

“One size does not fit all”

*En studie av språkferdigheter blant
flerspråklige barn og unge*

May-Britt Monsrud



PHD avhandling.
Institutt for spesialpedagogikk
Det utdanningsvitenskapelige Fakultet

Universitetet i Oslo

2022

© **May-Britt Monsrud, 2022**

*Doktoravhandlingar forsvart ved
Det utdanningsvitenskapelige fakultet, Universitetet i Oslo.
Nr. 344*

ISSN 1501-8962

Det må ikke kopieres fra denne boka i strid med åndsverkloven eller med avtaler om kopiering inngått med Kopinor, interesseorgan for rettighetshavere til åndsverk.

Omslag: Hanne Baadsgaard Utigard.
Grafisk produksjon: Representralen, Universitetet i Oslo.

«One size does not fit all»

En studie av språkferdigheter blant flerspråklige barn og unge

Sammendrag

Avhandlingen er skrevet innenfor det spesialpedagogiske fagområdet og er utført ved Institutt for spesialpedagogikk, Det Utdanningsvitenskapelige fakultet, Universitetet i Oslo, Oslo Met, Storbyuniversitetet og Statped.

Et utgangspunkt er at kunnskap om hva som karakteriserer språkferdigheter i flerspråklige utvalg fordrer å undersøke ferdigheter både på første- og andrespråket over et visst aldersspenn. Utvalgene representerer seks språkgrupper i aldersgruppene seks til tretten år. Fem måleinstrumenter undersøker norskspråklige ferdigheter og de samme måleinstrumentene er oversatt og tilpasset til seks førstespråk for å undersøke førstespråket.

Det samlede utvalget har stigende skårer på alle tester ved økende alder på både første- og andrespråket. Dette indikerer ikke bare at de har kompetanse på begge språk, men også at kompetansen på både første- og andrespråket øker med alderen.

Vi undersøkte vokabularferdigheter ut fra en forståelsesramme om at vokabularferdigheter er *distribuerte* mellom første- og andrespråket. Vi undersøkte derfor vokabular ut fra en prosedyre hvor barn og unge fikk uttelling for kompetanse dersom de kan ordet på ett språk: en konseptuell skåre. Forskjellene mellom flerspråklige og enspråklige normutvalg ble vesentlig mindre eller forsvinner helt når en konseptuell skåre beregnes for reseptive vokabularferdigheter. Forskjellene blir også vesentlig mindre for ekspressive vokabularferdigheter. Dette mønsteret er gjeldende i alle seks språkgrupper og i aldersgruppene fra seks til tretten år. Forskjellene mellom flerspråklige og enspråklige normutvalg kan derfor forklares utfra kartleggingsprosedyrer og ikke vokabularkompetanse.

I artikkel 2 blir tester for fonologisk minne (ordspenn og setningsrepetisjon) undersøkt. Utvalget viser signifikant høyere skårer for ordspenn på førstespråket sammenlignet med andrespråket. Denne relative styrken på førstespråket ble identifisert i alle seks språkgrupper. Resultatene for testen setningsrepetisjon avdekker langt på vei de samme forskjellene som vokabularferdigheter i artikkel 1. De to testene for fonologisk minne måler dermed svært ulik kompetanse.

I artikkel 3 undersøkes om testen setningsrepetisjon kan ha kvaliteter som en mulig screeningtest for språkferdigheter. Resultatene viser at setningsrepetisjon predikeres av vokabular, grammatikk og fonologisk minne på både første- og andrespråket. Ordspenn målt på førstespråket predikerer setningsrepetisjon også på andrespråket. Ordspenn synes derfor å fange opp underliggende kognitive ferdigheter av betydning for språkferdighetene både på første- og andrespråket. Resultatene viser at setningsrepetisjon gjennomført på begge språk

kan avdekke risiko for språkvansker, mens enspråklige normer kan innebære en vesentlig feilkilde. Avhandlingen viser at å undersøke andrespråket alene ikke gir tilstrekkelig innsikt i språkkompetansen til flerspråklige barn og unge.

Det er også et hovedfunn at det er signifikant forskjeller mellom språkgruppene hvorvidt de har en styrke på første- eller andrespråket.

Liste over publikasjoner

Artikkel I

The Advantages of Jointly Considering First and Second Language Vocabulary Skills among Emergent Bilingual Children, Monsrud, Rydland, Geva, Thurmann-Moe & Lyster, publisert i *International Journal of Bilingual Education and Bilingualism*, juni 2019.

Artikkel II

Assessing Phonological Memory in Dual Language Learners: A Study of Sentence Repetition and Word Span Tests in First and Second Languages. Monsrud, Lyster, Meyer Bjerkan & Rydland (sendt inn for vurdering i *Journal of Communication Disorders*, 07.09.21) .

Artikkel III

First and Second Language Sentence Repetition - A Potential Screening Measure in Dual Language Learners Monsrud, Rydland, Geva & Lyster (resubmitted i *Language & Education* juni 2021 og akseptert for publisering i august 2021).

Forord

Usikker identifisering av mulige utviklingsmessige språkvansker hos barn og unge med flerspråklig bakgrunn er årsaken til at denne avhandlingen ble realisert. En stor takk til alle som har bidratt inn i dette arbeidet!

Det har vært et privilegium å ha fått veiledning fra to forskere med stort engasjement for klinisk praksis. Jeg vil rette en særlig stor takk til min hovedveileder professor emeritus Solveig-Alma Halaas Lyster som har fulgt prosjektet mitt helt siden starten. Takk for generøs deling av kunnskap og erfaring. Takk for entusiasme og for utallige gjennomlesninger av mine manuskript og for at du har vært uendelig tålmodig med meg. En stor takk til medveileder professor Veslemøy Rydland for din faglige innsikt og klokskap. Din positive interesse har medført at avhandlingen ble slutført. Takk også til professor Esther Geva som er medforfatter på to av artiklene og gjorde det mulig å tilbringe tid i en multikulturell by som Toronto.

Det er mange år siden min dyktige leder Tove Lindqvist bidro til oppstart av prosjektet som er utgangspunkt for avhandlingen. Takk Tove! Takk også til Liv Bøyesen for samarbeid om utvikling av prosjektplaner. Takk til Kisten Meyer Bjerkan og Anne-Cathrine Thurmann-Moe for inspirerende samarbeid om prosjekt og bok. En særlig takk til spesialpsykolog Ernst Ottem for verdifull veiledning samt for tilgang til tester og normdata fra Språk 6-16.

Takk til Institutt for yrkesfaglærerutdanningen ved Oslo Met, ved tidligere instituttleder Sidsel Grande og forskningsleder Olav Eikeland som oppmuntret meg til å søke opptak som ekstern stipendiat ved Utdanningsvitenskaplig Fakultet. Takk for støtte og oppmuntring fra gode kollegaer ved instituttet: Sidsel Sandtrøen, Børge Skåland, Adam Vogt og Kristin Vonheim. Takk også til Nasjonal Forskerskole for lærerutdanning (NAFOL) for interessante seminarer og til NATED for gode forskerkurs. Takk til Anne Arnesen som bidro til at jeg holdt motet og humøret oppe både på forskerkurs og i kloster!

Takk til ledere i Statped, tidligere Bredtvet kompetansesenter som har tilrettelagt for en arbeidssituasjon som gjorde det mulig å slutføre arbeidet med avhandlingen: Anne-Berit Andreassen og Marit Bollingmo.

En stor takk for viktige oppklaringer og konstruktive innspill fra professor Øistein Anmarkrud som var midtveisleser og til 1. amanuensis Thor Arnfinn Kleven for gode råd ved gjennomlesingen av avhandlingen.

Mange elever, skoler og testledere har bidratt i innsamlingen av data og fortjener en takk. Det gjør også oversettere, tilbakeoversettere samt alle som bidro med datainnsamling.

Takk for støtte og kunnskapsdeling kjære kollegaer: Jannicke Karlsen, Marianne Klem, Helene Fulland, NAFO og Anne-Lise Rygvold ved Institutt for spesialpedagogikk. Avhandlingen ble slutført på grunn av dere fire. Takk også til Kirsten Meyer Bjerkan og Stig Bardal for bidrag i sluttfasen.

Familie, venner og kollegaer: takk for positiv støtte og for at dere aldri sluttet å etterspørre arbeidet mitt.

Tora, Ane, Trond, Erik og Live minner meg på at livet er mer enn denne «stilen». Kjæreste Jarl, takk for oppmuntrende «tilrop» og for tålmodighet med utallige utsettelse av levering

Innhold

«One size does not fit all»	iii
Sammendrag	v
Liste over publikasjoner	vii
Forord	viii
1 Innledning.....	1
1.1 Bakgrunn for avhandlingen.....	2
1.2 Kunnskapsbehov	3
1.3 Formål og problemstillinger.....	5
1.4 Begrepsavklaring og begrepsbruk i avhandlingen	6
1.5 Avhandlingens struktur og avgrensinger	9
1.6 Oversikt over artiklene	10
2 Tidligere forskning og teoretisk rammeverk	11
2.1 Faktorer av betydning for utvikling av språkferdigheter hos flerspråklige barn og unge	11
2.1.1 Særlige forhold ved førstespråkutviklingen	11
2.1.2 Andrespråkutviklingen	12
2.2 Distribuerte vokabularferdigheter på tvers av første – og andrespråket.....	14
2.3 «Dualisfjellmodellen» og hypotesen om felles kognitive prosesser.....	15
2.4 Fonologisk korttidsminne	16
3 Developmental Language Disorders - utviklingsmessige språkforstyrrelser.....	19
3.1 DLD i utvalg med flerspråklig bakgrunn	21
3.2 Kartleggingsprosedyrer og identifisering av DLD i utvalg med flerspråklig bakgrunn.....	23
4 Metodiske forhold.....	25
4.1 Utvalg.....	25
4.2 Rekruttering og bakgrunnsinformasjon.....	27
4.3 Tester for vokabular, grammatikk og fonologisk minne.....	32
4.4 Oversettelse og tilpasning til seks førstespråk	34
4.5 Kartleggingsprosedyrer.....	36
4.6 Statistiske analyser	37
5 Oppsummering av hovedfunn i de tre artiklene	41
5.1 Hovedfunn i artikkel 1.....	41
5.2 Hovedfunn i artikkel 2.....	42
5.3 Hovedfunn i artikkel 3.....	43
Kapittel 6 Testteori, reliabilitet og validitet – metodiske drøftinger	45

6.1 Testteori.....	46
6.2 Reliabilitet og systematiske målingsfeil.....	47
6.3 Validitetsutfordringer ved bruk av tester i flerspråklige utvalg.....	49
6.4 Etske overveielser.....	52
Kapittel 7 Oppsummerende diskusjon.....	54
Litteraturliste.....	62
Vedlegg.....	70

1 Innledning

Stor variasjon i språklig bakgrunn samt forskjeller i eksponering for første- og andrespråket gjør det vanskelig å si hva som er en typisk utvikling på første- og andrespråket hos ulike grupper med flerspråklig¹ bakgrunn i Norge. Mange års erfaring fra arbeid med utviklingsmessige språkforstyrrelser² (Developmental Language Disorder/DLD) innen det spesialpedagogiske fagområdet har vist meg at det er stor usikkerhet på ulike nivåer (barnehage, skole og pedagogisk-psykologisk tjeneste/PPT) og i forsknings- og utdanningsinstitusjoner om hva som kjennetegner en typisk språklig utvikling på første- og andrespråket. Noe av denne usikkerheten kan skyldes at et begrenset vokabular og mangelfulle ferdigheter i grammatikk kan opptre i sammenheng med både et andrespråk i utvikling og være symptomer på DLD (Bedore & Pena, 2008; Peña et al., 2016; Pena & Halle, 2011; Pratt et al., 2020). Det er bred enighet om at for flerspråklige barn må vansker påvises på førstespråket for å bekrefte DLD (for eksempel Bishop et al., 2016; Paul et al., 2018). Det er imidlertid en utfordring at det ikke finnes normerte tester for førstespråket. Samtidig er tester som vurderer andrespråkferdigheter utviklet og normert for majoritetsspråklige barn og unge og ikke nødvendigvis valide for flerspråklige barn og unge (Bedore et al., 2012; Geva & Herbert, 2013; Peña, 2008). I tillegg kan en delvis politisert og emosjonell debatt mellom og innad i fagmiljøer knyttet til hvilke faktorer som best mulig bidrar til utvikling av språkkompetanse blant flerspråklige barn og unge og problemstillinger i sammenheng med kartlegging, bidra til å skape ytterligere usikkerhet (Vik, 2017). Samlet er dette faktorer som kan medføre at elever med flerspråklig bakgrunn med spesialpedagogiske behov blir både over- og under-identifisert med DLD. Som følge av dette settes ofte hjelpen inn for sent, og den pedagogiske og spesialpedagogiske tilretteleggingen kan være mindre treffsikker enn tilbudet som gis til enspråklige elever (Paradis et al., 2011).

¹ Flerspråklig blir benyttet i avhandlingen når elever har et annet morsmål enn norsk eller samisk. Det redegjøres mer inngående for begrepet i punkt 1.4

² Den norske betegnelsen for det som tidligere oftest ble betegnet som spesifikke språkvansker er ikke er avklart ennå. Derfor benyttes den engelske betegnelsen Developmental Language Disorder/DLD i avhandlingen. Begrepsbruken i Norge er under utvikling gjennom prosjektet Catalise Norge, som har datainnsamling i et norsk ekspertpanel i 2020/2021. Artikkel om norsk terminologi blir publisert i Logopeden høsten 2021

1.1 Bakgrunn for avhandlingen

Andelen barn og unge som har et annet hjemmespråk enn skolespråket og får opplæring på et andrespråk er økende (OECD, 2020). Denne økningen kan henge sammen med migrasjon på grunn av globale demografiske trender skapt av krig, naturkatastrofer, arbeidsledighet og fattigdom. Det har også sammenheng med migrasjon på grunn av etterspørsel etter spesifikk kompetanse i mottakerlandet (McCabe et al., 2013). Andelen barn og unge med innvandrerbakgrunn øker også i Norge. Den prosentvise andelen (innvandrere og norskfødte med innvandrerforeldre) økte fra 11% i 2010 til 18.4% i 2020. Elevgruppen med innvandrerbakgrunn i Norge er en svært heterogen gruppe, både med hensyn til botid, språkbruksmønstre i familiene, hvor lenge de har vært eksponert for ulike språk, språkferdigheter på to eller flere språk samt sosioøkonomiske forhold (SSB, 2020, Rydland et al., 2020, Karlsen et al., 2019).

Manger studier viser at flerspråklige barn og unge har svakere språkferdigheter på andrespråket enn enspråklige barn og unge (August & Shanahan, 2010; Lervåg & Aukrust, 2010; Snow, 2010). Det er fortsatt usikkert hva som karakteriserer deres språklige ferdigheter og utvikling og det er særlig få studier som inkluderer førstespråket. Mangel på forskning om hva som kjennetegner typiske språkferdigheter på første – og andrespråket i flerspråklige utvalg, gjør at vi vet for lite om hvordan flerspråklige barn og unge vanligvis utvikler språkferdighetene sine. Derfor er det også ufullstendig kunnskap om hvilke språkferdigheter som bør medføre ekstra støtte og oppfølging av enkeltelever.

Det er en spesialpedagogisk utfordring at flerspråklige barn og unge er overrepresentert innenfor det spesialpedagogiske hjelpeapparatet. Det kan for eksempel skyldes at andrespråkferdigheter under utvikling mistolkes som DLD (Bedore & Pena, 2008; Geva & Farnia, 2017; Nordahl & Hausstätter, 2009; Solari et al., 2014). Både internasjonalt og i Norge rapporteres det imidlertid også om en *underrepresentasjon* av flerspråklige barn og unge innenfor det spesialpedagogiske hjelpeapparatet. Forskere hevder for eksempel at DLD ikke blir oppdaget, fordi vanskene feilaktig tolkes som usikre ferdigheter på andrespråket (Hakansson et al., 2003; Paul et al., 2018; Aagaard, 2011).

Et utgangspunkt for denne avhandlingen er at flerspråklige barn og unge utvikler seg på to språk samtidig, og at vi trenger å forstå hvordan dette spiller inn på dagens kartleggingspraksis knyttet til å identifisere barn og unge som trenger spesialpedagogisk hjelp.

Klinisk praksis gjorde meg oppmerksom på mangel på tester for å undersøke førstespråket. Med utgangspunkt i et opplevd behov i hjelpeapparatet og skolen tok jeg i 2006 initiativ til et prosjekt for å utvikle kartleggingsmateriell for flerspråklige elever. Dette initiativet resulterte i et tiltak i regi av Kunnskapsdepartementet «Språklige minoriteter med behov for særskilt tilrettelagt opplæring» (tiltak 13)

i forbindelse med strategiplanen «Likeverdig opplæring i praksis» (Kunnskapsdepartementet, 2007). Jeg utformet en prosjektbeskrivelse og ledet tiltaket som seniorrådgiver ved Bredtvet kompetansesenter. Tiltaket tok utgangspunkt i tester som fortsatt benyttes i forskning og klinisk praksis. Testene som ble valgt var: *British Picture Vocabulary Scale* (BPVS-II), (Dunn et al., 1997) og Begreper fra *Språk 6-16* (Ottem & Frost, 2005), en grammatikktest (*Test for Reception of Grammar* (Trog-2) (Bishop, 1989; Lyster & Horn, 2009) samt Setningsrepetisjon og Ordspenn fra *Språk 6-16* Ottem & Frost, 2005). Disse testene ble oversatt og tilpasset til seks førstespråk: albansk (Kosovo), somali, tamil (Sri Lanka), tyrkisk, urdu (Pakistan) og vietnamesisk. Testene ble prøvd ut på førstespråket for 648 elever (jenter = 344, gutter = 304) og for andrespråket norsk for 783 elever på (jenter = 399, gutter = 383). Ikke alle elever som ble testet på førstespråket ble testet på andrespråket, og motsatt. Det var derfor ikke testresultater for alle elever på begge språk (upublisert prosjektrapport, Monsrud, Bjerkan & Thurmann-Moe 2016). I avhandlingen inngår kun flerspråklige elever med testresultater fra både første- og andrespråket, i aldersgruppene 6 -13 år. En nærmere beskrivelse av tiltaket foreligger i vedlegg C. Til forskjell fra tiltaket i regi av Kunnskapsdepartementet, inkluderer avhandlingen testdata for et flerspråklig utvalg med utviklingsmessige språkforstyrrelser (DLD) og enspråklige normdata for de fem testene. Flerspråklige utvalg med typisk utvikling sammenlignes med enspråklig normutvalg i de tre artiklene.

1.2 Kunnskapsbehov

Avhandlingen bidrar på tre områder der det er særlig behov for mer kunnskap knyttet til flerspråklige barn og unges språkferdigheter. Jeg vil innledningsvis beskrive hvordan denne avhandlingen vil møte disse behovene.

1) Behov for studier som inkluderer elevers språkferdigheter på første- og andrespråket

Innenfor pedagogisk og spesialpedagogisk forskning, har det i de senere tiårene vært en økning av studier som inkluderer barn og unge med flere språk (Bedore & Pena, 2008; Bedore et al., 2012; Peña, 2016; Zebib et al., 2019). Mye av denne forskningen har ferdigheter på andrespråket som fokus, og bekrefter dermed oftest kunnskap vi har hatt i lengre tid, nemlig at når flerspråklige elever kartlegges på andrespråket, vil de som gruppe prestere svakere enn enspråklige elever (Gunnerud, 2021; Heller, 2020; Melby-Lervåg & Lervåg, 2014). Forskningen knyttet til hvordan flerspråklige elever utvikler språklige ferdigheter på tvers av sine to (eller flere) språk, er mer begrenset (Gross et al., 2014; Oh & Mancilla-Martinez, 2021). I tillegg har mange av de studiene som faktisk inkluderer både første- og andrespråkferdigheter små utvalg innenfor et kort aldersspenn (for eksempel Thordardottir et al., 2006).

Et utgangspunkt for denne avhandlingen er at for å få kunnskap om hva som karakteriserer språkferdigheter, må vi undersøke ferdigheter både på første- og andrespråket over et visst aldersspenn. Det kan gi oss en indikasjon på hvordan første- og andrespråkferdighetene varierer blant elever som får

sin opplæring på andrespråket. Jeg vil imidlertid understreke at den typen tverrsnittstudie jeg har gjennomført her, ikke kan si noe om individuelle utviklingsløp på samme måte som longitudinelle studier der det foretas mange målinger av de samme barna over tid. Imidlertid medførte designet at det var mulig å samle data fra større grupper og å undersøke variasjon i flerspråklige elevers språkkompetanse på ulike alderstrinn.

2) Behov for studier som ser på undergrupper av flerspråklige elever

En utfordring for forskning på språkkompetanse hos elever med flere språk er at det ofte ikke tas høyde for at elevene kan være svært ulike. Det er ofte slik at til tross for at flerspråklige elever tilhører språkgrupper med svært ulik immigrasjonshistorie og ulike språkbruksmønstre, blir de behandlet som én populasjon fordi de har samme andrespråk (Geva et al., 2019; O'Connor et al., 2019). Det finnes flere studier av noen få utvalgte språkgrupper, som for eksempel elever med spansk som førstespråk og engelsk som andrespråk i USA (Bedore & Pena, 2008; Bedore et al., 2012; Peña et al., 2016). Mange i denne populasjonen har mer krevende levekår i USA enn de fleste flerspråklige i Norge. Selv om studiene kan kaste lys over utvikling som også er relevant for grupper av norske elever med ulike førstespråk, kan det være visse begrensninger i overføringsverdi. De norske studiene som har undersøkt førstespråket i flerspråklige utvalg, har i hovedsak vært begrenset til to språkgrupper, tyrkisk (Fulland, 2016; Rydland et al., 2013) og urdu/panjabi (Fulland, 2016; Karlsen et al., 2016). I tillegg har en studie av barnehagebarn fra 2020 inkludert flere ulike førstespråk (Rydland et al. 2020). Det er et stort behov for mer kunnskap om språkferdigheter hos flerspråklige barn og unge i Norge med forskjellige førstespråk og fra ulike aldersgrupper. Dersom det er systematiske forskjeller mellom språkgrupper, for eksempel knyttet til om elevene har sin styrke på første- eller andrespråket, kan det være av betydning for både kartlegging- og tiltakstenkning.

3) Behov for studier med utgangspunkt i data fra kartleggingsprøver som kan bidra til å identifisere språkvansker.

Med mer kunnskap om flerspråklige barn og unges språkkompetanse, har vi også et bedre utgangspunkt for å forstå og også avdekke utviklingsmessige språkforstyrrelser som DLD. I norsk kontekst er dette spesielt viktig fordi Opplæringsloven skiller mellom elevers rettigheter til særskilt språkopplæring, hjemlet i § 2-8: *Særskild språkopplæring for elevar frå språklege minoritetar (Lov om grunnskolen og den vidaregåande opplæringa (opplæringslova) - Kapittel 2. Grunnskolen1 - Lovdata)*³ og rett til spesialundervisning basert på blant annet lærevansker som DLD, hjemlet i § 5-1. Basert på den store

³ « Elevar i grunnskolen med anna morsmål enn norsk og samisk har rett til særskild norskopplæring til dei har tilstrekkeleg dugleik i norsk til å følgje den vanlege opplæringa i skolen»

variasjonen i språkferdigheter blant flerspråklige barn og unge, er det komplisert å utvikle kartleggingsverktøy med normer tilpasset den store heterogeniteten. Det er derfor behov for mer kunnskap om hvilken informasjon tester for vokabular og grammatikk på første- og andrespråket gir i ulike språkgrupper. Det er også stort behov for mer kunnskap om hvordan tester som antas å være mindre avhengig av språkferdigheter slår ut i flerspråklige grupper. Det gjelder blant annet tester av fonologisk minnefunksjon og setningsrepetisjon.

1.3 Formål og problemstillinger

Det primære formålet med avhandlingen er å bidra til forskningsbasert kunnskap om språklige ferdigheter på første- og andrespråket blant flerspråklige elever i barneskolealder i Norge.

Studiens overordna problemstilling er:

Hva karakteriserer språkferdighetene på første- og andrespråket til flerspråklige barn og unge?

Samlet belyser de tre artiklene ulike aspekter ved den overordna problemstillingen og følgende forskningsspørsmål:

- 1) Hvilken kunnskap kan tester for vokabular, grammatikk, setningsrepetisjon og fonologisk minne på første- og andrespråket gi oss om flerspråklige elevers språkferdigheter?
- 2) Hvilke likheter og forskjeller i språkferdigheter identifiseres i de seks språkgruppene?
- 3) Hvilken betydning kan tospråklige kartleggingsprosedyrer ha for identifisering av utviklingsmessige språkforstyrrelser?

Artikkel 1 undersøker hvilke reseptive og ekspressive vokabularferdigheter elever i barneskolealder fra ulike språkgrupper har, på første- og andrespråket, samt om vokabularferdighetene er distribuert mellom første – og andrespråket. I likhet med annen forskning sammenlignes vokabularferdigheter på andrespråket med enspråklige normutvalg.

Kapasiteten i fonologisk minne antas å være sentral for utvikling av språkferdigheter, og kan undersøkes ved hjelp av ulike måleinstrumenter. I artikkel 2 benyttes to ulike måleinstrument: ett måleinstrument som antas å være et sammensatt mål for språk og minne -*Setningsrepetisjon*- og ett måleinstrument som anta å være et «renere» mål på fonologisk minne, -*Ordspenn*-. Artikkel 2 belyser resultater for slike tester når de benyttes på første- og andrespråket i flerspråklige utvalg. Resultatene for det flerspråklige utvalget sammenlignes med norske normer.

Artikkel 1 og 2 belyser også forskjeller mellom testene og likheter eller forskjeller mellom språkgruppene. Resultater fra de to artiklene belyser dermed også om språkferdighetene varierer avhengig av hvilket språklige domene (vokabular eller fonologisk minne) som blir undersøkt.

I artikkel 3 undersøker jeg et bredt spekter av språktester på første- og andrespråket. Basert på forskning som indikerer at setningsrepetisjon er en god markør for å identifisere DLD, undersøkes spesifikt hvilke språklige komponenter -vokabular, grammatikk og minne- som predikerer variasjon i resultater på denne testen. I artikkelen inngår alle de seks språkgruppene i *ett utvalg* som også sammenlignes med ett flerspråklig utvalg med identifisert DLD, fra de samme, men ikke alle, språkgruppene. Målsettingen er å undersøke om setningsrepetisjon er et sensitivt mål for å undersøke variasjon i språkferdigheter samt om den kan bidra i identifisering av mulige språkforstyrrelser hos flerspråklige barn og unge.

Tabell 1 Oversikt over artikler, utvalg, variabler og tester

Artiklene	Flerspråklig utvalg og variabler	Enspråklig normutvalg og variabler	Tester som inngår
The Advantages of Jointly Considering First and Second Language Vocabulary Skills among Emergent Bilingual Students	Reseptivt vokabular Ekspressivt vokabular (N=542)	Reseptivt vokabular (N=495) Ekspressivt vokabular (N= 900)	BPVS-II Begreper fra Språk 6-16
Assessing Phonological Memory in Dual Language Learners: A Study of Sentence Repetition and Word Span Tests in First and Second Languages.	Setningsrepetisjon Ordspenn (N=590)	Setningsrepetisjon Ordspenn (N= 900)	Setningsrepetisjon Ordspenn fra Språk 6-16
First and Second Language Sentence Repetition - A Potential Screening Measure in Dual Language Learners	Setningsrepetisjon Ordspenn Grammatikk Ekspressivt vokabular Reseptivt vokabular (N=546) (N= 14)	Setningsrepetisjon Ordspenn Ekspressivt vokabular (N=900) Grammatikk Reseptivt vokabular(N=495)	Setningsrepetisjon Ordspenn Ekspressivt vokabular/Begreper fra Språk 6-16 Trog-2 BPVS-II

1.4 Begrepsavklaring og begrepsbruk i avhandlingen

Begreper på norsk som favner en flerspråklig populasjon

Et mye benyttet begrep for grupper som har tilgang til flere språk er *bilingual/tospråklig* (Grosjean, 2008). Det synes å være enighet om at tospråklig inkluderer de som benytter seg av to (eller flere) språk i dagliglivet (for eksempel Grosjean, 1998). Selv om personer betegnet som tospråklige mestrer to (eller flere) språk, varierer både språkkompetanse og språkbruk for de to (eller flere) språkene

de benytter (Montrul, 2016). Språklige behov vil ikke være de samme på begge (alle) språk til enhver tid. Det språklige repertoaret kan dermed også endres over tid (De Houwer, 2007; Hoff, 2014). Dette resulterer ofte i svært ulike språklige profiler for kompetanse for første- og andrespråket (Cummins, 2000; Montrul, 2013, 2016). Det kan stilles spørsmål ved om tospråklige med samme kompetanse på begge språk, såkalt balansert tospråklige, i realiteten finnes (Baker, 2014). Slike forhold kan bidra til at det i noen tilfeller kan være vanskelig å avgjøre om enkeltindivider best kan beskrives som en-, to-, flerspråklig eller minoritetsspråklig.

Det er vanlig å skille mellom ulike typer tospråklige ut fra lengde for eksponering for aktuelle språk. Barn som blir eksponert for to språk fra fødselen betegnes som simultant tospråklige, mens barn som har vært eksponert for andrespråket fra 3-års alder ofte betegnes som sekvensielt tospråklige (De Houwer, 2007; Hoff, 2014). Et begrep som ofte benyttes synonymt med sekvensielt tospråklige er *andrespråkbruker*. Mange flerspråklige i Norge blir eksponert for andrespråket mer systematisk ved oppstart i barnehage og vil da, i tillegg til å betegnes som flerspråklige, også kunne betegnes som sekvensielt tospråklige. Likevel blir mange allerede fra fødsel eksponert for norsk tale, blant annet i butikken, på gata, i parken og fra media. Det kan derfor stilles spørsmål ved om eksponering for andrespråket startet ved 3-års alder eller tidligere, selv om den mer systematiske eksponeringen kanskje startet ved 3-års alder. Dette er et eksempel på at inndelingen i sekvensielt og simultant tospråklig på mange måter er kunstig og en forenkling. Det er også et eksempel på en forenklet begrepsbruk på et komplekst felt.

Med *minoritetsspråklige* elever menes elever med et annet morsmål enn norsk og samisk (Meld. St. 6 (2012-2013)). Begrepet minoritetsspråklige elever benyttes også i Opplæringsloven. Samtidig er begrepet foreslått erstattet av betegnelse to- eller flerspråklig (<https://www.regjeringen.no/no/dokumenter/stmeld-nr-23-2007-2008-/id512449/>). Begrunnelsen for dette er å synliggjøre at kompetanse i mer enn ett språk er en ressurs. Begreper som to- eller flerspråklig⁴ kan imidlertid også underkommunisere utfordringer ved å ha et andrespråk i en minoritetssituasjon. Begrepet minoritetsspråklig kommuniserer et asymmetrisk forhold mellom første- og andrespråket ved at tilegnelse av norsk er en forutsetning for læringsutbytte i norsk skole, og dagens lovverk og praksis kommuniserer langt på vei at andrespråket er mest sentralt å utvikle (Opplæringsloven § 2-8). Flerspråklig utvikling i en minoritetssituasjon er ikke det samme som flerspråklig utvikling for grupper hvor førstespråket har en høyere status eller har bedre sosioøkonomiske betingelser. Eksempler

⁴ For enkelthets skyld benytter jeg flerspråklighet om to språk (førstespråk og andrespråk), selv om det i realiteten kan være flere

på slike betingelser er tospråklige samfunn som i Canada hvor både engelsk og fransk er offisielle språk eller tospråklige i tospråklige familier med to høystatusspråk.

Flere forskere påpeker at betydningen av et godt førstespråk som en mulig ressurs i opplæringen underkommuniseres i både forskning og lovverk (Arntzen & Karlsen, 2019; Montrul, 2016). Samtidig blir språklige ferdigheter blant flerspråklige barn og unge vurdert ut fra et enspråklig grunnlag og sammenliknet med en enspråklig norm selv om både første – og andrespråket har andre utviklingsbetingelser for flerspråklige enn for enspråklige (blant annet (Bedore et al., 2012; Bohman et al., 2010). En slik vurderingspraksis synes også å være tilfelle i en norsk kontekst hvor språklige forskjeller i mange tilfeller blir tilskrevet vansker hos den enkelte, uten at samfunnsmessige faktorer som levekår blant innvandrere eller skolens enspråklige praksis inngår i vurderingen (Arntzen & Karlsen, 2019; Øzerk, 2008; Aagaard, 2011). I den senere tid er en transspråklig forståelse av flerspråklighet løftet fram (Wei & Auer, 2009). Det språklige repertoaret, som består av to eller flere språk, gjør det mulig for flerspråklige individer å tilpasse språkbruken til ulike situasjoner (Li Wei, 2007). Ved å forstå flerspråklighet som noe dynamisk vektlegges også at språk ikke utvikler seg lineært, og at språk ikke fungerer separat, men at språk kan utvikle seg og endres ut i fra hvilken situasjon det skal brukes i (García, 2009; García & Li Wei, 2019).

Begrepsbruken i kapp og i artiklene

Selv om *minoritetsspråklige* blir ofte benyttet i stortingsmeldinger og offentlige dokumenter, velger jeg å benytte begrepet *flerspråklig* som hovedbegrep i kapp, fordi litteraturen som er inkludert ikke bare omhandler de som har et førstespråk som er et minoritetsspråk. *Flerspråklig* overlapper derfor også i noen tilfeller med begrepet *tospråklighet* som jeg har redegjort for innledningsvis, men *flerspråklig* er valgt gjennomgående i hele kapp for å gjøre teksten mest mulig tilgjengelig, selv om mitt utvalg også kan karakteriseres som minoritetsspråklige. I kapp anvender jeg begrepet *førstespråk* (*first language, L1*) synonymt med hjemmespråk eller morsmål, slik det gjøres i Stortingsmelding nr. 23 « Språk bygger broer» (Kunnskapsdepartementet, 2008) og Stortingsmelding «Tett på---» (Kunnskapsdepartementet, 2010-2020). *Andrespråk* (*second language, L2*) anvendes synonymt med norsk når jeg refererer internasjonal forskning som belyser tilegnelse av andrespråket i etterkant av førstespråket (*jf. sekvensiell tospråklighet*).

Majoritetsspråklige eller *enspråklige* anvendes for individer som benytter det offisielle språket i det aktuelle landet samt for de norske normeringsutvalgene i avhandlingen.

Jeg har brukt «*emergent bilingual students*» og «*dual language learners*» om flerspråklige barn og unge i artiklene i avhandlingen. Begrepet «*second language learners*» ble ofte anvendt tidligere (se

August & Shanahan, 2010) men betegnelsen er i økende grad erstattet av «*emergent bilingual students*» (e.g. ref) eller «*dual language learners*» (for eksempel Hammer et al., 2014). Betegnelsene «*emergent bilingual students*» og «*dual language learners*» viser til at språkferdigheter på første- og andrespråket er under utvikling og reflekterer at utvalgene har tilgang til to språk uten at det er klart definert hvilken språkkompetanse språkferdighetene på første- eller andrespråket skal representere. Begrepene indikerer heller ikke at andrespråket kommer etter førstespråket og jeg oppfatter dem som mindre statiske begreper enn minoritetsspråklig og flerspråklig. Under arbeidet med artikkel 1 ble *emergent bilingual students* oftest benyttet i forskningslitteraturen, og derfor anvendt. Betegnelsen ble etter hvert som arbeidet skred fram erstattet med «*dual language learners*» av flere forskere og er benyttet i artikkel 2 og 3. Et slikt skifte i begrepsbruk kan reflektere det faktum at forskningsfeltet tospråklig på mange måter er et relativt ungt forskningsfelt der begrepsbruken fortsatt er i utvikling, men også at jeg har utviklet bedre kjennskap til forskningslitteraturen og funnet min posisjon i fagfeltet gjennom arbeidet med avhandlingen.

