle annoteras uppfattats olika mellan annoterarna, särskilt mellan annoterarna i Göteborg och Stockholm, då möjligheten var större att kontinuerligt diskutera samstämmighet med den person en annoterare utförde sin studie ihop med.

Det var av stor vikt att föreliggande studie utvärderade ARS-systemets validitet och reliabilitet vid mätning av acceptabilitet, då mätinstrumentet redan använts i flera logopediska studier avseende andra logopediska variabler. Föreliggande studie valde att jämföra två instrument som mäter samma variabel, det vill säga subjektiv acceptabilitet. Utifrån den starka korrelation som påvisades mellan resultatet från ARS-systemet och skattningen, vid undersökning med vuxna naiva lyssnares acceptabilitet, kan man dra slutsatsen att det finns en ganska god validitet hos ARS-systemet som mätinstrument, för just denna grupp av lyssnare och barn med SSD. Ett problem vid jämförelsen mellan mätinstrumenten är deras olikheter, då ett ARS-system mäter acceptabilitet dikotomt och ett skattningsformulär mäter helhetsintryck på en gradskala. För att överbrygga denna skillnad hanterades ARS-systemets insamlade data som en samlad klickfrekvens för en hel ljudfil, således blev ARS-systemets data ett samlat helhetsintryck. Skillnaden i mätsätten kvarstår på så sätt att gradering utifrån skattningsformulär sker då en deltagare lyssnat till hela ljudfilen medan bedömning med ARS-systemet sker steg för steg. Risk finns att klickerna via ARS-systemet gjorts slumpmässigt, något som diskuteras i studien av Strömbergsson och Tännander (2013) där lyssnare fick markera svårförståeligt tal med hjälp av ARS-system. Studien rapporterar att flera lyssnare likt i föreliggande studie upplevde det vara svårt att använda ARS-system korrekt och tyckte sig ha klickat slumpmässigt. Dock visades tydliga strukturerade mönster vid ARS-analysen i studien av Strömbergsson och Tännander (2013), vilket det även gjorde i föreliggande studie. Dessa mönster tyder på att klickerna troligen inte varit helt slumpmässiga, liksom det faktum att föreliggande studies talmaterial med typiskt tal hade betydligt lägre klickfrekvens än det

talmaterial påverkat av SSD samt att de fyra ljudfiler med högst klickfrekvens var det talmaterial som hade flest annoterade talavvikelser. Detta resultat kan sägas visa att ARS-systemet har god validitet då det kommer till mätning av acceptabilitet och barn med SSD. Tecknet på en ganska god validitet hos ARS-systemet som uppstått i föreliggande studie kan sägas bekräfta de goda resultaten från tidigare nämnda studier av Contardo et al. (2014) och av Johansson et al. (2017) men detta bör ses med stor försiktighet. Ett ifrågasättande av ARS-systemets validitet, rörande att mäta acceptabilitet på det sätt som gjorts i föreliggande studie, är huruvida måttet klickfrekvens verkligen kan sägas motsvara grad av acceptabilitet.

Skattningen i föreliggande studie skedde på en intervallskala. Att det var flera skalsteg kan ha varit negativt då det troligen var svårt för deltagarna att avgöra skillnaden mellan stegen och risken finns att samstämmigheten hos lyssnarna blev lidande, då de kan ha tolkat skalstegen olika. Dock sågs en god interbedömarreliabilitet mellan dessa lyssnare vid analys med ICC. Samstämmigheten kan dock ifrågasättas då ICC är ett parametriskt test och skalan inte ger möjlighet till stor spridning mellan lyssnarna. Skattning vid acceptabilitet är väl beprövat (Bagnall & David, 1988; Dagenais et al., 2006; Eadie & Doyle, 2004; Trudeau, 1987). I en studie av Stevens, Bressmann, Gong och Tompson (2011) användes en 4-gradig intervallskala vid mätning av acceptabilitet av tal, där varje skalsteg hade en beskrivning; *0, normal speech acceptability; 1, speech acceptability mildly affected; 2, speech acceptability moderately affected; and 3, speech acceptability severely affected.* Detta sätt att gradera varje steg, skulle kunnat vara till hjälp för både enskild tolkning och allmän samstämmighet hos lyssnarna. På grund av den större mängden skalsteg i föreliggande studie är det inte säkert att sådana beskrivningar skulle underlättat för lyssnarna, men troligen hade det varit hjälpsamt med en beskrivning på vartannat skalsteg. Positivt med ARS-systemet är att det inte finns någon övre gräns för hur mycket man kan klicka, spridning kan således tydligt uppträda mellan individer. Vid skattning erbjuds lyssnare utlåta sig om ett helhetsintryck, där den sammanlagda bilden av såväl mängd som typ av talavvikelser ingår. Således kan de olika mätsätten sägas komplettera varandra. Det kan vara passande att använda sig av ARS-system vid markeringar av enskilda händelser, i sammanhang där syftet inte är att gradera dessa enskilda fenomen eller att gradera helhetsupplevelsen. För sammanhang där sådana graderingar är önskvärda, kan en skala vara ett mer passande mätinstrument.

