



Eira Troneng Akselberg er logoped MNLL og spesialpedagog i Mølleplassen spesialbarnehage.

eira.akselberg@gmail.com



Jannicke Karlsen er logoped MNLL og jobber i Statped sørøst, Avdeling språk/tale. Hun jobbet som postdoktor ved Institutt for spesialpedagogikk, Universitetet i Oslo, da arbeidet med artikkelen startet.

jannicke.karlsen@statped.no

HVOR GODT KAN CELF-4 OG CCC-2 IDENTIFISERE UTVIKLINGSMESSIGE SPRÅKFORSTYRRELSER?

Sammendrag

Vi vet lite om hvor nøyaktig norske kartleggingsverktøy identifiserer hvilke barn som har, og hvilke barn som ikke har, språkforstyrrelser. I denne studien undersøkte vi sensitiviteten og spesifisiteten til språktesten CELF-4 og spørreskjemaet CCC-2. Utvalget besto av 21 barn i alderen 5 til 12 år med påviste utviklingsmessige språkforstyrrelser og 21 barn med typisk utvikling som var parvis matchet med barna med språkforstyrrelser på alder, kjønn og nonverbal IQ. Resultatene viste at CELF-4 møtte anbefalte verdier for både spesifisitet og sensitivitet, mens CCC-2 kun identifiserte to tredeler av deltakerne med utviklingsmessige språkforstyrrelser. En justering av terskelverdien for CCC-2 økte sensitiviteten betydelig uten å påvirke spesifisiteten. Den skandinaviske versjonen av CELF-4 ser derfor ut til å være godt egnet til identifisering av utviklingsmessige språkforstyrrelser i en norsk kontekst. Våre resultater tyder også på at CCC-2 ikke bør brukes alene til å avgjøre om barn med mistanke om utviklingsmessige språkforstyrrelser skal utredes videre.

Summary

There is a lack of evidence regarding how accurately Norwegian assessment instruments identify children with developmental language disorder. In the current study, we examined the sensitivity and specificity of the language test CELF-4 and the questionnaire CCC-2. The sample included 21 children aged 5 to 12 years with clinically verified DLD and 21 typically developing children who were pair-wise matched with the DLD sample on age, gender and nonverbal IQ. The results indicated that the CELF-4 met the current recommendations for specificity and sensitivity, while the CCC-2 only identified two thirds of the children with DLD. However, adjusting the cut-off scores for the CCC-2 increased the sensitivity without affecting the specificity in our sample. We conclude that the Scandinavian version of the CELF-4 appears to be well-suited for identifying DLD in Norwegian-speaking children. Our results also suggest that the CCC-2 is not sensitive enough to determine whether further assessment is necessary for children with suspected DLD.

Introduksjon

Mange barn og unge med språkforstyrrelser vil oppleve vedvarende vansker med språk og lesing (Tomblin et al., 2003), noe som blant annet kan påvirke muligheten for å lykkes sosialt (Mok et al., 2014) og med skole og utdanning (Conti-Ramsden et al., 2018). For å kunne sette inn tilpassede og målrettede tiltak som kan bedre mulighetene til barn og unge med språkforstyrrelser vil identifisering av vansken være et viktig første skritt. Fra internasjonale studier vet vi at mange anerkjente språktester ikke er sensitive nok, slik at barn med språkforstyrrelser risikerer å bli feilvurdert (Bishop et al., 2016;

Bishop & McDonald, 2009). I Norge har vi få kartleggingsverktøy som kan identifisere språkforstyrrelser. I tillegg vet vi lite om spesifisitet og sensitivitet for disse verktøyene. Vi vet med andre ord ikke om de mest kjente kartleggingsverktøyene gjør oss i stand til å identifisere hvilke barn som har språkforstyrrelser og hvilke barn som ikke har det. I denne studien vil vi undersøke hvor nøyaktig to av de mest brukte kartleggingsverktøyene i Norge identifiserer språkforstyrrelser. Kartleggingsverktøyene som undersøkes er språktesten *Clinical Evaluation of Language Fundamentals, Fourth Edition* (CELF-4; Semel et al., 2003; norsk utgave Semel et al.,

2013) og spørreskjemaet *Children's Communication Checklist, Second Edition* (CCC-2; Bishop, 2003; norsk utgave Bishop, 2011).

Utviklingsmessige språkforstyrrelser

Utviklingsmessige språkforstyrrelser (developmental language disorder; DLD; tidligere spesifikke språkvansker; SSV) viser til språklige vansker som er vedvarende og har en betydelig innvirkning på sosial interaksjon og læring, uten at de kan settes i sammenheng med en kjent biomedisinsk tilstand (Bishop et al., 2017). Begrepet utviklingsmessige språkforstyrrelser innebærer ikke krav om en gitt diskrepans mellom språklige og ikke-språklige evner, så lenge barnet ikke oppfyller kriteriene for utviklingshemning (Bishop et al., 2017). Dersom barnet har en utviklingshemning, vil det fortsatt kunne ha en språkforstyrrelse, men denne vil da klassifiseres som en «språkforstyrrelse assosiert med en gitt biomedisinsk tilstand» og ikke som en utviklingsmessig språkforstyrrelse (Bishop et al., 2017). Internasjonal litteratur rapporterer om en forekomst av utviklingsmessige språkforstyrrelser på omtrent 7 % (Norbury et al., 2016; Tomblin et al., 1997). De få studiene som er gjennomført i Skandinavia indikerer at forekomsten er like høy her (Hollund-Møllerhaug, 2010a; Westerland, 1994 i Laveson, Lödén & Hansson, 2018). Terminologien for DLD på norsk er foreløpig ikke avklart, men en pågående studie kalt «CATALISE Norge» arbeider med en avklaring (catalis norge.no). Det er ventet at resultatene fra denne studien vil komme i løpet av høsten 2021. I denne studien vil vi benytte begrepet *språkforstyrrelser/ utviklingsmessige språkforstyrrelser* når det er konkludert med at det foreligger vansker som er i tråd med det som tidligere ble referert til som SSV, eller som i nyere terminologi refereres til som *utviklingsmessige språkforstyrrelser*. I andre tilfeller vil vi bruke språkvansker.

Identifisering av språkforstyrrelser

Testresultater alene kan ikke avgjøre om et barn har språkforstyrrelser (Bishop et al., 2016; Leonard, 2014), men språktester spiller ofte en viktig rolle i identifiseringen (Bishop & McDonald, 2009; Denman et al., 2017).