1.5 Avhandlingens struktur og avgrensinger

Avhandlingen består av en kappe og tre artikler. Kappa belyser tospråklig utvikling, empiri om språkferdigheter og fonologisk minne og DLD i utvalg med flerspråklig bakgrunn samt problemstillinger i sammenheng med kartlegging. Kappa inkluderer også metodiske forhold som utvalg, problemstillinger i forbindelse med oversettelse av tester, begrunnelse for analysemetoder, samt metodiske drøftinger i forbindelse med testteori og validitet samt implikasjoner for klinisk praksis.

Den andre delen består av tre artikler. Den første og andre artikkelen belyser forskningsspørsmål 1 ved å undersøke reseptive og ekspressive vokabularferdigheter (artikkel 1) samt setningsrepetisjon og fonologisk minne (artikkel 2) i ulike populasjoner og alderskohorter, og sammenligner andrespråket med enspråklige normdata. Begge artiklene bidrar også til å belyse forskningsspørsmål 2: hvordan tester avdekker forskjeller og likheter mellom språkgrupper. Artikkel 3 besvarer forskningsspørsmål 3 ved å undersøke hvorvidt setningsrepetisjon predikeres av vokabular, grammatikk og fonologisk minne på første- og andrespråket, og hvorvidt et slikt instrument kan bidra til å avdekke at det grunn til bekymring for svake språkferdigheter og gi indikasjoner på DLD.

1.6 Oversikt over artiklene

I. The Advantages of Jointly Considering First and Second Language Vocabulary Skills among Emergent Bilingual Children (Monsrud, Rydland, Geva, Thurmman-Moe & Lyster publisert i *International Journal of Bilingual Education and Bilingualism* juni 2019).

II. Assessing Phonological Memory in Dual Language Learners: A Study of Sentence Repetition and Word Span Tests in First and Second Languages. Monsrud, Lyster, Meyer Bjerkan & Rydland (sendt inn for vurdering i *Journal of Communication Disorders*, 07.09.21).

III. First and Second Language Sentence Repetition - A Potential Screening Measure in Dual Language Learners Monsrud, Rydland, Geva & Lyster (resubmitted i *Language & Education* juni 2021 og akseptert for publisering i august 2021).

2 Tidligere forskning og teoretisk rammeverk

Utvalgene i avhandlingen har norsk som sitt andrespråk. I første del av dette kapitlet vil jeg redegjøre for faktorer av betydning for første- og andrespråktilegnelse i en minoritetssituasjon. Disse faktorene kan ha betydning for å tolke og forstå resultater på språktester på både første- og andrespråket. Deretter vil jeg gjøre rede for begrepet *distribuerte ferdigheter* (Cobo-Lewis et al., 2002; Oller & Eilers, 2002; Pearson et al., 1993). Dette begrepet belyser at språkferdigheter blant flerspråklige barn og unge består av mer enn andrespråkferdigheter og er sentralt i alle de tre artiklene. Deretter, presenteres «*dualisjellmodellen*» og «*The Underlying Common Cognitive Processes Hypothesis*», som viser til sammenhenger mellom første- og andrespråket via et felles språklig fundament (Cummins, 1984, 2000, 2017, 2021). Fonologisk korttidsminne er et sentralt tema for artikkel to og tre og blir gjort rede for avslutningsvis i kapitlet.

2.1 Faktorer av betydning for utvikling av språkferdigheter hos flerspråklige barn og unge

I dette avsnittet vil jeg belyse noen trekk ved språkutvikling og språkbruk hos personer som har hovedspråket i samfunnet som et andrespråk, slik som for barna og ungdommene i denne avhandlingen.

Spedbarn synes å ha en medfødt evne til å diskriminere språklyder uavhengig av språk. Denne sensitiviteten forutsetter at eksponeringen for ulike språk opprettholdes, og fra seks måneders alder reduseres sensitiviteten for språklyder som barnet ikke eksponeres for. Barn som eksponeres for to språk opprettholder imidlertid evnen til lydidentifikasjon på to språk, så lenge de blir eksponert for begge språk (Hammer et al., 2014; Pierce et al., 2015). Dette er en antatt fordel ved en tospråklig utvikling, og det spekuleres i hvorvidt lydidentifikasjon er en kompetanse som bidrar til større sensitivitet for språk blant tospråklige enn blant enspråklige (Barac et al., 2014).

Både språkutvikling generelt og et førstespråk- og et andrespråk i utvikling påvirkes av komplekse forhold som kvaliteter ved språkstimuleringen, antall år og timer med eksponering for hvert av språkene samt sosioøkonomisk status (for eksempel Hoff, 2014). Noen slike forhold ved første- og andrespråkutviklingen blir belyst videre.

2.1.1 Særlige forhold ved førstespråkutviklingen

Førstespråkutviklingen er først og fremst avhengig av dialogpartnere som gjør det relevant å benytte førstespråket (Grosjean, 2008). Studier fra norsk kontekst viser at språkbruksmønstre i flerspråklige familier ofte endres ved at førstespråket i mindre grad benyttes når andrespråket blir tilegnet i barnehagen eller på skolen (Arntzen & Karlsen, 2019; Rydland & Grøver, 2021). Intervjuer med 10-

åringene med pakistansk/norsk og tyrkisk/norsk språkbakgrunn (N=56) avdekket stor variasjon i språkbruksmønstre i familier med like lang botid i Norge og viser at familienes språkbruksmønstre besto av ulike kombinasjoner av førstespråket og andrespråket norsk (Fulland, 2016). For noen av informantene var førstespråket svært viktig, mens andre var mindre opptatt av førstespråkets betydning siden de bor i Norge (Fulland, 2016). Slike studier illustrerer en kompleksitet i språkbruksmønstre i flerspråklige familier i norsk kontekst. Å rapportere om selvopplevde språkbruksmønstre i familien og blant venner, er en vanlig metode for å beskrive førstespråket blant flerspråklige.

Språkbruksmønstre vil kunne endre seg i løpet av et livsløp, for eksempel ved at en person som er dominant på førstespråket i fire års alder, kan være dominant på andrespråket i ti-års alder (Grosjean, 2010). Utviklingen og opprettholdelse av førstespråket, og kompetansen på førstespråket, kan også i noen grad avta over generasjonene. Etter en eller to generasjoner vil ofte andrespråket ta over som familiens hovedspråk (Montrul, 2016), og få en stadig viktigere plass i takt med at familien blir mer aktive deltagere i samfunnet (skole, arbeid osv.). Det er ikke uvanlig at barn benytter førstespråket i kommunikasjon med foreldre og besteforeldre, og andrespråket i kommunikasjon med søsken. I en norsk studie som innbefattet språkbruksmønstre blant flerspråklige barnehagebarn, benyttet 90% av foreldrene førstespråket sitt overfor barna, samtidig som 60% av barna adresserte foreldrene på norsk (Rydland & Grøver, 2021)

Samtidig kan språkgruppetilhørighet ha betydning for hvorvidt bruk av førstespråket opprettholdes. Forskningsresultater fra andre europeiske land som Nederland og Tyskland, viste for eksempel at flerspråklige med tyrkisk som førstespråk, i større grad opprettholdt førstespråket enn andre språkgrupper (Extra & Yagmur, 2010).

2.1.2 Andrespråkutviklingen

Andrespråket følger i hovedsak samme utvikling som førstespråkutviklingen. De språklige milepælene nås i samme rekkefølge som i førstespråkutviklingen (Paradis, 2019b; Smithson et al., 2014). Likevel kan milepælene tilegnes senere på både første- og andrespråket enn hos enspråklige (Hoff, 2014). Det er ikke alltid sammenfall mellom hva personer kan uttrykke på sine to språk og enkelte strukturer kan mestres på det ene språket, men ikke på det andre. Bruk av førstespråkets fonologiske, syntaktiske og morfologiske regler kan forekomme i innlæringsfasen av andrespråket, fordi andrespråket tilegnes ut fra språklige milepæler og strukturer som allerede er tilegnet på førstespråket. Språkklydmønstre, grammatiske strukturer, morfologi og ordstilling kan være svært forskjellig i forskjellige språk, og utviklingen i de ulike språklige domeneene kan derfor variere også ut fra språkspesifikke forhold

(Leonard, 2014).

For å tilegne seg uttale og grammatikk på andrespråket på samme nivå som førstespråkbrukere, kan det være en fordel å tilegne seg et andrespråk tidlig (Svendsen 2021). En svensk studie av et spansk/svensk utvalg (N=195) som tilegnet seg svensk på ulike alderstrinn (<1- 47år), konkluderte med at det er vanskelig å lære seg «innfødtlik» uttale av andrespråket ved eksponering etter 12 års alder (Abrahamsson & Hyltenstam, 2009). Slike studier kan imidlertid ta fokus vekk fra at forståelsen av innhold er en sentral språklig komponent for læring. En god uttale kan bidra til å maskere utfordringer med språkforståelse og vokabular og medføre at flerspråklige eleveres læringsutfordringer ikke blir oppdaget. Eldre barn og voksne kan ha kommet lenger i sin kognitive og språklige utvikling og dermed ha et bedre grunnlag for å tilegne seg andrespråkets innhold og struktur enn yngre barn (for eksempel (García & Baker, 2007; Smithson et al., 2014)

Selv om alder for eksponering andrespråket ofte løftes fram som en sentral faktor for utvikling av andrespråkets fonologi, slik som studien til Abrahamsson og Hyltenstam (2009) viser, er det ikke ensidig støtte for at tidlig eksponering for andrespråket alltid er en fordel (Paradis et al., 2019; Smithson et al., 2014). En slik studie viste at barn av immigranter (N=169) som ble eksponert for andrespråket engelsk etter fire års alder, utviklet mer komplekse syntaktiske konstruksjoner sammenlignet med barn av immigranter som ble eksponert for andrespråket i to til fire års alder (Govindarajan & Paradis, 2019; Paradis et al., 2019). Slike studier viser, ifølge Paradis, at aldersfaktoren ikke er tilstrekkelig til å forklare andrespråkutviklingen.

Mange med flerspråklig bakgrunn vokser i tillegg opp i omgivelser med svakere sosioøkonomiske betingelser. Dette er også tilfelle i Norge hvor en nylig publisert rapport viste at 50% av fattige barn vokser opp i flerspråklige familier (Epland & Normann 2020/SSB). Svakere sosioøkonomiske betingelser kan i seg selv predikere risiko for svakere språkferdigheter sammenlignet med barn og unge med bedre levekår (Blom et al., 2014; de Abreu et al., 2012; Hammer et al., 2014; Vagh et al., 2009). Det er velkjent at barn og unge med utfordrende sosioøkonomiske betingelser kan ha mindre stimulerende språkmiljø og at slike betingelser har stor betydning for språkutviklingen (Hart & Risley, 1995). I avhandlingen undersøkes ikke demografiske forhold direkte. Det kan allikevel være viktig med en bevissthet om slike faktorer fordi de kan bidra til å forstå hvorfor eventuelle systematiske gruppeforskjeller i andrespråkferdigheter kan oppstå. Landbakgrunn kan ha betydning for resultatene på vokabulartester, generelt. En metaanalyse med fokus på språk- og leseferdigheter fant blant annet mindre forskjeller

mellom en- og flerspråklige utvalgs vokabularferdigheter i Canada sammenlignet med forskjeller mellom en – og flerspråklige grupper i Europa og USA (Melby-Lervåg & Lervåg 2014).

Utvalg med ulik språkbakgrunn blir ofte behandlet som en populasjon i forskning fordi de har samme andrespråk. Som vist over kan det være store variasjoner på tvers av ulike språkgrupper og resultater fra slike studier kan tilsløre mange feilkilder ved en slik tilnærming. Avhandlingen søker derfor å identifisere variasjon i språkferdigheter i både første – og andrespråket i seks språkgrupper med flerspråklige bakgrunn.

Det kunne også vært informativt å undersøke mulige språkspesifikke forskjeller mellom språkgruppene i form av kontrastive analyser som at språk som er svært forskjellige er vanskelige å lære sammenlignet med språk som har mange likhetstrekk. Men en spesifikk sammenligning mellom de seks språkene har ikke vært tema for denne avhandlingen. Videre blir ikke demografiske forhold undersøkt direkte, men bevissthet om hvordan slike faktorer kan bidra til eventuelle systematiske gruppeforskjeller i andrespråkferdigheter har vært sentralt i tolkningen av resultatene.

2.2 Distribuerte vokabularferdigheter på tvers av første – og andrespråket

Avhandlingen undersøker vokabularferdigheter på første – og andrespråket i flerspråklige utvalg. Det er imidlertid viktig å understreke at flerspråklige barn og unges språkferdigheter ikke består av to ulike enspråklige «sett» av vokabular, men av vokabular hvor begge språk, på ulik måte, inngår. Vokabularferdigheter er altså distribuert mellom første- og andrespråket ved at noen ord er kjent på førstespråket og andre ord er kjent kun på andrespråket, selv om mange også er kjent på begge språk. Begrepet distribuerte ferdigheter har bakgrunn i en studie av et spansk-amerikansk utvalg der ferdigheter på både første- og andrespråket ble undersøkt «Miami-studien» (Cobo-Lewis, 2002 ;Oller, 2007). Studien viste at elevene hadde kjennskap til en betydelig andel ord på kun ett av språkene, selv om kjennskap til flere ord til en viss grad overlappet mellom spansk og engelsk. Dersom vokabularferdigheter ble beregnet basert på testresultater på ett språk, skåret utvalget svakere enn normutvalget på både engelsk og spansk. Til forskjell fra studier som sammenligner flerspråklige barn og unges vokabularferdigheter på andrespråket med enspråklige tester og normer, ble reseptive vokabularferdigheter på både engelsk og spansk undersøkt ved tilnærmet like tester for begge språk; Peabody Picture Vocabulary Test (Dunn & Dunn 1997) og Test de Vocabularionen Imagenes Peabody (Dunn & Dunn 2002). Elevene hadde kjennskap til noen ord på spansk og andre ord på engelsk. Forskjellene mellom spansk/engelsk og enspråklig engelsk utvalg ble forklart ut fra at enspråklige kartleggingsprosedyrer ikke tar høyde for at vokabularferdigheter er distribuert mellom første– og andrespråket. «Miami-studien» har hatt stor

betydning ved å dokumentere at ferdigheter på ett språk ikke gir et fullstendig bilde av språkkompetanse i flerspråklige utvalg og at en sammenlikning med en enspråklig norm kan være en vesentlig feilkilde.

Distribuerte vokabularferdigheter er undersøkt i flere empiriske studier og de bekrefter i stor grad funnene fra Miamistudien; at forskjellene mellom vokabularferdigheten hos enspråklige og flerspråklige utvalg reduseres betydelig ved å ta høyde for språkferdigheter på begge språk (Bedore et al., 2005; Core et al., 2013; Gross et al., 2014; Mancilla-Martinez & Vagh, 2013; Oller et al., 2007a; Pearson et al., 1993; Peña et al., 2016). Distribuerte vokabularskårer er undersøkt ved foreldrevurderinger på spansk og engelsk (Mancilla-Martinez & Vagh, 2013) og med vokabulartester i typisk utviklede spansk/engelske utvalg (Bedore et al., 2005). Distribuerte ferdigheter er også undersøkt ved å sammenligne reseptive og ekspressive vokabularferdigheter i ulike utvalg; enspråklig engelsk, simultant flerspråklig spansk/engelsk og sekvensielt flerspråklig spansk/engelsk (Gross et al., 2014). Samlet viser disse studiene at distribuerte ferdigheter er signifikant bedre sammenlignet med å undersøke andrespråket. En nyere studie fra Sør-Afrika av to flerspråklige utvalg med ulik afrikansk/engelsk bakgrunn viste at distribuerte reseptive og ekspressive vokabularskårer var signifikant bedre sammenlignet med vokabularskårer beregnet ved hjelp av enspråklige testprosedyrer (Jordaan et al., 2021).

Samlet viser studiene at distribuerte reseptive og ekspressive vokabularferdigheter er betydelig bedre enn vokabularferdigheter beregnet utfra enspråklige kartleggingsprosedyrer. Felles for studiene er små utvalg i et lite aldersspenn og dermed kan resultatene innebære noe usikkerhet. Til forskjell fra de refererte studiene, undersøkes distribuerte ferdigheter i et relativt stort flerspråklig utvalg i aldersgruppen fra seks til tretten i år i artikkel 1.

2.3 «Dualisfjellmodellen» og hypotesen om felles kognitive prosesser

I pedagogisk sammenheng har Cummins modeller hatt stor innflytelse og bidratt med et rammeverk for forskning rundt flerspråklig utvikling (Cummins, 2000). Modellene er også benyttet i argumentasjonen for tospråklige undervisningsmodeller (Øzerk, 2008). I avhandlingen har «dualisfjellmodellen» og «The Underlying Common Cognitive Processes Hypothesis», heretter kalt hypotesen om felles kognitive prosesser, relevans. For det første kan modellene illustrere at under gitte forutsetninger har utvikling av førstespråket betydning for utvikling av andrespråket og visa versa. For det andre kan de illustrere at utvikling av språkferdigheter på første- og andrespråket har sammenheng med kognitive prosesser som er felles for begge språk. Det er imidlertid viktig å understreke at Cummins modeller er lite spesifikke når det gjelder sammenhengen mellom ulike språklige ferdigheter.

Ifølge «dualisfjellmodellen», har språkferdigheter utviklet på et førstespråk betydning for utvikling av et andrespråk fordi det er en gjensidig avhengighet mellom de to språkene. Språkferdigheter som er lært i en førstespråk-kontekst kan bidra til utvikling av andrespråket og motsatt. En gjensidig positiv utvikling mellom første- og andrespråket er imidlertid også avhengig av kvaliteter ved språkmiljøet, kvaliteter ved undervisning, samt motivasjon hos den som skal lære språk. Modellen er ofte tatt til inntekt for at førstespråkutviklingen har betydning for andrespråkutviklingen, selv om det motsatte også er tilfelle ifølge Cummins (Cummins, 2017). Ifølge Cummins kan gjensidig påvirkning mellom første- og andrespråket bare fullt ut realiseres innenfor tospråklige undervisningsmodeller og det er derfor begrensninger ved undervisningsmodeller som ensidig bidrar til å utvikle ett språk (Cummins, 2021).

En gjensidig avhengighet mellom første- og andrespråket indikerer ikke nødvendigvis overføring av språklige ferdigheter, men at førstespråket bidrar til utvikling av andrespråket gjennom et felles språklig og kognitivt fundament. All språklæring som forekommer via første- eller andrespråket, bidrar til å utvikle og styrke fundamentet, og ifølge Cummins inngår abstrakt tenkning, problemløsning, resonnering, metakognisjon og metaspråklig bevissthet i et felles språklig og kognitivt fundament (Cummins, 2017). «Dual- isfjellmodellen» og *hypotesen om felles kognitive prosesser* får fram at kognitive prosesser er nært knyttet til erfaringer med og eksponering for språk (Cummins, 2021). Hypotesen om felles kognitive prosesser er diskutert og også kritisert for å være lite eksplisitt på hvilke spesifikke faktorer som bidrar til å støtte opp under forholdet mellom første- og andrespråkutviklingen. Flere forskere har presisert at faktorer som er mindre avhengig av språkferdigheter slik som kapasitet i fonologisk korttidsminne, arbeidsminne og informasjonsprosesser er spesielt interessante for forskning i forbindelse med flerspråklige barn og unges språkferdigheter (August & Shanahan, 2008, 2010; Geva, 2014; Geva & Wiener, 2014).

2.4 Fonologisk korttidsminne

Det fonologiske korttidsminnet (eller fonologisk minne) representerer en kapasitet som støtter bearbeidelsen av og fastholder språklig informasjon i kortere tidsrom og er antatt å være døråpneren for lagring av fonologiske enheter. Denne komponenten som fastholder informasjon i kort tid antas å ha en sentral rolle i tilegnelse av nye ord. Begrenset kapasitet i denne komponenten er fremmet som en mulig forklaring på begrensninger i språklæring blant barn og unge med DLD (Gathercole & Baddeley, 1990). Mye forskning innenfor tospråklighetsfeltet har etter hvert undersøkt nettopp evnen til å fastholde fonologisk informasjon (fonologisk minne), men også evnen til å fastholde og manipulere informasjon (arbeidsminne) (Blom et al., 2014; Calvo & Bialystok, 2014; de Abreu et al., 2012; Engel et al., 2008; Gunnerud et al., 2020). Årsaken er at både fonologisk korttidsminne og arbeidsminne antas å være mindre

avhengig av språkferdigheter enn vokabular og grammatikk. Dette er også bekreftet empirisk, og på tester som måler fonologisk korttidsminne skårer flerspråklige elever tilsvarende, (for eksempel Calvo & Bialystok, 2014; de Abreu et al., 2013; Shany & Geva, 2012) eller høyere (Morales et al., 2013) enn enspråklige.

Selv om begrepene fonologisk korttidsminne og arbeidsminne antas å reflektere ulike kognitive funksjoner, er de til dels overlappende og det er ikke enkelt å trekke et klart skille mellom dem. Det er ikke tilstrekkelig avklart hvordan bakenforliggende faktorer som oppmerksomhet og prosesseringshastighet (hvor rask informasjon bearbeides), har innflytelse på både fonologisk korttidsminne og arbeidsminne (Aben et al., 2012). Likevel benyttes begrepet fonologisk korttidsminne oftest når det refereres til å huske språklig informasjon i kort tid, og begrepet arbeidsminne refererer oftest til å fastholde og i tillegg manipulere informasjon i kort tid (Aben et al., 2012). En arbeidsminnemodell som har hatt stor innflytelse på forskningen knyttet til fonologisk minne er Baddeley og Hitch sin modell (Baddeley & Hitch, 2000). Denne modellen er et teoretisk utgangspunkt for to av testene i avhandlingen; Ordspenn og Setningsrepetisjon. Modellen beskriver en eksekutiv funksjon som styrer tre undersystemer med begrenset lagringskapasitet. De tre undersystemene omtales som den visuospatiale skisseblokken, den episodiske buffer og den fonologiske løkken. Den visuospatiale skisseblokken er et midlertidig lager som fastholder og bearbeider visuell informasjon, som antas å ha mindre betydning i språklæring. Den episodiske buffer representerer den delkomponenten av hukommelsessystemet som integrerer fonologisk informasjon fra det midlertidige minnet med leksikalsk og semantisk informasjon fra langtidsmindet. Den fonologiske løkken omtales ofte synonymt med fonologisk korttidsminne eller fonologisk minne. Som nevnt fastholder det fonologisk korttidsminne informasjon i kort tid og antas å ha en sentral rolle i tilegnelse av nye ord (Gathercole & Baddeley, 1990). Kapasitet til å fastholde språklig informasjon i kort tid er blitt undersøkt både i typisk utviklede flerspråklige utvalg (de Abreu et al., 2013; Gunnerud et al., 2020) og blant flerspråklige med mulige språkvansker (Bedore & Pena, 2008; Kohnert, 2010; Kohnert et al., 2010) og resultatene samlet indikerer mindre eller ingen forskjeller mellom flerspråklige og enspråklige utvalg.

Det har vært mye forskning om mulige kognitive fordeler av en flerspråklig utvikling for kompleks problemløsning, som skifte av oppmerksomhet mellom konkurrerende stimuli presentert simultant eller sekvensielt og metaspråklig bevissthet (Calvo & Bialystok, 2014; de Abreu et al., 2012; Morales et al., 2013). Eventuelle fordeler antas å ha sammenheng med mestring og kontroll av ulike språklige systemer som kan ha effekt på eksekutive funksjoner. Det er imidlertid relativt få studier som finner en slik fordel. Det er også rettet kritikk mot disse studiene utfra at utvalgene ofte har vært små (Von Bastian et al., 2017). Det er et stort behov for mer forskning knyttet til sammenhengen mellom de ulike komponentene i

hukommelsesmodeller og hvordan de er relatert til språklæring (de Abreu et al., 2012; Kohnert, 2010; Paradis, 2016; Peña et al., 2018).

Tester som antas å være mindre avhengig av språkferdigheter

Det er flere typer tester som har til formål å undersøke kognitive prosesser som i ulik grad er relatert til øvrige språkferdigheter. Ordspenn og tallminne forlengs (innebærer at henholdsvis tall eller ord skal gjentas i den rekkefølge de er presentert) er eksempler på tester som gjerne benyttes for å undersøke fonologisk korttidsminne. Ordspenn tester er bygget opp av meningsbærende enkeltord og antas å være mer språkavhengig enn tester av tallminne forlengs (Buac et al., 2016; Hulme et al., 1997; Melby-Lervåg & Hulme, 2010; Ottem & Lian, 2008). Også tester for nonord: å gjenta ord uten betydning i rekkefølge, benyttes i mange studier av flerspråklige utvalg fordi slike tester vurderes å være lite avhengig av øvrige språkferdigheter (Bele et al., 2008; Messer et al., 2010). Ordspenn testene for første- og andrespråket blir diskutert i artikkel 2 og 3.

Setningsrepetisjonstester antas å være et sammensatt mål på mer generelle språkferdigheter som inkluderer flere språkkomponenter; blant annet grammatisk kunnskap og vokabular, i tillegg til fonologisk minne (se f.eks. Klem et al., 2015; Tuller et al., 2018). Tester for setningsrepetisjon vurderes imidlertid å være mindre avhengig av språkferdigheter enn språktester som direkte måler mer avgrensede språklige delkomponenter som vokabular eller grammatikk (Klem et al., 2015; Zebib et al., 2019). Selv om det ikke er tilstrekkelig avklart hva setningsrepetisjonstester måler, viser flere studier av flerspråklige utvalg at slike tester måler både en språkkomponent og en minnekomponent (Pratt et al., 2021; Riches, 2012; Zebib et al., 2019). Setningsrepetisjon på første – og andrespråket i en norsk kontekst er tema i artikkel 3.

Oppsummering – kapittel 2

I dette kapitlet har jeg redegjort for flere faktorer som kan ha betydning for både første og andrespråkutviklingen. Videre har jeg redegjort for begrepet som viser til at språkferdigheter i flerspråklige utvalg er distribuert mellom første og andrespråket. Forholdet mellom første- og andrespråket er belyst med utgangspunkt i to modeller av Cummins. Det er også referert til forskning som har undersøkt forhold ved fonologisk korttidsminne mer spesifikt.

I kapittel 3 vil jeg redegjøre nærmere for hvorfor et skille mellom utviklingsmessige språkforstyrrelser og et andrespråk i utvikling kan være komplisert-

3 Developmental Language Disorders - utviklingsmessige språkforstyrrelser

Med utgangspunkt i over- og underidentifisering av DLD hos flerspråklige barn og unge, vil jeg i dette kapitlet belyse DLD generelt og løfte fram særlige utfordringer ved identifisering blant flerspråklige samt gjøre rede for hvordan sentrale forskere avdekker DLD i flerspråklige utvalg.

Anslagsvis inntil 10% av alle barn har ved skolestart uventede eller uforklarte vansker med utvikling av språk og kommunikasjon sammenlignet med utviklingen på andre områder. DLD er imidlertid en relativt ukjent vanske sammenlignet med for eksempel dysleksi eller ADHD, som i likhet med DLD kan medføre vansker med læring (Bishop, 1998; Bishop et al., 2016; Bishop et al., 2017; Hulme & Snowling, 2009).

I mange år har det innenfor miljøer for forskning- og klinisk praksis vært diskutert hvordan uventede vansker med språket skal betegnes og identifiseres. Den internasjonale forskningslitteraturen har fram til 2016 benyttet delvis overlappende, men ikke helt identiske betegnelser for språkvansker, som for eksempel; «language difficulties», «language disorders», «language impairment», «language disability» eller «language delay», med prefikser som «developmental», «specific» or «primary» (Bishop, 2014; Reilly et al., 2014).

Med bakgrunn i den eksisterende uklarheten knyttet til terminologi og identifisering av språkvansker, er det gjennomført et omfattende forskningsarbeid som involverer sentrale engelskspråklige fagmiljøer og engelskspråklige forskere. Forskningsarbeidet skal utvikle felles terminologi, inklusjonskriterier samt kliniske kjennetegn for vasken. Dette forskningsarbeidet er kalt Catalise studiene og arbeidet er blant annet publisert i to konsensusartikler (Bishop et al., 2016; Bishop et al., 2017). Et argument for å igangsette Catalise-studiene er, at på tross av omfattende forskning, har mangel på enhetlig begrepsbruk medført uklarhet, særlig for identifisering av vanskene Det har vært uklarheter med hensyn til om forskningen har belyst samme fenomen. Mangel på entydige kriterier for identifisering av vansker kan også ha medført at barn og unge med behov for systematisk språkpedagogisk støtte ikke har mottatt et slikt tilbud.

I norsk kontekst har det mest brukte begrepet for slike språkvansker vært spesifikke språkvansker (SSV). Siden bakgrunnen for «Catalise» studiene nettopp har sammenheng med at vanskene med språk ikke er så spesifikke som tidligere betegnelser kan tilsi, har sentrale fagmiljøer i Norge etablert «Catalise Norge» for å utvikle norske begreper og kriterier. De første resultatene fra undersøkelsen blant tretti

sentrale forskere og klinikere publiseres høsten 2021. Som gjort rede for innledningsvis er det derfor ennå ikke avklart hvordan det engelske begrepet for DLD bør tilpasses norske forhold. En mulighet kan være at DLD innarbeides som begrep også på norsk på samme måte som andre innarbeidede engelske begrep, som for eksempel ADHD.

Det er enighet i de engelskspråklige miljøene om at språkforstyrrelser som medfører utfordringer med kommunikasjon, og læring og som vil ha betydning for og konsekvenser for livsutfoldelse og livskvalitet, kan kalles Developmental Language Disorder (DLD). Språkvanskene har langvarig og vedvarende betydning for språklig fungering og læring. Vanskene kan komme til uttrykk i ulike språklige domener som språkets form, innhold og bruk (Bloom & Lahey, 1978; Paul et al., 2018). Form-siden i språket kan rammes ved problemer med å oppfatte og etablere språklydmønstre eller problemer med å tolke og forstå grammatiske uttrykk, morfologiske og syntaktiske regler samt verbbygninger. Innholdssiden rammes ved at DLD kan innebære ordletingsvansker eller et begrenset vokabular. DLD kan medføre at regler for kommunikasjon ikke blir forstått eller gi utfordringer med å produsere strukturerte historier og kan slik også ramme bruken av språket. DLD kjennetegnes også av begrensninger i kapasiteten i fonologisk minne. Slike utfordringer kan komme til uttrykk ved vansker med å huske lyder eller ord i kort tid, problemer med å etablere assosiasjoner mellom ord og mening og utfordringer med statistisk læring. Kapasiteten i fonologisk minne antas å ha betydning for språktilegnelse (se kapittel 2). Når barn og unge skal utredes for DLD, er det nødvendig med kartlegging av de ulike språkkomponentene som barn med DLD kan streve med. Men også kartlegging av fonologisk minne som kan undersøkes ved å benytte tester som krever repetisjon av nonord og setninger (Bishop et al., 2016; Bishop et al., 2017). Barn med DLD vil ha behov for å få gjentatt språklige mønstre flere ganger enn barn uten språkforstyrrelser for at mønstrene skal lagres og læres (Paul et al., 2018).

Tidligere ble DLD identifisert gjennom å undersøke om resultatene på en eller flere språktester var svakere enn en «cut off» grense på -1.25 SD på én språktest eller en «cut off» grense på -1.00 SD på flere språktester (Bishop, 1998; Leonard, 2014). Hvis kun *en* test ble benyttet var «cut off» kriteriet altså strengere (-1.25 SD) enn om *flere* tester ble anvendt (-1.00 SD). Det er imidlertid oppnådd enighet i fagmiljøene om at de skisserte kriteriene har vært anvendt for restriktivt og medført at barn, med behov for språkpedagogisk tiltak, ikke har fått den hjelpen de har hatt behov for. Fagmiljøene anbefaler derfor at den enkeltes behov for tilpasning bør tillegges større vekt enn tidligere slik at tilrettelegging utfra individuelle forutsetninger blir iverksatt (Paul et al., 2018).

Språkvansker har også vært identifisert ved å avdekke et misforhold mellom språkferdigheter og kognitive forutsetninger (vurdert med evneprøver). På tross av uenighet i Catalise-panelet ble det etter hvert oppnådd enighet om å fjerne et slikt kriterium. Hovedargumentet er at svakere non-verbal kognisjon synes å være en del av det kliniske bildet ved DLD (Gallinat & Spaulding, 2014). Årsaken kan være at språkferdigheter støtter problemløsning også for såkalte nonverbale evneprøver og at språk ikke er løsrevet fra kognitiv utvikling (Snow, 2017).

Enkelte yngre barn har en sen språkutvikling uten at dette er en risiko i seg selv. Disse barna betegnes som «late-talkers». Det har vist seg vanskelig å predikere hvilke barn som vil få vedvarende vansker med språk fordi språklig normalutvikling blant barn under 5 år karakteriseres av svært stor variasjon i språkferdigheter. Denne store variasjonen i normalfordelingen blant de yngste barna bidrar til at det ikke er enkelt å avklare hvilke barn som er innenfor eller utenfor en språkvanskekategori. Mange barn som har en langsom språkutvikling i tidlig alder sammenlignet med jevnaldrende, har tatt igjen jevnaldringsgruppen ved skolestart. Et inklusjonskriterium er derfor at språkvanskene vedvarer utover 5 års alder (Paul et al., 2018).

Selv om språkmiljø og sosioøkonomiske forhold har innflytelse på språkutviklingen, ansees ikke dette til å være årsak til DLD (Paul et al., 2018). En stor utfordring er hvordan DLD kan avdekkes i utvalg hos flerspråklig barn, fordi, som nevnt i innledningen vil et andrespråk i utvikling, der ikke minst vokabular kan være begrenset, delvis ha liknende språklige strukturer som de man ser hos barn med DLD. De spesifikke vanskene kan derfor være komplisert å identifisere (Paradis, 2019b; Paradis et al., 2019). En tilnærming hvor både første- og andrespråket undersøkes antas med større sikkerhet å avklare eventuelle forskjeller på symptomer på DLD og et andrespråk i utvikling. Det er enighet om at det ikke er tilstrekkelig å undersøke bare andrespråket for å identifisere DLD.

3.1 DLD i utvalg med flerspråklig bakgrunn

Flere studier avdekker svært usikker identifisering av mulige språkforstyrrelser blant barn og unge med flerspråklig bakgrunn. I en studie i USA fant man en under-diagnostisering i utvalg med flerspråklig bakgrunn i tidlig skolealder, men en overidentifisering seinere i skoleløpet (Solari et al., 2014). Tilsvarende tendens er avdekket i Canada, mens resultatene fra en svensk studie viste at arabisk/svenske barn ble henvist senere (etter fylte 5 år) samt at de hadde mer alvorlige språkforstyrrelser enn enspråklige (Geva & Massey-Garrison, 2013; Hakansson et al., 2003). I Norge er det ikke gjennomført nyere undersøkelser av elever med flerspråklig bakgrunn henvist til spesialundervisning for språkrelaterte vansker. En mulig overrepresentasjon i spesialundervisning er dokumentert i en studie fra

Oslo. Der man fant at 10.2% elever med flerspråklig bakgrunn og 7.7% elever med enspråklig bakgrunn mottok spesialundervisning (Nordahl & Hausstätter, 2009).

Som tidligere påpekt er det komplisert å identifisere DLD hos flerspråklige barn fordi kjennetegn på et andrespråk i utvikling til dels kan overlapse med symptombildet på DLD. Svake skårer på språktester for vokabularferdigheter og grammatiske ferdigheter på andrespråket er ikke tilstrekkelig til å identifisere DLD. Mangel på tester for å undersøke førstespråket medfører også at det kan være komplisert å undersøke hva som er det sterkeste eller dominante språket dersom språkkompetansen består av to språk.

Catalise studiene konkluderer med at DLD rammer begge språk. En mulig avklaring av DLD må baseres på mangel på aldersadekvate ferdigheter på de språk som er i bruk. Det er enighet i de fleste fagmiljøer om at språklig kompetanse på to (eller flere språk) språk bør inngå i kartleggingsprosedyrer (Bishop et al., 2017; Kohnert et al., 2010; Paradis, 2011). Samtidig skal det påpekes at selv om svake språkferdigheter avdekket på begge språk kan være et uttrykk for en språkvanske, kan det også henge sammen med lite eksponering for begge språk sammenlignet med den eksponeringen enspråklige barn har for sitt ene språk.