Vid analys av ARS-systemets reliabilitet undersöktes lyssnarnas, som använt ARS-system, intrabedömarreliabilitet. Resultatet visade på en mycket låg sådan på såväl gruppnivå som sammantaget för alla lyssnare som använt ARS-systemet. Utifrån detta kan man ifrågasätta tillförlitligheten hos ARS-systemet som mätinstrument, på det sätt det använts i föreliggande studie. Vid beräkning av punkt för punkt samstämmighet hos dessa lyssnare tilläts endast en marginal på ett klick mellan första och andra lyssningen eftersom antal klick hos en individ ibland var mycket lågt och i vissa fall kunde vara noll klick på ett stimuli. Man bör dock påpeka att en marginal på ett klick är en sträng marginal, då det intuitivt inte är något anmärkningsvärt om det skiljer ett par klick mellan första och andra klicktillfället på samma stimuli, med tanke på att en del har över 30 klick på de dubblade ljudfilerna. Vid två klick som marginal ökade punkt för punkt samstämmighet betydligt inom varje grupp. Vid tre klick som marginal visade logopederna på god reliabilitet och barnen samt de vuxna naiva lyssnarnas samstämmighet låg strax under 80%, vilket tyder på att även de hade en ganska god

reliabilitet. Eftersom ljudfilerna är en minut långa och bygger på subjektiv bedömning är det inte oväntat att resultaten skilde sig från gång till gång, eftersom referensramen kan ha förändrats allteftersom lyssnarna hörde fler stimuli. Edlund et al. (2013) fastslog i sin studie att ARS-systemet hade god reliabilitet vad gällde att lyssnare klickade på samma tidpunkter, då de lyssnade på toner eller på en talad mening där de ombads klicka när meningen var slut. I det fallet fick lyssnarna veta vad de skulle få höra och blev ombuds att klicka vid en viss tydlig händelse, vilket inte var fallet i föreliggande studie. Eftersom föreliggande studie och studien av Edlund et al. (2013) mäter olika variabler, kan inte studierna bekräfta varandras resultat angående ARS-systemets reliabilitet.

En svårighet rörande föreliggande studies undersökning av reaktioner på talmaterialet, var att lyssnarna vid bedömningen inte hade någon gemensam given referensram rörande vad som var *normalt* respektive *mycket konstigt* tal. Således blev den första ljudfilen deltagarna fick lyssna på en slags referensram. Flera av deltagarna som skattade acceptabilitet ville ändra tidigare skattningar när de fått lyssna till fler ljudfiler, vilket inte var möjligt. Effekten detta kan ha gett på resultatet kunde i viss mån minskas av att ordningen på ljudfilerna varierades vid lyssningarna. Varje lyssnare hade sin egen referensram kring vad de uppfattade som *konstigt* och huvudmotivet till att instruktionen var så ospecificerad, var att den skulle vara enkel och kunna förstås på ett intuitivt sätt av lyssnare utan skolning. En oundviklig konsekvens av detta var att varje lyssnare lämnades till att själva definiera och tolka begreppet *konstigt*, och att ingen inblick i varje lyssnares tolkning var möjlig. Det är en begränsning i metoden, vilket gör att den inte kan användas till att förstå varför varje lyssnare klickar. Resultatet kan även ha påverkats av att vissa talare hade språkliga talavvikelser såsom grammatiska talavvikelser, vilket troligen har tagits med i lyssnarnas bedömningar men inte annoterats vid annoteringen av talavvikelser. Dock har det talmaterial som analyserats mycket få grammatiska avvikelser, detta har därför troligen inte påverkat resultatet i sin helhet. En svaghet i jämförelsen mellan ARS-bedömningarna och skattningarna av acceptabilitet var att instruktionerna skiljde sig åt. Dock gjordes ett förtydligande av instruktionerna vid pilottestningen i Stockholm, vilket gjorde att misstag troligen undveks vid ARS-bedömningarna i föreliggande studie. Vikten av att alla bedömningar inom det större projektet bör ha haft samma instruktion för att garantera god kvalitet i analyser, ansågs vara av stor betydelse.