Språktester brukes som regel for å få et detaljert bilde av barnets ferdigheter knyttet til språkets strukturside (fonologi, syntaks og semantikk), men internasjonalt er det også økt fokus på å utvikle tester av språk i bruk (pragmatikk og sosiale kommunikasjonsferdigheter) (Wilson & Bishop, 2021). For å sikre sammenlignbare betingelser for alle som tar språktester har testene standardisert prosedyre for administrering, skåring og tolkning. Språktester kan være gode verktøy for å bekrefte eller avkrefte en mistanke om språkforstyrrelse, spesielt for barn som har en mer «usynlig» vanske, som barn med forståelsvansker (Bishop et al., 2016).

Språktester har blitt kritisert for at de ikke fanger opp vanskene barnet har med språkbruk i sosiale hverdags situasjoner (Hollund-Møllerhaug, 2010a). Det argumenteres for at foreldrerapportering gir et bredere (Dockrell & Marshall, 2015) og mer representativt (Rydz et al., 2005) bilde av barns språkferdigheter. Spørreskjema og foreldrerapporter på sin side kritiseres for å være subjektive, og at foreldrenes ulike forutsetninger for å vurdere barnets språk vil kunne virke inn (Dockrell & Marshall, 2015; Saches & Suchodoletz, 2008). Flere studier har likevel funnet robuste sammenhenger mellom resultatene fra standardiserte språktester og foreldres rapportering av barns språkferdigheter (Ebert, 2017; Massa et al., 2008). Det er også blitt hevdet at foreldrerapporter kan brukes til å skille mellom barn med og uten språkforstyrrelser (Paradis et al., 2010).

Felles for både språktester og spørreskjema er at de har en terskelverdi (cut-off), hvor de som skårer høyere enn terskelverdien vurderes til å være innenfor normalområdet, og de som skårer lavere enn terskelverdien vurderes til å ha en språkforstyrrelse. En forutsetning for en slik tankegang er at språkferdigheter vurderes på en skala som går fra svake til sterke ferdigheter. Et problem med absolutte terskelverdier er at det ikke er noe kvalitativ forskjell på å skåre rett over eller rett under terskelverdien, men at svake ferdigheter går over til å bli en språkforstyrrelse på et punkt i det nedre normalområdet (Tomblin, 2019). Det er imidlertid ingen



Stian Valand
er logoped MNLL og jobber i Statped sørøst, Avdeling språk/tale.
stian.valand@statped.no



**Janne von Koss
Torkildsen**
er professor ved Institutt for spesialpedagogikk, Universitetet i Oslo.
janneto@isp.uio.no

klare kriterier for hvor denne terskelverdien bør settes (Bishop et al., 2016). Plasseringene av terskelverdien for språkforstyrrelser varierer derfor betydelig, men ligger stort sett mellom 16. og 7. persentil (McCauley, 2001). Den mest brukte terskelverdien ser ut til å være 10. persentil, som tilsvarer 1,25 standardavvik under gjennomsnittet for normeringsgruppen (Leonard, 2014; Paul, 2007). Terskelverdien vil imidlertid variere, ofte etter hva formålet med verktøyet er. Formålet med screening vil ofte være å oppdage vansker som bør utredes nærmere (Rydz et al., 2005), mens formålet med standardiserte tester ofte vil være å bestemme om vanskene møter diagnostiske kriterier, som eksempelvis språkforstyrrelser (Denman et al., 2017).

Testers nøyaktighet

Bishop og McDonald (2009) hevder at det er et misforhold mellom barna med språkforstyrrelse som identifiseres kun ved bruk av språktester og de som identifiseres klinisk. Dette kan ha sammenheng med at testene ikke er sensitive og spesifikke nok. Hvor nøyaktig en test måler det den skal måle, angis vanligvis gjennom testens sensitivitet og spesifisitet. Testens sensitivitet referer til andelen av barn som har en språkforstyrrelse som blir korrekt identifisert. Testens spesifisitet refererer til andelen av barn som ikke har en språkforstyrrelse som blir korrekt identifisert (Rydz et al., 2005; Sattler, 2008). Som Tabell 1 viser, kan testresultatene organiseres i grupper basert på om barnet skårer over eller under terskelverdien og barnets sanne gruppetilhørighet (språkforstyrrelse eller ikke). Barnets gruppetilhørighet omtales også som det uavhengige kriteriet (eller «kriteriet»), og viser til en vurdering gjort av fagpersoner eller resultater fra andre tester (Rydz et al., 2005; Sattler, 2008).

Når testen sees i forhold til kriteriet, kan den gi fire utfall. Det første er «sann positiv», hvor testen identifiserer språkforstyrrelse hos en som har det. Det andre er «sann negativ», hvor testen identifiserer noen uten en språkforstyrrelse som innenfor normalområdet. Det tredje er «falsk positiv» eller overidentifisering, hvor testen

identifiserer en språkforstyrrelse hos en som ikke har det. Det fjerde utfallet er «falsk negativ» eller underidentifisering, hvor testen identifiserer noen som innenfor normalområdet selv om de har en språkforstyrrelse.

Den prediktive verdien til en test angir prosentvis hvor mange korrekte klassifiseringer testen gjør i forhold til det uavhengige kriteriet (Sattler, 2008). Et nivå på 70 % for både sensitivitet og spesifisitet blir sett på som en minimumsstandard (Rydz et al., 2005), men ofte vurderes 80 % som et mer akseptabelt nivå for sensitivitet (Barnes, 1982), mens spesifisitet bør være nærmere 90 % (Rydz et al., 2005).

Kartleggingsinstrumentene CELF-4 og CCC-2

CELF-4 er en språktest som har som formål å vurdere språkferdigheter og avdekke språkforstyrrelser hos barn i alderen 5 til 12 år. Testen ble opprinnelig utviklet i USA i 2003, men er oversatt og tilpasset skandinaviske forhold, og normert på et skandinavisk utvalg i 2013 (Semel et al., 2013). Ifølge testmanualen gjør CELF-4 det mulig å foreta en fleksibel og bred språkvurdering, samt å vurdere tiltak på bakgrunn av testresultatene (Semel et al., 2013). Testen består av tretten deltester som kan sammenfattes i sju indekser: grunnleggende språkferdigheter, reseptivt språk, ekspressivt språk, språklig innhold, språklig struktur, språklig hukommelse og arbeidsminne (Semel et al., 2013). Hovedindeksen i CELF-4 er Indeks for grunnleggende språkferdigheter (IGS). Indeksen dekker et spekter av språkferdigheter, inkludert grammatikk, språklig hukommelse og ord- og begrepsforståelse og benyttes som et første steg i en vurdering av en mulig språkforstyrrelse. Gjennomsnittet i normeringsutvalget er 100 og standardavviket 15. Terskelverdien er 85 (1 standardavvik under gjennomsnittet), hvilket betyr at en IGS-skåre under 85 indikerer en språkforstyrrelse (Semel et al., 2013).