Studier av utvalg med DLD, både med flerspråklig og enspråklige bakgrunn, har vist at å vurdere morfologi, syntaks og spesielt verbbygninger kan bidra til å identifisere mulig DLD (Paradis, 2016; Simonsen & Bjerkan, 1998). Likevel er det viktig å merke seg at barn med DLD også kan ha gode skårer på slike prøver.

Svake vokabularferdigheter kan være en markør for DLD, men forskjeller i vokabular mellom utvalg med flerspråklige og enspråklig bakgrunn er også dokumentert i mange studier av typisk utviklede flerspråklige utvalg og vokabular er derfor den minst pålitelige språklige indikatoren for DLD hos flerspråklige barn (Lervåg & Aukrust, 2010; Rogde et al., 2016; Salameh, 2012). Som kapittel 2 viser til, er forskjellene i vokabular mellom utvalg med flere- og enspråklig bakgrunn betydelig mindre eller forsvinner helt ved en kontekstuell skåre (Buac et al., 2016; Gross et al., 2014; Oller et al., 2007b).

Et sentralt symptom for identifisering av språkforstyrrelser som DLD i tillegg til forenklete grammatiske konstruksjoner og begrenset vokabular, er begrensninger i kapasiteten i fonologisk minne. Dette er ansett som særlig sentralt i utredning av flerspråklige utvalg fordi kapasiteten i fonologisk minne antas å være mindre avhengig av språkferdigheter og sosiale forhold slik kapittel 2 viser (Thordardottir & Brandeker, 2013; Thordardottir et al., 2006). Nonordrepetisjonsprøver er antatt å være et relativt «rent»

mål på fonologisk minne og kan bidra til å avdekke en mulig underliggende vansker. Flere studier viser at resultater fra en slik prøve kan bidra til skille mellom utvalg med flerspråklig bakgrunn med og uten DLD (Chiat et al., 2013; Chiat & Polišenská, 2016; Thordardottir & Brandeker, 2013).

Setningsrepetisjon er ofte inkludert i utredningsprosedyrer for å avdekke mulig DLD også i utvalg med flerspråklige bakgrunn og vurdert som en klinisk markør for å identifisere mulig DLD. Samtidig viser flere empiriske studier at tester for setningsrepetisjon er mindre presise for å avklare og identifisere mulig DLD i utvalg med flerspråklig bakgrunn sammenlignet med å benytte en slik test for å avklare DLD i utvalg med enspråklig bakgrunn (Chiat et al., 2013; Polišenská et al., 2015; Pratt et al., 2021; Tuller et al., 2018; Zebib et al., 2019). Årsaken kan være at setningsrepetisjon reflekterer en generell språkkompetanse i tillegg til fonologisk minne og dermed avdekker språkferdigheter mer enn å avdekke fonologisk minne kapasitet (Kohnert, 2013).

3.2 Kartleggingsprosedyrer og identifisering av DLD i utvalg med flerspråklig bakgrunn

Sen språkutvikling, språkvansker i familien og svak mestring av førstespråket kan være faktorer som kan gi grunnlag for bekymring og for at potensielle språkforstyrrelser blir gjenstand for nærmere kartlegging (Geva & Wiener, 2014; Paradis, 2016; Thordardottir & Brandeker, 2013). En bekymringsfull eller en forsinket utvikling på ett eller flere områder avdekket av foreldre eller lærer, vil derfor ofte være utgangspunkt for henvisning til hjelpetjenesten, som pedagogisk -psykologisk tjeneste (PPT) og/eller andre instanser.

Kartlegging forutsetter en bred tilnærming hvor flere typer informasjonskilder, som strukturerte intervjuer, observasjoner av læringsmiljø og sosialt samspill, inngår. Også mer uformelle kartleggingsprosedyrer, som vurdering av skriftlige arbeider, spørreskjema til foreldre og lærere og ikke minst samtale med barnet selv, bør inngå, i tillegg til normerte tester (Sattler, 2008). En slik bred tilnærming blir redegjort for og anbefalt i Catalise studiene (Dorothy Bishop et al, 2017).

Informasjon fra foreldre om førstespråkutviklingen ved hjelp av intervju er i mange tilfeller det nærmeste en kan komme evaluering av førstespråket. Foreldreintervju er prøvd ut systematisk i flere studier. (Armon-Lotem et al, 2015; Paradis et al, 2010). Disse gir empirisk støtte for at foreldreintervju gir valid informasjon om førstespråkkompetanse (Bishop et al, 2017). Informasjon fra foreldre er også sentralt også for å avdekke en mulig arvelig komponent for språkforstyrrelser.

Kvaliteten på tester og kartleggingsmateriale er avgjørende for et valid og reliabelt utgangspunkt

for å vurdere elevenes mulige språkforstyrrelser. Mange og relativt små språkgrupper gjør at det er svært krevende å utvikle prøver for ulike førstespråk. Oversettelse av tester slik at de måler de samme ferdighetene på begge språk er komplisert og dette blir drøftet videre i kapittel 4 og 6.

I mangel av tester på førstespråket benyttes ofte tester for andrespråket tilpasset til flerspråklige med ulike prosedyrer. I noen studier undersøkes grammatiske ferdigheter på andrespråket dersom andrespråket benyttes av barna mer enn 40% av tiden (Paradis, 2016). I en svensk studie benyttes andrespråktester for grammatikk dersom det aktuelle barnet har vært eksponert for andrespråket i minst to år (Hakansson et al., 2003). I en annen studie har utvalget vært eksponert for andrespråket engelsk i tre år (Thordardottir, 2015). Kartlegging av grammatikk kompliseres også ved at grammatikk kartlagt på førstespråket påvirkes av andrespråkets grammatikk. Språktypologiske forskjeller kan ha innflytelse på grammatisk utvikling på andrespråket (Leonard, 2014). Som vist har det vært benyttet svært ulike måter å tilpasse tester for grammatikk på og dette medfører usikkerhet med hensyn til om studiene avdekker det samme.

Siden resultater på andrespråktester er vanskelig å tolke for flerspråklige barn og unge, er det foreslått at en mulig bekymringsgrense kan være en cut -off på -1.75 når tester for andrespråket blir benyttet (Thordardottir, 2015). Thordardottir argumenterer også naturlignok for at det dominante språket skal være utgangspunkt for å vurdere språkkompetansen (Thordardottir, 2016; Thordardottir & Topbaş, 2021). Denne tilnærmingen er likevel komplisert å realisere fordi en hovedproblemstilling vil ofte å være å avdekke hvilke språk som er det dominante. En slik tilnærming er derfor også svært kritisert for å være unøyaktig (Geva & Farnia, 2017; Paul et al., 2018).

Som gjort rede for tidligere i kapitlet 3 er det konsensus om at DLD diagnosen forutsetter symptomer på begge språk. Det betyr at identifisering basert på svake andrespråkferdigheter ikke er tilstrekkelig for å avdekke mulige vansker med språk (Thordardottir, 2016; Vagh et al., 2009). Som gjort rede for er dette kapitlet er det viktig å utvikle tester som kan benyttes for førstespråket og andrespråket, når mulig DLD skal identifiseres hos flerspråklige barn og unge.

4 Metodiske forhold

I dette kapitlet presenteres utvalgene som inngår i avhandlingen, bakgrunnsinformasjon om utvalgene, testene som er benyttet, tilpasningen av testene til førstespråket, prosedyrer for innsamling av data samt begrunnelse for de statistiske analysene som er benyttet i artiklene. Reliabilitet og validitet blir diskutert mer inngående i kapittel 6.

Språkgruppene i avhandlingen

Som nevnt er det inkludert data fra 6 ulike språkgrupper. Språkgruppene i avhandlingen med lengst botid i Norge har bakgrunn fra Pakistan og Tyrkia. Gruppene kom til Norge om arbeidsinnvandrere på 1960 -tallet. Begge gruppene har relativt god kontakt med opprinnelseslandet ved hyppige besøk. Det er ikke uvanlig at ekteskap inngås med personer fra opprinnelseslandet (Østby 2016). Et foreldrepar kan dermed representere første, andre eller tredje generasjon immigranter og bidra til at førstespråket opprettholdes som familiespråk.

De første båtflyktningene fra Vietnam ankom Norge i 1975 og har også lang botid her i landet. Mange med bakgrunn fra Vietnam har opprettholdt førstespråket gjennom sosialt og språklige fellesskap, blant annet i den katolske kirke og sommerskole for barn og unge i regi av kirken (Pettersen & Østby, 2014).

Krigen på Balkan på 80-tallet førte til at albanske flykninger ankom Norge. Deres mulighet til å opprettholde førstespråket er også gjennom sosiale og religiøse møteplasser.

Krisen på Sri Lanka førte til at tamilske flykninger ankom i den samme perioden. I denne språkgruppen er det sterk tradisjon for opprettholdelse av førstespråket og det er opprettet tamilske ressursentre over hele Norge. Sentrene tilbyr lørdagsskole med undervisning på førstespråket, fritidstilbud hvor barn og unge blir eksponert for tamilsk førstespråk samt leksehjelp for norsk.

Språkgruppen med bakgrunn fra Somalia og norskfødte med innvandrerforeldre fra Somalia, er den tredje største gruppen innvandrere i Norge og er den gruppen i avhandlingen med kortest botid (Engebrigtsen & Fuglerud, 2009; Østby, 2016). Helt siden Somalia ble selvstendig i 1960 har i særlig grad den sørlige delen av landet vært preget av konflikt og krig (Leerand, 2014).

4.1 Utvalg

Det er to utvalg flerspråklige elever i avhandlingen; et utvalg med antatt typisk språkutvikling og et utvalg med DLD. De flerspråklige utvalgene sammenlignes med to ulike enspråklige normutvalg; utvalget fra normeringen av BPVS-II og Trog-2 (N= 495) og utvalget fra normeringen av Språk 6-16 (N=900), (Bishop, 1989; Dunn et al., 1997; Karlsen et al., 2016; Lyster, 2009; Ottem & Frost, 2005). Tre

av testene i avhandlingen- Setningsrepetisjon, Ordspenn og Begreper- er fra Språk 6-16, en screeningtest for å avdekke mulige språkvansker. Språk 6-16 er mye benyttet innenfor det norske skolesystemet. Begreper blir i avhandlingen og artiklene kalt ekspressivt vokabular (N=900).

Flerspråklig utvalg med antatt typisk utvikling er inkludert i artikkel 1 (N=542), artikkel 2 (N=590) og artikkel 3 (N= 546) og består av elever med testresultater på begge språk. Forskjeller i N i artiklene kan forklares ut fra at var ønskelig å inkludere størst mulig utvalg på de aktuelle variablene og også utfra sykdom eller mangel på testledere på det aktuelle språket og i noen få tilfeller flytting til en annen skole. Førstespråkversjonene for tyrkisk for prøvene ordspenn og ekspressivt vokabular var noe forsinket på førstespråket og det var derfor noen i utvalget som ikke ble testet med de to prøvene. Utvalget med tyrkisk som førstespråk var imidlertid såpass stort at vi ikke vurderte dette som et problem. Flerspråklig utvalg med påvist DLD er inkludert i artikkel 3 (N=14).

Utvalget med typisk utvikling benytter, ut fra lærervurderinger, både første- og andrespråket aktivt. Elever som er henvist til hjelpeapparatet eller elever hvor det er mistanke om spesialpedagogiske behov, er ikke inkludert i utvalget. Utvalget er fra skoler med høy andel elever med flerspråklig bakgrunn generelt og fra bomiljøer og nabolag med et større antall elever med samme førstespråk sammenlignet med naboskolene. Generell informasjon om demografiske forhold i de aktuelle oppvekstmiljøer gjør det sannsynlig at utvalgene har mer krevende sosiale og økonomiske oppvekstforhold enn normutvalgene de sammenlignes med (Bakken & Hyggen, 2018; Statistics Norway, 2016). Utvalget er dermed ikke et tilfeldig utvalg, og funnene som rapporteres i avhandlingen kan ikke uten videre generaliseres til disse språkgruppene i Norge eller til populasjonen flerspråklige utvalg mer generelt (se også kapittel 6.). Tabellen viser oversikt over utvalg alder og kjønn.

Tabell 4.1 Alder og kjønn for utvalget med typisk utvikling

Alder	Jenter	Gutter	TOTAL
6-0-6.11	28	20	48
7.0-7.11	42	44	86
8.0-8.11	47	37	84
9.0-9.11	40	3	93
10.0-10.11	66	45	111
11.0-11.11	37	23	60
12.0-12.11	34	30	64
TOTAL	294	252	546

Kjønnforskjeller ble undersøkt ved å benytte uavhengige t-tester for reseptivt og ekspressivt vokabular. Det er ikke signifikante kjønnforskjeller for reseptivt vokabular for førstespråket ($t = -1.32, p = 0.431$) eller for andrespråket ($t = 0.31, p = 0.753$). Det er en signifikant forskjell i favør av jentene for

ekspressivt vokabular ($t = -1.77$, $p = 0.000$), men det er ikke signifikante forskjeller mellom gutter og jenter for ekspressivt vokabular for andrespråket ($t = 0.314$, $p = 0.753$).

Et utvalg flerspråklige elever henvist til og utredet av pedagogisk-psykologisk tjeneste inngår i artikkel 3 ($n=14$). Elevene ble henvist til PPT på grunn av bekymring for språkferdigheter fra foreldre eller lærer. Pedagogisk- psykologisk tjeneste (PPT) i to store kommuner på Østlandet ble invitert til å benytte de aktuelle førstespråktestene for elever henvist PPT for mistanke om DLD fra de samme språkgruppene som det typisk utviklede utvalget (vedlegg G). Utvalget ble undersøkt med de aktuelle testene på første- og andrespråket, i tillegg til ordinær utredning i regi av PPT.

Utvalget med DLD, som vist i Tabell 4.2, består av 5 jenter og 9 gutter og har bakgrunn fra fem av språkgruppene som er inkludert i avhandlingen: Pakistan = 3, Somalia =1, Sri Lanka = 5, Tyrkia 4 og Vietnam = 1. Som tabellen viser er elevene henvist PPT relativt seint, etter fylte 9 år.

Tabell 4.2 Aldersgruppe og kjønn for utvalget med DLD

Alder	Jenter	Gutter	TOTAL
9.0-10.11	3	3	6
11.0-12.11	2	6	8
TOTAL	5	9	14

4.2 Rekruttering og bakgrunnsinformasjon

Det typisk utviklede utvalget ble rekruttert fra skoler med relativt stor andel flerspråklige elever fra de aktuelle språkgruppene. Skolene ble kontaktet på telefon og spurt om de ønsket å delta i utprøving av testene. Skoler som var positive mottok nærmere informasjon og foreldre ble informert på foreldremøter. Foreldre som var positive til at barna deltok mottok deretter skriftlig informasjon samt samtykkeerklæring via skolene (vedlegg E). Informasjon og samtykkeerklæring er utarbeidet på norsk og oversatt og tilpasset aktuelle førstespråk (vedlegg E). Det ble understreket at det var frivillig å delta og at deltakerne kunne trekke seg underveis. Samtykkeerklæring ble utarbeidet og innsamling av data er godkjent av konsesjon samt Norsk samfunnsvitenskapelig datatjeneste (vedlegg A og B). Utvalget som var henvist for mulig DLD ble rekruttert gjennom pedagogisk-psykologisk tjeneste (PPT). PPT ble forespurt om de ønsket å delta på utprøvingen av BPVS II, Trog-2, Ordspenn, Setningsrepetisjon og Ekspressivt vokabular på førstespråket og PPT fikk da mulighet til å inkludere tester på både første – og andrespråket i sin ordinære utredning (vedlegg B). Tester for førstespråket var tilleggsinformasjon til PPTs øvrige kartlegging og observasjoner. Foreldre ble forespurt om de ønsket at barnet skulle delta i utprøvingen og foreldre som var positive mottok samtykkeerklæring. Disse var ikke tilpasset aktuelle

førstespråk fordi en henvisning til PPT forutsetter nært samarbeid med foreldrene. PP-rådgivere involverte tolk eller morsmåls lærer t i de tilfeller det var nødvendig (vedlegg B og G).

Bakgrunnsinformasjon om utvalget med antatt typisk utvikling

Det ble utviklet et spørreskjema til foreldre om botid i Norge, språkbruksmønstre i familien, deltakelse i barnehagen samt spørsmål om foreldres arbeid og utdanning (Se vedlegg F). Spørreskjema, på førstespråk og norsk ble formidlet til foreldrene gjennom skolene. Spørreskjema via organisasjoner er forventet å ha lavere svarprosent enn surveys til enkeltpersoner og svarprosenten på 60% er dermed akseptabel, selv om det hadde vært ønskelig at den var høyere (Baruch & Holtom, 2008). Samtidig varierer svarprosenten avhengig av hvilken informasjon som etterspørres. Svarprosenten på spørsmål om hvor barnet er født eller deltakelse i barnehage er høy (60%), mens mer sensitive spørsmål om foreldrenes utdanning ble besvart av langt færre. Svarprosenten på 60% (N=325) varierte med språkbakgrunn, men det ble innhentet bakgrunnsinformasjon fra foreldre fra alle språkgrupper (albansk=27, tamil, 83, Tyrkisk=81, Somali = 32, Urdu = 57, vietnamesisk=45).

Forskjeller i språkferdigheter mellom gruppen hvor foreldre besvarte spørreskjemaet og gruppen som hadde foreldre som ikke besvarte spørreskjemaet, ble undersøkt med uavhengige t-tester. Reseptive og ekspressive vokabularferdigheter ble benyttet som språkmål. Det var ikke signifikante forskjeller mellom resultatene for vokabulartestene på andrespråket norsk mellom gruppene (reseptivt vokabular: $t = 3.330$, $p = .274$; ekspressivt vokabular: $t = .571$, $p = .972$). Gruppen med foreldre som besvarte spørreskjemaet hadde imidlertid signifikant bedre skåret på reseptivt og ekspressivt vokabular på førstespråket (reseptivt vokabular: $t = .207$, $p = .000$; ekspressivt vokabular: $t = 3.569$, $p = .005$) sammenlignet med gruppen som ikke besvarte spørreskjemaet. En forklaring på denne forskjellen er den store gruppen foreldre med språkbakgrunn fra Sri Lanka (N=83) og Tyrkia (N=81) som besvarte spørreskjemaet sammenlignet med foreldre fra de andre språkgrupper (albansk=27, Somali, 32, Urdu, 57, vietnamesisk=45). Resultatene fra to av artiklene har vist at språkgruppene med bakgrunn fra Sri Lanka og Tyrkia har signifikant bedre resultater på førstespråket enn de andre språkgruppene

Språkbruk i hjemmet

283 av utvalget som har besvart spørreskjemaet er født i Norge og 42 er født i andre land; Albania = 4, Somalia = 24, Sri Lanka (Tamil) = 1, Tyrkia = 5, Pakistan (Urdu) = 5 og Vietnam = 3. Dette bekrefter at mange i utvalget har lang botid i Norge. På spørsmål om barnet benytter førstespråket, andrespråket eller flere språk, rapporterte 324 av foreldrene at barna benyttet førstespråket (N= 324) hjemme.

Eksposering for første – og andrespråket

I tillegg til språkbruk i hjemmet var både første – og andrespråket i aktiv bruk i flere kontekster. Tabell 4.3 viser eksponering for førstespråket og tabell 4.4. viser eksponering for andrespråket.

Førstespråket brukes særlig i lek med venner. Informasjon fra foreldrene styrker dermed generell demografisk informasjon om at det typisk utviklede utvalget er fra bomiljøer og nabolag med et større antall elever med samme førstespråk. Relativt mange rapporterer at de også får opplæring på førstespråket og deltar i aktiviteter der førstespråket inngår.

Tabell 4.3 Eksponering for førstespråket

	Daglig	Ukentlig	Månedlig	ubesvart
Sammen med venner	179	80	18	56
Opplæring hvor førstespråket inngår	76	110	13	134
Andre aktiviteter hvor førstespråket inngår n=189	55	79	32	167

Mange i utvalget ble eksponert for norsk i barnehagen ($n = 276$). 214 av barna i utvalget har deltatt i fulltids barnehagetilbud og 62 i barnehagetilbud på deltid. 19 av barna har ikke gått i barnehage (38 er ubesvart). Gjennomsnittstid i barnehagen for utvalget er to år.

Som tabell 4.4 viser blir de fleste i utvalget eksponert for andrespråket i opplæring og organiserte aktiviteter og uformelle situasjoner som lek og aktiviteter daglig eller ukentlig

Tabell 4.4 Eksponering for andrespråket

	Daglig	Ukentlig	Månedlig	ubesvart
Lek sammen med venner	260	31	5	38
Kulturelle aktiviteter	74	120	23	116
Andre aktiviteter utenfor hjemmet hvor andrespråket inngår	70	77	28	158

Foreldres vurdering av barnets språkferdigheter

Tabell 4.5 viser foreldrenes vurdering av barnas språkferdigheter på første og andrespråket. Mange foreldre vurderer barnas språkferdigheter som svært gode eller gode.

Tabell 4.5 Foreldrenes vurdering av barnas språkferdigheter

	Svært godt	Godt	Ganske godt	svakt	Ikke svart
Vurdering av førstespråket sammenlignet med jevnaldrende med samme morsmål N=320	137	124	39	6	27
Vurdering av andrespråket sammenlignet med jevnaldrende med samme morsmål N=319	146	135	21	6	25

Bakgrunnsinformasjon om foreldrene

Tabellen 4.6 viser at mange av foreldrene har lang botid Norge, men botiden viser også at alle ikke er født i Norge.

Tabell 4.6 Foreldrenes botid i Norge i antall år og måneder

	Albania	Paki stan	Somalia	Sri Lanka	Tyrkia	Vietnam	ubesvart
Botid mor	13.01 N=24 Sd 5.4 Min -maks 24	17.1 N=59 Sd 6.9 min-maks 3-34	7.6 N=34 Sd 4.6 min-maks 0-20	14.0 N=83 Sd 3.8 min -maks 7-24	18.6 N=74 Sd 6.2 min-maks 2-32	16.7 N=45 Sd 5.6 min-maks 2-28	14
Botid far	17.0 N=25 Sd 7.8 Min-maks 6-32	22.0 N=58 Sd 7.6 Min-maks 6-37	9.5 N=24 Sd 5.5 Min-maks 0-20	18.9 N=83 Sd 4.2 Min-maks 7-25	18.8 N=76 Sd 8.3 Min-maks 2-38	20.6 N=44 Sd 7.8 Min-maks 0-31	23

Det er en tendens til at fedre har lengre botid i Norge enn mødrene, og det er større variasjon i botid blant fedre. Mindre standardavvik i botid blant språkgrupper med bakgrunn fra Somalia og Sri Lanka kan indikere at mødre og fedre i disse språkgruppene kom til Norge mer samlet enn fedre og mødre fra de andre språkgruppene. I samsvar med andre undersøkelser, er utvalget med bakgrunn fra Somalia språkgruppen med kortest botid (Bakken & Hyggen 2019, SSB 2010). Tabell 4.7 viser foreldrenes deltakelse i arbeidslivet. Tallene indikerer at fedre i noe større grad er i arbeid enn mødrene som er i arbeid.

Tabell 4.7 Foreldre i arbeid

	Albania	Pakistan	Somalia	Sri Lanka	Tyrkia	Vietnam
Arbeid mor	20 N=25	38 N=60	10 N=32	69 N=81	44 N=79	40 N=45
Arbeid far	24 N= 25	56 N=60	25 N=31	81 N=83	69 N=79	42 N=45

De fleste fedre med språkbakgrunn fra Pakistan og Tyrkia er i arbeid, til forskjell fra mødre fra de samme språkgruppene. De fleste fedre og mødre med språkbakgrunn fra Albania og Vietnam er også i arbeid. Mange med språkbakgrunn fra Sri Lanka er i arbeid, men færre mødre enn fedre.

104 mødre og 96 fedre besvarte spørsmål om skolegang. 49 mødre og 63 fedre har mer enn tretten års skolegang (fra hjemland, Norge eller andre land). Dette spørsmålet er ubesvart av relativt mange (N=133).

En mulighet til å øke svarprosenten kunne ha vært å ansvarliggjøre skolene i større grad. Skolene fikk merarbeid gjennom å samle inn samtykkeskjema og spørreskjemadata fra foreldre samt organisere timeplan for eksterne testledere til gjennomføring av kartleggingen. Det var derfor et bevisst valg å ikke gi skolene ytterligere ansvar for datainnsamlingen. Andre forskningsprosjekter med relativt store utvalg har hatt direkte kontakt med foreldre gjennom telefonintervjuer, men dette var ikke mulig innenfor rammen av dette prosjektet.

Siden det er signifikante forskjeller i førstespråkferdighetene for gruppene med og uten spørreskjemadata valgte jeg ikke å inkludere spørreskjemadata som variabler i analysene, selv om andre typer analyser enn de som er gjennomført kunne tatt høyde for dette. Noen av variablene som for eksempel deltagelse i barnehage og bruk av førstespråket, ble forsøkt inkludert i enkelte analyser, men bidro ikke å forklare varians. Variabler som hvor barnet er født og deltagelse i barnehage er benyttet i beskrivelsen av utvalgene i alle de tre artiklene. Spørreskjemadata bekrefter imidlertid at alle de seks språkgruppene er aktive brukere av førstespråket og dette styrker lærervurderingene om at barn og unge i avhandlingen aktivt benytter to språk.

Bakgrunnsinformasjon om utvalget med DLD

Samtlige i utvalget som har påvist DLD er født i Norge, og de er fra samme nabolag som det typisk utviklede utvalget. De er også fra de samme språkgruppene. Barna benytter førstespråket med foreldre, men andrespråket med søsken og venner. Elevene har hatt barnehagetilbud i gjennomsnitt i tre år. I tillegg til informasjon fra foreldre, observasjon i klassen og tester som er ledd i kartleggingen fra PPT (kognitive forutsetninger, lesetester, sosialt samspill) og informasjon fra syn og hørsel, ble elevene kartlagt på første- og andrespråket med de samme testene som ble benyttet i det flerspråklige utvalget med typisk utvikling. Det ble også gjennomført en samtale med foreldrene basert på et spørreskjema utviklet på bakgrunn av Alberta Language and Developmental Questionnaire (ALDeQ Questionnaire, Paradis 2011). Dette er et eksempel på foreldrevurderingsskjema henvisning til i kapittel 3. I samtalen ble det lagt vekt på hvordan det aktuelle barnet er eksponert for første- og andrespråket, språkbruksmønstre samt kvalitative vurderinger fra foreldre om språkutviklingen på førstespråket. Jeg gjennomførte også workshops med de aktuelle PPT kontorene om kartlegging av flerspråklige og unge generelt. Samtalene med foreldre styrket bekymringen for språkferdighetene og bekreftet at årsaken til henvisning til PPT ikke var manglende andrespråkferdigheter.

4.3 Tester for vokabular, grammatikk og fonologisk minne

Testene som er inkludert i avhandlingen er mye benyttet i forskning samt av rådgivere i PPT, logopeder og spesialpedagoger, både i utredning av mulige språkvansker hos enspråklige og flerspråklige barn og unge.

Testen for reseptivt vokabular, BPVS-II (Dunn et al., 1997; Lyster et al., 2010) ble normert til norsk samtidig med Trog-2 (Bishop, 1989; Lyster & Horn, 2009). Barn og unge fra alle deler av landet er representert i normeringen og barn med språkvanske, andre innlæringsvansker eller med flerspråklig bakgrunn er ekskludert fra normeringsutvalget. Reliabiliteten for BPVS-II målt med Cronbach Alpha varierer mellom .88-.92 i aldersgruppene (3.0 – 16.1). For Trog-2 er reliabiliteten, Cronbachs alpha, 95.

Normeringsdata for BPVS-II inngår i artikkel 1 og 3 og normeringsdata fra Trog-2 inngår i artikkel 3 (n=495). Språk 6-16 (Ottem & Frost, 2005) er utviklet i Norge. Tre tester; Setningsrepetisjon, Ordspenn (serial recall of words/word span) og Ekspressivt vokabular, er inkludert (Ottem & Frost, 2005). Ordspenn er en test for fonologisk minne, Begreper undersøker ekspressivt vokabular og Setningsrepetisjon er utviklet som en minnetest. Som vist i kapittel 2 og 3, er imidlertid setningsrepetisjon langt mer enn en minnetest. Den måler også språkferdigheter mer generelt.

Språk 6-16 ble normert i 2005 (N= 1080), revidert i 2007 (N=1214) og i 2010. Utvalget er fra hele landet og 70 logopeder, psykologer og studenter deltok i normeringen. I avhandlingen har jeg benyttet en versjon samt normeringsdata fra 2005 fordi antall oppgaver/item i deltesten Begreper/Ekspressivt vokabular ble vesentlig redusert i en revidert utgave fra 2010 (fra 26 oppgaver til 17 oppgaver). Normeringsdata i aldersgruppen 6.0-12.11 inngår i analysene i de tre artiklene i avhandlingen (N=900).

Cronbach alpha for de norske deltestene fra Språk 6-16 er høy; Setningsrepetisjon er .81, for Ordspenn .75 og for Begreper .90. Språk 6-16 er evaluert ut fra internasjonale standarder beskrevet i den Europeiske Federation of Psychology Association (EFPA), (Evers et al., 2013). Rasjonale for testen er «excellent» og normer, reliabilitet og validitet er vurdert som «good». Konklusjonen er dermed at testen har gode psykometriske egenskaper (Arnesen et al., 2019).

Samlet gir de fem testene et godt utgangspunkt for å undersøke reseptivt og ekspressivt vokabular, grammatisk forståelse og fonologisk minne på første— og andrespråket for de seks språkgruppene som inngår i avhandlingen.

Reseptivt vokabular (benyttes i artikkel 1 og 3)

BPVS-II er en reseptiv vokabulartest hvor eleven skal velge ett av fire bilder som skal passe til ordet som testleder leser opp. Den norske versjonen består av 144 item delt inn i 12 blokker med økende vanskegrad. Elevene starter på ulike blokker avhengig av alder. Basert på stor variasjon i

vokabularferdigheter blant elever med flerspråklig bakgrunn startet imidlertid alle i utvalget på oppgave 1 i blokk 1 (2.5 - 3 år) på både første- og andrespråket. Stoppkriteriene er imidlertid det samme som for normeringsutvalget og stoppkriteriet er å avslutte testen etter 8 feil i en blokk.

Utfra at vokabularferdigheter kan være distribuert mellom første- og andrespråket i flerspråklige utvalg, ble en vokabularvariabel hvor vokabular på begge språk inngår, konstruert. Variabelen er kalt konseptuell skåre. Skåre på førstespråket gir ett poeng, skåre på andrespråket gir ett poeng og skåre på begge språk gir ett poeng. Det er ikke mulig å oppnå mer enn ett poeng på ett item, men det gis poeng så lenge ordet forstås på ett av språkene slik dette er framstilt under.

	Førstespråk	Andrespråk (norsk)	Konseptuell skåre
Elev 1	1	0	1
Elev 2	0	1	1
Elev 3	1	1	1

I artikkel 1 inngår også data fra normeringsutvalget for reseptivt vokabular, BPVS-II (N=495) og data fra normeringsutvalget for ekspressivt vokabular Språk 6-16 (N=900) i deskriptiv statistikk og analyser.

Ekspressivt vokabular (Begreper (benyttes i artikkel 1 og 3))

Den ekspressive vokabulartesten fra Språk 6-16 består av 14 antonymer (*hva er det motsatte av vått*) og 12 definisjonsoppgaver (*Hva er en seng?*). Alle oppgaver blir gitt muntlig til eleven. Hvert item gis skåre 1 eller 0 og testen stoppes etter tre feil i deltesten antonymer og etter tre feil på rad i deltesten definisjoner. Skåren for ekspressivt vokabular er summen av korrekte item på antonymer og definisjonsoppgaver. Samme prosedyre er benyttet for førstespråket og andrespråket. Konseptuell skåre ble beregnet for ekspressivt vokabular utfra samme prosedyre som for reseptivt vokabular.

Setningsrepetisjon (benyttes i artikkel 2 og 3)

Som gjort rede for i kapittel 3 er setningsrepetisjon sammensatt mål hvor både vokabular, grammatikk og fonologisk minne inngår.

Setningsrepetisjon består av 16 setninger med økende lengde og kompleksitet. Setningen blir lest høyt en gang for eleven. Eleven skal gjenta setningen helt ordrett for å oppnå ett poeng (skåring 1 eller 0).

Testingen stoppes etter 3 nullskårer på rad og samme prosedyre er benyttet for første- og andrespråket.

Ordspenn (benyttes i artikkel 2 og 3)

Ordspenn måler hvor mange enheter som huskes i kort tid og ordspenn (også fra Språk 6-16) antas å være et mål på språklig korttidsminne på linje med tallminne forlengs. Testen ordspenn består av 12 deloppgaver med serier på tre, fire eller fem enstavelsesord og fire oppgaver i hver serie. Seriene blir lest opp og skal gjentas av eleven. Hver ordserie skåres 1 eller 0 og testen stoppes etter tre feil. Ordene i den

norske utgaven av ordspenn er tilpasset fra en liste av ord brukt i en studie av korttidsminne (Lian et al., 2004). Samme prosedyre ble benyttet for førstespråket og andrespråket.

Grammatikk (benyttes i artikkel 3)

Test for Reception of Grammar (Trog-2) (Bishop 2003; Lyster & Horn 2010) undersøker grammatisk forståelse. Trog-2 er en bilde-basert test hvor eleven får presentert en setning muntlig og skal velge ett av fire bilder som passer til setningen. Testen består av 80 setninger fordelt på 20 blokker med grammatiske konstruksjoner med økende kompleksitet og vanskegrad. I de første blokkene i TROG-2 inngår leksikalske distraktorer (det vil si at bildet skiller seg fra de andre ved at innholdsordet varierer). De siste blokkene har grammatiske distraktorer hvor bøyning av et ord eller ordrekkefølge endrer meningen. Selv om det var relativt uproblematisk å tilpasse oppgavene i TROG-2 til ulike språk, kan vanskegraden på konstruksjonene variere mellom språkene, Alle 80 oppgaver ble derfor gjennomført på både første – og andrespråket. Stoppkriterier i den norske og engelske utgaven av TROG-2 ble derfor ikke benyttet.

4.4 Oversettelse og tilpasning til seks førstespråk

Hovedmålsettingen med avhandlingen var å undersøke første- og andrespråkerferdigheter med de samme testene og med det samme bildematerialet. Det er derfor lagt stor vekt på systematisk samarbeid med oversettere for alle førstespråk for å kvalitetssikre oversettelsene i størst mulig grad. Oversetterne hadde det aktuelle førstespråket som sitt morsmål og alle hadde gode norskspråklige ferdigheter. I tillegg var pedagogisk kompetanse godt representert iblant oversetterne. Dette er forhold som bidro til at oversetterne kunne vurdere særlige utfordringer ved oversettelsene og også vurdere relevans av alle item ut fra elevenes spesifikke flerspråklige kontekst.

De fem testene er oversatt og tilpasset til seks språk som representerer relativt store språkgrupper i norsk kontekst; albansk, somali, tamil, tyrkisk, urdu/punjabi og vietnamesisk. Oversettelse og tilpasning av tester er gjort ved bruk av anerkjente prinsipper for oversettelse av tester (Peña, 2007; Peña et al., 2011; Pena & Halle, 2011; Peña, 2008; The American- Speech-Language-Hearing- Association, 2012). To oversettere fra hver av språkgruppene, med kulturell, språklig og/eller pedagogisk kompetanse og erfaring, har oversatt alle oppgaver fra norsk og til de forskjellige førstespråkene uavhengig av hverandre. Ved ulik oversettelse vurderte oversetterne og en norsk prosjektgruppe som besto av lingvist samt fagpersoner med testkompetanse og utredningserfaring⁵ alle item for å komme fram til enighet om hvilken oversettelse som tilsvarte det norske på best måte (vedlegg C viser oversikt over oversettere og tilbake oversettere). Deretter ble testene gitt til en annen oversetter som oversatte testene fra førstespråket tilbake til norsk for å vurdere hvorvidt tilbakeføringen til norsk ble lik eller tilnærmet lik norske testitem.