Det bör i framtiden undersökas om tal med låg grad av acceptabilitet i sin tur påverkar omgivningens beteende gentemot talaren, för att kunna förebygga negativt beteende. Om det finns ett kausalt samband mellan låg självbild hos personer med talavvikelser och omgivningens reaktioner, bör också undersökas. Rörande vilka talavvikelser som tycks vara vanligt förekommande i tal med låg grad av acceptabilitet, vore det önskvärt att i framtida forskning jämföra föreliggande studies resultat med en studie av vilka talavvikelser barn med SSD själva upplever vara mest socialt stigmatiserande samt vilka specifika talavvikelser vid SSD som olika lyssnare reagerar mest på. Detta för att kunna ge goda riktlinjer för var fokus bör ligga inom logopedisk behandling av barn med SSD. Angående användningen av ARS-systemet som mätinstrument vid bedömning av acceptabilitet, bör denna fortsätta utvärderas i kommande forskning. Detta för att veta hur tillförlitliga resultaten är vid sådana mätningar, då instrumentet kan betraktas som ett ganska nytt och obeprövat mätinstrument inom bedömning av acceptabilitet. Eftersom ARS-system kräver mycket teknisk utrustning såsom analysprogram, kontroller samt

förberedelser och efteranalys, kan systemet ses som något osmidigt att använda inom klinisk logopedi och forskning för mätning av acceptabilitet. En förhoppning är att KDE-analysen hos ARS-systemet ska kunna börja användas inom en snar framtid, då skulle ARS-systemet kunna användas i undersökningar där exakta tidsanalyser krävs. I logopedisk klinisk verksamhet skulle ARS-systemet kunna innebära ett sätt att undvika att logopeder använder invanda skattningsskalor för att bedöma acceptabilitet. Logopeder skulle kunna markera avvikelser med ARS-system i en patients inspelade tal som därefter kan analyseras. Ännu ett potentiellt användningsområde för ARS-systemet inom klinisk logopedi är vid utförandet av standardiserade test vid mätning av antal avvikelser, såsom PCC och Percentage of syllables stuttered, med ARS-systemets analys skulle man snabbt få en överblick av antal markerade avvikelser.

Sammanfattningsvis kan följande konstateras: Hur acceptabilitet av tal varierar beroende på lyssnare tycks vara ganska utforskat. Skillnaden som framkom mellan barn och vuxna naiva lyssnare vad gällde grad av acceptabilitet bör tolkas med försiktighet då resultatet skulle behöva stärkas av ytterligare studier. Ett skäl att fortsätta utforska olika lyssnares grad av acceptabilitet av tal, är att attityd gentemot personer med SSD potentiellt skulle kunna formas utifrån acceptabilitet, huruvida det finns ett kausalt samband mellan dessa bör också undersökas. I föreliggande studie presenterades ett antal talavvikelser som var vanligt förekommande i tal med låg grad av acceptabilitet, det behövs dock fler studier som kan bekräfta dessa resultat. Önskvärt är även att denna forskningsfråga i framtiden kan studeras med metoder som ger tydligare resultat. ARS-systemet har i ett fåtal tidigare studier utvärderats rörande mätningar inom olika logopediska forskningsfält, tydligt blir att detta mätinstrument bör utvärderas angående varje ny variabel man anser sig vilja mäta. ARS-systemet har potential för att komma till användning i såväl klinisk verksamhet som inom forskning, så länge det utvärderas utifrån dess tänkta uppgift.