CCC-2 er et spørreskjema for vurdering av strukturelle og pragmatiske språkferdigheter hos barn og ungdom. Instrumentet er oversatt fra engelsk, tilpasset og normert

Tabell 1

De fire mulige utfallene av en språktest som er ment for å identifisere språkforstyrrelser

	Testresultat	
	Språkforstyrrelse	Normalvariasjon
Kriteriet	Sann positiv	Falsk negativ
	Kontrollgruppe	Sann negativ

for norske barn i alderen 4 til 17 år (Bishop, 2011; Helland et al., 2009). Formålet med skjemaet er tredelt: 1. Å fange opp barn med språkvansker som bør henvises til grundigere språkutredning. 2. Å identifisere pragmatiske vansker hos barn med språkvansker. 3. Å bidra til å identifisere barn som bør utredes for autismespekterforstyrrelser (Bishop, 2011; 9). For å få et realistisk bilde av hvordan barnet kommuniserer, skal CCC-2 besvares av en voksen som har regelmessig kontakt med barnet. Skjemaet består av 70 utsagn med flervalgssvar. Den voksne blir spurt om å vurdere hvor ofte de observerer den språklige eller kommunikative atferden beskrevet i utsagnet. Utsagnene er inndelt i 10 skalaer som skal vurdere hvert sitt aspekt innenfor språk og kommunikasjon. Skalaene er: tale (A), syntaks (B), semantikk (C), sammenheng (D), samtaleinitiativ (E), stereotypisk språk (F), bruk av kontekst (G), ikke-verbal kommunikasjon (H), sosiale relasjoner (I) og interesser (J). Råskårene på hver skala konverteres til aldersrelaterede skalerte skårer med gjennomsnitt 10 og standardavvik 3. Den ene av de to hovedindeksene i CCC-2 kalles Generell kommunikasjonsindeks (GKI). Denne gir et sammensatt mål på generell språklig fungering, og kan indikere om barnet har en språkforstyrrelse eller ikke (Bishop, 2011). CCC-2 GKI består av summen av skalerte skårer fra skalaene A-H, der A-D måler strukturelle språkferdigheter og E-H måler pragmatiske språkferdigheter. En GKI-skåre på 55 (10. persentil) kan ifølge den norske CCC-2 manualen «ses som en «kritisk» verdi og brukes som cut-off-grense når man skal bestemme om vanskenivået er innenfor det kliniske området eller ikke» (Bishop, 2011;33). Det kan forstås som at en GKI-skåre under 55 indikerer en språkforstyrrelse. Den andre hovedindeksen kalles Indeks for avvik i sosial interaksjon (IASI), og gir et sammensatt mål på språkbruk i sosiale situasjoner. Denne skåren kan bidra til å fange opp barn med en kommunikasjonsprofil som er karakteristisk for autismespekterforstyrrelser og pragmatiske språkvansker (Bishop, 2011).

Studier som har undersøkt kvaliteten på språkkartleggingsverktøy i engelskspråklige land har konkludert med at både CELF-4 (Spaulding et al. 2006) og etterkommeren, CELF-5 (Denman et al., 2017), er blant få kartleggingsverktøy som rapporterer akseptable nivåer for spesifisitet og sensitivitet i sin manual og at de har god nok psykometrisk kvalitet til å anbefales for diagnostisk kartlegging. Siden testutviklere har kommersielle interesser i verktøyene, er det imidlertid nødvendig med uavhengige studier for å avgjøre om testene har god nok kvalitet (Denman et al., 2017).

I en britisk valideringsstudie av CCC-2 fant Norbury et al. (2004) signifikante forskjeller i GKI-skåre (GCC) mellom kontrollgruppen (n=21) og kliniske grupper som «Specific Language Impairment» (n = 19) (spesifikke språkvansker), PLI (n = 14) (pragmatiske språkvansker), PLI+ (n = 21) (pragmatiske språkvansker med autistiske trekk), HFA (n = 17) (høyt fungerende autisme) og ASP (n= 16) (Asperger syndrom). De konkluderte med at CCC-2 er et godt screeningverktøy. Kvaliteten på CCC-2 er også undersøkt i en norsk kontekst. Helland et al. (2009) undersøkte den psykometriske kvaliteten samt sensitiviteten og spesifisiteten til den norske oversettelsen av CCC-2 hos 153 barn, hvorav 45 hadde språkforstyrrelser. Forfatterne, som baserte analysene på det britiske normgrunnlaget, konkluderte med at verktøyet hadde en tilfredsstillende reliabilitet. Når det gjaldt sensitivitet, fant de imidlertid at bare 69,8% av utvalget med forhåndsdiagnostiserte språkforstyrrelser skåret under terskelverdien. Derimot var spesifisiteten svært god, på 98,1 %. Ved å flytte terskelverdien fra skåre 55 (terskelverdien fra den engelske normeringen) til skåre 64, steg sensitiviteten til CCC-2 fra 69,8 % til 86 %. På bekostning av økt sensitivitet, ble spesifisiteten redusert fra 98,1 % til 90,7%. Imidlertid anses en spesifisitetsverdi over 90 % fremdeles som god (Barnes, 1982; Rydz et al., 2005). Hollund-Møllerhaug (2010b) sammenlignet GKI-skårer og persentilfordeling i et norsk utvalg (n=655) basert på det britiske normeringsgrunnlaget (Bishop, 2003). Hun fant at en GKI-skåre tilsvarende 10. persentil i det norske utvalget var 60, ikke 55, som Bishop (2003) anbefalte som terskelverdi for britiske barn.

I dag vet vi lite om nøyaktigheten til kartleggingsinstrumentene som har formål om å identifisere barn med språkforstyrrelser i Norge. Med denne studien ønsker vi å undersøke nøyaktigheten til to mye brukte kartleggingsverktøy, CELF-4 (Semel et al., 2013) og CCC-2 (Bishop, 2011). Med bakgrunn i litteraturgjennomgangen over stiller vi følgende forskningsspørsmål:

1. Tilfredsstillende de norske versjonene av CELF-4 og CCC-2 gjeldende krav til sensitivitet og spesifisitet for identifisering av språkforstyrrelser hos barn i skolealder?
2. Hvordan vil en justering av terskelverdien til CCC-2 påvirke instrumentets sensitivitet og spesifisitet?

Metode

Denne studien er knyttet til forskningsprosjektet «Undersøkelse av taleoppfattelse, språkferdigheter og livskvalitet

hos personer med cochleaimplantat (CI), høreapparat, språkvansker og typisk utvikling» ved Universitetet i Oslo og Oslo universitetssykehus.