⁵ Kirsten Meyer Bjerkan, Anne -Cathrine Thurmann-Moe og May-Britt Monsrud

For å etterstrebe at vokabulartester er ekvivalente på tvers av språk påpekes det ofte at oversettelser av tester bør baseres på frekvensordlister. Slike frekvensordlister finnes ikke for de aktuelle førstespråk i norsk kontekst. Det finnes frekvensordlister for tyrkisk i Tyrkia, men det er min vurdering at disse ikke er gyldige for tyrkiske barn med hele sin oppvekst i Norge. Kompetanse om den kontekst språket tilegnes i er derimot en svært sentral kompetanse ved oversettelse av tester. Et eksempel er normeringen av ordforrådsprøven Peabody til spansk. Selv om ordet på spansk er det samme, varierer resultatene for det samme item mellom utvalg fra Spania, Puerto Rico og Mexico, (The Spanish test Vocabulary en Imagenes Peabody TVIP, Dunn, Padila, Lugo & Dunn 1986). Dette viser at samme ord ikke har samme frekvens i ulike språklige kontekster. For BPVS-II må 8 av 12 item være feil for å stoppe testen.

Det var en grundig diskusjon mellom kvalifiserte oversettere samt en prosjektgruppe om tilpasningen av de ulike grammatiske konstruksjonene og bøyningsformene for å oversette Trog-2. Det var mulig å tilpasse de grammatiske konstruksjonene til alle førstespråk, og det var de samme grammatiske forhold som ble testet i de to språkene; entall/flertall, nåtid/framtid/fortid, innskutte relativsetninger. Trog-2 testet dermed spesielle grammatiske forhold slik disse kommer til uttrykk i de to språkene. Slik sett er testene ekvivalente, selv om grammatiske uttrykk dannes på ulike måter. Vi valgte å gjennomføre alle 80 item for alle elever på begge språk.

I samarbeid med oversetterne ble prøven Setningsrepetisjon oversatt til setninger med tilnærmet like mange informasjonsenheter, grammatisk kompleksitet, struktur og lengde målt i stavelser i alle språk. For å unngå typisk «norskulturelt» «innhold som «*Mor og far reiste ut på tur i en liten kano*» ble innholdet noe endret, men antall informasjonsenheter ble beholdt. Også egennavn ble erstattet med navn som var vanlige i den aktuelle språkgruppen.

Ved tilpasning av Ordspenn til førstespråket ble ikke ordene oversatt, men de samme kriterieriene som i den norske versjonen er benyttet for å lage ordserier; ordseriene består av høyfrekvente og meningsfulle enstavelserord.

Prosentandelen i hver av språkgruppene som skåret rett på første- og andrespråket på de fem testene, er beregnet. For BPVS-II er vanskegrad beregnet ut fra hvor mange prosent som mestret hver blokk (12 blokker med 12 item i hver blokk) og for Ordspenn ble vanskegrad beregnet ut fra hvor mange prosent som skåret korrekt i hver av de tre ordseriene som alle besto av fire ordserier (4 ordserier som besto av 3 ord, 4 ordserier som besto av 4 ord og 4 ordserier som besto av 5 ord). På oppgavene for Setningsrepetisjon, Ekspressivt vokabular og Trog-2 ble det regnet ut hvor mange prosent fra hver språkgruppe som skåret rett på hver enkelt oppgave på første og andrespråket. For alle fem tester var det synkende skårer ved økende vanskegrad på item.

4.5 Kartleggingsprosedyrer

Det er ikke uvanlig at elever med et annet førstespråk enn majoritetsspråket ikke ønsker å anvende førstespråket i majoritetsspråklige miljøer. Vegring eller angst for å benytte førstespråket i en majoritetssetting er ikke et ukjent fenomen (Sevinç & Backus, 2017). Fordi vi var sensitive for denne muligheten og fordi det var vesentlig for å få et valid bilde av elevens førstespråkferdigheter, rekrutterte vi testledere fra den aktuelle språkgruppen. Testlederne for førstespråket var PP-rådgivere, studenter med fordypning i pedagogisk -psykologisk rådgiving, rådgivere i Statped samt morsmåslærere.

Svært mange testledere var involvert i kartlegging av de to utvalgene (typisk utviklet) og DLD utvalget (N=14). Det ble lagt vekt på å etablere en ensartet kartleggingsprosedyre gjennom opplæring. Kurs og veiledning samt prosedyrer for gjennomføring av testene ble gjennomført for testledere for utvalget med typisk utvikling. For utvalget med DLD ble det gjennomført fagmøter med pedagogisk-psykologiske rådgivere.

Kartleggingen tok ca 1 time på hvert språk. Halvparten av utvalget ble testet på førstespråket først og halvparten på andrespråket først. I klinisk kartleggingspraksis skal individuelle språkferdigheter undersøkes mer presist. For BPVS-II er det «considered poor practice for standardized assessments to be repeated within six months». Dette for å unngå en test-retest effekt når samme test blir gjentatt innenfor et kort tidsrom (Dunn et al., 1997). For TROG-2 anbefales det å vente 9 måneder før retesting. I samarbeid med referansegruppen til prosjektet i regi av Utdanningsdirektoratet ble det avgjort at det skulle være 6 måneder mellom de to testtidspunktene. Dette ble bestemt utfra en målsetting om å utvikle aldersnormer for første- og andrespråket for de fem testene. Dette var særlig sentralt i vurderingen av de visuelt baserte testene BPVS-II og Trog-2. Denne vurderingen kompliserer imidlertid variabelen for alder (og aldersgruppeinndelingene) i de tre artiklene. Dette blir diskutert nærmere i artiklene samt under statistiske analyser (4.7) og i kapittel 6. Denne avstanden i testtidspunktene må vurderes med tanke på den mulige språkutviklingen som finner sted i perioden mellom tidspunktene. I den grad standardskårer benyttes, vil alder og utvikling ivaretas, men når råskåre benyttes er utfordringen større.

Uavhengige t-tester viste imidlertid at det ikke var signifikante forskjeller mellom gruppene som ble testet først og sist for ekspressivt vokabular for førstespråket ($t=-1.51$, $p=.131$) eller andrespråket ($t=.40$, $p=.691$); Det var imidlertid signifikante forskjeller for reseptivt vokabular ($t=-2.23$, $p=0.25$) i favør av gruppen som ble testet sist og signifikante forskjeller i favør av andrespråket for gruppen som ble testet først ($t=1.13$, $p=.26$). Slikt sett kan det ikke påvises en systematisk bedring for den gruppen som ble testet sist.

Skåringsprosedyrene er relativt enkle med skåre 0 eller 1.

Som nevnt ble spesielle prosedyrer benyttet for BPVS-II hvor alle flerspråklige barn og unge startet på sett 1, oppgave 1 og på TROG_2 hvor alle item ble gjennomført både på første – og andrespråket. Siden det for flere tester ikke ble brukt stoppkriterier, vil dette med vanskegrad på enkeltitem derfor ikke påvirke elevenes totalskåre på testen.

4.6 Statistiske analyser

I avhandlingen blir den overordna problemstillingen og forskningsspørsmålene besvart ved hjelp av tre analytiske tilnærminger; deskriptive analyser, variansanalyser og complex-analyser i Mplus for å belyse avhandlingens overordna problemstilling; *Hva karakteriserer språkferdighetene på første – og andrespråket blant flerspråklige barn og unge?* Complexanalyser i artikkel 3 undersøker Setningsrepetisjon (Muthén & Muthén, 2017). Begrunnelsen for dette er at mange empiriske studier indikerer at en slik test har kvaliteter som kan bidra til å gi oss innsikt i språkferdigheter i en- og flerspråklige utvalg.

I de tre artiklene presenteres deskriptive data med resultatfordelingene for første og andrespråket i råskårer, standardavvik og aldersgrupper for de fem testene. Til forskjell fra mange studier som avdekker store forskjeller mellom andrespråket og enspråklige utvalg, er en sentral problemstilling i avhandlingen å få mer kunnskap om både første- og andrespråket i ulike språkgrupper i Norge – og dermed få bedre innsikt i disse barnas totale språkkompetanse.

MANOVA og ANOVA analyser i artikkel 1 og 2 ble valgt fordi slike analyser gjør det mulig å undersøke vokabularferdigheter og fonologisk minne samtidig som betydningen av alder og språkgruppetilhørighet kan inngå. Selv om store utvalg gir et godt utgangspunkt for mer avanserte SEM-analyser, var dette ikke ønskelig utfra forskningsspørsmålene for artikkel 1 og 2. Bruk av latente variabler er en stor fordel med tanke på målingsfeil, men gitt de svært høye reliabilitetstallene som foreligger for alle testene, synes målefeilene til en viss grad å være under kontroll. I tillegg, var det ikke aktuelt å slå sammen minnevariablene eller vokabularvariablene fordi det var ønskelig å få spesifikk informasjon om barnas mestring på de enkelte testene og informasjon om testenes nytteverdi i det norske hjelpeapparatet.

Utvalg med flerspråklig bakgrunn sammenlignes også med enspråklige normutvalg for å få kunnskap om forskjeller mellom flerspråklig og enspråklig utvalg har sammenheng med om vokabular eller fonologisk minne undersøkes.

MANOVA og ANOVA analyser benyttes i artikkel 1. Reseptivt og ekspressivt vokabular er avhengig variabel og aldersgruppe og språkgruppe er uavhengige variabler i en MANOVA analyse. En toveis ANOVA for reseptivt vokabular og en for ekspressivt vokabular sammenlignet andrespråket med to normutvalg, henholdsvis for BPVS-II og Ekspressivt vokabular fra Språk 6-16. Det ble gjennomført to separate analyser fordi forskjellen i N mellom normutvalgene for de to vokabulartesten var svært store; reseptivt vokabular (N=495) og ekspressivt vokabular (N= 900). Reseptivt og ekspressivt vokabular var avhengige variabler i begge analysene og uavhengig variabel var språkgruppe (flerspråklige vs normgruppe). Basert på at vokabularferdigheter i flerspråklige utvalg er distribuert mellom første og andrespråket ble det laget en ny variabel, en konseptuell skåre, hvor både første- og andrespråket inngår. Den konseptuelle skåren ble også sammenlignet med norske normdata ved å benytte ANOVA analyser.

Basert på de store forskjellene i standardavviket for vokabularferdigheter mellom utvalgene med flerspråklig bakgrunn og enspråklig normeringsutvalg er forskjeller i effektstørrelsen beregnet ved hjelp av Cohens d for tydeligere å kunne vurdere forskjeller mellom gruppene (Tabachnick & Fidell, 2014; Tabachnick & Fidell, 2001).

I Artikkel 1 ble det kontrollert for alder ved å benytte det siste testtidspunktet som utgangspunkt for å beregne konseptuell skåre. Svakheten med denne måten å beregne alder på, som drøftet i artikkel 1, er at siden den kontekstuelle skåre består av både første- og andrespråkets vokabular kan ett av de to språkene som inngår i konseptuell skåre være kartlagt da det flerspråklige utvalgte var yngre enn normutvalget det sammenlignes med. Vokabularferdighetene utvikles ved økende alder, og ved at det flerspråklige utvalget er yngre en normutvalget kan en slik måte å beregne alder på medføre lavere skåre på grunn av alder i det flerspråklige utvalget. Som også gjort rede for i artikkel 1, kan et slikt valg impliserer en disfavør for utvalget med flerspråklig bakgrunn fordi konseptuell skåre til dels er beregnet utfra at utvalget er yngre enn utvalget med majoritetsbakgrunn. Samtidig har utvalget med flerspråklig bakgrunn større sjanse for å få rett ved gjetting fordi de fikk to muligheter til å få korrekt skåre på den bildebaserte testen BPVS-II. Det flerspråklige utvalget hadde dermed økt sjanse for å oppnå ett poeng på grunn av mulighet til å få rett på både første- og andrespråket. Samtidig har utvalget også større mulighet for å få feil fordi det flerspråklige utvalget startet på oppgave 1 til forskjell fra normutvalget som startet utfra sitt alderstrinn. Normutvalget fikk rett skåre for alle oppgaver i forkant av startpunktet. Dette medførte ikke en trussel for den konseptuelle skåren for ekspressivt vokabular. Den konseptuelle skåren for ekspressivt vokabular er svakere enn den for reseptivt vokabular, men det er som forventet utfra annen forskning som kan vise til at reseptivt vokabular generelt er sterkere enn ekspressivt vokabular (Ouellette, 2006; Schmitt, 2014).

I artikkel 2 ble forskjellene i resultater mellom flerspråklig utvalg og enspråklig normutvalg analysert ved å benytte ANCOVA analyser. Setningsrepetisjon og Ordspenn var avhengige variabler og språkgruppene (de seks flerspråklige gruppene andrespråk samt normutvalget) som uavhengige variablene. Alder ble brukt som kovariat. En follow-up Bonferroni test ble gjennomført for å undersøke hvorvidt forskjeller i skåre hadde sammenheng med språktype og alder. Det var ingen interaksjonseffekter for i disse analysene.

Uavhengige t-tester undersøkte om betydningen av testene Setningsrepetisjon ble undersøkt på første eller andrespråket og resultatene avdekket i hovedsak de samme forskjellene mellom språkgruppene som for vokabular. Alle språkgruppene skåret imidlertid signifikant bedre på førstespråket sammenlignet med andrespråket for Ordspenn.

I både artikkel 1 og artikkel 2 ble det lagt inn bakgrunnsvariabler som hvilke land barnet er født i samt tid i barnehagen. Men slike variabler bidro ikke til å forklare varians signifikant og ble tatt ut igjen av analysene siden de kun bidro til å påvirke kraften (power) i analysene på en negativ måte.

Setningsrepetisjon er vurdert som en sentral klinisk markør for å avdekke potensielle språkvansker. Samtidig er det stor variasjon i resultater på testen når den benyttes i utvalg med flerspråklig bakgrunn. Setningsrepetisjon inkluderer både en språklig komponent (grammatikk og vokabular) og en minnekomponent. Forskningen som undersøker denne testen, er ikke entydig med hensyn til om grammatikk eller vokabular forklarer mest av variasjonen i setningsrepetisjon. Det er heller ikke entydig hva testen måler i flerspråklige utvalg. Valget av den analysemessige tilnærmingen i artikkel 3 ble gjort på en slik bakgrunn, og tilnærmingen i artikkel 3 ble gjort uavhengig av analysene i artikkel 1 og 2. Årsaken var til dels at problemstillingen var annerledes og til dels fordi data og innledende analyser i denne studien ga noen føringer som antydte at complexanalysene var hensiktsmessig å bruke. Målsettingen i artikkel 3 var å undersøke potensiale for Setningsrepetisjon som screeningtest for språkferdigheter. Vi undersøkte derfor hvorvidt resultater fra andre språktester som fonologisk minne, grammatikk og alder og vokabular predikerte resultater for setningsrepetisjon. Foreløpige analyser viste at Setningsrepetisjon på førstespråket ikke forklarte variasjon i Setningsrepetisjon på andrespråket etter at prediktorvariablene for førstespråket ble lagt inn. Det samme var tilfelle for Setningsrepetisjon på andrespråket. Det ble derfor gjennomført en analyse for førstespråket og en for andrespråket.

Siden reseptivt og ekspressivt vokabular korrelerte svært høyt inkluderte vi ekspressivt vokabular som vokabularvariabelen. En annen mulighet som ble vurdert var å lage en variabel av de to vokabularvariablene, men siden ekspressivt vokabular bygger på det lagrede impressive vokabularet,

ønsker vi å undersøke den spesifikke forklarings-effekten som lå i det ekspressive vokabularet. Det er også gjort en vurdering av hvilke analyser (stegvise regresjonsanalyser eller complex analyser i Mplus) som gir den mest hensiktsmessige informasjonen for å undersøke hvilke variabler som forklarer varians i setningsrepetisjon for første og andrespråket. Resultatene for regresjonsanalyser for første – og en for andrespråket og complex analysen i M-plus ligger tett opp til hverandre og peker i samme retning og det er en styrke at funnene er robuste på tvers av analysemetode. (Muthén & Muthén, 1998 – 2017). Complex analysen ble likevel valgt fordi analysen gir mulighet til å kontrollere for skoletilhørighet, og for hver av de observerte variablene er det kontrollert for alder i måneder. Dette er helt klart en styrke ut fra tidsforskjellen mellom de to testtidspunktene

I artikkel 3 er de seks språkgruppene i utvalget slått sammen. Begrunnelsen var å sammenligne elever med antatt typisk utvikling, men svake ferdigheter på første- og andrespråket (- 1SD for begge språk) med et utvalg med påvist DLD. DLD gruppen representerer de samme språkgruppene, men ikke alle språkgruppene, med typisk utvikling. Siden gruppen med DLD ikke var representert i de yngstealdersgruppene og utvalget var svært lite, valgte vi å dele både gruppen med DLD og gruppen identifisert med svake språkferdigheter basert på første- og andrespråket for setningsrepetisjon, i to aldersgrupper. Vi valgte å presentere resultatene som deskriptiv statistikk fordi begge gruppene var relativt små (N=14) og (N=26) og resultatene viste også stort standardavvik for alle tester.

5 Oppsummering av hovedfunn i de tre artiklene

Avhandlingen belyser språkferdigheter i flerspråklige grupper gjennom å undersøke hvordan et utvalg med antatt typisk utvikling presterer på språktester på andrespråket norsk og på de samme testene på sine respektive førstespråk. Økt kunnskap om flerspråklig utvikling kan gi et bedre utgangspunkt for å identifisere flerspråklig barn og unge som har språklige vansker. Et utvalg identifisert med DLD er inkludert i artikkel 3

I artikkel 1 undersøker vi reseptive og ekspressive vokabularferdigheter. I artikkel 2 undersøker vi setningsrepetisjon og ordspenn. I artikkel 3 undersøker vi hvilke andre tester som predikerer skåren for setningsrepetisjon på første- og andrespråket. I de tre artiklene inngår også to enspråklige normutvalg.

5.1 Hovedfunn i artikkel 1

Engelsk tittel: The Advantages of Jointly Considering First and Second Language Vocabulary Skills Among Emergent Bilingual Children (Monsrud, Rydland, Thurmann-Moe & Halaas Lyster 2019), publisert i *International Journal of Bilingual Education and Bilingualism*

Det er stor forskningsproduksjon og interesse for vokabularferdigheter fordi vi vet at vokabularferdigheter har stor betydning for læringsutbytte. Mange studier viser at barn og unge med flerspråklig bakgrunn har svakere vokabularferdigheter på andrespråket enn enspråklige på samme alder. Denne artikkelen har tatt utgangspunkt i reseptive og ekspressive vokabulartester på første- og andrespråket for å undersøke vokabular blant barn og unge fra 6-13 år. Vi undersøkte forskjeller mellom alderskohorter og språkgrupper ved å bruke ANOVA og MANOVA analyser.

Flerspråklige barn og unges språkferdigheter er undersøkt på to språk. Det samlede utvalget har stigende skårer for på reseptivt og ekspressivt vokabular ved økende alder på både første- og andrespråket. Dette viser at kompetansen på både første- og andrespråket øker med alderen. Flerspråklige barn og unges vokabularferdigheter er distribuerte mellom første- og andrespråket. Vi undersøkte derfor vokabularferdigheter ved å konstruere en konseptuell skåre for reseptivt og ekspressivt vokabular, og sammenlignet den konseptuelle skåren med skårene til de enspråklige normutvalgene på de respektive testene.

Et hovedfunn i artikkel 1 er at forskjeller mellom utvalg med flerspråklig bakgrunn og de enspråklige normutvalgene forsvinner eller blir vesentlig mindre ved en konseptuell skåre for reseptive vokabularferdigheter. Forskjellene blir også vesentlig mindre når den konseptuelle skåren blir benyttet

for å undersøke ekspressive vokabularferdigheter. Forskjellene mellom flerspråklige og enspråklige barn og unge kan derfor forklares med prosedyrer for kartlegging, ikke vokabularkompetanse.

Forskjeller mellom utvalg med flerspråklig og enspråklig bakgrunn er godt dokumentert i mange empiriske studier. I likhet med mange andre studier oppnår utvalget med flerspråklig bakgrunn svakere resultater for både reseptivt og ekspressivt vokabular på andrespråket sammenlignet med de enspråklige normutvalgene. Selv om forskjellene er mindre blant de eldste barna, er forskjellene fortsatt signifikante i de to eldste aldersgruppene. Dette viser at mange flerspråklige barn og unge henger etter på andrespråket og ikke har tatt igjen sine enspråklige klassekameraters vokabularferdigheter etter mange år i norsk skole. Analysene viser i tillegg at det er systematiske forskjeller mellom språkgruppene når det gjelder hvorvidt første- eller andrespråket er det sterkeste språket. Dette viser at undersøkelser av språkferdigheter i flerspråklige utvalg uten en presis beskrivelse av utvalgene, kan kamuflere store ulikheter mellom språkgrupper.

Studien har implikasjoner for kartlegging ved å dokumentere at vokabularferdigheter er distribuerte ferdigheter i utvalg med flerspråklig bakgrunn og at det å vurdere ett språk ikke er tilstrekkelig for å identifisere en vokabularkompetanse som består av to språk.

5.2 Hovedfunn i artikkel 2

Engelsk tittel: Assessing Phonological Memory in Dual Language Learners: A Study of Sentence Repetition and Word Span Tests in First and Second Languages (Monsrud, Halaas Lyster, Bjerkan & Rydland, 2020). Manuskript innsendt til vurdering i *Journal of Communication Disorders*.

Tester som undersøker fonologisk minne, er mindre avhengig av språklig og leksikalsk kompetanse og blir derfor ofte inkludert i studier av barn og unge med flerspråklig bakgrunn. Slike tester er sentrale for å avklare forutsetninger for språkkompetanse og blir derfor også inkludert i forbindelse med kartlegging av potensiell DLD. Vi undersøkte testene *ordspenn* og *setningsrepetisjon* på første- og andrespråket i ulike alderskohorter ved å benytte ANCOVA analyser og t-tester.

Et hovedfunn i artikkel 2 er at elevene skåret signifikant bedre på førstespråket enn på andrespråket på testen *ordspenn*. Denne relative styrken på førstespråket ble identifisert i alle seks språkgruppene.

For det andre var ikke resultatene for *ordspenn* på andrespråket signifikant forskjellig fra det enspråklige normutvalget for tre av språkgruppene, og en språkgruppe skåret signifikant bedre på *ordspenn* for andrespråket enn det enspråklige normutvalget. Dette viser at kapasiteten i fonologisk minne er en styrke i det flerspråklige utvalget. Slik fonologisk minne kartlegges med *ordspenn*, vil vi

anta at utvalget har gode forutsetninger for språklæring.

For det tredje viser resultatene for *setningsrepetisjon* langt på vei de samme forskjellene mellom første- og andrespråket som vokabularferdigheter på første- og andrespråket i artikkel 1 ved at noen språkgrupper har en relativ styrke på førstespråket, mens andre språkgrupper har en relativ styrke på andrespråket. Setningsrepetisjon synes dermed å fange opp forskjeller i språkbruksmønstre som er karakteristiske for ulike språkgrupper.

Studien har implikasjoner for kartlegging ved at *setningsrepetisjon* og *ordspenn* avdekker ulike ferdigheter og ulik kompetanse, selv om begge testene antas å være mindre avhengig av språkferdigheter enn tester for vokabular og grammatikk.

5.3 Hovedfunn i artikkel 3

Engelsk tittel: First and Second Language Sentence Repetition Measures. A Potential Screening Measure for Dual Language Learners? (Monsrud, Rydland, Geva & Halaas Lyster 2020). Manuskript er resubmitted til *Language and Education* i juni 2021 og akseptert for publisering.

Vi undersøkte hvorvidt setningsrepetisjon på første- og andrespråket predikeres av andre språkmål som vokabular, grammatikk og fonologisk minne målt med ordspenn på første- og andrespråket. Complexanalysene viste at setningsrepetisjon på første- og andrespråket langt på vei predikeres av ekspressivt vokabular og grammatikk i tillegg til fonologisk minne. Dette indikerer at setningsrepetisjon reflekterer språkkompetanse både på første- og andrespråket i tillegg til fonologisk minne slik det kartlegges av ordspenn. Analysene viser også at for førstespråket blir setningsrepetisjon i stor grad forklart av ekspressivt vokabular på førstespråket.

Et interessant funn er at ordspenn målt på førstespråket i tillegg til i bidra til å predikere setningsrepetisjon på førstespråket også predikerte setningsrepetisjon på andrespråket. Dette indikerer at ordspenn måler en felles underliggende kognitiv komponent for begge språk.

Man kan anta at i et så stort utvalg ville det også være noen elever med svake språkferdigheter, uten at de har kjente spesialpedagogiske behov. En slik gruppe ble identifisert ved at de hadde svake skårer på både første- og andrespråket på setningsrepetisjon (N=26). Denne gruppen hadde tilsvarende svake skårer på setningsrepetisjon og ordspenn som utvalget henvist og utredet av pedagogisk-psykologisk tjeneste for DLD. Dette bekrefter at språklige risikosymptomer ikke alltid blir oppdaget og medfører at svake språkferdigheter da heller ikke blir vurdert og ikke alltid vil bli identifisert.

Vi identifiserte også en gruppe i det flerspråklige utvalget ved å benytte samme cutoff på setningsrepetisjon på andrespråket som for setningsrepetisjon i det enspråklig normutvalget. Vi fant da en stor overrepresentasjon av flerspråklige barn og unge i risiko. Det viser at å benytte enspråklige normer kan innebære en vesentlig feilkilde ved kartlegging av barn og unge med flerspråklig bakgrunn.

Implikasjoner for kartlegging er at setningsrepetisjon er et sensitivt mål for språkferdigheter og derfor har kvaliteter som en mulig screeningtest.

Kapittel 6 Testteori, reliabilitet og validitet – metodiske drøftinger

Forskningsspørsmålene i avhandlingen er i stor grad knyttet til målte ferdigheter på et første- og et andrespråk, og til analysene av de målte ferdighetene. Svarene på forskningsspørsmålene og gyldigheten for analysene som er gjort og svarene som vi har fått, er derfor avhengige av at de målte ferdighetene er riktige eller «sanne». I tillegg benyttes tester som er oversatt fra norsk til en rekke språk og norske tester som er oversatt fra engelsk. Resultatenes troverdighet er derfor sterkt knyttet til troverdigheten til de observasjoner (tester) som er gjennomført. Denne troverdigheten vurderes gjerne ut fra testenes reliabilitet, om testenes skårer har store målefeil eller ikke. I tillegg må testene vurderes ut fra sin validitet, i hvilken grad de måler den ferdigheten de er ment å måle.

Språkferdigheter er ikke direkte observerbare og for å kunne studere språkferdigheter kreves indirekte observasjonsmetoder. Spørsmålet er hvor godt målemetodene er begrunnet i teori og hvorvidt de undersøker den virkeligheten de er konstruert for å beskrive og undersøke. Språkferdigheter er i avhandlingen operasjonalisert ved hjelp av måleinstrumenter for reseptivt og ekspressivt vokabular, grammatikk, setningsrepetisjon og fonologiskminne. Disse testene eller måleinstrumentene er redskaper som benyttes mye både i forskning og klinisk praksis for å undersøke språkferdigheter blant enspråklige barn og unge. Det er en klar tendens i mange empiriske studier til at måling av språkferdigheter blant flerspråklige barn og unge avgrenses til måling av andrespråkferdigheter som sammenlignes med en enspråklig norm. En slik tilnærming innebærer et brudd på en av grunnforutsetningen ved bruk av tester; alle skal ha hatt samme mulighet til å tilegne seg det som etterspørres og måles i testen for at testene skal gi mulighet til å være sammenlignbare. Det Messick kaller «comparability» (Messick, 1989a, 1989b). Som vist til i avhandlingens drøfting av empiri og teori, er det store forskjeller mellom en- flerspråklige utvalg med hensyn til å tilegne seg det som etterspørres i tester utviklet og normert i en enspråklig kontekst og på tester som er tilpasset forskjellige førstespråk.

Et utgangspunkt for avhandlingen er at kartlegging av språkkompetanse i flerspråklige utvalg må gjennomføres på begge språk. Avhandlingen belyser språkferdigheter på førstespråket ved å benytte de samme testene og ved hjelp av det samme bildemateriale som er normert for enspråklige elever. Det er begrensede tillatelser til å gjøre endringer i et testmateriale. Selv om det ikke er antatt som «best practise» å kartlegge barn og unge med tester som ikke er utviklet for dem, gjøres dette likevel i klinisk praksis. Det er derfor viktig å undersøke hvordan flerspråklige barn og unge skårer på tester som er normert for enspråklige utvalg. Fonologiske minneprøver er særlig interessante fordi slike prøver antas å

avdekke kognitiv kapasitet som er viktig for språkutviklingen, samtidig som resultatene på slike prøver er mindre avhengige av ferdigheter på det aktuelle språket (de Abreu et al., 2013; Pena & Halle, 2011).

Avhandlingen belyser utfordringer som oppstår når flerspråklige barn og unge kartlegges med tester utviklet i og basert på enspråklige majoritetsutvalg. Selv om bruk av standardiserte språktester generelt kan bidra til å redusere antall feilkilder, er det åpenbare validitetsutfordringer ved å benytte tester for grupper de ikke er normert for eller utviklet for. Det er også trusler i sammenheng med systematiske og tilfeldige målefeil basert på oversettelse av tester samtidig som oversettelser kan føre til at ferdigheter som måles på førstespråket ikke er de samme som man har målt på andrespråket. Problemstillinger i sammenheng med reliabilitet og validitet blir diskutert i dette kapitlet som tar utgangspunkt i testteori.

6. 1 Testteori

Testteori omhandler vurderinger av testers kvalitet og psykometriske egenskaper i form av reliabilitet (pålitelighet), validitet (gyldighet) og normering (Kleven 2018). Både klassisk og moderne testteori angir prosedyrer for å analysere variabler og undersøke testers sensitivitet og spesifisitet (Crocker & Algina, 1986). Testteori bidrar til å undersøke om tilgjengelige kartleggingsverktøy og måleinstrumenter samsvarer med det testene er utviklet for å måle.

Det psykometriske eller testteoretiske grunnlaget for språktester har i stort omfang tatt utgangspunkt i klassisk testteori der oppgaver på en test skåres uavhengig av vanskegrad eller tar høyde for hvordan oppgavene diskriminerer mellom grupper av elever. En sentral faktor i den klassiske testteorien er knyttet til stabilitet og målefeil, til at den poengsummen en person får ville vært den samme om testingen fant sted på en annen dag eller med forskjellige testere. I dag benyttes stadig oftere en såkalt IRT- teori (Item Response Theory) ved utvikling av tester og prøver innenfor det pedagogiske området. En slik teori er blant annet benyttet ved utviklingen av de nasjonale prøvene fra 2014. Med utgangspunkt i IRT kan man estimere oppgavens vanskegrader på en helt annen måte enn ved klassisk testteori der alle oppgaver teller likt, uavhengig av vanskegrad. IRT gir mer presise opplysninger på alle ferdighetsnivåer. Med utgangspunkt i IRT vil sannsynligheten for målefeil beregnes ut fra den enkelte elevs skåre. Det gir bedre kontroll på målefeilene for den enkelte elev selv om målefeilene er noe større for yttergruppene, de med de beste og de med de dårligste ferdighetene. Når barn/elever skal følges over år er det spesielt viktig å kunne ta utgangspunkt i IRT, som gir en sikrere mulighet for sammenligning enn hva klassisk testteori gir mulighet for. I tillegg gir et slikt utgangspunkt en bedre mulighet for å diskriminere mellom elevenes ferdigheter og kan også begrense omfanget (antall item) av testen samtidig som høy reliabilitet og validitet beholdes.

Når det allikevel i denne avhandlingen er benyttet en klassisk testteori i arbeidet med å oversette tester, skyldes det delvis at alle testene i utgangspunktet er utviklet med utgangspunkt i klassisk testteori, at noen data er hentet fra en normeringsgruppe der data forelå på forhånd og fordi skolesystemet og hjelpeapparatet er kjent med disse testene. Siden studiene som inngår er tverrsnittsstudier, skal elevene heller ikke testes på nytt med de samme prøvene over flere år.

Det å ta utgangspunkt i en klassisk test teori innebærer imidlertid at vi må være nøye med å vurdere resultatene ut fra de målefeil som foreligger (reliabiliteten). I tillegg er validitetsspørsmål sentrale i denne avhandlingen, ikke minst knyttet til validitetsutfordringer som oppstår ved oversettelse av tester.

6.2 Reliabilitet og systematiske målingsfeil

Systematiske målefeil er i første rekke knytte til påliteligheten til målingene i form av reliabilitet. Reliabilitet gir informasjon om hvor stabilt og reproduserbart måleinstrumentet er (målesikkerhet) Reliabilitet er en forutsetning for validitet, og kartleggingsresultater er ikke valide uten å være reliable (Kleven, 2018). Det er vanlig å benytte tre hovedkategorier av reliabilitet: indre konsistens, test- retest reliabilitet og inter-reliabilitet. *Indre konsistens* har sammenheng med i hvilken grad ulike item som skal måle det samme frembringer de samme resultatene og peker i samme retning målt med for eksempel Cronbach alpha (Shadish et al., 2002). Alfaverdien er en funksjon av gjennomsnittlig korrelasjon mellom to og to oppgaver og antall oppgaver i testen. Det siste gjenspeiler at tilfeldige feil jevner seg ut i løpet av testen. En høy Cronbach alpha innebærer en høy indre korrelasjon mellom de enkelte oppgavene i en test og avdekker om deloppgavene måler en ferdighet som kan etterspørres av andre. En $\alpha = .70$ blir ofte betraktes som en akseptabel verdi, men er sett på som et minstekrav til bruk i forskning, mens en $\alpha = .80$ generelt akseptabel til klinisk bruk (Kerlinger et al., 2000). Større alpha enn .90 er betegnet som utmerket (bla Feilberg 2010). Det kan være ulike årsaker til at reliabiliteten varierer, men generelt er det vanlig at reliabiliteten blir lavere i tester med få item. Det er i hovedsak ikke vesentlige forskjeller i reliabiliteten mellom de norske versjonene og førstespråkversjonene av testene som er benyttet i denne avhandlingen. Ordspenn har noe lavere Cronbach Alpha i begge versjoner sammenlignet med de andre testene, men det er ikke overraskende gitt at Ordspenn består av kun 12 item.

Tabellene 6.1 og 6.2 viser reliabilitet målt som Cronbachs Alpha for alle testene for begge språk.

Tabell 6.1 Cronbach Alpha for Setningsrepetisjon, Ordspenn, grammatikk, reseptivt og ekspressivt vokabular for de seks **førstespråkene**

Språk	Setningsrepetisjon	Ordspenn	Grammatikk	Reseptivt vokabular	Ekspressivt vokabular
Albansk	$\alpha.87$	$\alpha.83$	$\alpha.98$	$\alpha.98$	$\alpha.90$
Somali	$\alpha.88$	$\alpha.80$	$\alpha.96$	$\alpha.99$	$\alpha.92$
Tamil	$\alpha.85$	$\alpha.80$	$\alpha.64$	$\alpha.96$	$\alpha.86$
Tyrkisk	$\alpha.85$	$\alpha.68$	$\alpha.92$	$\alpha.96$	$\alpha.92$
Urdu	$\alpha.85$	$\alpha.80$	$\alpha.76$	$\alpha.98$	$\alpha.89$
Vietnamesisk	$\alpha.88$	$\alpha.82$	$\alpha.64$	$\alpha.98$	$\alpha.86$

Tabellen viser at ordspenn på tyrkisk og grammatikk på tamil har svakere verdier sammenlignet med de andre førstespråkene. Det samme er tilfelle for andrespråket hvor ordspenn på somali og tyrkisk har lavere, reliabilitetstall som vist i tabell 6.2.