Referenser:

- Alcorn, S., Griffith, S., & Miner, L. E. (1974). Comparison of articulation severity ratings of /s/ and /r/ by lower-, middle- and upper socioeconomic groups. *Journal of Communication Disorders*, 7, 79-87. doi: 10.1016/0021-9924(74)90009-4
- Bagnall, A. D., & David, D. J. (1988). Speech results of cleft palate surgery: Two methods of assessment. *British Journal of Plastic Surgery*, 41(5), 488-495. doi:10.1016/0007-1226(88)90005-7
- Burroughs, E. L., & Tomblin, J. B. (1990). Speech and language correlates of adult judgments of children. *Journal of Speech and Hearing Disorders*, 55, 485-494. doi:10.1044/jshd.5503.485
- Cicchetti, D. V. (1994). Guidelines, criteria, and rules of thumb for evaluating normed and standardized assessment instruments in psychology. *Psychological Assessment*, 6(4), 284-290. doi: 10.1037/1040-3590.6.4.284
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences (2nd ed)*. Hillsdale, N.J: L. Erlbaum Associates.
- Contardo, I., McAllister, A., & Strömbergsson, S. (2014). Real-time registration of listener reactions to unintelligibility in misarticulated child speech. *Proceedings from FONETIK 2014. Swedish Phonetics Conference, June 9-11 2014* (ss. 127-132). Stockholm, Sweden.
- Cordes, A. K. (1994). The reliability of observational data: I. Theories and methods for speech-language pathology. *Journal of Speech and Hearing Research*, 37(2), 264-278.
- Crowe, B. (1987). *Attitudes of adults and children toward children with mild articulation disorders* (ProQuest Dissertations and Theses). The Ohio State University. Tillgänglig: <https://search-proquest.com.ezproxy.ub.gu.se/docview/303619192?accountid=11162>
- Dagenais, P.A., Adlington L.M., & Evans, K.J. (2011). Intelligibility, comprehensibility, and acceptability of dysarthric speech by older and younger listeners. *Journal of Medical Speech-Language Pathology*, 19(4), 37-48.
- Dagenais, P. A., Brown, G. R., & Moore, R. E. (2006). Speech rate effects upon intelligibility and acceptability of dysarthric speech. *Clinical Linguistics & Phonetics*, 20(2-3), 141. doi: 10.1080/02699200400026843
- Diagnostic and statistical manual of mental disorders: DSM-5*. (2013). Arlington, Va: American Psychiatric Association, cop.
- Eadie, T. L., & Doyle, P. C. (2004). Auditory-perceptual scaling and quality of life in tracheoesophageal speakers. *Laryngoscope*, 114(4), 753-759. doi:10.1097/00005537-200404000-00030
- Eadie, T. L., Rajabzadeh, R., Isetti, D. D., Nevdahl, M. T., & Baylor, C. R. (2017). The effect of information and severity on perception of speakers with adductor spasmodic dysphonia. *American Journal of Speech-Language Pathology*, 26(5), 327-341. doi: 10.1044/2016_AJSLP-15-0191
- Eadie, T., Otero, D., Cox, S., Johnson, J., Baylor, C., Yorkston, K., & Doyle, P. (2015). The relationship between communicative participation and postlaryngectomy speech outcomes. *Head & Neck*, 38(S1), E1955-E1961. doi: 10.1002/hed.24353
- Edlund, J., Al Moubayed, S., Tännander, C. & Gustafson, J. (2013). Temporal precision and reliability of audience response system based annotation. *Proc. of multimodal Corpora 2013*. Edinburgh, UK