Deltakere

Utvalget i denne studien består av 21 barn med språkforstyrrelser, hvorav 11 er jenter og 10 er gutter, og en kontrollgruppe av 21 barn som er parvis matchet med barna i språkforstyrrelsegruppen på alder, kjønn og nonverbale evner. Språkforstyrrelsegruppen har en gjennomsnittsalder på 9;4 med et standardavvik (SD) på 25,59, og kontrollgruppen har en gjennomsnittsalder på 9;5 med et standardavvik på 26,24. Inklusjonskriterier for språkforstyrrelsegruppen (n=21) var: 1) Henvist til pedagogisk-psykologisk-tjeneste (PPT), Statped, eller annen relevant institusjon med språket som hovedvanske. 2) Skårer som lå -1,25 SD eller mer under gjennomsnittet på en test som måler nonordrepetisjon (Gathercole et al., 1994). 3) Skårer som lå -1 SD eller mer under normativt gjennomsnitt på testen The British Picture Vocabulary Scale II (BPVS-II; Dunn et al., 1997). 4) Nonverbale evner på minst 70, målt med Ravens fargede matriser (Raven et al., 2003). 5) Norsk som førstespråk. 6) Ingen andre diagnostiserte utviklingsforstyrrelser. 7) Tilstedeværelse av otoakustiske emisjoner, som indikerer normal funksjon i det indre øret og normal hørsel med hørselsterskler bedre enn 30 dB. Barna i kontrollgruppen gjennomgikk samme kartlegging som barna med språkforstyrrelser. Inklusjonskriteriene for barna i kontrollgruppen var at de ikke hadde noen familiehistorie med språkforstyrrelser eller dysleksi, at de fikk skårer over -1,25 SD på nonordrepetisjon og over -1 SD på BPVS, samt kriteriene 4-7 over.

Barna i språkforstyrrelsegruppen ble utredet av enten PPT, Barne- og ungdomspsykiatrisk poliklinikk (BUP) eller Statped mellom 2009 og 2016, og kun én av deltakerne i språkforstyrrelsegruppen hadde tidligere gjennomgått testing med CELF-4 som en del av utredningen. Imidlertid lå dette barnets skåre på CELF-4 IGS på samme nivå som skårene på de andre språktestene, som barnet ikke hadde gjennomført tidligere. Vi vurderte det derfor som sannsynlig at skåren på CELF-4 var lite påvirket av den tidligere testen, og valgte derfor å beholde dette barnet i utvalget.

Innsamling av data

Prosjektet som denne studien inngår i, er godkjent av Regionale komiteer for medisinsk og helsefaglig forskningsetikk.

Barna i språkforstyrrelsegruppen ble rekruttert ved at PPT-kontor i alle landsdeler og Statped sørøst distribuerte informasjon om studien til aktuelle deltakere. Barna i kontrollgruppen ble rekruttert ved at skoler i Sørøst-Norge distribuerte informasjon om studier til barn i den aktuelle aldersgruppen. Deltakerne ble kartlagt med et testbatteri som inkluderte tester av språk- og leseferdigheter, kognitive evner, oppmerksomhet og hørsel. Kartleggingen foregikk individuelt. Både kartleggingen og skåringen ble utført av prosjektledere, forskningsassistenter og masterstudenter som hadde gjennomgått trening i testbatteriet. Kartleggingen varte fra halvannen til to timer og ble fordelt over to kartleggingsøkter over en to-ukersperiode. CCC-2 ble fylt ut av barnas foresatte på samme tid som kartleggingen med testbatteriet.

Kartleggingsinstrumenter

I tillegg til CELF-4 og CCC-2, som er beskrevet over, benyttet vi tre bakgrunnstester som en del av inklusjonskriteriene for studien.

Ravens matriser (Raven et al., 2003) måler nonverbal IQ. Versjonen med fargede matriser ble brukt med barn i alderen 5;9-8;11, mens Ravens standard progressive matriser ble brukt for barn i alderen 9;0-12;11. Begge versjonene har 36 oppgaver. Barnas oppgave er å identifisere det manglende elementet som fullfører et matrisemønster. Testen inneholder kun visuelle, ikke-språklige stimuli, og er organisert slik at matriseoppgavene har stigende vanskelighetsgrad.

The British Picture Vocabulary Scale (BPVS) II (Dunn et al., 1997; norsk versjon og normering av Lyster et al., 2010) måler reseptivt ordforråd. Testen har 12 blokker med 12 oppgaver i hver. For hver oppgave får deltakerne se et sett med fire tegninger og skal peke på den tegningen som tilsvarende et ord testlederen sier. Testen avsluttes dersom deltakeren får 8 feil i en blokk.

Nonordrepetisjonstesten (Gathercole et al., 1994; norsk versjon av Furnes & Samuelsson, 2009) måler fonologisk korttidsminne. Testen består av 28 nonord med to, tre, fire eller fem stavelser. Barna får høre et opptak av hvert nonord og blir bedt om å gjenta det. Hvis barnet ikke svarer, kan ordet gjentas én gang. Bare fullstendige og korrekte repetisjoner gir poeng. Det finnes ikke normer for den norske versjonen av denne testen, og derfor er råskårer brukt i alle utregninger.

Tabell 2

Deskriptiv statistikk for bakgrunnsvariabler for barna i språkforstyrrelsegruppen (n= 21) og barna i den matchede kontrollgruppen (n=21)

	Språkforstyrrelsegruppen			Kontrollgruppen			t-test (df)	Effektstr.	
	M	SD	Var.bredde	M	SD	Var.bredde	t(40)	p	d
Alder i mnd.	113.12	25.59	69-155	114.09	26.24	71-155	0.12	.91	0.04
Nonverbal IQ	88.33	11.55	70-120	95.71	10.87	70-115	2.13	.04	0.66
Reseptivt vokabular	76.57	17.89	16-100	105.29	14.11	79-126	5.78	<.001	1.78
Nonordsrepetisjon	12.76	6.16	0-20	23.38	2.69	18-27	7.24	<.001	2.23

Note: M= gjennomsnitt, SD= standardavvik, Var.bredde= variasjonsbredde, (df)= frihetsgrader. Nonverbal IQ målt med Ravens matriser. Reseptivt vokabular målt med British Picture Vocabulary Scale II. Nonordrepetisjon målt med Test av nonordrepetisjon for barn (Furnes & Samuelsson, 2009). Skårene for nonverbal IQ og reseptivt vokabular er standardskårer. Skårene for nonordrepetisjon er råskårer (det finnes ikke norske normer for denne testen). Effekttørrelsen er Cohens *d*.

Resultat

Deskriptiv statistikk

Som Tabell 2 viser, var det ingen forskjell i alder mellom språkforstyrrelsegruppen og kontrollgruppen. Det var derimot en moderat gruppeforskjell i nonverbal IQ og store forskjeller i reseptivt vokabular og nonordrepetisjon. Merk imidlertid at testene av reseptivt vokabular og nonordrepetisjon ble brukt som inklusjonskriterier for gruppene (jf. metodedel), og at en derfor kunne forvente gruppeforskjeller på disse variablene.