Tabell 6.2 Cronbach Alpha for Setningsrepetisjon, Ordspenn, grammatikk, reseptivt og ekspressivt vokabular for de seks språkgruppene for **andrespråket norsk**

Språk	Setningsrepetisjon	Ordspenn	Grammatikk	Reseptivt vokabular	Ekspressivt vokabular
Albansk	$\alpha.79$	$\alpha.75$	$\alpha.69$	$\alpha.98$	$\alpha.90$
Somali	$\alpha.80$	$\alpha.68$	$\alpha.95$	$\alpha.98$	$\alpha.88$
Tamil	$\alpha.82$	$\alpha.74$	$\alpha.96$	$\alpha.98$	$\alpha.88$
Tyrkisk	$\alpha.74$	$\alpha.64$	$\alpha.89$	$\alpha.98$	$\alpha.86$
Urdu	$\alpha.80$	$\alpha.76$	$\alpha.76$	$\alpha.98$	$\alpha.88$
Vietnamesisk	$\alpha.76$	$\alpha.76$	$\alpha.77$	$\alpha.98$	$\alpha.85$
Norsk	$\alpha.81$	$\alpha.73$	$\alpha.95$	$\alpha.83-93$	$\alpha.90$

Korrelasjoner mellom ulike mål er presentert i vedlegg H. Det er signifikante korrelasjoner mellom alle testene, og dette indikerer at alle testene undersøker en felles faktor, språkferdigheter. Det er særlig sterke sammenheng mellom variablene innad i førstespråket og innad i andrespråket.

Korrelasjonene mellom ordspenn for første- og andrespråket er imidlertid sterke og bidrar til å styrke tolkningen av resultatene i artikkel 2 og 3 om at ordspenn måler en kognitiv ferdighet som er mindre avhengig av språkferdigheter slik disse måles med de andre språktestene.

Svært mange testledere var involvert i kartlegging av utvalgene. Dette kan påvirke resultatene (*inter-rater reliability*). Få vurderinger ble imidlertid overlatt til testledere. Skåringsprosedyrene var relativt enkle, med skåre 0 eller 1, og medførte at lite ble overlatt til skjønnsmessige vurderinger.

Tydelige avvik fra start- og stoppkriterier ble avtalt på forhånd med testledere.

I tillegg ble kurs og veiledning samt prosedyrer for gjennomføring av testene gjennomført for testledere for utvalget med typisk utvikling og for rådgivere i pedagogisk -psykologisk tjeneste.

6.3 Validitetsutfordringer ved bruk av tester i flerspråklige utvalg

Testene som er inkludert i avhandlingen brukes mye av fagpersoner i ulike instanser som skole, pedagogisk-psykologisk tjeneste, barne- og ungdomspsykiatri og Statlig spesialpedagogisk støttesystem. De er også særlig sentrale i utredning av barn og unge med mistanke om DLD. Tester er redskaper for å undersøke hvilke ferdigheter som mestres. I klinisk sammenheng er tolkning av testresultater vesentlig for å vurdere videre utredning og omfang av tilrettelegging. En vurdering av hvorvidt flerspråklige barn og unge har behov for å styrke norskspråklige ferdigheter utfra § 2- 8 i Opplæringsloven eller har spesialpedagogiske behov utfra §5- 1 i den samme loven, blir vurdert i lys av blant annet testresultater. Dette betyr at testresultatene kan få store konsekvenser for personen som blir kartlagt. Messick er opptatt av at tolkning av testresultater alltid har en sosial og etisk konsekvens (Messick, 1989a, 1993). En overordnet validitetstrussel er at tester utviklet for enspråklige barn og unge kan innebære et brudd med rettferdighetsprinsippet for flerspråklige barn. Flerspråklige barn kan ha ferdigheter og ord de kan på førstespråket, men ikke på andrespråket. De blir også vurdert utfra enspråklige normer selv om språkferdighetene kan være distribuert mellom to (eller flere) språk.

Validitet refererer både til i hvilken grad det er evidens for og teoretisk støtte for de *slutningene* som trekkes basert på testresultater. Resultater fra en test er derfor ikke valide i seg selv, og slutningen som trekkes på bakgrunn av testen kan være mer eller mindre valid (Messick, 1989a). Validitet er viktig uansett, men det kan være større mulighet for feilkilder ved forskning på grupper hvor valide måleinstrumenter ikke er utviklet. Kane påpeker også at validitet har sammenheng med analyse av testresultater i større grad enn testene i seg selv (2006). De valgte operasjonaliseringene av språkferdigheter er undersøkt ved hjelp av anerkjente språktester for vokabular, grammatikk, setningsrepetisjon og fonologisk minne. Begrepsvaliditet har sammenheng med i hvilken grad et testresultat er et uttrykk for en faktisk ferdighet eller er en valid representasjon av begrepene. En test anses å ha høy begrepsvaliditet dersom den måler et intendert teoretisk begrep eller en intendert egenskap. Operasjonalisering av språkferdigheter innebærer å velge ut hvilke aspekter ved et fenomen som anses å være mest sentralt. Det vurderes som en styrke for begrepsvaliditeten at testene antas å måle sentrale aspekter ved språkferdigheter når testene er internasjonalt anerkjente redskaper for å måle språk. Samtidig er avhandlingen, både teoretisk og forskningsmessig, et argument for at andrespråktester ikke gir oss tilstrekkelig innsikt for å avdekke flerspråklige individers språkferdigheter.

Det er særlig to trusler for begrepsvaliditet som kan gjøre seg gjeldende (Messick, 1995). Det ene er at måleinstrumentet er for smalt eller ikke reflekterer det som det har til hensikt å måle. Utfra avhandlingens anliggende om å undersøke språkferdigheter på første- og andrespråket samt belyse et

mulig skille mellom typisk utvikling og DLD mener jeg utvalget av tester er relevante (kapittel 1- 3). Den andre trusselen er «construct-irrelevant difficulty» (Messick, 1995). En slik trussel har sammenheng med at måleinstrumentet innebærer en vesentlig feilkilde for individer eller grupper, så som at instrumentet ikke er tilpasset gruppen det blir benyttet for. Andrespråktester kan bidra til urettferdighet («unfairness») sammenlignet med en enspråklig norm fordi testene er normert utfra en mer enhetlig eksponering for språk enn mange flerspråklige har erfart. Førstespråktester kan også bidra til «unfairness» fordi, førstespråket har svært ulike betingelser for utvikling i flerspråklige grupper i norsk kontekst og for noen vil andrespråket være det sterkeste språket. Dette kommer også fram i kapittel 2 og av resultatene i artikkel 1 og 2.

Prinsippet om rettferdighet forutsetter at både lingvistiske, funksjonelle og målemessige faktorer ivaretas (Peña et al., 2011; Peña & Halle, 2011; Peña, 2008). Lingvistisk rettferdighet har sammenheng med å benytte anerkjente prosedyrer for oversettelse, slik det er gjort i denne avhandlingen (4.4). Men det handler også om at alle oversettelser av måleinstrumenter utfordres av kulturelle koder og språklige forskjeller. Det kan føre til systematiske forskjeller som også kan true begrepsvaliditeten fordi måleinstrumentene ikke har samme betydning, relevans og vanskegrad på tvers av kultur og språk (Maxwell, 2013). Funksjonell rettferdighet handler om at oppgaver i testene skal være relevante på tvers av grupper og forutsetter ikke bare lingvistisk kompetanse, men kompetanse om den særskilte flerspråklige virkeligheten som flerspråklige barn og unge lever i. Særskilt kompetanse hos oversetterne om både førstespråket og norsk språk samt den norskspråklige konteksten utvalget lever i, er nødvendig for å sikre kvaliteten ved oversettelsene (Messick, 1995; Peña, 2007). Metrisk rettferdighet har sammenheng med at vanskegraden skal være den samme på begge språk. Dette er ikke vurdert for hvert enkelt item, men en økende vanskegrad i testene på begge språk og for alle språkgrupper ved at prosentandelen (se 4.4) som mestret oppgavene sank ved stigende antall item, støtter validiteten i studiene i avhandlingen (Pena 2008). Samlet synes alle fem tester og alle fem item å dekkerelevante aspekter og et bredt spekter ved det konstruert som måles. Selv om et sentralt validitetsspørsmål er hvorvidt de samme språklige ferdighetene blir målt på førstespråket som på andrespråket.

Den store variasjonen i resultater og signifikante forskjeller mellom språkgrupper med hensyn til om første- eller andrespråket er det sterkeste språket (artikkel 1 og 2), demonstrerer kompleksiteten ved måling av språkferdigheter i flerspråklige utvalg. Alle tre artiklene trekker imidlertid slutninger om språkferdigheter. Det trekkes slutninger basert på måling av vokabular; reseptive oppgaver, antonymer og definisjonsoppgaver og basert på teori (Oller) og empiri (bla Gross et al) som indikerer at kartleggingsprosedyrer i flerspråklige utvalg ikke er de samme som for enspråklige. Men artikkel 1 viser

at « Conceptual scoring provides a more sensitive and positive picture of bilingual children's overall range of vocabulary. The findings demonstrate a need for more nuanced assessment procedures of semantic knowledge across children's L1 and L2" (Monsrud et al., 2019).

Testen som er valgt for grammatikk undersøker et bredt spekter av grammatiske strukturer. Det er den variabelen som viser til minst variasjon mellom aldersgruppene i de seks språkgruppene både på første- og andrespråket. Det kan ha årsak i for enkle oppgaver i forhold til alder, på begge språk, eller i at de fleste grammatiske strukturer mestres i relativt lav alder, slik at oppgavene av den grunn blir enkle for de høyeste aldersgruppene. Denne variabelen er benyttet for å predikere setningsrepetisjon, men det er utenfor denne avhandlingens ramme å trekke slutninger om grammatiske ferdigheter i gruppen.

I artikkel 2 og 3 trekkes slutninger om at setningsrepetisjon er et språkmål og i tillegg består av en minnekomponent på første- og andrespråket. Som kapittel 2 og 3 viser er denne slutningen i samsvar med empiri fra spansk-engelske utvalg. Samtidig er det behov for mer forskning om hvilke underliggende faktorer som kan bidra til å forklare resultater på en slik test (for eksempel Riches, 2012). Det er få item i testen (N=16) så det er begrensede muligheter til å få fram et differensiert bilde av elevens ferdighetsnivå med denne testen. Det er likevel mulig å argumentere for at det er en egnet screeningtest for å identifisere språkferdigheter i flerspråklige utvalg. Identifisering av DLD derimot fordrer at begge språk inngår i vurderingen.

Selv om det er mange studier som undersøker fonologisk minne i flerspråklige utvalg, er det færre studier som benytter ordspenn som måleinstrument. Slutningen som det er gjort nærmere rede for i artikkel 2 og 3 er likevel at Ordspenn er et korttidsminnemål og reflekterer en kognitiv kapasitet som er felles for første- og andrespråket.

De sterke korrelasjonene mellom variablene på førstespråket og mellom testene for andrespråket underbygge at testene måler språkferdigheter på hvert av språkene

Validiteten av for testene styrkes av to forhold; det er en styrke for validiteten at det er progresjon i skåre ved økende alder. Videre at prosentandelen som mestrer oppgavene synker med testenes progresjon. Det er dermed liten sjans for at elevene ville klare en senere oppgave om de ikke har mestret de siste som er gitt. Resultatene indikerer en stigende vanskegrad på første- og andrespråket i alle tester.

Ytre validitet har sammenheng med hvorvidt funnene fra undersøkelser kan overføres til andre individer og situasjoner enn de som er undersøkt. Som nevnt kan ikke funnene fra språkgruppene som er inkludert i avhandlingen generalisere til andre flerspråklige utvalg fordi utvalgene ikke ble trukket

tilfeldige, men er fra nabolag med høy andel barn og unge med flerspråklig bakgrunn fra samme språkgruppe. De seks språkgruppene har ulik historie i Norge og årsaken til at utvalgene kom til Norge og hva språkgruppene har flyttet fra ikke er sammenlignbare.

Utvalget er stort, men inklusjonskriteriene som «et typisk utviklet utvalg» og «beherske første og andrespråket» er ikke presise. Det er dermed også en usikkerhet relatert til hvor presise lærervurderingene har vært for utvalget. Generell informasjon om demografiske forhold i utvalgenes aktuelle oppvekstmiljøer (Statistics Norway, 2016) samt bakgrunnsinformasjon fra foreldre sannsynliggjøre at utvalgene har mer krevende sosioøkonomiske oppvekstforhold enn de enspråklige utvalgene som utgjør sammenligningsgrunnlaget.

Tester på både første- og andrespråket kan imidlertid gi støtte til validiteten i kartleggingen av barnas språk. Begrepet «fairness» i validitetssammenheng innebærer også to aspekter; riktighet (fairness (riktighet)) og rettferdighet (justice). Begge begrepene har betydning for slutninger basert på testresultater. Riktighet er forstått som en psykometrisk egenskap ved tester og som teoretisk og analytiske modeller for å undersøke egenskaper ved tester. Slike egenskaper ved tester får ofte stor oppmerksomhet. Det er et paradoks at slike egenskaper ved tester kan komme til å maskere rettferdighet som valg av test eller testspråk. Som vist i forrige avsnitt argumenterer jeg for at testene for både første- og andrespråket har god reliabilitet og imøtekommer krav til «fairness»/riktighet. Samtidig belyser avhandlingen utfordringer i sammenheng med begrepet «justice» i forbindelse med validitet når flerspråklige barn og unges språkferdigheter skal undersøkes.

6.4 Ethiske overveielser

Forskningsetikk bidrar til å regulere forskningsvirksomhet og ”.. viser til et mangfoldig sett av verdier, normer og institusjonelle ordninger for å konstituere og regulere vitenskaplig virksomhet» (Norsk forskningsetisk komite, 2011, s5). Forskningsetikk har derfor sammenheng med normative regler og instruksjoner samt refleksjoner over spesifikke utfordringer relatert til forskningen. Krav om informert og fritt samtykke krever tilstrekkelig informasjon til at det er mulig å vurdere konsekvenser. Informasjon og samtykkeerklæring er utarbeidet på aktuelle morsmål og norsk, for å unngå at manglende norskspråklige ferdigheter skulle være til hinder for «informert samtykke». Prinsippet om frivillighet ble understreket. For barna henviset til pedagogisk-psykologisk tjeneste kom forespørselen om deltakelse fra PPT.

Foreldrenes *opplevelse* av frivillighet har likevel betydning og det innebærer et særlig ansvar overfor utsatte grupper og ikke ta for gitt at vanlige prosedyrer for informasjon og samtykke

nødvendigvis sikrer selvbestemmelse. Det var vesentlig at forståelig informasjon om valgmuligheter og eventuelle valg ikke skulle få negative konsekvenser hverken i skole eller pedagogisk-psykologisk tjeneste. Det stilles særlige krav til informasjon og samtykke fordi politisk og skolepolitisk debatter kan føre til at store grupper kan fremstå i et uheldig lys i offentligheten.

Barn har særlige krav om beskyttelse og prinsippet om informert og fritt samtykke kan utfordres av varierende språklige forutsetninger hos foreldrene. For å ikke belaste deltakerne unødig var alle som deltok i kartlegging av utvalgene fagpersoner med erfaring med barn og unge som studenter eller morsmåslærere, PP-rådgivere, rådgivere fra kompetansesenter eller studenter med tekstkompetanse og kartleggingen ble gjennomført på skolene i en kjent setting.

Behovet for å verne personer eller grupper på den ene siden og behovet for vitenskapelig kunnskap på den andre siden, kan bidra til å skape utfordrende dilemmaer (Tangen 2013). Den politiserte, polariserte og emosjonelle offentlige debatten om flerspråklige barn og unges språkferdigheter og læringsutbytte, kan bidra til stigmatisering og negative forventninger. Derfor kreves en særlig varsomhet ved rapportering av resultater Samfunnsforskere deltar aktivt i konstruksjonen av sosiale virkeligheter. Kategoriseringen som benyttes er en viktig del av denne virkeligheten. Kategoriseringen skjer gjennom bruk av språk og begreper og som numeriske uttrykk relatert til begreper. Språk og begreper er derfor ikke nøytrale avspeilinger av virkeligheten, men kan også være med på å skape den (Nerland 2013). Respektfull beskrivelse av barn og unge som kan være i en sårbar opplærings situasjon er derfor svært sentralt i et slikt forskningsfelt som denne avhandlingen tilhører.

Kapittel 7 Oppsummerende diskusjon

Avhandlingen bidrar på flere områder hvor det er behov for mer forskningsbasert kunnskap om flerspråklige barn og unges språkferdigheter. Avhandlingens bidrag til ny kunnskap kan oppsummeres i tre punkter. For det første bidrar avhandlingen med økt innsikt for vurdering av flerspråklige barns samlede språkferdigheter. Resultatene viser at det er viktig med kartlegging på både første- og andrespråket for å få en reell kunnskap om flerspråklige barns språkkompetanse. En slik tospråklig tilnærming gir mer nyansert kunnskap om flerspråklige barns språkferdigheter. Denne tilnærmingen gir også viktig informasjon om både de forutsetningene som ligger til grunn for en god språkutvikling og gir et bedre utgangspunkt for å forstå flerspråklige elevers språkkompetanse. For det andre bidrar avhandlingen til å belyse forskjeller mellom ulike undergrupper flerspråklige barn og unge i Norge og synliggjør store variasjoner mellom ulike språkgrupper. For det tredje belyser avhandlingen hvorvidt tospråklige kartleggingsprosedyrer kan bidra til sikrere identifisering barn og unge som trenger spesialpedagogisk hjelp.

Med utgangspunkt i det overordnede spørsmålet; *hva karakteriserer språkferdigheter på første- og andrespråket blant flerspråklige barn og unge?* vil jeg i dette kapitlet løfte fram noen sentrale funn utfra følgende tre forskningsspørsmål:

- 1) Hvilken kunnskap kan tester for vokabular, grammatikk, setningsrepetisjon og fonologisk minne på første- og andrespråket gi oss om flerspråklige barn og unges språkferdigheter?
- 2) Hvilke likheter og forskjeller i språkferdigheter identifiseres i språkgrupper med førstespråk med bakgrunn fra Albania (Kosovo), Pakistan (urdu), Somalia, Sri Lanka (tamil), Tyrkia og Vietnam?
- 3) Hvilken betydning kan tospråklige kartleggingsprosedyrer ha i identifisering av DLD?

Hvilken kunnskap kan tester for vokabular, grammatikk, setningsrepetisjon og fonologisk minne på første- og andrespråket gi oss om flerspråklige elevers språkferdigheter?

Flerspråklige barn og unges språkferdigheter er undersøkt på to språk. Det samlede utvalget har stigende skårer på alle tester ved økende alder på både første- og andrespråket. Dette indikerer ikke bare at de har kompetanse på begge språk, men også at kompetansen på både første- og andrespråket øker med alderen. Det er et overraskende funn at førstespråkferdighetene er såpass sterke i de eldste aldersgruppene når man vet at det kan være krevende å opprettholde et førstespråk, som er et minoritetsspråk, i et majoritetssamfunn der andrespråket i økende grad også vil benyttes med foreldre og søsken etter hvert som barna eksponeres for andrespråket i barnehage og skole (Grosjean, 2008, 2010; Montrul, 2016; Rydland & Grøver, 2021). Det er imidlertid et særlig stort standardavvik på førstespråket for testene

vokabular og setningsrepetisjon, som viser en større variasjon i språkferdigheter på førstespråket sammenlignet med andrespråket (artikkel 1, 2 og 3). Dette kan være en konsekvens av at språkstimulering gjennom barnehage og skole har vært mer enhetlig for andrespråket enn for førstespråket.

Vi undersøkte vokabularferdigheter ut fra en forståelsesramme om at vokabularferdigheter er *distribuerte* mellom første- og andrespråket. Vi undersøkte derfor vokabular ut fra en prosedyre hvor barn og unge fikk uttelling for kompetanse dersom de kan ordet på ett språk: en konseptuell skåre. Forskjellene mellom flerspråklige og enspråklige normutvalg blir vesentlig mindre eller forsvinner helt når en konseptuell skåre beregnes for reseptive vokabularferdigheter. Forskjellene blir også vesentlig mindre for ekspressive vokabularferdigheter. Dette mønsteret er gjeldende i alle seks språkgrupper og i aldersgruppene fra seks til tretten år. Forskjeller i vokabularferdigheter mellom de seks språkgruppene og de to enspråklige normutvalgene (BPVS-II og ekspressivt vokabular) kan dermed, i likhet med resultater fra «Miami-studien» og andre empiriske studier, forklares ut fra kartleggingsprosedyrer som benyttes og ikke vokabularkompetanse (Hwang et al., 2020; Mancilla-Martinez & Vagh, 2013; Oller & Eilers, 2002). Det er i hovedsak spansk/amerikanske utvalg som er undersøkt tidligere, og blant annet i «Miami-studien» inngår et relativt homogent utvalg. Det er derfor interessant at distribuerte vokabularferdigheter også er et funn i svært ulike språkgrupper i norsk kontekst.

Reseptive vokabularferdigheter er bedre enn ekspressive vokabularferdigheter i studier av enspråklige utvalg (Schmitt, 2014). Forskjellene mellom reseptivt og ekspressivt vokabular er imidlertid større i flerspråklige utvalg. Slike forskjeller er også avdekket i artikkel 1. Når de flerspråklige sammenlignes med enspråklige er forskjellene i konseptuell skåre større for ekspressivt enn for reseptivt vokabular (Gibson et al., 2011). Det er ikke grunnlag for å finne årsak til slike forskjeller i avhandlingen. Årsak til bedre reseptive enn ekspressive vokabularferdigheter kan likevel ha sammenheng med at førstespråket i mindre grad er i aktiv bruk selv om språkforståelsen er god (Paradis et al., 2017). Det samme kan gjelde for andrespråket (Ribot et al., 2017). Det kan også ha sammenheng med at reseptive vokabulartester så som BPVS-II og Peabody oftere benyttes overfor flerspråklige fordi de er enkle å administrere, er bilde-basert og konstruert som flervalgsoppgaver. Slike prøver kan indikere resonneringsevne, i tillegg til vokabularkunnskap (Cummins, 2012)

Et annet funn på tvers av språkgruppene er at ordspenn er en test som ser ut til å fange opp underliggende kognitive ferdigheter av betydning for språkferdighetene på første- og andrespråket. Ordspenn målt på førstespråket er signifikant bedre enn ordspenn målt på andrespråket. Dette gjelder for

alle seks språkgrupper. Ordspenn målt på andrespråket viser at testresultatene for fire av språkgruppene ikke er signifikant forskjellig fra det enspråklige normutvalget. Ordspenn er bygd opp av meningsbærende enkeltord og derfor kunne det forventes at resultatene varierte avhengig av om første- eller andrespråket er det sterkeste språket, men dette er ikke tilfelle i dette datamaterialet. En mulig forklaring kan være, som diskutert i artikkel 2, at den tidlige eksponeringen for førstespråket har bidratt fonotaktisk kunnskap, som gir støtte til å mestre kravene i ordspenntesten på førstespråket, slik som studier referert i artikkel 2 viser til (Buac et al., 2016; Messer et al., 2010).

Ordspenn predikerte setningsrepetisjon på første- og andrespråket.

Videre er det svært interessant at resultatene for ordspenn målt på førstespråket også predikerte setningsrepetisjon på andrespråket. Dette kan igjen indikere at ordspenn måler en felles underliggende kognitiv komponent i begge språk fordi språkferdigheter, slik det måles ved setningsrepetisjon på andrespråket, forklares signifikant av kapasiteten i fonologisk minne målt med ordspenn på førstespråket (Cummins, 2021).

Selv om setningsrepetisjon i likhet med ordspenn synes å tappe en fonologisk minnekomponent, viste funnene i artikkel 2 og 3 at setningsrepetisjon i større grad er mål på mer tradisjonelle språkferdigheter (vokabular og grammatikk) sammenlignet med ordspenntesten. Analysene av setningsrepetisjon i artikkel 3 viste at vokabular, grammatikk og fonologisk minne målt med ordspenn predikerte resultater for setningsrepetisjon både for første- og andrespråket. Disse resultatene understøtter tidligere funn fra andre studier blant enspråklige i norsk kontekst som argumenterer for at setningsrepetisjon er et sammensatt språkmål (Klem, 2014; Klem et al., 2015) og fra studier av flerspråklige utvalg i Europa (Zebib et al., 2019) og i USA (Pratt et al., 2021).

Når det gjelder grammatikk viser resultatene at alle språkgruppene i utvalget små forskjeller mellom de yngste og de eldste aldersgruppene. Dette kan tyde på at testen i liten grad differensier mellom aldersgruppene og at grammatikkprøven er lite sensitiv for variasjon i språkferdigheter både på første- og andrespråket.

Hvilke likheter og forskjeller i språkferdigheter identifiseres i de seks språkgruppene?

Det første funnet jeg vil løfte fram er systematiske forskjeller mellom språkgruppene når det gjelder hvorvidt første- eller andrespråket er det sterkeste språket. Utvalget med bakgrunn fra Tyrkia skilte seg ut ved at de som gruppe opprettholdt førstespråket i større grad enn de andre språkgruppene. Studier om nabolagets betydning for førstespråkutviklingen blant flerspråklige tyrkisk/norske barn og unge (Rydland et al., 2013) og studier som undersøker forskjeller i språkbruksmønstre mellom grupper

flerspråklige i Nederland, har funnet samme tendens (Extra & Yagmur, 2010). Dette kan delvis forklares utfra en tendens til å opprettholde god kontakt med opprinnelseslandet, blant annet ved at ekteskap inngås med personer fra derfra. Førstespråket opprettholdes dermed av dialogpartnere som gjør det relevant å benytte språket. For to av språkgruppene, gruppen med vietnamesisk som første språk og gruppen med somalisk som førstespråk, synes botid i Norge å ha betydning for språkferdighetene på både første- og andrespråket. Språkgruppen med vietnamesisk som førstespråk, med lang botid i Norge, viste en styrke på andrespråket. Dette til forskjell fra utvalget med språkbakgrunn fra Somalia som har en styrke på førstespråket. Årsaken kan være relativt kortere botid i Norge sammenlignet med de andre språkgruppene. Utvalget med tamil som førstespråk har derimot opprettholdt førstespråkkompetansen og har i tillegg også gode vokabularferdigheter på andrespråket sammenlignet med alle de andre språkgruppene. En mulig forklaring kan være at mange i utvalget deltar i tilbud som gir en mer systematisk eksponering for førstespråket gjennom undervisningstilbud, fritidstilbud i regi av tamilsk ressurscenter, samt leksehjelp på andrespråket og regelmessig tilgang til tospråklige lærere. Det pakistansk-norske og albansk -norske utvalget viser mer like skårer på første- og andrespråket enn de øvrige utvalgene som er representert i denne avhandlingen. Som beskrevet over viser resultatene fra artikkel 1 og 2 systematiske forskjeller mellom språkgruppen hvorvidt første- eller andrespråket er det sterkeste språket.

Det andre funnet jeg vil løfte frem, er at resultatene for *setningsrepetisjon* i artikkel 2 langt på vei bekrefter tilsvarende forskjeller mellom første- og andrespråkferdighetene som vokabularferdigheter på første- og andrespråket i artikkel 1. Setningsrepetisjon synes dermed å fange opp forskjeller i språkbruksmønstre som er karakteristiske for ulike språkgrupper. I Norge er det stort behov for tester som kan bidra til å identifisere språkferdigheter blant flerspråklige barn og unge. Studier av dette instrumentet viser at testen er en lovende kandidat til å ha en slik funksjon.

Det tredje funnet viser at undersøkelser av språkferdigheter i flerspråklige utvalg uten en presis beskrivelse av utvalgene, kan kamuflere store ulikheter mellom språkgrupper når det gjelder blant annet ulike språkkombinasjoner og språkprofiler. I mange studier får vi lite informasjon om hvilke inklusjons- og eksklusjonskriterier som er benyttet. Samtidig er det vanskelig å finne en klar og entydig norm for begrepsbruk innenfor det flerspråklige fagfeltet. Ulike kontekster, som fransk/engelske språkomgivelser i Canada, har behov for andre begreper enn for flerspråklige i Norge eller Skandinavia, hvor nettopp forholdet mellom majoritet og minoritet kan være en vesentlig faktor for å fortolke resultatene (Giæver 2021). Det er en svakhet ved mange studier at slutninger om flerspråklige barn og unges språkferdigheter trekkes uten at utvalgskriteriene er tilstrekkelig avklart. I norsk kontekst refereres ofte til forskning fra

Canada og USA uten å eksplisitt ta høyde for svært ulike kontekster og betingelser for barn og unge fra disse landene, sammenliknet med norsk kontekst. Denne avhandlingen tydeliggjør nettopp betydningen av å være bevisst mulige ulikheter mellom språkgrupper i norsk kontekst.

Hvilken betydning kan tospråklige kartleggingsprosedyrer ha for identifisering av utviklingsmessige språkforstyrrelser?

Flerspråklige barn og unge må ha vansker på begge sine språk at diagnosen DLD kan stilles. Som gjort rede for i kapittel 3, kan DLD likevel være komplisert å avdekke i flerspråklige utvalg fordi lite vokabular og usikre ferdigheter i grammatikk kan opptre i sammenheng med både et andrespråk i utvikling eller være symptomer på DLD (Bedore & Pena, 2008; Kohnert, 2013; Paradis, 2019a). Et annet sentralt symptom som ofte vektlegges i forbindelse med identifisering av DLD, er begrensninger i kapasiteten i fonologisk minne (Paul et al., 2018).

I avhandlingen er ordspenn benyttet som et mål på fonologisk minne, og resultatene tyder på at dette er en test som egnet for å identifisere begrensninger i fonologisk minne hos et flerspråklig utvalg. Svake resultater i ordspenn for «at risk»-gruppen og gruppen henvist til PPT, understøtter dette (se artikkel 3). En slik kognitiv kapasitet hevdes å være mindre avhengig av erfaringer med språk, og dette er derfor en særlig aktuell test for å undersøke fonologisk minne blant barn og unge med flerspråklige bakgrunn (Blom et al., 2014; de Abreu et al., 2013; Girbau & Schwartz, 2008; Morales et al., 2013)-

Setningsrepetisjon omtales ofte som en klinisk markør for å identifisere mulig DLD (Polišenská et al., 2015; Pratt et al., 2021). Analysene både på første- og andrespråket i artikkel 3 bekrefter at setningsrepetisjon predikeres av vokabular og grammatikk i tillegg til fonologisk minne både før første- og andrespråket. Ekspressivt vokabular predikere en større del av setningsrepetisjon på førstespråket enn på andrespråket, og viser at setningsrepetisjon på førstespråket for en stor del forklares av en aktiv bruk av førstespråket.

Resultatene for begge språk viser at setningsrepetisjon kan bidra til å identifisere variasjon i språkferdigheter og dermed også identifisere individer med svake språkferdigheter på både første- og andrespråket. Da setningsrepetisjon på andrespråket ble sammenlignet med det enspråklige normutvalget i artikkel 3, ble en vesentlig større gruppe flerspråklige enn enspråklig identifisert til å være i risiko for språkvansker. Dette viser at det er en feilkilde å benytte enspråklige normer og cut off på tester for flerspråklige barn og unge.

I likhet med mange andre studier sammenlignes utvalg med flerspråklig bakgrunn med en enspråklig norm. Dette gir oss kunnskap om ulike språkgruppers andrespråkferdigheter målt med enspråklige tester. En styrke er at sammenlikningen er gjort til andre fra samme språkgruppe. Med mer kunnskap om hva vi kan forvente av flerspråklige barn og unges språkkompetanse, har vi også et bedre utgangspunkt for å forstå og også avdekke DLD. Dette er også vesentlig for å identifisere mulige avvik (Geva & Wiener, 2014).

Det primære formålet med avhandlingen var å bidra til forskningsbasert kunnskap om språklige ferdigheter på første- og andrespråket blant flerspråklige barn og unge og belyse den overordnede problemstillingen: **Hva karakteriserer språkferdighetene på første- og andrespråket til flerspråklige barn og unge?**

Avhandlingen viser at flerspråklige barn og unges språkferdigheter preges av variasjon, og at ferdigheter på ett språk ikke gir et fullstendig bilde av språkkompetanse i flerspråklige utvalg. Dette synliggjøres både av at reseptive og ekspressive vokabularferdigheter er *distribuert* mellom første- og andrespråket og at systematiske forskjeller mellom språkgruppene også er avdekket med det sammensatte målet setningsrepetisjon.

For det andre viser de systematiske forskjellene mellom gruppene at å behandle flerspråklige barn og unge som ett utvalg gir et skjevt og lite presist bilde av gruppen. Dette kan begrunnes ut fra at systematiske forskjeller også er tydelig for språkgruppene på andrespråket.

For det tredje viser den minst språkavhengige testen, Ordspenn, at utvalget har god kapasitet for språklæring. Dette kan ha sammenheng med kognitive prosesser som er felles for begge språk slik modellene til Cummins illustrer i *Dual-isjfellmodellen* og *hypotesen om felles kognitive prosesser*.

Avhandlingen kan ha betydning for kartleggingspraksis og for forskning knyttet til flerspråklige barn og unge ved at den gir en mer spesifisert kunnskap om flerspråklige barn og unges språkferdigheter. Den belyser også at det er begrensninger ved kartlegging på andrespråket når flerspråklige barn og unges språkkompetanse skal vurderes.

Implikasjoner for forskning og klinisk praksis

For det første viser resultatene at Setningsrepetisjon er et bredt språkmål som avdekker vokabular, grammatikk og fonologisk korttidsminne på både første- og andrespråket. Setningsrepetisjon har vist seg

å være et sensitivt mål for å identifisere variasjon i språkferdigheter og vil derfor kunne være egnet som screeningtest for å avdekke eventuell risiko for språkforstyrrelser i flerspråklige utvalg.

For det andre fremstår Ordspenn som et mål som kan identifisere begrensninger i fonologisk korttidsminne på første- og andrespråket og vil kunne bidra ved identifisering av eventuelle vansker på individnivå.

For det tredje viser avhandlingen at å undersøke andrespråket alene ikke gir tilstrekkelig innsikt i språkkompetansen til flerspråklige barn og unge som i varierende grad har kompetanse på første- og andrespråket.

For at det skal være mulig med de tospråklige kartleggingsprosedyrene som er anbefalt ved identifikasjon av DLD, og som resultatene fra avhandlingen støtter, er det nødvendig med valide måleinstrumenter på både første- og andrespråket. Avhandlingen belyser også at et stort og ressurskrevende arbeid ligger til grunn for å utvikle måleinstrumenter for å vurdere språkferdigheter og fonologisk minne på første- og andrespråket.

Det er bekymringsfullt at svært store ressurser både i skole og PPT benyttes for å vurdere om barn og unge skal ha rett til spesialundervisning eller rett til særskilt språkopplæring når tilgjengelig tester ikke er tilstrekkelig utviklet eller egnet for å avklare hvilke rettigheter de har. Kartlegging av flerspråklige barn og unge forutsetter også kompetanse om tospråklig utvikling, kartlegging og spesialpedagogikk hos de som skal foreta vurderingen. Det å benytte tester på begge språk, men samtidig være kritisk til hva testene kan gi av informasjon, er et mulig utgangspunkt.

Gitt mangel på valide kartleggingsprøver for førstespråket, er informasjon fra foreldre ofte den viktigste kilden til kunnskap om førstespråket. Systematisk innhenting av informasjon om førstespråkkompetanse er sentralt. Slik informasjon i kombinasjon med relevante tester som kan identifisere begrensninger i fonologisk korttidsminne, så som tester for ordspenn, setningsrepetisjon og nonord og også prøver for ekspressiv grammatikk, er sentralt for å identifisere DLD.

Forskningsfeltet, som omhandler flerspråklige utvalg med mulige spesialpedagogiske behov, er et ungt forskningsfelt og inkluderer mange problemstillinger som bør undersøkes nærmere. Mitt bidrag er et lite skritt på veien mot bedre kartleggingsprosedyrer for flerspråklige elever. Avhandlingen gir likevel en mer spesifisert kunnskap om flerspråklige elevers språkutvikling og om at kartlegging på andrespråket ikke er tilstrekkelig når flerspråklige elevers språkkompetanse skal vurderes.

Begrensninger

Det ville vært en styrke om avhandlingen hadde blitt utvidet med longitudinelle data eller om utvalget henvist til pedagogisk-psykologisk tjeneste for DLD hadde vært større.

Siden usikker grammatikk er et sentralt symptom på DLD, hadde det også vært ønskelig med et mål for grammatikk som var mer sensitivt for variasjon i grammatiske ferdigheter enn det som ble valgt.