- ELAN (Version 5.1) [Computer software]. Nijmegen: Max Planck Institute for Psycholinguistics. Hämtad från <https://tla.mpi.nl/tools/tla-tools/elan/>
- Ellis, L. W., Spiegel, B., & Benjamin, B. (2002). Effects of speakers augmented characteristics and listeners sex on intelligibility and acceptability of synthesized speech'. *Perceptual and Motor Skills*, *94*, 1081-1088. doi: 10.2466/pms.2002.94.3c.1081
- Finizia, C., Lindström, J., & Dotevall, H. (1998). Intelligibility and perceptual ratings after treatment for laryngeal cancer: Laryngectomy versus radiotherapy. *Laryngoscope*, *108*(1), 138-143.
- Henningsson, G., Kuehn, D. P., Sell, D., Sweeney, T., Trost-Cardamone, J. E. & Whitehill, T. L. (2008). 'Universal parameters for reporting speech outcomes in individuals with cleft palate'. *Cleft Palate Craniofac Journal*, *45*(1), 1-17. doi: 10.1597/06-086.1
- Hodson, B. W. & Paden, E.P. (1981). Phonological processes which characterize unintelligible and intelligible speech in early childhood. *Journal of Speech and Hearing Disorders*, *46*, 369-373. doi: 10.1044/jshd.4604.369
- Hodson, B. W. & Paden, E. P. (1983). *Targeting intelligible speech: A phonological approach to remediation*. London: College-Hill Press
- Johansson, K., Strömbergsson, S., Robieux, C. & McAllister, A. (2017). Perceptual detection of subtle dysphonic traits in individuals with cervical spinal cord injury using an audience response systems approach. *Journal of Voice*, *31*(1), 126.e7-126.e17. doi: 10.1016/j.jvoice.2015.12.015
- Liu, H., Wan, M. & Wang, S. (2005). Features of listeners affecting the perceptions of mandarin electrolaryngeal speech. *Folia Phoniatrica Et Logopaedica*, *57*(1), 9-19. doi:10.1159/000081957
- The Language Archive, Max Planck Institute for Psycholinguistics. (2017). *How-to guide*. Hämtad 180122 från <https://tla.mpi.nl/tools/tla-tools/elan/>
- McGrath, L., Pennington, B., Willcutt, E., Boada, R., Shriberg, L., & Smith, S. (2007). Gene × Environment interactions in speech sound disorder predict language and preliteracy outcomes. *Development and Psychopathology*, *19*(4), 1047-1072. doi: 10.1017/S0954579407000533
- Moreno, J., Niklasson, E. (2016) *Samband mellan kommunikativ delaktighet, grad av stamning, förståelighet och acceptabilitet hos vuxna personer som stammar* (Magisteruppsats). Göteborg: Institutionen för neurovetenskap och fysiologi, Göteborgs universitet. Tillgänglig: <http://hdl.handle.net/2077/47526>
- Turban, J., (2009). The audience response system: a modality for course evaluation. *Medical Education*, *43*(5), 488 -489. doi:10.1111/j.1365-2923.2009.03348.x
- Nettelbladt, U. (2007). Fonologiska problem hos barn med språkstörning. I U. Nettelbladt & E.-K. Salameh (Red.), *Språkutveckling och språkstörning hos barn: Del I - Fonologi, grammatik, lexikon* (s. 95-134). Lund: Studentlitteratur.
- Nyberg, J. & Havstam, C. (2016). 'Speech in 10-year-olds born with cleft lip and palate: What do peers say?'. *Cleft Palate Craniofac Journal*, *53*(5), 516-526. doi: 10.1597/15-140
- Ode, C., & Öster Cattu Alves, M. (2017) *Pingu och PSC: språkljudsproduktion hos barn med språkljudsstörning vid fyra olika taluppgifter* (Magisteruppsats). Uppsala: Institutionen för neurovetenskap - enheten för logopedi, Uppsala Universitet. Tillgänglig: <https://uu.diva-portal.org/smash/get/diva2:1177156/FULLTEXT01.pdf>

- Overby, M., Carrell, T., & Bernthal, J. (2007). Teachers' perceptions of students with speech sound disorders: A quantitative and qualitative analysis. *Language, Speech & Hearing Services in Schools, 38*(4), 327-41.
- Stevens, K., Bressmann, T., Gong, S., & Tompson, B.D. (2011). The impact of a rapid palatal expander on speech articulation. *American Journal of Orthodontics & Dentofacial Orthopedics, 140*(2), E67-75. doi: 10.1016/j.ajodo.2011.02.017
- Strömbergsson, S. & Tännander, C. (2013). Correlates to intelligibility in deviant child speech – comparing clinical evaluations to audience response system-based evaluations by untrained listeners. *Proceedings of the Annual Conference of the International Speech Communication Association. INTERSPEECH, August 25-29 2013* (ss. 3717-3721). Lyon, France.
- Trudeau, M. D. (1987). A comparison of the speech acceptability of good and excellent esophageal and tracheoesophageal speakers. *Journal of Communication Disorders, 20*(1), 41-49. doi:10.1016/0021-9924(87)90042-6.
- Waring, R., & Knight, R. (2013). How should children with speech sound disorders be classified? A review and critical evaluation of current classification systems. *International Journal of Language & Communication Disorders, 48*(1), 25-40. doi: 10.1111/j.1460-6984.2012.00195.x