Tabell 3 viser skårer på de to kartleggingsverktøyene som var i fokus i denne studien: CELF-4 og CCC-2. Gruppeforskjellene mellom språkforstyrrelsegruppen og kontrollgruppen på de to variablene som måler strukturelle språkferdigheter, CELF-4 IGS og CCC-2 GKI, var svært store, omtrent 3 standardavviksenheter. Forskjellene mellom gruppene på CCC-2 IASI var også store, men

omtrent halvparten av effekttørrelsen for de to målene på strukturelle språkferdigheter.

Kartleggingsinstrumentenes nøyaktighet (forskningsspørsmål 1)

Basert på inndelingen i språkforstyrrelsegruppe eller kontrollgruppe («kriteriet») og resultatene fra CELF-4 og CCC-2 ble verktøyenes sensitivitet (sann positiv/(sann positiv+falsk negativ)) og spesifisitet (sann negativ/(sann negativ + falsk positiv)) beregnet (Rydz et al., 2005; Sattler, 2008).

Som det kommer frem av Tabell 4 hadde CELF-4 en nøyaktighet på 95,2 % og CCC-2 en nøyaktighet på 80,9 %. CELF-4 identifiserte 90,4 % (n = 19) av barna i språkforstyrrelsegruppen korrekt med en IGS skåre < 85. To barn i språkforstyrrelsegruppen ble underidentifisert med IGS-skårer innen normalområdet. Samtlige av barna i

Tabell 3

Skårer på CELF-4 for språkforstyrrelsegruppen (n=21) og barna i den matchede kontrollgruppen (n=21)

	Språkforstyrrelsegruppen			Kontrollgruppen			t-test (df)	Effektstr.	
	M	SD	Var.bredde	M	SD	Var.bredde	t(40)	p	d
CELF-4 IGS	63.90	14.25	40-95	103.19	11.74	86-135	9.51	<.001	3.01
CCC-2 GKI	40.52	17.89	14-73	86.95	12.11	48-101	9.85	<.001	3.04
CCC-2 IASI	11.43	8.52	-7-27	-0,86	8.04	-30-7	-4.81	<.001	1.48

Note: M= gjennomsnitt, SD= standardavvik, Var.bredde= variasjonsbredde, (df)= frihetsgrader. CELF-4 = Clinical Evaluation of Language Fundamentals 4. IGS = Indeks for Grunnleggende Språkferdigheter. CCC-2 = Childrens Communication Checklist 2. GKI = Generell Kommunikasjonsindeks. IASI = Indeks for Avvik i Sosial Interaksjon. Skårene for CELF-4 IGS er standardskårer og skårene for CCC-2 GKI og CCC-2 IASI er skalerte skårer. Effekttørrelsen er Cohens *d*.

kontrollgruppen ($n = 21$) hadde en IGS skåre > 85 og ble korrekt identifisert. CCC-2 identifiserte 66,6 % ($n = 14$) av barna i språkforstyrrelsegruppen korrekt med en GKI < 55 . Syv av barna i språkforstyrrelsegruppen ble underidentifisert med skårer > 55 , mens ett barn i kontrollgruppen ble overidentifisert med skårer < 55 .

Tabell 4

Sensitivitet og spesifisitet for CELF-4 IGS og CCC-2 GKI

	Sensitivitet	Spesifisitet	Total nøyaktighet
CELF-4 IGS	90,4 %	100 %	95,2 %
CCC-2 GKI	66,6 %	95,2 %	80,9 %

De to verktøyene var svært samstemte i klassifiseringen av deltakerne i kontrollgruppen, men langt mindre samstemte i klassifiseringen av deltakerne i språkforstyrrelsegruppen. Kun tolv av barna i språkvanskegruppen fikk identifisert språkforstyrrelsen med begge verktøyene, men samtlige blir identifisert når man tar begge verktøyene i betraktning. Ti av barna får motstridende resultater med de to verktøyene og ni av disse barna tilhører språkforstyrrelsegruppen.

Justering av terskelverdi (forskningsspørsmål 2)

Vi undersøkte effekten av å flytte terskelverdien til GKI fra 55 til 64 slik Helland (2009) gjorde i sin studie basert på det britiske normeringsgrunnlaget. Med den justerte terskelverdien ble fem av syv deltakere som tidligere ble kategorisert som falske negative nå identifisert korrekt ut ifra det uavhengige kriteriet. Sensitiviteten endret seg fra 66,6 % til 90,4 %. Endringen i terskelverdi hadde ingen konsekvenser for spesifisiteten som fremdeles var på 95,2 %. Med denne justeringen ville den totale nøyaktigheten av CCC-2 GKI økt fra 80,9 % til 92,8 % i vårt utvalg.

Diskusjon

Hovedfunnet i denne studien er at språktesten CELF-4 identifiserer barn med språkforstyrrelser med stor nøyaktighet i et utvalg som er uavhengig av den skandinaviske normeringen. Verdier for sensitivitet og spesifisitet var tilfredsstillende. Spørreskjemaet CCC-2 identifiserer kun to tredjedeler av barna med språkforstyrrelser, og har derfor ikke tilfredsstillende nøyaktighet.

Tilfredsstillende de norske versjonene av CELF-4 og CCC-2 gjeldende krav til sensitivitet og spesifisitet?

Spesifisiteten til CELF-4 var optimal i dette utvalget på 42 barn, da samtlige av deltakerne som ifølge det uavhengige kriteriet var i kontrollgruppen ble korrekt identifisert med

CELF-4. Sensitivitetsverdien på over 90 indikerer også en akseptabel identifisering av deltagere som ifølge det uavhengige kriteriet var i språkforstyrrelsegruppen (Barnes, 1982; Rydz et al., 2005). Merk at identifiseringen kun baserte seg på den generelle kommunikasjonsindeksen (IGS), som består av fire deltester avhengig av aldersgruppe. Resultatene tyder altså på at under standardiserte betingelser gir bruk av disse deltestene fra CELF-4 en akseptabel identifisering av utviklingsmessige språkforstyrrelser. Imidlertid var det to av de 21 barna i språkforstyrrelsegruppen som ikke ble identifisert med CELF-4. Det ene barnet hadde en skåre like over terskelverdien til CELF-4 IGS (skåre 86), som indikerer at barnet ligger i grenseområdet for klassifisering av utviklingsmessige språkforstyrrelser med denne testen. Begge disse barna fikk skårer godt under gruppens gjennomsnitt på CCC-2 GKI, og ble identifisert med språkforstyrrelse ifølge spørreskjemaet.