Det hadde også vært ønskelig med mer presis bakgrunnsinformasjon om foreldrenes utdanningsnivå, deres deltagelse i arbeidslivet, utvalgets fritidsaktiviteter samt språkbruksmønstre i familien. Det er derfor også begrensninger med hensyn til å kunne inkludere slike bakgrunnsvariabler i analysene for å forklare mulige *årsaker* til forskjeller i språkferdigheter. Selv om mangelfull bakgrunnsinformasjon om informantene utgjør metodiske begrensninger, bidrar likevel generell informasjon om demografiske forhold i utvalgenes oppvekstmiljøer til å sannsynliggjøre at utvalgene har mer krevende oppvekstforhold enn utvalgene som utgjør sammenligningsgrunnlaget.

Litteraturliste

- Aben, B., Stapert, S. & Blokland, A. (2012). About the distinction between working memory and short-term memory. *Frontiers in psychology*, 3, 301.
- Abrahamsson, N. & Hyltenstam, K. (2009). Age of onset and nativelikeness in a second language: Listener perception versus linguistic scrutiny. *Language Learning*, 59(2), 249-306.
- Armon-Lotem, S., de Jong, J. & Meir, N. (2015). *Assessing multilingual children: Disentangling bilingualism from language impairment*. Multilingual matters.
- Arnesen, A., Bræken, J., Ogden, T. & Melby-Lervåg, M. (2019). Assessing children's social functioning and reading proficiency: A systematic review of the quality of educational assessment instruments used in Norwegian elementary schools. *Scandinavian Journal of Educational Research*, 63(3), 465-490.
- Arntzen, R. & Karlsen, J. (2019). Språkressurser, språkvalg og språkskifte: En studie av tospråklige barn fra fem-til tolvårsalder. *NOA-Norsk som andrespråk*, (1).
- August, D. & Shanahan, T. (2008). *Developing reading and writing in second-language learners: Lessons from the report of the National Literacy Panel on Language-Minority Children and Youth*. Routledge.
- August, D. & Shanahan, T. (2010). Response to a review and update on developing literacy in second-language learners: Report of the national literacy panel on language minority children and youth. *Journal of Literacy Research*, 42(3), 341-348.
- Baddeley, A. D. & Hitch, G. J. (2000). Development of working memory: Should the Pascual-Leone and the Baddeley and Hitch models be merged? *Journal of Experimental Child Psychology*, 77(2), 128-137.
- Baker, C. (2014). *A parents' and teachers' guide to bilingualism* (Bd. 18). Multilingual Matters.
- Bakken, A. & Hyggen, C. (2018). *Trivsel og utdanningsdriv blant minoritets elever i videregående. Hvordan forstå karakterforskjeller mellom elever med ulik innvandrerbakgrunn?* (8278946507).
<http://3.121.66.0/bitstream/handle/20.500.12199/1254/Nett-utgave-NOVA-Rapport-1-18.pdf?sequence=1&isAllowed=y>
- Barac, R., Bialystok, E., Castro, D. C. & Sanchez, M. (2014). The cognitive development of young dual language learners: A critical review. *Early Childhood Research Quarterly*, 29(4), 699-714.
- Baruch, Y. & Holtom, B. C. (2008). Survey response rate levels and trends in organizational research. *Human relations*, 61(8), 1139-1160.
- Bedore, L. M. & Pena, E. D. (2008). Assessment of bilingual children for identification of language impairment: Current findings and implications for practice. *International Journal of Bilingual Education and Bilingualism*, 11(1), 1-29.
- Bedore, L. M., Peña, E. D., Garcia, M. & Cortez, C. (2005). Conceptual Versus Monolingual Scoring When Does It Make a Difference? *Language, Speech, and Hearing Services in Schools*, 36(3), 188-200.
[https://doi.org/10.1044/0161-1461\(2005/020\)](https://doi.org/10.1044/0161-1461(2005/020))
- Bedore, L. M., Peña, E. D., Summers, C. L., Boerger, K. M., Resendiz, M. D., Greene, K., Bohman, T. M. & Gillam, R. B. (2012). The measure matters: Language dominance profiles across measures in Spanish-English bilingual children. *Bilingualism: Language and Cognition*, 15(03), 616-629.
<https://doi.org/10.1017/S1366728912000090>
- Bele, I., Helland, T., Ottem, E., Lian, A., Wold, A. H., Lyster, S.-A. H., Hundeide, K., Rommetveit, R., Rommetveit, K. & Solvang, P. K. (2008). *Språkvansker: teoretiske perspektiver og praktiske utfordringer*. Cappelen akademisk.
- Bishop, D. (1989). Test for reception of grammar. *London: Medical Research Council*.
- Bishop, D. (1998). *Uncommon understanding*. Psychology Press.
- Bishop, D., Snowling, M., Thompson, P. & Greenhalgh, T. (2016). CATALISE: A multinational and multidisciplinary Delphi consensus study. Identifying language impairments in children. *PloS one*, 11(7), e0158753.
- Bishop, D., Snowling, M. J., Thompson, P. A., Greenhalgh, T., Consortium, C., Adams, C., Archibald, L., Baird, G., Bauer, A. & Bellair, J. (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), 1068-1080.

- Bishop, D. V. (2014). Ten questions about terminology for children with unexplained language problems. *International Journal of Language & Communication Disorders*, 49(4), 381-415.
- Blom, E., Küntay, A. C., Messer, M., Verhagen, J. & Leseman, P. (2014). The benefits of being bilingual: Working memory in bilingual Turkish–Dutch children. *Journal of Experimental Child Psychology*, 128, 105-119.
- Bloom, L. & Lahey, M. (1978). Language development and language disorders.
- Bohman, T. M., Bedore, L. M., Peña, E. D., Mendez-Perez, A. & Gillam, R. B. (2010). What you hear and what you say: Language performance in Spanish–English bilinguals. *International Journal of Bilingual Education and Bilingualism*, 13(3), 325-344.
- Buac, M., Gross, M. & Kaushanskaya, M. (2016). Predictors of processing-based task performance in bilingual and monolingual children. *Journal of Communication Disorders*, 62, 12-29.
- Calvo, A. & Bialystok, E. (2014). Independent effects of bilingualism and socioeconomic status on language ability and executive functioning. *Cognition*, 130(3), 278-288.
- Chiat, S., Armon-Lotem, S., Marinis, T., Polisenska, K., Roy, P. & Seeff-Gabriel, B. (2013). The potential of sentence imitation tasks for assessment of language abilities in sequential bilingual children.
- Chiat, S. & Polišká, K. (2016). A framework for crosslinguistic nonword repetition tests: Effects of bilingualism and socioeconomic status on children's performance. *Journal of Speech, Language, and Hearing Research*, 59(5), 1179-1189.
- Cobo-Lewis, A. B., Pearson, B. Z., Eilers, R. E. & Umbel, V. C. (2002). *Effects of bilingualism and bilingual education on oral and written English skills: a multifactor study of standardized test outcomes; Effects of bilingualism and bilingual education on oral and written Spanish skills: a multifactor study of standardized test outcomes; Interdependence of Spanish and English knowledge in language and literacy among bilingual English children*. na.
- Core, C., Hoff, E., Rumiche, R. & Señor, M. (2013). Total and conceptual vocabulary in Spanish–English bilinguals from 22 to 30 months: implications for assessment. *Journal of Speech, Language, and Hearing Research*, 56(5), 1637-1649. [https://doi.org/10.1044/1092-4388\(2013/11-0044\)](https://doi.org/10.1044/1092-4388(2013/11-0044))
- Crocker, L. & Algina, J. (1986). *Introduction to classical and modern test theory*. Holt, Rinehart, and Winston.
- Cummins, J. (1984). Bilingualism and cognitive functioning. *S. Shapson &*, 500, 55-70.
- Cummins, J. (2000). *Language, power, and pedagogy: Bilingual children in the crossfire* (Bd. 23). Multilingual Matters.
- Cummins, J. (2012). The intersection of cognitive and sociocultural factors in the development of reading comprehension among immigrant students. *Reading and Writing*, 25(8), 1973-1990.
- Cummins, J. (2017). Teaching for transfer in multilingual school contexts. *Bilingual and multilingual education*, 3, 103-115.
- Cummins, J. (2021). *Rethinking the Education of Multilingual Learners: A Critical Analysis of Theoretical Concepts*. Multilingual Matters.
- de Abreu, P. M. E., Baldassi, M., Puglisi, M. L. & Befi-Lopes, D. M. (2013). Cross-linguistic and cross-cultural effects on verbal working memory and vocabulary: Testing language-minority children with an immigrant background. *Journal of Speech, Language, and Hearing Research*, 56(2), 630-642.
- de Abreu, P. M. E., Cruz-Santos, A., Tourinho, C. J., Martin, R. & Bialystok, E. (2012). Bilingualism enriches the poor enhanced cognitive control in low-income minority children. *Psychological science*, 23(11), 1364-1371. <https://doi.org/10.1177/0956797612443836>
- De Houwer, A. (2007). Parental language input patterns and children's bilingual use. *Applied Psycholinguistics*, 28(03), 411-424. <https://doi.org/10.1017/S0142716407070221>
- Dunn, L. M., Dunn, L. M., Whetton, C. & Burley, J. (1997). *The British Picture Vocabulary Scale* (2nd. utg.). nferNelson.
- Engebretsen, A. & Fuglerud, Ø. (2009). *Kultur og generasjon: tilpasningsprosesser blant somaliere og tamiler i Norge*. Universitetsforl.
- Engel, P. M. J., Santos, F. H. & Gathercole, S. E. (2008). Are working memory measures free of socioeconomic influence?
- Evers, A., Muñiz, J., Hagemester, C., Høstmælingen, A., Lindley, P., Sjöberg, A. & Bartram, D. (2013). Assessing the quality of tests: Revision of the EFPA review model. *Psicothema*, 25(3), 283-291.

- Extra, G. & Yagmur, K. (2010). Language proficiency and socio-cultural orientation of Turkish and Moroccan youngsters in the Netherlands. *Language and Education*, 24(2), 117-132. <https://doi.org/10.1080/09500780903096561>
- Fleckstein, A., Prévost, P., Tuller, L., Sizaret, E. & Zebib, R. (2018). How to identify SLI in bilingual children: A study on sentence repetition in French. *Language Acquisition*, 25(1), 85-101.
- Fulland, H. (2016). *Language minority children's perspectives on being bilingual-On" bilanguagers" and their sensitivity towards complexity* [Doktorgradsavhandling, Universitetet i Oslo].
- Gallinat, E. & Spaulding, T. J. (2014). Differences in the performance of children with specific language impairment and their typically developing peers on nonverbal cognitive tests: A meta-analysis. *Journal of Speech, Language, and Hearing Research*, 57(4), 1363-1382.
- García, O. & Baker, C. (2007). *Bilingual education: An introductory reader* (Bd. 61). Multilingual matters.
- Gathercole, S. E. & Baddeley, A. D. (1990). Phonological memory deficits in language disordered children: Is there a causal connection? *Journal of memory and language*, 29(3), 336-360.
- Geva, E. (2014). Introduction : The cross-language transfer journey – a guide to the perplexed. *Written Language & Literacy*, 17(1), 1-15.
- Geva, E. & Farnia, F. (2017). Issue Editor Foreword. *Topics in Language Disorders*, 37(2), 109-113.
- Geva, E. & Herbert, K. (2013). Assessment and interventions for English language learners with learning disabilities. *Learning about learning disabilities*, 271-198.
- Geva, E. & Massey-Garrison, A. (2013). A comparison of the language skills of ELLs and monolinguals who are poor decoders, poor comprehenders, or normal readers. *Journal of learning disabilities*, 46(5), 387-401.
- Geva, E. & Wiener, J. (2014). *Psychological Assessment of Culturally and Linguistically Diverse Children and Adolescents: A Practitioner's Guide*. Springer Publishing Company.
- Geva, E., Xi, Y., Massey-Garrison, A. & Mak, J. Y. (2019). Assessing Reading in Second Language Learners: Development, Validity, and Educational Considerations. I *Reading Development and Difficulties* (s. 117-155). Springer.
- Gibson, T. A., Oller, D. K., Jarmulowicz, L. & Ethington, C. A. (2011). The receptive–expressive gap in the vocabulary of young second-language learners: Robustness and possible mechanisms. *Bilingualism: Language and Cognition*, 15(1), 102-116. <https://doi.org/10.1017/S1366728910000490>
- Girbau, D. & Schwartz, R. G. (2008). Phonological working memory in Spanish–English bilingual children with and without specific language impairment. *Journal of Communication Disorders*, 41(2), 124-145.
- Govindarajan, K. & Paradis, J. (2019). Narrative abilities of bilingual children with and without Developmental Language Disorder (SLI): Differentiation and the role of age and input factors. *Journal of Communication Disorders*, 77, 1-16.
- Grosjean, F. (1998). Studying bilinguals: Methodological and conceptual issues. *Bilingualism: Language and Cognition*, 1(02), 131-149.
- Grosjean, F. (2008). *Studying bilinguals*. Oxford University Press, USA.
- Grosjean, F. (2010). *Bilingual*. Harvard university press.
- Gross, M., Buac, M. & Kaushanskaya, M. (2014). Conceptual scoring of receptive and expressive vocabulary measures in simultaneous and sequential bilingual children. *American Journal of Speech-Language Pathology*, 23(4), 574-586. https://doi.org/10.1044/2014_AJSLP-13-0026
- Gunnerud, H. L. (2021). Bilingualism: Advantages and Disadvantages in Cognitive Processing, Language and Reading Comprehension.
- Gunnerud, H. L., Ten Braak, D., Reikerås, E. K. L., Donolato, E. & Melby-Lervåg, M. (2020). Is bilingualism related to a cognitive advantage in children? A systematic review and meta-analysis. *PSYCHOLOGICAL BULLETIN*.
- Hakansson, G., Salameh, E.-K. & Nettelbladt, U. (2003). Measuring language development in bilingual children: Swedish-Arabic children with and without language impairment. *Linguistics*, 41(2; ISSU 384), 255-288.
- Hammer, C. S., Hoff, E., Uchikoshi, Y., Gillanders, C., Castro, D. C. & Sandilos, L. E. (2014). The language and literacy development of young dual language learners: A critical review. *Early Childhood Research Quarterly*, 29(4), 715-733. <https://doi.org/10.1016/j.ecresq.2014.05.008>
- Hart, B. & Risley, T. R. (1995). *Meaningful differences in the everyday experience of young American children*. Paul H Brookes Publishing.

- Heller, M. C. (2020). Supporting second-language learning in the early elementary years: A researcher-practitioner approach.
- Hoff, E. (2014). *Language development* (5th. utg.). Wadsworth Cengage learning.
- Hulme, C., Roodenrys, S., Schweickert, R., Brown, G. D., Martin, S. & Stuart, G. (1997). Word-frequency effects on short-term memory tasks: Evidence for a reintegration process in immediate serial recall. *Journal of Experimental Psychology: Learning, memory, and cognition*, 23(5), 1217.
- Hulme, C. & Snowling, M. J. (2009). *Developmental disorders of language learning and cognition*. Wiley-Blackwell. <http://search.ebscohost.com/login.aspx?direct=true&scope=site&db=nlebk&db=nlabk&AN=613441>
- Hwang, J. K., Mancilla-Martinez, J., McClain, J. B., Oh, M. H. & Flores, I. (2020). Spanish-speaking English learners' English language and literacy skills: The predictive role of conceptually scored vocabulary. *Applied Psycholinguistics*, 41(1), 1-24.
- Jordaan, H. L., van Zyl, A. & O'Brien, T. (2021). Conceptual vocabulary in Grade 2 isiZulu-English and Grade 1 Afrikaans-English bilinguals. *South African Journal of Childhood Education*, 11(1), 1-10.
- Karlsen, J., Lyster, S.-A. & Lervåg, A. (2016). Vocabulary development in Norwegian L1 and L2 learners in the kindergarten–school transition. *Journal of Child Language*, 1-25.
- Kerlinger, F. N., Lee, H. B. & Bhanthumnavin, D. (2000). Foundations of behavioral research: The most sustainable popular textbook by Kerlinger & Lee (2000). I.
- Klem, M. (2014). *Språkkartlegging som grunnlag for identifisering av førskolebarn i risiko for språkvansker: En modellbasert tilnærming* [Doktoravhandling, Det utdanningsvitenskapelige fakultet, Universitetet i Oslo]. Oslo.
- Klem, M., Melby-Lervåg, M., Hagtvet, B., Lyster, S. A. H., Gustafsson, J. E. & Hulme, C. (2015). Sentence repetition is a measure of children's language skills rather than working memory limitations. *Developmental science*, 18(1), 146-154.
- Kleven, T. A. H., Finn. (2018). *Innføring i forskningsmetode. En hjelp til kritisk tolkning og vurdering*. Fagbokforlaget.
- Kohnert, K. (2010). Bilingual children with primary language impairment: Issues, evidence and implications for clinical actions. *Journal of Communication Disorders*, 43(6), 456-473.
- Kohnert, K. (2013). *Language disorders in bilingual children and adults*. Plural Publ.
- Kohnert, K., Kan, P. F. & Conboy, B. T. (2010). Lexical and grammatical associations in sequential bilingual preschoolers. *Journal of Speech, Language, and Hearing Research*, 53(3), 684-698. [https://doi.org/10.1044/1092-4388\(2009/08-0126\)](https://doi.org/10.1044/1092-4388(2009/08-0126))
- Kunnskapsdepartementet. (2007). *Likeverdig opplæring i praksis! : Strategi for bedre læring og større deltakelse av språklige minoriteter i barnehage, skole og utdanning 2007–2009*. Kunnskapsdepartementet. https://www.regjeringen.no/globalassets/upload/kd/vedlegg/grunnskole/strategiplaner/udir_likeverdig_opplaering2_07.pdf
- Kunnskapsdepartementet. (2008). *Språk bygger broer — Språkstimulering og språkopplæring for barn, unge og voksne* (St.meld. nr. 23 (2007-2008)). <https://www.regjeringen.no/no/dokumenter/stmeld-nr-23-2007-2008-/id512449/>
- Kunnskapsdepartementet. (2010-2020). *Tett på – tidlig innsats og inkluderende fellesskap i barnehage, skole og SFO*. <https://www.regjeringen.no/no/dokumenter/meld.-st.-6-20192020/id2677025/>
- Leerand, D. (2014). Den arabiske våren. *Store Norske Leksikon*.
- Leonard, L. B. (2014). *Children with specific language impairment*. MIT press.
- Lervåg, A. & Aukrust, V. G. (2010). Vocabulary knowledge is a critical determinant of the difference in reading comprehension growth between first and second language learners. *Journal of Child Psychology and Psychiatry*, 51(5), 612-620. <https://doi.org/10.1111/j.1469-7610.2009.02185.x>
- Lian, A., Karlse, P. J. & Eriksen, T. B. (2004). Opposing effects of phonological similarity on item and order memory of words and nonwords in the serial recall task. *Memory*, 12(3), 314-337.
- Lyster, S. & Horn, E. (2009). *Test for Reception og Grammar (TROG-2)*. Norsk Versjon. I. Stockholm: Pearson Assessment.

- Lyster, S., Horn, E. & Rygvold, A.-L. (2010). Ordforråd og ordforrådsutvikling hos norske barn og unge. *Spesialpedagogikk*, 9, 35-43.
- Lyster, S. H., E. (2009). Test for Reception og Grammar (TROG-2). Norsk Versjon. I. Stockholm: Pearson Assessment.
- Mancilla-Martinez, J. & Vagh, S. B. (2013). Growth in toddlers' Spanish, English, and conceptual vocabulary knowledge. *Early Childhood Research Quarterly*, 28(3), 555-567. <https://doi.org/10.1016/j.ecresq.2013.03.004>
- Maxwell, J. A. (2013). *Qualitative research design: an interactive approach*. Sage.
- McCabe, A., Bornstein, M. H., Guerra, A. W., Kuchirko, Y., Páez, M., Tamis-LeMonda, C. S., Cates, C. B., Hirsh-Pasek, K., Melzi, G. & Song, L. (2013). Multilingual Children: Beyond Myths and Toward Best Practices and commentaries. *Social Policy Report*, 27(4), 1-37.
- Melby-Lervåg, M. & Hulme, C. (2010). Serial and free recall in children can be improved by training: Evidence for the importance of phonological and semantic representations in immediate memory tasks. *Psychological science*, 21(11), 1694-1700.
- Melby-Lervåg, M. & Lervåg, A. (2014). Reading comprehension and its underlying components in second-language learners: A meta-analysis of studies comparing first- and second-language learners. *Psychological Bulletin* *Psychological Bulletin*, 140(2), 409-433. <https://doi.org/10.1037/a0033890>
- Messer, M. H., Leseman, P. P., Boom, J. & Mayo, A. Y. (2010). Phonotactic probability effect in nonword recall and its relationship with vocabulary in monolingual and bilingual preschoolers. *Journal of Experimental Child Psychology*, 105(4), 306-323.
- Messick, S. (1989a). Meaning and values in test validation: The science and ethics of assessment. *Educational Researcher*, 18(2), 5-11.
- Messick, S. (1989b). Validity. I *Educational measurement, 3rd ed.* (s. 13-103) (The American Council on Education/Macmillan series on higher education.). American Council on Education.
- Messick, S. (1993). Foundations of validity: Meaning and consequences in psychological assessment. *ETS Research Report Series*, 1993(2), i-18.
- Messick, S. (1995). Validity of psychological assessment: Validation of inferences from persons' responses and performances as scientific inquiry into score meaning. *American psychologist*, 50(9), 741.
- Monsrud, M.-B., Rydland, V., Geva, E., Thurmann-Moe, A. C. & Halaas Lyster, S.-A. (2019). The advantages of jointly considering first and second language vocabulary skills among emergent bilingual children. *International Journal of Bilingual Education and Bilingualism*, 1-17.
- Montrul, S. (2013). How "native" are heritage speakers. *Heritage Language Journal*, 10(2), 15-39.
- Montrul, S. (2016). *The acquisition of heritage languages*. Cambridge University Press.
- Morales, J., Calvo, A. & Bialystok, E. (2013). Working memory development in monolingual and bilingual children. *Journal of Experimental Child Psychology*, 114(2), 187-202.
- Muthén, B. & Muthén, L. (2017). *Mplus*. Chapman and Hall/CRC.
- Nordahl, T. & Hausstätter, R. (2009). Spesialundervisningens forutsetninger, innsatser og resultater. Situasjonen til elever med særlige behov under Kunnskapsløftet [Special education under the knowledge promotion reform]. I. Hamar: Høgskolen i Hedmark/Utdanningsdirektoratet.
- O'Connor, M., Geva, E. & Koh, P. W. (2019). Examining Reading Comprehension Profiles of Grade 5 Monolinguals and English Language Learners Through the Lexical Quality Hypothesis Lens. *Journal of learning disabilities*, 52(3), 232-246.
- OECD. (2020). *Early Learning and Child Well-being*. <https://doi.org/doi:https://doi.org/10.1787/3990407f-en>
- Oh, M. H. & Mancilla-Martinez, J. (2021). Comparing Vocabulary Knowledge Conceptualizations Among Spanish-English Dual Language Learners in a New Destination State. *Language, Speech, and Hearing Services in Schools*, 52(1), 369-382.
- Oller, D. K. & Eilers, R. E. (2002). *Language and literacy in bilingual children* (Bd. 2). Multilingual Matters.
- Oller, D. K., Pearson, B. Z. & Cobo-Lewis, A. B. (2007a). Profile effects in early bilingual language and literacy. *Applied Psycholinguistics*, 28(02), 191-230.
- Oller, D. K., Pearson, B. Z. & Cobo-Lewis, A. B. (2007b). Profile effects in early bilingual language and literacy. *Applied Psycholinguistics*, 28(2), 191-230.

- Ottem, E. & Frost, J. (2005). Språk 6-16; screening test: manual. *Bredtvet kompetansesenter*.
- Ottem, E. & Lian, A. (2008). Spesifikke språkvansker I. I (s. s.31-42). Cappelen akademisk.
- Ouellette, G. P. (2006). What's meaning got to do with it: The role of vocabulary in word reading and reading comprehension. *Journal of Educational Psychology, 98*(3), 554.
- Paradis, J. (2011). Individual differences in child English second language acquisition: Comparing child-internal and child-external factors. *Linguistic Approaches to Bilingualism, 1*(3), 213-237.
- Paradis, J. (2016). The development of English as a second language with and without specific language impairment: Clinical implications. *Journal of Speech, Language, and Hearing Research, 59*(1), 171-182.
- Paradis, J. (2019a). English Second Language Acquisition from Early Childhood to Adulthood: The Role of Age, First Language, Cognitive, and Input Factors.
- Paradis, J. (2019b). English second language acquisition from early childhood to adulthood: The role of age, first language, cognitive, and input factors. Proceedings of the BUCLD,
- 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.
- Paradis, J., Genesee, F. & Crago, M. B. (2011). *Dual language development and disorders: a handbook of bilingualism and second language learning* (2. utg.). Brookes.
- Paradis, J., Marinova-Todd, S., Soto-Corominas, A., Zhang, X., Rusk, B. & Day, K. (2019). Language and literacy skills of English Language Learners in middle school (grades 7-9): How do they compare to their monolingual peers?
- Paradis, J., Rusk, B., Duncan, T. S. & Govindarajan, K. (2017). Children's second language acquisition of English complex syntax: The role of age, input, and cognitive factors. *Annual Review of Applied Linguistics, 37*, 148-167.
- Paul, R., Norbury, C. & Gosse, C. (2018). Language disorders from infancy through adolescence: Listening, Speaking. *Reading, Writing, and Communicating, 4*, 756.
- Pearson, B. Z., Fernandez, S. C. & Oller, D. K. (1993). Lexical development in bilingual infants and toddlers: Comparison to monolingual norms. *Language Learning, 43*(1), 93-120. <https://doi.org/10.1111/j.1467-1770.1993.tb00174.x>
- Peña, E. D. (2007). Lost in translation: Methodological considerations in cross-cultural research. *Child development, 78*(4), 1255-1264.
- Peña, E. D. (2016). Supporting the home language of bilingual children with developmental disabilities: From knowing to doing. *Journal of Communication Disorders, 63*, 85-92.
- Peña, E. D., Bedore, L. M. & Kester, E. S. (2016). Assessment of language impairment in bilingual children using semantic tasks: Two languages classify better than one. *International Journal of Language & Communication Disorders, 51*(2), 192-202.
- Peña, E. D., Gillam, R. B., Bedore, L. M. & Bohman, T. M. (2011). Risk for poor performance on a language screening measure for bilingual preschoolers and kindergarteners. *American Journal of Speech-Language Pathology, 20*(4), 302-314.
- Peña, E. D., Gutiérrez-Clellen, V. F., Iglesias, A., Goldstein, B. A. & Bedore, L. M. (2018). Bilingual English Spanish Assessment (BESA). *Baltimore, MD: Brookes*.
- Pena, E. D. & Halle, T. G. (2011). Assessing preschool dual language learners: Traveling a multiforked road. *Child Development Perspectives, 5*(1), 28-32.
- Peña, I. (2008). *Getting lost in translation : the dangers in literal translation*. <https://digitalcommons.hsc.unt.edu/theses/772/>
- Pettersen, S. V. & Østby, L. (2014). *Immigrants in Norway, Sweden and Denmark*. <https://www.ssb.no/en/befolkning/artikler-og-publikasjoner/immigrants-in-norway-sweden-and-denmark>
- Pierce, L. J., Chen, J.-K., Delcenserie, A., Genesee, F. & Klein, D. (2015). Past experience shapes ongoing neural patterns for language. *Nature Communications, 6*, 10073.
- Polišenská, K., Chiat, S. & Roy, P. (2015). Sentence repetition: What does the task measure? *International Journal of Language & Communication Disorders, 50*(1), 106-118.

- Pratt, A. S., Adams, A. M., Peña, E. D. & Bedore, L. M. (2020). Exploring the Use of Parent and Teacher Questionnaires to Screen for Language Disorders in Bilingual Children. *Topics in Early Childhood Special Education*, 0271121420942308.
- Pratt, A. S., Peña, E. D. & Bedore, L. M. (2021). Sentence repetition with bilinguals with and without DLD: differential effects of memory, vocabulary, and exposure. *Bilingualism: Language and Cognition*, 24(2), 305-318.
- Reilly, S., Bishop, D. V. & Tomblin, B. (2014). Terminological debate over language impairment in children: forward movement and sticking points. *International Journal of Language & Communication Disorders*, 49(4), 452-462.
- Riches, N. G. (2012). Sentence repetition in children with specific language impairment: An investigation of underlying mechanisms. *International Journal of Language & Communication Disorders*, 47(5), 499-510.
- Rogde, K., Melby-Lervåg, M. & Lervåg, A. (2016). Improving the General Language Skills of Second-Language Learners in Kindergarten: A Randomized Controlled Trial. *Journal of Research on Educational Effectiveness*, 9(sup1), 150-170. <https://doi.org/10.1080/19345747.2016.1171935>
- Rydland, V., Aukrust, V. G. & Fulland, H. (2013). Living in neighborhoods with high or low co-ethnic concentration: Turkish–Norwegian-speaking students' vocabulary skills and reading comprehension. *International Journal of Bilingual Education and Bilingualism*, 16(6), 657-674. <https://doi.org/10.1080/13670050.2012.709224>
- Rydland, V. & Grøver, V. (2021). Language use, home literacy environment, and demography: Predicting vocabulary skills among diverse young dual language learners in Norway. *Journal of Child Language*, 48(4), 717-736.
- Salameh, E.-K. (Red.). (2012). *Flerspråkighet i skolan: Språklig utveckling och undervisning*. Natur & Kultur.
- Sattler, J. M. (2008). *Assessment of children : cognitive foundations* (5th ed. utg.). Jerome M. Sattler Publ.
- Schmitt, N. (2014). Size and depth of vocabulary knowledge: What the research shows. *Language Learning*, 64(4), 913-951. <https://doi.org/10.1111/lang.12077>
- Sevinç, Y. & Backus, A. (2017). Anxiety, language use and linguistic competence in an immigrant context: A vicious circle? *International Journal of Bilingual Education and Bilingualism*.
- Shadish, W. R., Cook, T. D. & Campbell, D. T. (2002). Experimental and quasi-experimental designs for generalized causal inference.
- Shany, M. & Geva, E. (2012). Cognitive, language, and literacy development in socio-culturally vulnerable school children—the case of Ethiopian Israeli children. I M. Leikin, M. Schwartz & Y. Tobin (Red.), *Current Issues in Bilingualism: Cognitive and Socio-linguistic Perspectives* (s. 77-117). Springer.
- Simonsen, H. G. & Bjerkan, K. M. (1998). Testing past tense inflection in Norwegian: a diagnostic tool for identifying SLI children? *International journal of applied linguistics*, 8(2), 251-270.
- Smithson, L., Paradis, J. & Nicoladis, E. (2014). Bilingualism and receptive vocabulary achievement: Could sociocultural context make a difference? *Bilingualism: Language and Cognition*, 17(04), 810-821. <https://doi.org/10.1017/S1366728913000813>
- Snow, C. E. (2010). Academic language and the challenge of reading for learning. *science*, 328(5977), 450-452.
- Snow, C. E. (2017). The role of vocabulary versus knowledge in children's language learning: a fifty-year perspective/El papel del vocabulario frente al conocimiento en el aprendizaje lingüístico de los niños: una perspectiva de cincuenta años. *Infancia y Aprendizaje*, 40(1), 1-18.
- Solari, E. J., Petscher, Y. & Folsom, J. S. (2014). Differentiating literacy growth of ELL students with LD from other high-risk subgroups and general education peers: Evidence from grades 3–10. *Journal of learning disabilities*, 47(4), 329-348.
- Statistics Norway. (2016). *Immigrants and Norwegian-born to immigrant parents, 1 January 2016*. Hentet 09.06.2016 fra <https://www.ssb.no/en/befolkning/statistikker/innvbef/aar/2016-03-03>
- Tabachnick, B. G. & Fidell, L. S. (2014). *Using multivariate statistics*. Pearson.
- Tabachnick, B. & Fidell, L. (2001). *Using multivariate statistics* (Sth ed.). I. United States of America: Allyn & Bacon.
- The American- Speech-Language-Hearing- Association. (2012). *ASHA's Recommended to the DSM-5*. <http://www.asha.org/uploadedFiles/DSM-5-Final-Comments.pdf>

- Thordardottir, E. (2015). Proposed diagnostic procedures for use in bilingual and cross-linguistic contexts. *Assessing multilingual children: Disentangling bilingualism from language impairment*, 13, 331.
- Thordardottir, E. (2016). Grammatical morphology is not a sensitive marker of language impairment in Icelandic in children aged 4–14 years. *Journal of Communication Disorders*, 62, 82-100.
- Thordardottir, E. & Brandeker, M. (2013). The effect of bilingual exposure versus language impairment on nonword repetition and sentence imitation scores. *Journal of Communication Disorders*, 46(1), 1-16.
- Thordardottir, E., Rothenberg, A., Rivard, M.-E. & Naves, R. (2006). Bilingual assessment: Can overall proficiency be estimated from separate measurement of two languages? *Journal of Multilingual Communication Disorders*, 4(1), 1-21.
- Thordardottir, E. & Topbaş, S. (2021). How aware is the public of the existence, characteristics and causes of language impairment in childhood and where have they heard about it? A European survey. *Journal of Communication Disorders*, 89, 106057.
- Tuller, L., Hamann, C., Chilla, S., Ferré, S., Morin, E., Prevost, P., Dos Santos, C., Abed Ibrahim, L. & Zebib, R. (2018). Identifying language impairment in bilingual children in France and in Germany. *International Journal of Language & Communication Disorders*, 53(4), 888-904.
- Vagh, S. B., Pan, B. A. & Mancilla-Martinez, J. (2009). Measuring growth in bilingual and monolingual children's English productive vocabulary development: The utility of combining parent and teacher report. *Child development*, 80(5), 1545-1563. <https://doi.org/10.1111/j.1467-8624.2009.01350.x>
- Vik, N. E. (2017). Språkkartlegging av flerspråklige barn i barnehagen-fra kontrovers til kompromiss? *Nordisk barnehageforskning*, 14.
- Von Bastian, C., De Simoni, C., Kane, M., Carruth, N. & Miyake, A. (2017). Does being bilingual entail advantages in working memory? A meta analysis. 58th Annual Meeting of Psychonomic Society, Vancouver,
- Wei, L. & Auer, P. (2009). *Handbook of multilingualism and multilingual communication*. Mouton de Gruyter.
- Zebib, R., Tuller, L., Hamann, C., Abed Ibrahim, L. & Prévost, P. (2019). Syntactic complexity and verbal working memory in bilingual children with and without Developmental Language Disorder. *First Language*, 0142723719888372.
- Østby, L. (2016). Bedre integrert enn mor og far. *Plan*, 49(01), 18-25.
- Øzerk, K. (2008). Tospråklig opplæring og funksjonell tospråklighet. *Flerkulturell virkelighet i skole og samfunn*. Oslo: Cappelen Damm Akademisk.
- Aagaard, K. E. (2011). Den språklige faktor: Pedagogisk-psykologisk utredning av barn med minoritetsspråklig bakgrunn.

Vedlegg

Vedlegg A: Kongsesjon

Vedlegg B: Tilråding fra Norsk Samfunnsvitenskapelige Datatjeneste

Vedlegg C: Informasjon om forarbeidet til avhandlingen

Vedlegg D: Informasjon til skolene

Vedlegg E 1: Informasjonsskriv til foreldre på skolene

Vedlegg E2: Eksempel på informasjonsskriv, vietnamesisk

Vedlegg F: Bakgrunnsinformasjon fra foreldre

Vedlegg G Informasjonsskriv til foreldre til elever tilmeldt pedagogisk-psykologisk tjeneste

Vedlegg: Korrelasjonstabell for tester for første – og andrespråket

Datatilsynet



12012

K	UFD
Saknr.	15 MARS 2004 200306534
Arkivkode:	

Utdannings- og forskningsdepartementet
Avdeling for kompetanse og arbeidsliv
v/ Laila Gaupset
Postboks 8119 Dep
0032 OSLO

KOPI

Deres ref
200306534

Vår ref (bes oppgitt ved svar)
2002/922-9 MOF/-

Dato
12.03.2004

KONSESJON TIL Å BEHANDLE PERSONOPPLYSNINGER

Vi viser til Deres brev av 08.03.2004 om endring av behandlingsansvarlig for statlig spesialpedagogisk støttesystem og tidligere gitt konsesjon av 02.10.2002. Det utstedes herved ny konsesjon, og tidligere gitt konsesjon (ref 2002/922-4) trekkes tilbake med virkning fra og med 15.06.04. Det bes om at konsesjon av 02.10.2002 (ref 2002/922-4) oversendes tilsynet eller makuleres med bekreftelse til tilsynet.