Med unntak av én deltaker ble kontrollgruppen korrekt identifisert i forhold til det uavhengige kriteriet med CCC-2, noe som i likhet med CELF-4 tyder på en nær optimal spesifisitet. Det ser med andre ord ut til at foreldre gjør en valid vurdering av barnets språk- og kommunikasjonsferdigheter når ferdighetene er innenfor normalvariasjonen (jf. Hauerwas & Stone, 2000). Det var kun ett tilfelle av overidentifisering, det vil si en «falsk positiv». Keegstra et al. (2007) hevder det er mer vanlig at overidentifisering observeres ved foreldrerapportering enn ved bruk av språktester. Profilen til deltakeren som var «falsk positiv» indikerte pragmatiske språkvansker.

Sensitivitetsverdien indikerte imidlertid at en tredjedel av deltakerne som hadde språkforstyrrelser ifølge det uavhengige kriteriet ikke ble fanget opp av CCC-2. En så lav sensitivitet kan ikke anses som tilfredsstillende (Barnes, 1982; Rydz, 2005). Benyttes kun CCC-2 GKI med terskelverdien som er anbefalt i den norske manualen, for vurdering av barns språkferdigheter, vil vanskene til en betydelig andel av barna forbli uoppdaget og de vil sannsynligvis ikke få den oppfølgingen de trenger. Det samme gjelder om man bruker CCC-2 for å avgjøre om det er hensiktsmessig å gjøre ytterligere kartlegginger av barnets språk. Selv om det understrekes i manualen at CCC-2 alene ikke gir grunnlag for å sette diagnose fordi det er utviklet som et screeningverktøy, står det også at GKI kan brukes til å identifisere barn som har språkforstyrrelser (Bishop, 2011;21). Dette kan være en kilde til usikkerhet og bidra til at resultatene fra spørreskjemaet tillegges større vekt enn det er grunnlag for

ved mistanke om språkforstyrrelse. En høy andel falske negativer, som følge av underidentifisering, er i tråd med tidligere studier av samme instrument (Bishop & McDonald, 2009; Helland et al., 2009). En mulig forklaring på underidentifiseringen kan være knyttet til foreldrenes forutsetninger for å rapportere eget barns språk. For det første kan underrapporteringen være forårsaket av at foreldrene har for lite kunnskap om eget barns språk eller om språk generelt og dermed vurderer utsagnene feil (Dockrell & Marshall, 2015; Ebert, 2017). I tillegg blir barns språk mer komplekst med økende alder, noe som gjør det enda vanskeligere for foreldre å vurdere barns språk fra skolealder og oppover (Tetzchner et al., 1993). For det andre kan det hende at foreldrene ikke opplever barnets språkforstyrrelse som alvorlig, for eksempel fordi de ikke har lagt merke til vansker på skolen eller vansker med sosial fungering. Foreldre kan være gode til å forstå egne barns språk og kommunikasjon til tross for barnas språkforstyrrelse. Dette er i tråd med Sachse og Suchodoletz' (2008) bemerkninger om at foreldres vurdering ikke er objektiv. På en annen side har en COST-action-pilotstudie av Jensen de López et al. (under trykking) funnet at foreldre ofte har god forståelse for barnets vansker og innsikt i de funksjonelle konsekvensene språkforstyrrelsen har for barnet i hverdagen. Konklusjonene fra studien fremhever viktigheten av å anerkjenne foreldrenes kompetanse, og at denne er viktig for å forstå barnet og gi et relevant tjenestetilbud. Jensen de López et al. (under trykking) intervjuet foreldre til barn som allerede var identifisert med en språkforstyrrelse. Dette kan forklare forskjellene i resultatene mellom studiene.

En annen forklaring på den lave sensitivitetsverdien til CCC-2 kan være knyttet til operasjonaliseringen av begrepet «generell kommunikasjon» i CCC-2. Indeksen GKI har som ambisjon å identifisere barn med språkforstyrrelser, men inkluderer indikatorer som måler områder som ikke nødvendigvis er berørt hos barn med språkforstyrrelser. For de sju barna i språkforstyrrelsegruppen som vurderes innenfor normalområdet, kan det se ut til at ikke-språklige aspekter som måles med skalaen ikke-verbal kommunikasjon (H), trekker den totale CCC-2 GKI-skåren opp og fører til de blir vurdert innenfor normalområdet. I tillegg har CCC-2 et sterkt fokus på ekspressivt språk, og inneholder færre utsagn rettet mot reseptivt språk. Det kan bidra til å gjøre det vanskelig å oppdage barn som har mer subtile vansker med reseptivt språk.

Hvordan vil en justering av terskelverdien til CCC-2 påvirke sensitivitet og spesifisitet?

En justering av terskelverdien for GKI fra 55 til 64 bidro til en betydelig økning i andelen barn med språkforstyrrelser som ble korrekt identifisert med spørreskjemaet. Totalt sett virket derfor flytting av terskelverdi svært hensiktsmessig for instrumentets nøyaktighet i vårt utvalg. Dette funnet er parallelt med funnene til Helland et al. (2009) og Hollund-Møllerhaug (2010b) som fant at en justering av terskelverdiene i disse studiene, som er basert på det britiske normeringsgrunnlaget, resulterte i høyere sensitivitet. Disse funnene understreker viktigheten av en grundig validering av de psykometriske egenskapene, og særlig sensitivitet og spesifisitet, ved instrumenter i ulike kulturelle og språklige kontekster.

Praktiske implikasjoner

Vår studie indikerer at det er tilstrekkelig å gjennomføre kjernetestene i CELF-4 IGS for å kunne avgjøre med stor nøyaktighet om et barn har en utviklingsmessig språkforstyrrelse. Dette kan være nyttig for klinikere som møter barn som det av ulike grunner kan være vanskelig å gjennomføre hele testen med. Våre resultater tyder også på at CCC-2 ikke bør brukes alene til å avgjøre om barn med mistanke om utviklingsmessige språkforstyrrelser skal utredes videre. Til dette har ikke instrumentet god nok nøyaktighet med manualens terskelverdier, og man kan derfor ende opp med å ikke utrede barn som har en utviklingsmessig språkforstyrrelse. Spørreskjemaet kan brukes som supplement til annen utredning for å vurdere ulike aspekter ved barnets språk. Det må i tillegg tas i betraktning at ved klinisk bruk skal den totale CCC-2 profilen tolkes, ikke kun CCC-2 GKI.

Det er verdt å merke seg at samtlige av barna med utviklingsmessige språkforstyrrelser blir identifisert dersom man bruker begge instrumentene, det vil si CELF-4 IGS og CCC-2 GKI. Dette støtter anbefalingen om å ta i bruk minst to kartleggingsinstrumenter i utredning for utviklingsmessige språkforstyrrelser (Paul, 2007). Det bør imidlertid påpekes at det heller ikke anses som tilstrekkelig å benytte testskårer alene for å identifisere språkforstyrrelser, men at intervju og observasjoner i naturlige kontekster bør inngå i utredningen. Til sammen vil kombinasjonen av minst to standardiserte og normerte kartleggingsinstrumenter, intervjuer og observasjoner ofte være tilstrekkelig til å gjøre en god vurdering av barnets språkferdigheter (Bishop et al., 2016).