Datatilsynet har vurdert søknaden og gir Dem med hjemmel i personopplysningsloven § 33, jf. § 34, konsesjon til å behandle personopplysninger til følgende formål: statlig spesialpedagogisk støttesystem.

Behandlingsansvarlig er Utdanningsdirektoratet. Det daglige ansvaret ligger hos virksomhetens øverste leder. Det daglige ansvaret kan delegeres.

Konsesjonen er gitt under forutsetning av at behandlingen foretas i henhold til søknaden og de bestemmelser som følger av personopplysningsloven med forskrifter.

Dersom det skjer endringer i behandlingen i forhold til de opplysninger som er gitt i søknaden, må dette fremmes i ny konsesjonssøknad.

I medhold av personopplysningsloven § 35, fastsettes i tillegg følgende vilkår for behandlingen:

1. Den behandlingsansvarlige skal hvert tredje år sende Datatilsynet bekreftelse på at behandlingen skjer i overensstemmelse med søknaden og personopplysningslovens regler.

Med hilsen
[signature removed]

Guro Slettemark (e f)
seniorrådgiver

[signature removed]

Monica Fornes
rådgiver (saksbehandler, tel 22 39 69 00)

Postadresse:
Postboks 8177 Dep
0034 OSLO

Kontoradresse:
Tollbugt 3

Telefon:
22 39 69 00

Telefaks:
22 42 23 50

Org.nr:
974 761 467

Hjemmeside:
www.datatilsynet.no



May-Britt Monsrud
Institutt for yrkesfaglærerutdanning
Høgskolen i Oslo og Akershus
Kunnskapsveien 55
2007 KJELLER

Vår dato: 22.10.2012

Vår ref:31522 / 3 / MSI

Deres dato:

Deres ref:

TILBAKEMELDING PÅ MELDING OM BEHANDLING AV PERSONOPPLYSNINGER

Vi viser til melding om behandling av personopplysninger, mottatt 18.09.2012. All nødvendig informasjon om prosjektet forelå i sin helhet 16.10.2012. Meldingen gjelder prosjektet:

31522 *Flerspråklige elevers språkferdigheter*
Behandlingsansvarlig *Høgskolen i Oslo og Akershus, ved institusjonens øverste leder*
Daglig ansvarlig *May-Britt Monsrud*

Personvernombudet har vurdert prosjektet, og finner at behandlingen av personopplysninger vil være regulert av § 7-27 i personopplysningsforskriften. Personvernombudet tilrår at prosjektet gjennomføres.

Personvernombudets tilråding forutsetter at prosjektet gjennomføres i tråd med opplysningene gitt i meldeskjemaet, korrespondanse med ombudet, eventuelle kommentarer samt personopplysningsloven og helseregisterloven med forskrifter. Behandlingen av personopplysninger kan settes i gang.

Det gjøres oppmerksom på at det skal gis ny melding dersom behandlingen endres i forhold til de opplysninger som ligger til grunn for personvernombudets vurdering. Endringsmeldinger gis via et eget skjema, http://www.nsd.uib.no/personvern/forsk_stud/skjema.html. Det skal også gis melding etter tre år dersom prosjektet fortsatt pågår. Meldinger skal skje skriftlig til ombudet.

Personvernombudet har lagt ut opplysninger om prosjektet i en offentlig database, <http://pvo.nsd.no/prosjekt>.

Personvernombudet vil ved prosjektets avslutning, 01.01.2017, rette en henvendelse angående status for behandlingen av personopplysninger.

Venlig hilsen
[signature removed]

Vigdís Namtvedt Kvalheim

[signature removed]

Marte Sivertsen

Kontaktperson: Marte Sivertsen tlf: 55 58 33 48
Vedlegg: Prosjektvurdering

«Språklige minoriteter med behov for særskilt tilrettelagt opplæring»,

et tiltak i «Strategiplanen Likeverdig opplæring i praksis» (Kunnskapsdepartementet 2007)

Tiltaket, «Språklige minoriteter med behov for særskilt tilrettelagt opplæring» var et samarbeid mellom NAFO (Nasjonalt senter for flerkulturell opplæring), Torshov kompetansesenter og Bredtvet kompetansesenter. Jeg var prosjektleder ved Bredtvet kompetansesenter. For å kunne gjennomføre det store prosjektet ble etter hvert en prosjektgruppe ved Bredtvet kompetansesenter etablert. Prosjektmedarbeidere var Kirsten Meyer Bjerkan og Anne-Cathrine Thurmann-Moe. Målsettingen for prosjektet var å utvikle kartleggingsverktøy og prosedyrer for å kartlegge mulige spesialpedagogiske behov i utvalg med flerspråklig bakgrunn. NAFO utviklet prøver for å kartlegge fonologiske ferdigheter, minnefunksjonen og prosesseringsferdigheter på morsmål, <http://nafo.oslomet.no> (Bøyese, 2008; Bøyese, 2017). En fagbok om utredningsarbeid ble utarbeidet ved Torshov kompetansesenter (Egeberg, 2007). Ved Bredtvet kompetansesenter ble en reseptiv (BPVS-II) og en ekspressiv vokabulartest (Begreper fra Språk 6-16), to minne tester; Setningsrepetisjon og Ordspenn (fra Språk 6-16) samt en test for grammatisk forståelse; Trog-2 valgt ut, oversatt og prøvd ut i et stort utvalg barn og unge med flerspråklig bakgrunn. Utvalget inkluderte data for førstespråket for 648 elever (jenter = 344, gutter= 304) og norsk for 783 elever på (jenter= 399, gutter= 383) (upubliseret prosjektrapport, Monsrud, Bjerkan & Thurmann-Moe 2016). Noen av elevene var bare teste på førstespråket sitt og noen bare på andrespråket. Utvalget med testresultater fra begge språk er inkludert i avhandlingen min. Det er publisert populærvitenskapelige artikler og en fagbok med utgangspunkt i prosjektet ved Bredtvet kompetansesenter. I forbindelse med prosjektet ble det utnevnt en referansegruppe som besto av Morten Hendis ved Torshov kompetansesenter, Olaf Huseby NTNU, Lise Iversen Kulbranstad, Høgskolen i Hedmark og Steven von Tetzner, Psykologisk institutt ved Universitet i Oslo. Det var også mange oversettere og tilbakeoversettere involvert i arbeidet med oversettelse og tilpasning av tester

Språk	Oversettere og tilbakeoversettere
Albansk	Tolk: Nufri Lekaj, pedagog/tolk: Basri Trena, forsker: Anila Nauni, Universitetet i Oslo
Somali	Statsautorisert tolk Osman Abdi Guled, pedagog Hassan Balbaal, pedagog Bashir Jama, Elverum
Tamil	Journalist Chelliah Rajan, pedagog Sivathas Sivabalasingaun, forsker Dhayalan Velauthapillai
Tyrkisk	Seniorrådgiver Meral Özerk, Statped/ Torshov kompetansesenter, professor Kamil Özerk, IPED forsker Emel Türker, Multiling
Urdu/punjabi	Tolk/forsker Jamshed Iqbal Rana, pedagog Rukhsana Arshad, canded Noshila Riaz, lærer Usmana Sahar, universitetslektor Øyvind Johan Stenberg,
Vietnamesisk	Tolk Duc Thinh Vu, antropolog Whyn Lam, I.amanuensis Olaf Husby

Publikasjoner

- Bjerkan, K. M., Monsrud, M.-B. & Thurmann-Moe, A. C. (2013). *Ordforråd hos flerspråklige barn: pedagogiske og spesialpedagogiske utfordringer*. Gyldendal Akademisk.
- Bøyese, L. (2008). Systematisk lærerarbeid om fagopplæring og lese- og skriveopplæring. I S. Aamodt & A.-M. Hauge (Red.), *Familielæring og andre modeller innenfor flerkulturelt arbeid*. Fagbokforlaget.
- Bøyese, L. (2017). Kartlegging i et minoritetsspråklig perspektiv. I M. Lunde, & S. Aamot (Red.). *Inkluderende og flerspråklig opplæring*, 108-129.
- Egeberg, E. (2007). *Minoritetsspråklige med særskilte behov: en bok om utredningsarbeid*. Cappelen akademisk.
- Monsrud, M.-B., Bjerkan, K. M. & Thurmann-Moe, A. C. (2011). Minoritetsspråklige barn og unges grammatiske utvikling. *Psykologi i kommunen*, 46(4), 3-16.
- Monsrud, M.-B., Thurmann-Moe, A. C. & Bjerkan, K. M. (2010). Minoritetsspråklige barns ordforråd og ordforrådsutvikling. *Spesialpedagogikk*, 9, 44-53.

Thurmann-Moe, A.C. (2017) Prediksjon av leseforståelse hos flerspråklige elever i longitudinelt perspektiv, *Spesialpedagogikk* 01/2017.

<https://www.utdanningsnytt.no/files/2019/06/27/Spesialpedagogikk%201%202017.pdf>

UTPRØVING AV KARTLEGGINGSMATERIELL

Til foresatte

På oppdrag fra Utdanningsdirektoratet utarbeider Bredtvet kompetansesenter oversettelser av kartleggingsverktøy for å vurdere ulike sider av språk og minne på flere morsmål. Dette gjelder språkene albansk (kosovo), somalisk, tamilsk, tyrkisk, urdu/punjabi og vietnamesisk

Det foreligger nå en utgave på albansk som vi ønsker å prøve ut på en gruppe barn i skolealder.

Elevene vil bli kartlagt både på albansk og norsk. Hensikten med dette er å vurdere om oversettelsen fungerer for elevgruppen. Prøveresultatene vil bli behandlet anonymt.

Utprøvingen vil foregå i løpet av 2008 og 2009. Det kommer til å gå 6 mnd mellom kartleggingen på hvert av språkene. Kartleggingen tar ca 45 minutter på hvert språk.

Vi håper dere gir tillatelse til at deres barn kan delta i utprøvingen.

Vi har også et spørreskjema som vi ønsker at foreldre skal fylle ut.

Takk for hjelpen!

May-Britt Monsrud,
prosjektleder

klipp:

Jeg gir tillatelse til at mitt barn kan delta i utprøvingen.

Foresatte:.....

Dato.....

ÁP DỤNG THỬ CÁC DỮ KIỆN DÙNG ĐỂ KHẢO SÁT

Kính gửi : Quý phụ huynh

Trung tâm thẩm định khả năng Bredtvet hiện đang soạn thảo bản dịch về những dữ kiện dùng để khảo sát nhằm nhận xét các lãnh vực khác nhau về ngôn ngữ và trí nhớ qua tiếng mẹ đẻ.

Hiện nay chúng tôi có một ấn bản bằng tiếng Việt mà chúng tôi muốn áp dụng thử với một nhóm trẻ em trong tuổi đi học.

Các em học sinh sẽ được những cộng tác viên làm việc trong dự án khảo sát cả bằng tiếng Việt lẫn tiếng Na uy. Mục đích của việc này là để nhận xét xem sự phiên dịch có hiệu quả đối với nhóm học sinh này hay không.

Việc áp dụng thử sẽ diễn ra trong niên khoá 2007/2008. Các em học sinh sẽ được khảo sát bằng tiếng Na uy và tiếng Việt. Thời gian kéo dài giữa những lần khảo sát ngôn ngữ này sẽ cách nhau 6 tháng. Việc khảo sát này sẽ dùng thời gian khoảng 45 phút cho mỗi ngôn ngữ. Chúng tôi cũng có phiếu trả lời các câu hỏi dành cho phụ huynh.

Cám ơn sự giúp đỡ của quý vị !

May-Britt Monsrud
Dự án trưởng

Cắt ở phần này khi trả lời

Tôi cho phép con của tôi là : có thể tham dự vào việc áp dụng thử này.

.....
Chữ ký của phụ huynh

.....
Ngày

ID-skjema tiltak 10 / elev

IDNR: _____

FØDSELSDATO: _____

KJØNN: _____

Hvem er informant?

- Mor
 Far
 Begge

Hvem bor barnet sammen med:

	Ja	Nei
Mor		
Far		
Søsken		
Andre voksne		

Hvor er barnet født?

- I Norge
 Annet land, spesifiser _____



Hvor gammelt var barnet da det kom til Norge? _____

Har barnet gått på skole eller fått andre pedagogiske tilbud i et annet land enn Norge?

- Nei
 Ja, i så fall hvor mange år og måneder: _____



Hvis ja, i hvilke land: _____

Hvilke språk snakker barnet?

- Morsmål Ja Nei
Norsk Ja Nei
Flere språk Ja Nei

Deltakelse i pedagogisk tilbud før skolestart. Sett kryss:

	I Norge	Antall måneder
Barnehage fulltid		
Halv tid		
Mindre enn halv tid		

Bakgrunnsinformasjon om foreldrene

Botid og skolegang. Skriv antall år:

	Mor	Far
Botid i Norge (antall år)		
Skolegang i hjemland		
Skolegang i Norge		
Skolegang i andre land		

Arbeid

Har mor arbeid i Norge?

Nei

Ja



Hvis ja, hva? _____

Har far arbeid i Norge?

Nei

Ja



Hvis ja, hva? _____

Forespørsel om å delta i forskningsprosjektet « Flerspråklige barns språkferdigheter»

Prosjektet skal studere variasjon i flerspråklige barns språkferdigheter. Målet er å bidra til økt forståelse av språkutvikling i en flerspråklig situasjon.

Ditt barn er henvist til pedagogisk-psykologisk tjeneste i XXXXX kommune og det er ønskelig å benytte testresultater og informasjon om ditt barns språkutvikling i studien.

Det er bare ansatte på PPT og jeg som vil få tilgang til personidentifiserbare opplysninger. Resultatene fra studien vil bli publisert som gruppedata, uten at enkelte kan gjenkjennes. Prosjektet forventes å være ferdig i begynnelsen av 2017. Etter prosjektet er avsluttet vil opplysningene bli anonymisert. Dersom du tillater ditt barn å delta, er det fint om du signerer samtykkeerklæringen og returnerer til din kontaktperson i pedagogisk-psykologisk tjeneste i XXXX.

Din identitet er ukjent for meg helt til du eventuelt samtykker i å delta i studien ved å returnere samtykkeerklæringen. Det er frivillig å delta i prosjektet og du kan når som helst trekke barnet ditt fra undersøkelsen, uten å begrunne dette nærmere. Det vil ikke få noen innvirkning på forholdet til PPT om barnet blir trukket fra undersøkelsen.

Hvis du har spørsmål i forbindelse med denne henvendelsen, kan du gjerne ta kontakt med meg!

Prosjektet er meldt til Personvernombudet for forskning, Norsk samfunnsvitenskapelig datatjeneste (NSD).

Vennlig hilsen

May-Britt Monsrud

Høgskolen i Oslo og Akershus

P.b 4, St Olavs plass, 0130 Oslo eller mobil 41806397

Artikkel III.

First and Second Language Sentence Repetition - A Potential Screening Measure in Dual Language Learners

Monsrud, Rydland, Geva & Lyster (resubmitted i Language & Education juni 2021 og akseptert for publisering i august 2021).

First and Second Language Sentence Repetition: A Screening Measure for Dual Language Learners?

May-Britt Monsrud, Veslemøy Rydland, Esther Geva & Solveig- Alma Halaas Lyster

Abstract

The sentence repetition (SR) test is considered a promising diagnostic tool for detecting language proficiency in monolingual samples, but less is known about its potential to identify language proficiency of dual language learners (DLLs). Since challenges with language learning among DLLs, such as a developmental language disorder (DLD), need to be evident in both the first- and second languages (L1 and L2), it is important to use procedures that assess language functions in both languages wherever feasible.

We found that students' L1 and L2 SR performance levels draw upon their phonological short-term memory capacity, vocabulary, and grammar skills in the language being assessed. We also found a positive relationship between students' L1 phonological memory and L2 SR scores. This suggests that DLLs' phonological short-term memory capacity assessed in their L1 captures individual differences in short-term verbal memory that support performance in SR, as assessed in the L2.

The current results highlight a challenge in both research and clinical settings, which often use monolingual assessment tools and associated norms to identify DLLs who might be at risk of DLDs or other learning difficulties.

Key words: Assessment, Dual Language Learners, Developmental Language Disorder, Sentence Repetition

First and Second Language Sentence Repetition: A Screening Measure for Dual Language Learners?

Introduction

In many countries, an increasing number of children are learning more than one language, and the first and home language is often not the majority language. The large individual variations/heterogeneity of these children's language skills is widely documented (Hammer et al., 2014). A major challenge for assessment in educational and clinical settings is distinguishing the large variety of competencies in the first (L1) and the competencies in second language (L2) from developmental language disorder (DLD) (Bedore & Pena, 2008). Increasing our knowledge of dual language learners' (DLL's) L1 and L2 competencies and shedding light on how assessment procedures can capture language capacities among DLLs is a great need (Pratt et al., 2021).

It is widely acknowledged that L1 assessments are important in identifying potential DLD (American-Speech-Language-Hearing Association [ASHA] 2021). Few assessment tools are available in the language combinations spoken by the multitude of DLLs in Norway and other European countries. Also, in situations in which assessments do exist in both languages, the norms in the target language may not be reliable for DLLs who have typically experienced more multifaceted L1 and L2 exposures in their language environments compared to their monolingual peers.

This study provides an investigation of L1 and L2 language skills in a sample of 546 DLLs (age range 6–13), with no known special needs and a sample of 14 DLLs with DLD (N = 14). The DLLs represented in this study are children of immigrants. Their parents arrived in Norway as working immigrants in the 1970s (Turkish and Urdu/Panjabi) or arrived as refugees between 1970 and 2000 (Albanian, Somali, Tamil and Vietnamese).

Focusing on a widely used screening measure, the sentence repetition (SR) test, the present study seeks to investigate whether and how assessment in both the L1 and L2 adds information about the language proficiency and potential language difficulties of DLLs, beyond an assessment of L2 skills only.

Developmental Language Disorder (DLD)

Approximately 10% of all children start school with language difficulties that no known diagnoses can explain them (Paul, Norbury, & Gosse, 2018). The term specific language impairment (SLI) has been broadly used for the definition of language impairments in children whose cognitive skills are within normal limits and where there is no identifiable reason for the impairment (Reilly et al., 2014). Recently, a consensus has been reached suggesting that the somewhat broader term developmental language disability (DLD) better covers language difficulties with unknown cause (Gallinat & Spaulding, 2014; Paul et al., 2018). The identification of DLDs should incorporate multiple sources and the language learning difficulties and evaluate the difficulties impact on everyday life and affect functional academic attainment. However, no clear cutoff point distinguishes DLD from the lower end of normal variations in language abilities (Bishop et al., 2017; Paul et al., 2018).

Identifying DLD in DLLs is challenging because capturing vocabulary distributed between L1 and L2 is often possible and grammatical structures commonly seen in second language (L2) acquisition—and attributed to less exposure to the L2 societal language—overlap to some extent with the errors made by monolingual children with DLD (Bedore & Pena, 2008; Paradis, 2016; Tuller et al., 2018).

DLD among DLLs needs to be evident in both the L1 and L2 ; therefore, wherever feasible, it is important to use procedures that assess language functions in both the L1 and L2 (Paradis, 2016; Thordardottir & Brandeker, 2013). A study of Swedish-Arabic preschool

children with and without DLD showed that DLLs' without DLD demonstrated strong language development in at least one language (Hakansson et al., 2003). Since DLLs may be less exposed to both languages compared to their monolingual peers, it is difficult to differentiate between language variety caused by a lack of exposure and language challenges caused by DLD (Hammer et al., 2014). This is a major concern in the assessment procedures of possible DLD in DLLs (Geva & Farnia, 2017).

The systematic group differences between DLLs and monolingual learners typically seen in the L2 measures of vocabulary (e.g. August et al., 2005) and grammar (e.g. Hakansson et al., 2003) are less pronounced in assessments of phonological short-term memory, such as non-word repetition, the recall of digit span and the recall of words (word span) (Calvo & Bialystok, 2014; de Abreu et al., 2013; Peña et al., 2018). Non-word repetition tests are, however, not entirely considered independent of language experience, but may be influenced by phonological proximity to real words (Estes et al., 2007). Word span tests and forward digit span tests require children to recall lists of words or numbers; however, word span tests are considered to rely more on linguistic knowledge than digit span tests (Buac et al., 2016; Melby-Lervåg & Hulme, 2010). Altogether, phonological short-term memory tests are considered to be more distinct and less influenced by language skills compared to grammar, vocabulary and SR tests.

Using the Sentence Repetition Test (SR) to Detect Potential DLD in DLLs

SR tasks are considered relatively less influenced by prior knowledge compared to the tasks of grammar and vocabulary (Hesketh & Conti-Ramsden, 2013) and SR are therefore considered a clinical marker of DLD in DLLs and are also included in most diagnostic batteries, used to identify DLDs in students (Riches, 2012; Tuller et al., 2018; Zebib et al., 2019). Even though SR is perceived to be a promising diagnostic tool, the skills that underpin the performance of the tasks is less understood (Riches, 2012). According to Riches, two main

factors predict SR in Baddeley's multicomponent model of working memory: predictors that implicate memory and those that implicate language (Baddeley et al., 1998; Riches, 2012). In this model, SR taps the capacity of the episodic buffer. Research with monolingual samples has demonstrated that SR performance draws on a broad set of related skills, including vocabulary, grammar and phonological short-term memory (Frizelle et al., 2017; Klem et al., 2015). A study of monolingual samples provided an examination of the potential underlying mechanisms of SR, such as short-term memory, syntactic knowledge and working memory, showed that the included DLD group relied more on short-term memory compared to students without DLD (Riches, 2012). This was also the case of DLL (English/Spanish speaking) samples with and without DLD, when the contribution of short-term memory, vocabulary and language exposure in SR was investigated (Pratt et al., 2021). The results showed that SR relies on skills from all these domains in both groups, but DLLs with DLD rely most heavily on short-term memory compared to DLL without DLD.

Since SR is widely used in screening for potential DLD, we wanted to look more closely at what information SR gives in both L1 and L2. We investigate language components that predict L1 and L2 SR performance in DLLs. Even though SR differentiate between DLLs with and without a DLD, the results shows larger variability in DLL samples compared to results in monolingual samples (Chiat et al., 2013). More studies are needed, including investigations of non-English speaking populations with diverse L1 backgrounds.

Present Study

In the present cross-sectional study, we investigate to what extent phonological memory, grammar and vocabulary predict SR in L1 and L2 in a large sample of DLLs without DLD. We also investigate whether L1 and L2 SR identify DLLs at risk in this sample in amore explorative analyses and compare the results of L1 and L2 SR, word span, grammar and vocabulary of DLLs at risk to a sample of DLLs with DLD.

The following research questions (RQs) guided this study:

RQ1. To what extent are students' L1 and L2 sentence repetition scores predicted by their L1 and L2 word span, L1 and L2 vocabulary, and L1 and L2 grammar?

RQ2. How do students identified as being at risk of DLD based on their L1 and L2 sentence repetition perform relative to a sample of DLLs identified as having DLD?

Method and Design

Participants

The first sample of participants consisted of 546 students without DLD (girls = 294, boys = 252, and age range = 6.0–12.11). These students were all recruited for the study by their teachers, who were instructed to only include students with no special education needs and who did not receive special needs education. The teachers recruited students who used their L1 with parents, siblings and friends.

All the students came from multiethnic neighborhoods in the greater Oslo area. We recruited students whose first languages were Albanian (n = 62), Somali (n = 66), Tamil (n = 89), Turkish (n = 116), Urdu (n = 136) and Vietnamese (n = 77). A questionnaire completed by the students' parents provided information about each student's country of birth, home language background, language use at home and preschool attendance.¹ The participating schools distributed the questionnaires (in Norwegian and in the respective L1s), and 60% of the parents (n = 325) returned the completed questionnaire. The questionnaire data from the respondents showed that 87% of the students were born in Norway (n = 291) and that 13% (n = 42)² were born in other countries; 85% (n = 276) attended preschool. The questionnaire data showed that N = 324 parents reported that students use their L1 at home.

Table 1 displays the distribution of participants by gender and age group.

¹ Children in Norway generally begin nursery/preschool at the age of 1 and go to nursery/preschool until they start school at age 6, but preschool attendance is not compulsory.

² Albania = 4, Somalia = 24, Sri Lanka (Tamil) = 1, Turkey = 5, Pakistan (Urdu) = 5, and Vietnam = 3.

----- *Table 1 approximately here* -----

The second sample consisted of DLLs who had been identified by the special need services (i.e., school psychologists or speech therapists) as having DLD (n = 14; girls = 5, boys = 9, and age range = 8.0–12.1). The participants with DLD were recruited from the same neighborhoods as the DLLs without DLDs. They were all born in Norway, representing the following home language groups: Tamil = 5, Turkish = 4, Somali = 1, Urdu = 3, and Vietnamese = 1. The parents of the students in the DLD group reported that their students used their L1 to communicate with family members but used their L2 (i.e., Norwegian) to interact with siblings and friends. Students in this sample had attended preschool for one to three years. We were unable to recruit students who had been referred to special need services in the youngest age group. Table 2 displays information about the DLLs with DLD, by gender and age group.

----- *Table 2 approximately here* -----

Measures³

The following standardized tests are tailored to the Norwegian context and normed on monolingual Norwegian students. These tests were translated and adopted to the six home languages.

Translation and Adaptation of Measures

To identify the students' abilities in their L1 and L2, we translated and adapted the Norwegian tests into the first languages of the students in line with the principles for translation and adaption of tests (American Educational Research Association et al., 2014; Peña, 2007). For each language, two qualified translators with native competence in the language independently translated all the test items. The translators had high levels of L1 and L2 competence and relevant professional/academic backgrounds as linguists, interpreters,

³ Permission for translation for research purposes in this project was provided by the authors E. Ottem and J. Frost for sentence repetition, word span and expressive vocabulary by the GL Assessment for the British Picture Vocabulary Scale (BPVS-II) and by Pearson for TROG-2 Test for the reception of grammar.

teachers, or researchers and had expertise in the participants' bilingual and bicultural background. The two translators and a Norwegian project group consisting of professionals and research experts in the field of language development discussed the translations of each item. Once the draft list of translated items was complete, six independent translators who were not involved in the previous phase (with the same relevant qualifications) back-translated the home-language versions to Norwegian to determine whether the original version of the test was retained.

It was important to evaluate whether the words, utterances, grammatical syntax, and complexity were appropriate for the students in their specific bicultural and bilingual contexts (American Educational Research Association et al., 2014).

Given that translation is not the best practice for language assessments, we also calculated the percentage of participants who correctly responded to each item on the tests. We noted a steady decline in the percentage of correctly identified items as the tests progressed and we observed this trend in both L1 and L2 assessment procedures.

Sentence Repetition. The way SR tasks are constructed vary as to sentence length, syntactic and grammatical complexity, the number of items and the scoring process, the correct repetition of the target constructions or the number of target words repeated correctly. This SR test consists of 16 sentences of increasing length, syllables syntactic , grammatical complexity and information units (Ottem & Frost, 2005).

The test administrator read the sentences to the student one at a time, and the student was asked to repeat the sentence verbatim. Correct repetition of each test item had a score of 1

(correct) or 0 (incorrect); the administrator terminated the test after three consecutive incorrect answers.⁴

Word Span (Serial Word Recall). The word span test consists of a series of three, four, or five monosyllabic words. There are 4 series at each length, for a total of 12 test items (Ottem & Frost, 2005). The words in each item are read to the student, and the student is asked to repeat them verbatim and in the right order. Like the SR test, each item is scored as 1 or 0, with testing terminated after three consecutive incorrect answers.⁵

Grammar. The Test for Reception of Grammar (TROG) assesses the students' grammatical comprehension (Bishop, 1989; Lyster & Horn, 2009). The TROG-2 is a picture test in which the student is asked to identify one out of four pictures that is consistent with the sentence read. The test consists of 80 items, divided into 20 blocks of four sentences with the similar structures. Because the syntactic structures are not equally difficult in each language, the complete test (80 items) was administered to children in their L1 and L2.

Expressive Vocabulary. The test of expressive vocabulary consists of antonyms (14 items) and word definitions (12 items) (Ottem & Frost, 2005). In the antonym subtests, the researcher asks the child to name the opposite of a given word (e.g., "What is the opposite of *cold*?"). In the definition section, the researcher asks the student to define words (e.g., "What is a *bed*?"). Each item receives a score of correct or incorrect (1 or 0). In each section, testing is terminated after three consecutive incorrect responses.

Receptive Vocabulary. The Norwegian version of the British Picture Vocabulary Scale (BPVS-II) contains 144 items (12 blocks) of increasing difficulty (Dunn et al., 1997; Lyster et al., 2010). The student is asked to identify one out of four drawings that corresponds to a word spoken by the test administrator. In this study, each student started with the first item given in

⁴ Following is an example of a sentence repetition task: "The girl kicked the soccer ball over the roof" ("Jenta sparket ballen over hustaket").

⁵ Following is an example of the word span test: "green, can, cut, wind" ("grønn, kan, kutt, vind").

both languages because of the large variation that exists in vocabulary knowledge in general in DLLs. We used the termination criterion of the BPVS-II manual: the test ends when eight or more of the 12 items in a block are answered incorrectly. Each correct answer receives 1 point.

Procedure

All the children were individually assessed in their L1 and L2 at their respective schools. The assessors had professional backgrounds as speech and language therapists, school psychologists, or trained graduate students. The L1 was assessed by assessors from the relevant L1 backgrounds, and the L2 was evaluated by native Norwegian assessors. The order of assessments in the L1 and the L2 was counterbalanced; half the bilingual children in each age group were assessed in their L1 first and were tested in Norwegian (their L2) later. The other half of the bilingual sample was assessed in Norwegian first and later in their L1.

All the measures were scored on the spot. There was some attrition due to absence from school (missing completely at random) and due to delay of test material in L1.

Analysis

To answer RQ1, we examined the extent to which phonological memory, vocabulary and grammar predicted L1 and L2 SR scores in the larger sample of students (N=546). We investigated this in two separate models because the preliminary analyses revealed that the children's L1 and L2 SR scores were not significantly related once the predictor variables were included in the analyses. We used the Mplus program (Muthén & Muthén, 1998–2017) to perform these analyses to achieve two goals: First, as students were nested within schools, the Complex option in the Mplus program allowed us to account for the cluster effect on estimated standard errors. Second, because of the relatively large variability in students' ages, we included a covariate (age in months) for all the variables to increase the precision of the estimates. The predictor variables were allowed to co-vary in both predictive models.

To predict the students' scores on L2 SR, we first regressed L2 SR on L2 vocabulary (expressive), L2 grammar and L2 word span. As a second analytic step, we entered the

equivalent L1 measures as predictors. Only the significant predictor variables were included in the figure presented here. Note that, due to the high correlation between expressive and receptive vocabulary, we only included expressive vocabulary because it also represents receptive vocabulary when it comes to building a model that predicts the results of SR (appendix 1). We employed the same procedures to predict students' L1 SR. L1 SR was first regressed on L1 vocabulary, L1 grammar and L1 word span before we entered the L2 predictors.

We evaluated the model fit against the following guidelines. The Root Mean Squared Error of Approximation (RMSEA) estimate should be lower than .06. The Comparative Fit Index (CFI) and the Tucker Lewis Index (TLI) should be higher than 0.90. Both the models presented below had good model fit.

To answer RQ2, we needed to identify students who might be at risk of DLD in the sample without DLD. We identified students who might be at risk of DLD by identifying those in this sample who achieved low SR scores (i.e., -1.25 SD or -1.00 SD) in both their L1 and L2 (in the 6.0–12.11 age range). This group was defined as an *at-risk* group.

To investigate the percentage of students in a potential at-risk group when comparing DLLs to monolingual norms, we investigated the percentage of students who had scores at or below -1.25 SD in L2 SR, compared to monolingual norms ($n = 900$; Appendix 1) in the selected age groups (6.0–12.11).

Results

The descriptive statistics of the typically achieving sample, in terms of the language task battery in the L1 and L2 (by age group), are summarized in Table 3. Cronbach's alphas are included.

-----Table 3 approximately here-----

Table 3 shows the expected age-related increases in raw scores for all the variables. At the group level, the students' L1 test scores for SR, word span and vocabulary (receptive and expressive) were stronger compared to their scores on the equivalent L2 assessments. However, the difference between the students' L1 and L2 scores on the grammar assessment was minor.

Predicting L1 and L2 Sentence Repetition

The prediction of L2 SR is displayed in Figure 1. As can be seen, L2 expressive vocabulary had a relatively strong standardized partial regression on L2 SR ($\beta = 0.30, t = 5.57, p < 0.001$). L2 SR was also predicted by L2 grammar ($\beta = 0.20, t = 4.33, p < 0.001$), L2 word span ($\beta = 0.25, t = 5.69, p < 0.001$) and L1 word span ($\beta = 0.13, t = 3.33, p < 0.001$). It is noteworthy that L1 word span explained significant variance in L2 SR beyond the same-language predictors. Moreover, as seen in Figure 1, relatively strong intercorrelations were observed between the predictor variables, particularly between L2 expressive vocabulary and L2 grammar ($r = 0.49, t = 14.20, p < 0.001$) and between L1 and L2 word span ($r = 0.42, t = 13.06, p < 0.001$).

As expected, the predictor variables were related to the students' age in months at the time of the assessments. This age effect was particularly strong for the L2 expressive vocabulary ($\beta = 0.66, t = 24.77, p < 0.001$) and L2 grammar ($\beta = 0.59, t = 21.88, p < 0.001$), but it was also evident in the students' scores on L2 word span ($\beta = 0.37, t = 11.99, p < 0.001$) and L1 word span ($\beta = 0.32, t = 5.76, p < 0.001$). Students' age in months did not predict additional variance in their L2 SR scores once the age effects on the predictor variables were accounted for ($\beta = 0.00, t = 0.03, p = 0.978$). This model explained 48% of the variance in L2 SR.

-----*Figure 1 approximately here*-----

Figure 2 presents a prediction of L1 SR. As seen in the figure, L1 expressive vocabulary had a strong standardized partial regression on L1 SR ($\beta = 0.52, t = 11.01, p < 0.001$), followed by L1 word span ($\beta = 0.23, t = 4.39, p < 0.001$) and grammar ($\beta = 0.14, t = 2.32, p < 0.001$). L2 word span did not have a standardized partial regression on L1 SR ($\beta = -0.02, t = -0.35, p = 0.725$). Similar to the L2 model, we found strong intercorrelations between L1 expressive vocabulary and L1 grammar ($r = 0.45, t = 6.38, p < 0.001$) and between L1 word span and L2 word span ($r = 0.42, t = 13.76, p < 0.001$). It is worth noting that L1 expressive vocabulary and grammar also significantly correlated with students' L2 word span scores (respectively, $r = 0.20, t = 4.43, p < 0.001$, and $r = 0.23, t = 3.75, p < 0.001$).

As one might expect, all the predictor variables were related to students' age in months: L1 expressive vocabulary ($\beta = 0.47, t = 9.16, p < 0.001$), L1 grammar ($\beta = 0.44, t = 9.72, p < 0.001$), L1 word span ($\beta = .30, t = 5.22, p < 0.001$) and L2 word span ($\beta = 0.36, t = 10.88, p < 0.001$). Similar to what we found in the L2 model, students' age in months did not predict additional variance in students' L1 SR score once the predictor variables were included ($\beta = -0.06, t = -1.16, p = 0.245$). This model explained 49% of the variance in L1 SR.