Flere har også tatt til orde for at barn som skårer i nedre del av normalområdet bør få det samme tilbudet om intervensjon som barn som skårer like under terskelverdi (Bishop et al., 2016). Vår oppfordring er derfor å ha en pragmatisk tilnærming til bruk av testresultater i praksis.

Styrker og begrensninger

En begrensning ved denne studien er at utvalgsstørrelsen er mindre enn hva som er ideelt i en studie av testers sensitivitet og spesifisitet. Samtidig er dette én av svært få studier som har undersøkt et utvalg som allerede er identifisert med utviklingsmessig språkforstyrrelse av PPT eller andre instanser, og som i tillegg har verifisert språkforstyrrelsen gjennom en uavhengig språkkartlegging (i dette tilfellet BPVS og nonordrepetsjon). Det er også en styrke at studien inkluderte en kontrollgruppe som var matchet med språkforstyrrelsegruppen på kjønn, alder og non-verbal IQ. Ett av barna i utvalget hadde tidligere tatt CELF-4 som en del av utredningen hos PPT, men bortsett fra dette barnet var andre tester, i kombinasjon med klinisk skjønn, blitt brukt til å identifisere språkforstyrrelser i den opprinnelige utredningen. Resultatene av vår studie tyder derfor på at CELF-4 kan identifisere utviklingsmessige språkforstyrrelser som er påvist gjennom en utredning med andre instrumenter. Utover informasjonen om CELF-4 var tatt tidligere eller ikke, mangler vi informasjon om hvilke kartleggingsinstrumenter som ble brukt i utredningen hos PPT, BUP og Statped. Vi vet derfor ikke om CCC-2 har blitt brukt som en del av utredningen.

Deltakerne i denne studien hadde norsk som førstespråk. Resultatene kan derfor ikke overføres til å gjelde identifisering av utviklingsmessige språkforstyrrelser hos flerspråklige barn.

Videre bør det understrekes at formålet med studien var å vurdere instrumentenes nøyaktighet i identifisering av barn med utviklingsmessige språkforstyrrelser. Av den grunn brukte vi kun indeksen for generelle språkferdigheter (GKI) til å vurdere sensitivitet og spesifisitet, og resultatene kan derfor ikke si noe om CCC-2s nøyaktighet i identifisering av pragmatiske eller sosiale kommunikasjonsvansker eller mulige autismespekterforstyrrelser. Avslutningsvis skal det også nevnes at CELF-4 og CCC-2 ikke er de eneste testene som benyttes i identifisering av språkvansker i Norge, og at fremtidige studier også bør undersøke sensitiviteten og spesifisiteten til andre språktester som TROG-II (Bishop, 2003) og BPVS-II (Dunn et al., 1997).

Konklusjon

For å korrekt identifisere barn med utviklingsmessige språkforstyrrelser kreves valide og nøyaktige instrumenter. Denne undersøkelsen bidrar med konkrete resultater om identifisering av språkforstyrrelser hos norske barn med verktøyene CELF-4 og CCC-2 i utvalg som er uavhengige av normeringsutvalgene.

Barnas språkferdigheter målt med CELF-4 IGS klassifiserer deltakerne korrekt i språkforstyrrelsegruppe og kontrollgruppe i svært stor grad, og tilfredsstillende dermed de gjeldende krav til nøyaktig vurdering av barns språkferdigheter. CCC-2 identifiserer barnas plassering i språkforstyrrelsegruppe og kontrollgruppe i mindre grad enn CELF-4, og med manualens terskelverdier tilfredsstillende ikke instrumentet sitt formål som screeninginstrument for å fange opp barns språkforstyrrelse i tilstrekkelig grad i vårt utvalg. CCC-2 kan imidlertid brukes som supplement til andre standardiserte verktøy.