----- *Figure 2 approximately here* -----

Comparing the Low-Performing At-Risk Group to the DLD Group

When identifying at-risk students within the DLL sample using a cutoff of -1.25 SD in both L1 and L2, a few students (1%; $n = 5$) scored that low. However, some of the students scored -1.25 on L1 SR ($n = 56$), and others ($n = 58$) scored that low on L2 SR.⁶ Because we were unable to detect a sizable low-performing group using the -1.25 SD cutoff, we applied a

⁶ Students scoring low on L1: Albanian = 5, Somali = 4, Tamil = 9, Turkish = 0, Urdu = 5, and Vietnamese = 22. Students scoring low on L2: Albanian = 2, Somali = 10, Tamil = 2, Turkish = 6, Urdu = 15, and Vietnamese = 6. Students scoring low on both L1 and L2 sentence repetition: Somali = 1, Tamil = 2, and Vietnamese = 2.

-1.00 SD cutoff in both the L1 and L2. Twenty-six participants had a score that was -1.00 SD or lower below the mean on the L1 and L2 SR tests.

Table 4 presents descriptive statistics of the at-risk group (n = 26) and the identified DLD group (n = 14) on L1 and L2 SR, word span, grammar, expressive vocabulary and receptive vocabulary. We excluded the youngest age group among the at-risk students because no students in this age group were in the identified DLD group.

-----*Table 4 approximately here*-----

The comparison between the at-risk group and the identified DLD group showed small differences between the groups. However, the L2 expressive and receptive vocabulary was stronger in the DLD group compared to the at-risk group.

Identifying Students Who Might Be at Risk Using Monolingual Norms

Next, we used monolingual norms for SR assessment to uncover possible at-risk students in the DLL group (n = 900; Appendix 1). Only the students with no previously known special needs (n = 546) who scored at least 1.25 SD below the mean on L2 SR were considered. This resulted in a considerable overrepresentation of DLLs who might be considered at risk of DLD. The percentage of students scoring below the cutoff point were 45%, 58% and 40% in the DLL group, compared to 9%, 12% and 9% in the same monolingual age groups, respectively.

Discussion

This study offered an investigation of the language components that predict L1 and L2 SR performance in DLLs. We also identified a group of DLLs at risk by low scores on L1 and L2 SR and compared this group to DLLs with DLD.

First, we found that students' L1 and L2 SR performance levels draw upon their language (vocabulary and grammar) and memory capacities (phonological short-term memory capacity) in the language being assessed. The SR performance assessed in the L2 predicted by students' L2 vocabulary, grammar and word-span scores and not by their L1 vocabulary and

grammar skills. However, we detected a positive relation between the students' L1 word span scores and their L2 SR scores, suggesting that DLLs' phonological short-term memory capacity assessed in their home language captures individual differences in short-term verbal memory, which relates to the students' performance on the SR when assessed in the L2. The reverse was not found, however, as a similar relation between L2 word span and L1 SR was not significant. Although the relation between L1 word span and L2 SR should be interpreted with caution, these findings may suggest that a specific cross-language effect may operate from the language that children are exposed to from birth to the societal language that children typically encounter later in life. Furthermore, the fact that we detected cross-language correlations between students' word span scores in one language and vocabulary and grammar scores in the other language also indicates that word span is an important marker of language aptitude among DLLs. The present study builds on a concurrent correlational design with limited ability to address questions of transfer between students' L1 and L2 over time. Nonetheless, the relationship between students' L1 word span and L2 performance suggests/indicates that cognitive processing impacts language skills. Future longitudinal studies should investigate the relationship between dual language learners' word span and SR scores over time.

Although the patterns of relationships were relatively similar in our predictive models of L1 and L2 SR, it is noteworthy that vocabulary explained most of the variance on L1 SR. The large standard deviations in the students' L1 vocabulary scores across age groups indicate that some of the students in the sample did not have a strong command of the L1. It is thus interesting to note that vocabulary appears to be a strong predictor of L1 SR. These results suggest that vocabulary explains L1 SR more than grammar.

Secondly, a surprising finding when we used 1.00 SD on L1 and L2 SR to identify an at-risk group for the DLL without DLD was the fact that minor differences between the DLD

and the at-risk group were detected. However, the low scores on SR and word span indicate that undetected language difficulties exist in the DLLs without DLD. This finding suggests that SR in both L1 and L2 is a promising clinical marker for identifying language difficulties, and the test may be a good steppingstone for further decision makers to uncover language difficulties or other learning disabilities.

The results of comparing the L1 and L2 SR of DLLs without DLD, using a -1.25 SD cutoff point criteria suggested that some students are L1 dominant, while others are L2 dominant. In other words, the L1 and L2 language proficiency levels in the sample had large variations. This finding may also reflect the fact that the DLLs receive language exposure in two different languages, and L1 and L2 SR detect the variability in their relative language proficiency and not DLD.

The current results also underscore the systematic bias related to DLLs as having potential language difficulties generated by monolingual norms. These results underscore a challenge in both in research and in clinical settings, often using monolingual assessment tools and associated norms to investigate DLL's language skills to identify DLLs who might be at risk for DLD or other learning difficulties.

Implications for Assessment and Practice

Many dual language learners need increased language intervention or support to reach their academic potential in schools. This study provided an examination of a group of students not identified and not referred to as special need services by their teachers even though some test scores are similar to those of the identified DLD group. The distinction between low-performing L2 skills and DLDs is, to a large extent, arbitrarily understood by special need services and teachers. Although poor L2 skills among DLLs is not necessarily indicative of a DLD, the low performances should be identified to gain access to systematic and qualified language assessment that may provide a platform for developing interventions.

Given that SR uncovers a broad range of language competencies in both L1 and L2, our study indicates that SR may also identify language difficulties. To avoid over-identification of DLLs at risk for DLD, it is important to compare students' performance in relation to their linguistic reference group and not to monolingual norms. Moreover, both L1 and L2 assessments are needed to uncover potential DLD.

What is new in this study is that SR detected variations in language competencies in L1 and L2 to a large extent. Low scores on L1 and L2 SR may identify language difficulties, but an assessment in L2 only using monolingual norms is a source of error. The result in this study also suggests that vocabulary – in addition to memory and grammar – is important to include in the analysis of what SR measures.

Limitations

Because of limited background information, we could not for sure determine whether the students who scored below the cutoff on both the L1 and L2 SR tests were at risk for DLD or if their scores reflected insufficient exposure to L1 and/or L2 language. We also should admit that the DLD group was small, which might have resulted in a biased result. However, bearing in mind that this group was identified by language experts and school psychologists, we should trust the assessment and believe this group to be a valid group to use for comparison.

References

- American Educational Research Association, American Psychological Association & National Council on Measurement in Education. (2014). *Standards for educational and psychological testing*. American Educational Research Association.
- August, D., Carlo, M., Dressler, C. & Snow, C. (2005). The critical role of vocabulary development for English language learners. *Learning Disabilities Research & Practice, 20*(1), 50-57.
- Baddeley, A., Gathercole, S. & Papagno, C. (1998). The phonological loop as a language learning device. *Psychological review, 105*(1), 158.

- Bedore, L. M. & Pena, E. D. (2008). Assessment of bilingual children for identification of language impairment: Current findings and implications for practice. *International Journal of Bilingual Education and Bilingualism*, 11(1), 1-29.
- Bishop, D. (1989). Test for reception of grammar. London: Medical Research Council.
- Bishop, D., Snowling, M. J., Thompson, P. A., Greenhalgh, T., Consortium, C., Adams, C., Archibald, L., Baird, G., Bauer, A. & Bellair, J. (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), 1068-1080.
- Buac, M., Gross, M. & Kaushanskaya, M. (2016). Predictors of processing-based task performance in bilingual and monolingual children. *Journal of Communication Disorders*, 62, 12-29.
- Calvo, A. & Bialystok, E. (2014). Independent effects of bilingualism and socioeconomic status on language ability and executive functioning. *Cognition*, 130(3), 278-288.
- Chiat, S., Sarmon-Lotem, S., Marinis, Theo, Polinska, K., Roy, P. & Seeff-Gabriel, B. (2013). The Potential of Sentence Imitation Tasks for Assessment of Language.
- de Abreu, P. M. E., Baldassi, M., Puglisi, M. L. & Befi-Lopes, D. M. (2013). Cross-linguistic and cross-cultural effects on verbal working memory and vocabulary: Testing language-minority children with an immigrant background. *Journal of Speech, Language, and Hearing Research*, 56(2), 630-642.
- Dunn, L. M., Dunn, L. M., Whetton, C. & Burley, J. (1997). *The British Picture Vocabulary Scale* (2nd utg.). nferNelson.
- Estes, K. G., Evans, J. L. & Else-Quest, N. M. (2007). Differences in the nonword repetition performance of children with and without specific language impairment: A meta-analysis. *Journal of Speech, Language, and Hearing Research*.
- Frizelle, P., O' & Bishop, D. V. (2017). Assessing understanding of relative clauses: A comparison of multiple-choice comprehension versus sentence repetition. *Journal of Child Language*, 44(6), 1435-1457.
- Gallinat, E. & Spaulding, T. J. (2014). Differences in the performance of children with specific language impairment and their typically developing peers on nonverbal cognitive tests: A meta-analysis. *Journal of Speech, Language, and Hearing Research*, 57(4), 1363-1382.
- Geva, E. & Farnia, F. (2017). Issue Editor Foreword. *Topics in Language Disorders*, 37(2), 109-113.
- Hakansson, G., Salameh, E.-K. & Nettelblatt, U. (2003). Measuring language development in bilingual children: Swedish-Arabic children with and without language impairment. *Linguistics*, 41(2; ISSU 384), 255-288.
- Hammer, C. S., Hoff, E., Uchikoshi, Y., Gillanders, C., Castro, D. C. & Sandilos, L. E. (2014). The language and literacy development of young dual language learners: A critical review. *Early Childhood Research Quarterly*, 29(4), 715-733. <https://doi.org/10.1016/j.ecresq.2014.05.008>
- Hesketh, A. & Conti-Ramsden, G. (2013). Memory and language in middle childhood in individuals with a history of specific language impairment. *PloS one*, 8(2), e56314.
- Klem, M., Melby-Lervåg, M., Hagtvet, B., Lyster, S. A. H., Gustafsson, J. E. & Hulme, C. (2015). Sentence repetition is a measure of children's language skills rather than working memory limitations. *Developmental science*, 18(1), 146-154.
- Lyster, S. & Horn, E. (2009). Test for Reception og Grammar (TROG-2). Norsk Versjon. I. Stockholm: Pearson Assessment.
- Lyster, S., Horn, E. & Rygvold, A. (2010). Ordforråd og ordforrådsutvikling hos norske barn og unge. Resultater fra en utprøving av British picture vocabulary scale, (BPVS II). *Spesialpedagogikk*, 9, 35-43.
- Melby-Lervåg, M. & Hulme, C. (2010). Serial and free recall in children can be improved by training: Evidence for the importance of phonological and semantic representations in immediate memory tasks. *Psychological science*, 21(11), 1694-1700.
- Ottem, E. & Frost, J. (2005). Språk 6-16; screening test: manual. *Bredtvet kompetansesenter*.

- Paradis, J. (2016). The development of English as a second language with and without specific language impairment: Clinical implications. *Journal of Speech, Language, and Hearing Research, 59*(1), 171-182.
- Paul, R., Norbury, C. & Gosse, C. (2018). Language disorders from infancy through adolescence: Listening, Speaking. *Reading, Writing, and Communicating, 4*, 756.
- Peña, E. D. (2007). Lost in Translation: Methodological Considerations in Cross-Cultural Research. *Child development, 78*(4), 1255-1264.
- Peña, E. D., Gutiérrez-Clellen, V. F., Iglesias, A., Goldstein, B. A. & Bedore, L. M. (2018). Bilingual English Spanish Assessment (BESA). *Baltimore, MD: Brookes*.
- Pratt, A. S., Peña, E. D. & Bedore, L. M. (2021). Sentence repetition with bilinguals with and without DLD: differential effects of memory, vocabulary, and exposure. *Bilingualism: Language and Cognition, 24*(2), 305-318.
- Riches, N. G. (2012). Sentence repetition in children with specific language impairment: An investigation of underlying mechanisms. *International Journal of Language & Communication Disorders, 47*(5), 499-510.
- Thordardottir, E. & Brandeker, M. (2013). The effect of bilingual exposure versus language impairment on nonword repetition and sentence imitation scores. *Journal of Communication Disorders, 46*(1), 1-16.
- Tuller, L., Hamann, C., Chilla, S., Ferré, S., Morin, E., Prevost, P., Dos Santos, C., Abed Ibrahim, L. & Zebib, R. (2018). Identifying language impairment in bilingual children in France and in Germany. *International Journal of Language & Communication Disorders, 53*(4), 888-904.
- Zebib, R., Tuller, L., Hamann, C., Abed Ibrahim, L. & Prévost, P. (2019). Syntactic complexity and verbal working memory in bilingual children with and without Developmental Language Disorder. *First Language, 0142723719888372*.

Table 1 Typical developing DLL's by age group and gender

Age	Girls	Boys	TOTAL
6-0-6.11	28	20	48
7.0-7.11	42	44	86
8.0-8.11	47	37	84
9.0-9.11	40	3	93
10.0-10.11	66	45	111
11.0-11.11	37	23	60
12.0-12.11	34	30	64
TOTAL	294	252	546

Table 2 DLL with DLD (N=14) by age group and gender

Age	Girls	Boys	TOTAL
9.0-10.11	3	3	6
11.0-12.11	2	6	8
TOTAL	5	9	14

Table 3: Means, Standard Deviations, and Cronbach's Alpha Values for the Variables of L1 and L2 Test Scores of Typical Developing DLLs (N = 546)

Age	Sentence Repetition		Word Span		Grammar		Expressive Vocabulary		Receptive Vocabulary	
	L1	L2	L1	L2	L1	L2	L1	L2	L1	L2
	α .87	α .80	α .79	α .73	α .91	α .91	α .90	α .91	α .99	α .93
6.0-6.11	4.42 (SD=26)	3.23 (SD=1.7)	7.21 (SD=2.4)	4.05 (SD=2.0)	52.09 (SD=15.6)	52.81 (SD=11.2)	6.23 (SD=3.5)	4.27 (SD=1.8)	61.33 (SD=28.6)	48.04 (SD=16.9)
	n=24	n=48	n=14	n=39	n=22	n=47	n=13	n=40	n=24	n=48
7.0-7.11	4.55 (SD=2.9)	4.02 (SD=1.9)	6.67 (SD=2.4)	5.03 (SD=2.1)	57.97 (SD=11.2)	58.40 (SD=10.7)	6.70 (SD=3.6)	5.57 (SD=2.3)	69.00 (SD=29.8)	59.49 (SD=18.4)
	n=74	n=86	n=66	n=78	n=72	n=86	n=67	n=79	n=69	n=86
8.0-8.11	5.13 (SD=2.9)	4.70 (SD=2.2)	7.25 (SD=2.5)	5.61 (SD=2.0)	62.36 (SD=10.1)	62.74 (SD=9.7)	8.08 (SD=3.9)	7.49 (SD=3.7)	76.03 (SD=30.7)	69.71 (SD=17.0)
	n=89	n=84	n=80	n=79	n=85	n=84	n=79	n=80	n=86	n=84
9.0-9.11	5.61 (SD=3.2)	4.56 (SD=1.8)	7.86 (SD=2.5)	5.67 (SD=2.1)	64.44 (SD=9.9)	67.01 (SD=7.0)	9.47 (SD=5.4)	8.45 (SD=3.6)	83.76 (SD=28.8)	74.60 (SD=17.6)
	n=97	n=93	n=91	n=88	n=93	n=92	n=91	n=88	n=97	n=93
10.0-10.11	6.86 (SD=3.8)	5.90 (SD=2.3)	8.01 (SD=2.2)	6.43 (SD=1.8)	67.49 (SD=9.6)	69.95 (SD=6.8)	10.88 (SD=4.8)	10.70 (SD=3.5)	92.79 (SD=27.7)	85.02 (SD=18.1)
	n=65	n=111	n=90	n=106	n=100	n=110	n=89	n=107	n=100	n=110
11.0-11.11	7.76 (SD=3.3)	6.42 (SD=2.8)	8.86 (SD=2.2)	6.47 (SD=1.9)	72.222 (SD=5.6)	72.53 (SD=5.0)	12.56 (SD=4.87)	12.85 (SD=4.5)	109.98 (SD=16.6)	96.12 (SD=17.2)
	n=68	n=59	n=64	n=59	n=64	n=60	n=64	n=59	n=63	n=60
12.0-12.11	5.13 (SD=2.9)	6.86 (SD=2.8)	8.93 (SD=2.1)	6.76 (SD=1.9)	70.96 (SD=8.5)	72.14 (SD=6.3)	13.75 (SD=5.1)	13.52 (SD=3.9)	103.33 (SD=22.0)	97.75 (SD=16.1)
	n=89	n=64	n=70	n=62	n=69	n=64	n=69	n=62	n=70	n=64

Table 4: Means, Standard Deviations, and *N* for Variables of L1 and L2 Test Scores of DLLs without (*N* = 26) and DLL with DLD (*N* = 14)

At-risk	Age	Sentence Repetition		Word Span		Grammar		Expressive Vocabulary		Receptive Vocabulary	
		L1	L2	L1	L2	L1	L2	L1	L2	L1	L2
	9.0-10.11 mnd 110.50 (<i>SD</i> = 2.83)	2.21 (<i>SD</i> = 2.1) <i>N</i> = 14	2.57 (<i>SD</i> = 1.2) <i>N</i> = 14	5.71 (<i>SD</i> = 1.8) <i>N</i> = 14	4.14 (<i>SD</i> = 2.3) <i>N</i> = 14	58.46 (<i>SD</i> = 10.8) <i>N</i> = 14	57.64 (<i>SD</i> = 12.5) <i>N</i> = 14	4.86 (<i>SD</i> = 2.9) <i>N</i> = 14	5.79 (<i>SD</i> = 3.7) <i>N</i> = 14	76.00 (<i>SD</i> = 29.3) <i>N</i> = 14	65.57 (<i>SD</i> = 19.2) <i>N</i> = 14
	11.0-12.11 mnd 133.50 (<i>SD</i> = 1.00)	2.36 (<i>SD</i> = 1.8) <i>N</i> = 12	3.64 (<i>SD</i> = 1.5) <i>N</i> = 12	6.36 (<i>SD</i> = 2.2) <i>N</i> = 12	4.73 (<i>SD</i> = 1.7) <i>N</i> = 12	65.08 (<i>SD</i> = 9.2) <i>N</i> = 12	68.36 (<i>SD</i> = 4.2) <i>N</i> = 12	7.0 (<i>SD</i> = 3.0) <i>N</i> = 12	7.83 (<i>SD</i> = 6.0) <i>N</i> = 12	82.3 (<i>SD</i> = 21.3) <i>N</i> = 12	80.5 (<i>SD</i> = 15.3) <i>N</i> = 12
	9.0-10.11 mnd 116.33 (<i>SD</i> = 9.56)	3.33 (<i>SD</i> = 2.9) <i>N</i> = 6	3.67 (<i>SD</i> = 0.8) <i>N</i> = 6	5.83 (<i>SD</i> = 2.2) <i>N</i> = 6	5.33 (<i>SD</i> = 2.5) <i>N</i> = 6	60.17 (<i>SD</i> = 10.5) <i>N</i> = 6	59.00 (<i>SD</i> = 14.2) <i>N</i> = 6	5.33 (<i>SD</i> = 4.2) <i>N</i> = 6	4.83 (<i>SD</i> = 2.2) <i>N</i> = 6	66.50 (<i>SD</i> = 27.0) <i>N</i> = 6	66.33 (<i>SD</i> = 12.8) <i>N</i> = 6
	11.0-12.11 mnd 145.50 (<i>SD</i> = 4.75)	2.71 (<i>SD</i> = 2.1) <i>N</i> = 8	4.14 (<i>SD</i> = 1.1) <i>N</i> = 8	6.00 (<i>SD</i> = 2.5) <i>N</i> = 8	4.57 (<i>SD</i> = 1.7) <i>N</i> = 8	62.88 (<i>SD</i> = 10.4) <i>N</i> = 8	67.13 (<i>SD</i> = 4.2) <i>N</i> = 8	6.25 (<i>SD</i> = 2.6) <i>N</i> = 8	4.88 (<i>SD</i> = 4.1) <i>N</i> = 8	74.0 (<i>SD</i> = 20.8) <i>N</i> = 8	73.38 (<i>SD</i> = 7.5) <i>N</i> = 8

Appendix 1: Monolingual Norms ($N = 900$) by Age Group Sentence Repetition, Cronbach Alpha, 1 Raw Scores and SD

Age Group	Sentence Repetition α	1 Raw Scores	SD
6.0-6.11	6.45 ($SD = 2.26$) $n = 432$		
9.0-10.11	8.28 ($SD = 2.52$) $n = 223$		
11.0-12.11	9.43 ($SD = 2.86$) $n = 245$		
TOTAL		$N = 900$	

Appendix 2: Correlations Between L1 and L2 Sentence Repetition, Word Span, Grammar, and Expressive and Receptive Vocabulary

	L1 Sentence Repetition	L2 Sentence Repetition	L1 Word Span	L2 Word Span	L1 Grammar	L2 Grammar	L1 Expressive Vocabulary	L2 Expressive Vocabulary	L1 Receptive Vocabulary
L1 Sentence repetition									
L2 Sentence Repetition	.279**								
L1 Word Span	.464**	.412**							
L2 Word Span	.326**	.544**	.494**						
L1 Grammar	.400**	.285**	.292**	.314**					
L2 Grammar	.239**	.552**	.281**	.446**	.335**				
L1 Expressive Vocabulary	.649**	.301**	.392**	.350**	.523**	.347**			
L2 Expressive Vocabulary	.328**	.603**	.368**	.472**	.415**	.618**	.468**		
L1 Receptive Vocabulary	.383**	.250**	.263**	.321**	.538**	.393**	.523**	.435**	

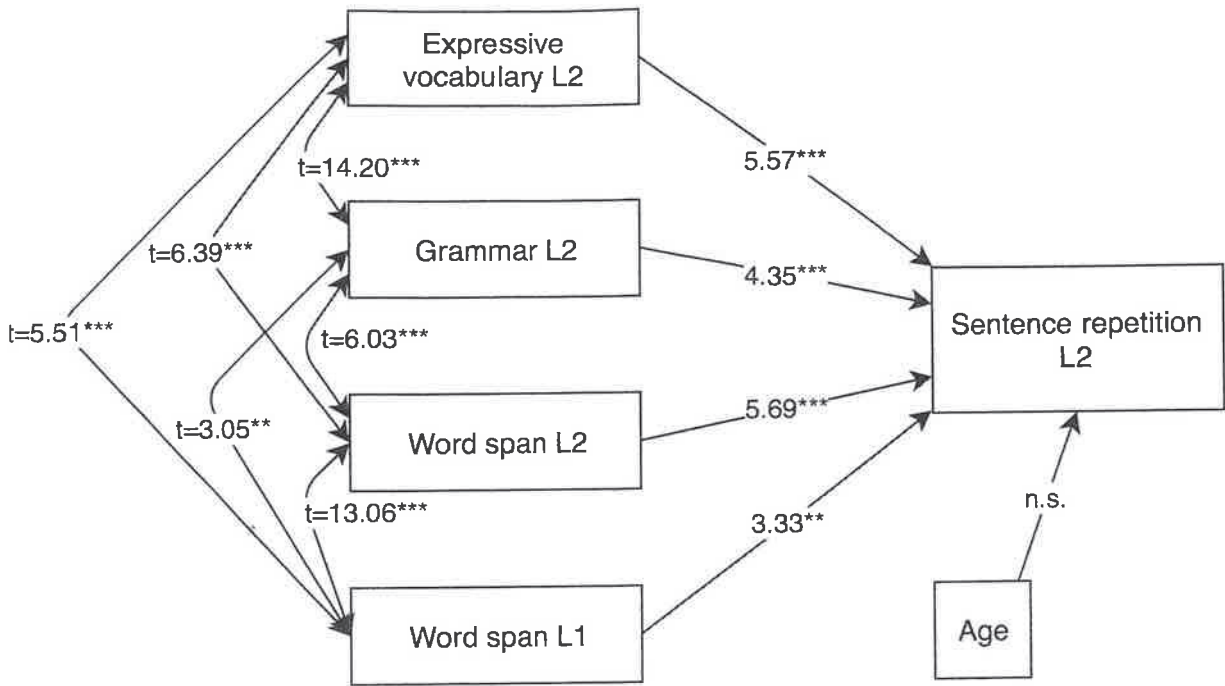


Figure 1 Model of L2 sentence repetition and linguistic (expressive vocabulary and grammar) and phonological short- term memory (word span)

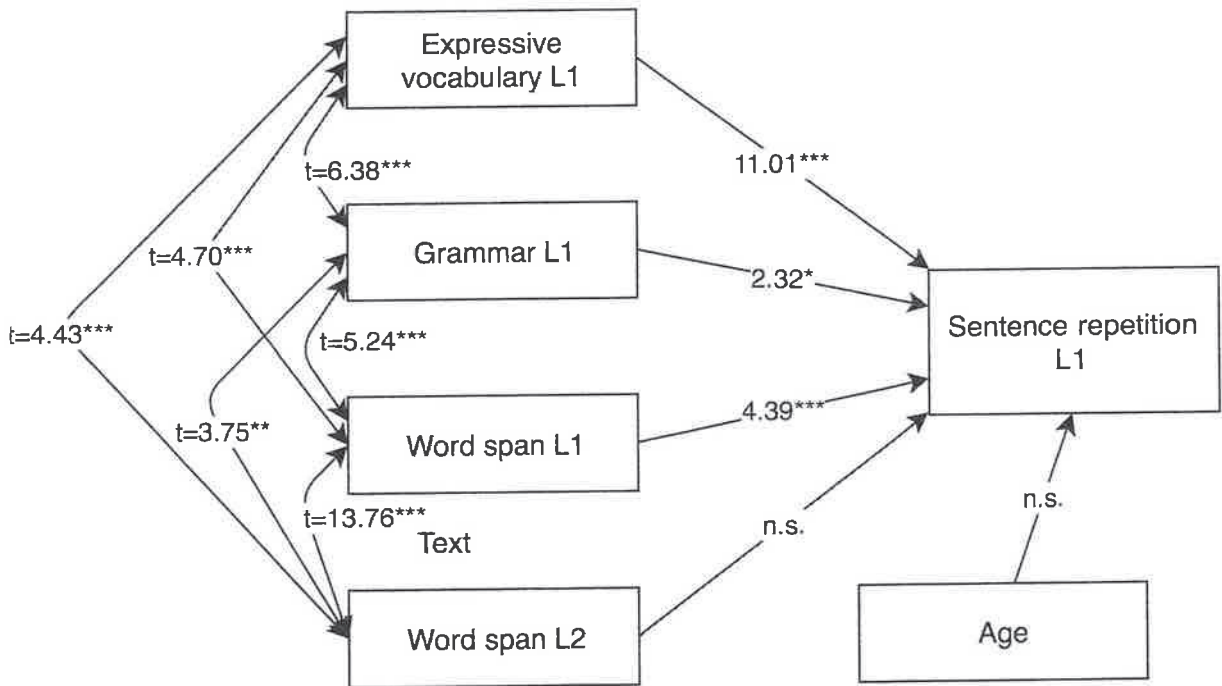


Figure 2 Model of L1 sentence repetition and linguistic (expressive vocabulary and grammar) and phonological short- term memory (word span)

Additional Information: Measures

Sentence Repetition, Word Span, and Expressive Vocabulary are subtests from an individual administered language screening test: *Language 6-16 (Språk 6-16)* (Ottem & Frost, 2005). The screening test is a non-profit test distributed by Statped, a National Service for Special Needs Education, Department of Speech and Language Disorders, in Norway.

The *Language 6-16* is frequently used by special needs services and schools by speech therapists, schools, psychologists, and special needs teachers to assess students with suspected DLD. The norming data is 1,214 monolingual students with no suspected special needs from all over Norway.

As stated in the article, qualified translators with native competencies in the L1 as well as the students specific bilingual and bicultural context translated all the items in the measures included.

Sentence repetition (Language 6-16)

The content, grammar, and syntax complexity of the Norwegian versions of the sentences were changed in the L1 versions to keep the difficulty and complexity of the sentences. The length of the sentences, number of syllables, syntactic and grammatical complexity, and information units in Norwegian were translated and adapted to L1 versions of the sentences. The structures tested in each language thus reflect their relative difficulty in that language, not the relative difficulty of the tested structures in Norwegian. These adaptations were needed to keep the relative difficulty in each L1 version. Features that were culturally biased in the Norwegian sentences were changed, and personal names were independently chosen for each language.¹

Item no	English translation	Norwegian wording
1	He ran out again	Han løp ut igjen
4	The girl kicked the football over the roof	Jenta sparket fotballen over hustaket
6	Pears in my garden are better than those in the shops	Pærene i hagen min er bedre enn de i butikken
9	She is so good to play guitar that she may play in a rock band	Hun er så flink til å spille gitar at hun kan bli med i et rockeband
13	The boy who skipped the training, was not allowed to play at the team before a week was passed	Gutten som skulka treningen, fikk ikke lov å spille på laget før det var gått en uke

We calculated the percentage of students who responded correctly to each item and noted a steady decrease in the percentage of correct responses in L1 and L2 as the complexity increased. One item in the L1 and L2 test was more difficult than the rest²; however, it did not lead to a situation in which students were forced to stop the test.

Word span (serial recall by words)

The words in this test were monosyllabic and high-frequency words in the age group included, similar to the Norwegian version of the test. We calculated the percentage of students who correctly identified the items of each series and noted a steady decline in the percentage of correctly identified series as the test progressed. The researchers observed this trend in both L1 and L2 assessment procedures.

Grammar (TROG-2, Test for Reception of Grammar) The complexity of the 20 structures (4 items in each block; total 80 items) increases as the test progresses. For this study, the researchers translated the sentences into the six (L1) home languages with corresponding syntactic structures found in all languages. It was not problematic to translate the different structures to the six home languages. However, in some of the languages the structure was more common than in others, and in some languages one structure was more difficult compared to other structures. Accordingly, the sample with and the sample without DLD were assessed by all 20 structures, and 80 items in L1 and L2.

Expressive vocabulary (Language 6-16)

The translated versions of this test contained words that were as close as possible to the original in terms of degree of difficulty. Translating vocabulary tests can pose challenges because words tend to differ in concreteness and frequency across languages. The expertise of the translators native competencies in the languages and also the expertise in the participants specific bilingual and bicultural background was vital. We calculated the percentage of students who responded correctly to each item of the antonym and definition tests. A reasonable decrease in the percentage of correct items for L1 and L2 vocabulary items, was observed. However, two items in the antonym subtest did not show a steady decrease.³ Nonetheless, this did not cause any student to stop the test. The definition section showed a steady decrease in the percentage of correct items.

Receptive vocabulary (British Picture Vocabulary Test -II)

Each student started at the first item of the test (item 1, block 1) given in both languages because of the large variation that exists in vocabulary knowledge in general in DLLs. We therefore calculated the percentage of students who passed each block. Researchers noted a steady decrease in the percentage of children who could progress to the end of the 12 items in the 12 blocks in each of the six languages assessed. The researchers observed this trend in both the L1 and L2 assessments. The findings indicate that it is unlikely for a student to master a subsequent block if he/she failed the previous one.

¹ E.g., "Ali" in Urdu, "Arani" in Tamil, and "Chau" in Vietnamese.

² Item 5

³ Item 3 was more difficult than items 4 and 5, and item 6 was more difficult than items 7 and 8 in the Norwegian version and for all L1 language groups.

Errataliste

Side	Original tekst	Oppdatert tekst	Korreksjon
Innholdsoversikten er oppdatert			
s. 1, linje 18	Utvikling av språkkompetanse flerspråklige barn	Utvikling av språkkompetanse blant flerspråklige barn	Satt inn ordet <i>blant</i>
s. 1, linje 1	1.3 bakgrunn for avhandlingen	1.1 Bakgrunn for avhandlingen	1.1 erstatter 1.3
s.3, linje 22	Appendix 3	Vedlegg C	<i>Vedlegg C</i> erstatter <i>appendix 3</i>
s. 6, linje 18	-- forskjeller mellom språkgruppene mellom testene er likheter	... forskjeller mellom testene og likheter	Fjernet <i>mellom språkgruppene</i> , og erstatter <i>er</i>
s. 8, linje 1	...avgjøre om e enkeltindivider	...avgjøre om enkeltindivider..	Fjernet <i>e</i>
s. 8, linje 12	...tidligere, f selv	... tidligere, selv...	Fjernet <i>f</i>
s.12, linje 13 t.o.m.29	Punkt 2.1 15 linjer tas ut		Teksten 2.1 s. 12 satt inn to ganger ..
s. 12 linje 30 og 31 og s.13, linje 1 t.o.m. linje 11	Punkt 2.1.1.12 linjer tas ut		Teksten 2.1.1 satt inn to ganger
s.15, linje 19	,Smithson et al 2014) 2019.	,Smithson et al 2014)	Fjernet 2019).
s.16, linje 13	Det kunne også kunne vært	Det kunne også vært	Fjernet <i>kunne</i>
s.17, linje 19	Samlet viser disse studien at	Samlet viser studiene at	Fjernet <i>disse</i>
s.19, linje 26	... Baddeley and Hitch sin modell ...	Baddeley og Hitch sin modell	og erstattet <i>and</i>
s.22 linje 13	I mange år har den innenfor ...	I mange år har det innenfor.....	<i>det</i> erstatter <i>den</i>
s.23, linje 23	.. kan slik også ramme brukes av språket	... kan slik også ramme bruken av språket	<i>bruken</i> erstatter <i>brukes</i>
24 linje 4	...cutoff grense på -1.25 SD	... cutoff-grense på SD -1.00	- 1.00 SD erstatter - 1.25 SD
s.26, linje 6	...bakgrunn, gar vist at åbakgrunn, har vist at å ..	<i>har</i> erstatter <i>gar</i>
s.26, linje 11	Vokabular kan derfor den	Vokabular er derfor den	<i>er</i> erstatter <i>kan</i>
s. 26, linje 24	...et relativt ? «rent «mål	...et relativt «rent» mål	Spørsmålsteget fjernet
s.28, linje 9	.. grammatikk kartlagt på førstespråket påvirkes	grammatikk kartlagt på andrespråket påvirkes	<i>andrespråket</i> erstatter <i>førstespråket</i>
s.31, linje 14	Tabell XX	tabell 4.2	4.2 erstatter XX
s.32, linje 5	... rekruttert gjennom til pedagogisk-psykologisk tjeneste	... rekruttert gjennom pedagogisk-psykologisk tjeneste	Fjernet <i>til</i>
s. 33, linje 23	Tabell XX viser eksponering for førstespråket og tabell XX viser	Tabell 4.3 viser eksponering for førstespråket og tabell 4.4 viser	4.3 erstatter XX
s.39, linje 27	(vedlegg 3)	(vedlegg C)	<i>C</i> erstatter <i>3</i>
s.42, linje 15	... å gi oss innsikt (undersøke) i	... å gi oss innsikt i...	Fjernet (<i>undersøke</i>)
s.43, linje 23	... andrespråkets vokabular kan det ett av de to	... andrespråkets vokabular kan ett av de to	Fjernet <i>det</i>
s.47, linje 10	I likhet med mange andre studier har	I likhet med mange studier	Fjernet <i>har</i>
s.48, linje 11	...benytte ANVOVA analyser benytte ANCOVA analyser	<i>ANCOVA</i> erstatter <i>ANVOVA</i>
s. 53,	Korrelasjoner mellom ulike mål er presentert i appendix xx	Korrelasjoner mellom ulike mål er presentert i vedlegg H	<i>vedlegg H</i> erstatter <i>appendix xxx</i>
s.56, linje 25	Validiteten av for testene av to forhold	Validiteten for testene styrkes av to forhold	Tatt ut <i>av</i> og satt inn <i>styrkes</i>
s.57, linje 16	... at testene for bae første – og andrespråket..	...at testene for første – og andrespråket...	Tatt ut <i>bae</i>
s. 63, linje 1	... at diagnosen DLD kan stilles DLD.	...at diagnosen DLD kan stilles.	Tatt ut <i>DLD</i>
Doble parenteser: s.3,8,9,15,25,26,27,52,54			Fjernet paranteser s.3,8,9,15,25,26,27,52,54
Punktum mangler s.10, 22, 28, 51			Satt inn punktum s.10,22, 28, 51,64