Referanser

- Barnes, K. E. (1982). *Preschool Screening: The Measurement and Prediction of Children at Risk*. Springfield: Charles C. Thomas Publishers.
- Bishop, D. V. M. (2003). *The children's Communication Checklist - Second Edition (CCC2)*. London: Psychological Corporation.
- Bishop, D. (2003). *Test for Reception of Grammar II*. London: Harcourt Assessment.
- Bishop, D. V. M. (2011). *CCC-2. The Children's Communication Checklist. Manual (Norsk versjon) (2 utg.)*. Stockholm: Pearson Assessment.
- Bishop, D. V. M., & McDonald, D. (2009). Identifying language impairment in children: combining language test scores with parental report. *International Journal of Language & Communication Disorders*, 44(5), 600-615.
- Bishop, D. V. M., Snowling, M. J., Thompson, P. A., & Greenhalgh, T. (2016). CATALISE: A Multinational and Multidisciplinary Delphi Consensus Study. Identifying Language Impairments in Children. *PLoS ONE*, 11 (7).
- Bishop, D. V. M., Snowling, M., Thompson, P. A., & Greenhalgh, T. (2017). Phase 2 of CATALISE: A Multinational and Multidisciplinary Delphi Consensus Study of Problems with Language Development: Terminology. *Journal of Child Psychology and Psychiatry*, 58(10).
- Conti-Ramsden, G., Durkin, K., Toseeb, U., Botting, N., & Pickles, A. (2018). Education and employment outcomes of young adults with a history of developmental language disorder. *International Journal of Language & Communication Disorders*, 53(2), 237-255.
- Dale, P. S., McMillan, A. J., Haiyou-Thomas, M. E., & Plomin, R. (2014). Illusory recovery: Are recovered children with early language delay at continuing elevated risk? *American Journal of Speech-Language Pathology*, 23(3), 437-447.
- Denman, D., Speyer, R., Munro, N., Pearce, W. M., Chen, Y-W. & Cordier, R. (2017). Psychometric properties of language assessments for children aged 4-12 years: A systematic review. *Frontiers in Psychology*. 07. September 2017.
- Dockrell, J. E., & Marshall, C. R. (2015). Measurement Issues: Assessing language skills in young children. *Child and Adolescent Mental Health*, 20(2), 116-125.
- Dunn, L. M., Dunn, L. M., Whetton, C., & Byrley, J. (1997). *The British Picture Vocabulary Scale (2 utg.)*. London: Nelson Publishing Company.
- Ebert, K. D. (2017). Convergence between parent report and direct assessment of language and attention in culturally and linguistically diverse children. *PLoS ONE*, 12(7).
- Furnes B., Samuelsson S. (2010). Predicting reading difficulties in transparent and opaque orthographies: a comparison between Scandinavia and U.S./Australian children. *Dyslexia* 16 119-142.
- Gathercole, S. E, Willis, C. S., Baddeley, A. D. & Emslie, H. (1994). The children's test of nonword repetition: A test of phonological working memory, *Memory*, 2:2, 103-127.
- Hauerwas, L. B., & Stone, C. A. (2000). Are parents of school-age children with specific language impairments accurate estimators of their child's language skills? *Child Language Teaching and Therapy*, 16(1), 73-86.
- Helland, W. A., Biringner, E., Helland, T., & Heimann, M. (2009). The usability of a Norwegian adaptation of the Children's Communication Checklist Second Edition (CCC-2) in differentiating between language impaired and non-language impaired 6- to 12-year-olds. *Scandinavian Journal of Psychology*, 50(3), 287-292.
- Hollund-Møllerhaug, L. (2010). Forekomst av språkvansker hos norske barn. *Tidsskrift for norsk psykologforening*, 47(7), 608-610.
- Hollund-Møllerhaug, L. (2010b). Children's Communication Checklist (CCC-2): En sammenligning av GCC-skårenes persentilfordeling i et britisk og norsk utvalg. *Tidsskrift for norsk psykologforening*, 47(9), 824-826.
- Jensen de López, K. M., Lyons, R., Novogrodsky, R., Baena, S., Feilberg, J., Harding, S., Kelić, M., Klatte, I. S., Mantel, T. C., Tomazin, M. O., Ulfssdottir, T. S., Zajdó K. & Rodriguez-Ortiz, I. R. (under trykking). Exploring Parental Perspectives of Childhood Speech and Language Disorders Across 10 Countries: A Pilot Qualitative Study. *Journal of Speech, Language, and Hearing Research*.
- Keegstra, A. L., Knijff, W. A., Post, W. J., & Goorhuis-Brouwer, S. M. (2007). Children with language problems in a speech and hearing clinic: Background variables and extent of language problems. *International Journal of Pediatric Otorhinolaryngology*, 71(5), 815-821.
- Laveson, A., Lödén, M. & Hansson, K. (2018) Development of a language screening instrument for Swedish 4-year-olds. *International Journal of Communication Disorders*, 53(3), 605-614.
- Leonard, L. B. (2014). *Children with Specific Language Impairment (2 utg.)*. Cambridge, Mass: MIT Press.
- Lyster, S. A. H., Horn, E., & Rygvold, A. L. (2010). Ordforråd og ordforrådsutvikling hos norske barn og unge. Resultater fra en utprøving av British picture vocabulary scale, (BPVS II). *Spesialpedagogikk* 9, 35-43.
- Massa, J., Gomes, H., Tartter, V., Wolfson, V., & Halperin, J. M. (2008). Concordance Rates between Parent and Teacher Clinical Evaluation of Language Fundamentals Observational Rating Scale. *International Journal of Language & Communication Disorders*, 43(1), 99-110.
- McCauley, R. J. (2001). *Assessment of language disorders in children*. Mahwah, New Jersey: Lawrence Erlbaum Associates, Inc.
- Mok, P.L.H., Pickles, A., Durkin, K. and Conti-Ramsden, G. (2014), *Longitudinal trajectories of peer relations in children with specific language impairment*. *Journal of Child Psychology and Psychiatry*, 55, 516-527.
- Norbury, C. F., Nash, M., Baird, G., & Bishop, D. V. (2004). Using a parental checklist to identify diagnostic groups in children with communication impairment: A validation of the Children's Communication Checklist-2. *International Journal of Language & Communication Disorders*, 39(3), 345-364.
- Norbury, C. F., Gooch, D., Wray, C., Baird, G., Charman, T., Simonoff, E. mfl. (2016). The impact of nonverbal ability on prevalence and clinical presentation of language disorder: evidence from a population study. *Journal of Child Psychology and Psychiatry*, 57(11), 1247-1257.
- Paradis, J., Emmerzael, K., & Duncan, T. S. (2010). Assessment of English language learners: using parent report on first language development. *Journal of Communication Disorders*, 43(6), 474-497.
- Paul, R. (2007). *Language Disorders from Infancy through Adolescence: assessment & intervention (3 utg.)*. St. Louis, Mo: Mosby.
- Raven, J., Raven, J. C. & Court, J. H. (2003). *Manual for Raven's Progressive Matrices and Vocabulary Scales*. San Antonio, TX: Harcourt Assessment.
- Rydz, D., Shevell, M. I., Majnemer, A., & Oskoui, M. (2005). Developmental Screening. *Journal of Child Neurology*, 20(1), 4-21.
- Sachse, S. v., & Suchodoletz, W. v. (2008). Early Identification of Language Delay by Direct Language Assessment or Parent Report? *Journal of Developmental & Behavioral Pediatrics*, 29(1), 34-41.
- Sattler, J. M. (2008). *Assessment of children: Cognitive foundations*. San Diego, CA: JM Sattler.
- Semel, E., Wiig, E. H. & Secord, W. A. (2003). *Clinical Evaluation of Language Fundamentals (CELF-4)*. San Antonio, TX: The Psychological Corporation.
- Semel, E., Wiig, E. H. & Secord, W. A. (2013). *Clinical Evaluation of Language Fundamentals (CELF-4) manual (Norsk versjon)*. Stockholm: Pearson Assessment.
- Spaulding TJ, Plante E, Farinella KA. (2006). Eligibility criteria for language impairment: Is the low end of normal always appropriate? *Language, Speech, and Hearing Services in Schools*. 37(1), 61-72.
- Tetzchner, S. v., Feilberg, J., Hagtvet, B., Martinsen, H., Mjaavatt, P. E., Simonsen, H. G., & Smith, L. (1993). *Barns språk (2 utg.)*. Oslo: Ad Notam Gyldendal.
- Tomblin, J.B. (2019). *Developmental language disorder*. I: J.S. Horst & J.v.K. Torkildsen (red.), *International Handbook of Language Acquisition*: Routledge.
- Tomblin, J.B., Records, N.L., Buckwalter, P., Zhang, X., Smith, E. & O'Brien, M. (1997). Prevalence of specific language impairment in kindergarten children. *Journal of Speech, Language, and Hearing Research*, 40(6), 1245-1260.
- Tomblin, J. B., Zhang, X., Buckwalter, P., & O'Brien, M. (2003). The stability of primary language disorder. *Journal of Speech, Language, and Hearing Research*, 46(6), 1283-1296.
- Wilson, A. C., & Bishop, D. V. M. (2021). A novel online assessment of pragmatic and core language skills: An attempt to tease apart language domains in children. *Journal of Child Language*, 1-22.