

Tospråklig ordforråd i tospråklig kontekst

*En kvantitativ studie av ordforrådet til barn
som snakker norsk, urdu og panjabi*

Liv Therese Norberg Haga



Masteroppgave i spesialpedagogikk
Institutt for spesialpedagogikk
Det utdanningsvitenskapelige fakultet

UNIVERSITETET I OSLO

Høst 2012

Tospråklig ordforråd i tospråklig kontekst

*En kvantitativ studie av ordforrådet til barn
som snakker norsk, urdu og panjabi*

Liv Therese Norberg Haga

Masteroppgave i spesialpedagogikk
Institutt for spesialpedagogikk
Det utdanningsvitenskapelige fakultet

UNIVERSITETET I OSLO

November 2012

© Liv Therese Norberg Haga

2012

Tospråklig ordførråd i tospråklig kontekst

Liv Therese Norberg Haga

<http://www.duo.uio.no/>

Trykk: Nilz & Otto Grafisk AS, Youngs gate 11, 0181 Oslo.

Sammendrag

Tospråklige kontekster for språkutvikling kan medføre at språkkompetansen til tospråklige kan skille seg fra språkkompetansen til enspråklige. Studier viser for eksempel ofte at tospråklige har lavere vokabularstørrelser enn enspråklige. Barn med pakistansk bakgrunn i Norge har en relativt kompleks språksituasjon, og bruker flere språk i det daglige. Derfor er det spesielt interessant og viktig å undersøke denne gruppens språksituasjon og språkkompetanse.

Problemstillingen for prosjektet er: Hva kjennetegner det reseptive breddeordforrådet til barn med pakistansk bakgrunn, og hvordan kan dette ordforrådet forstås i lys av aspekter ved barnas språksituasjon? *Reseptive breddeordforrådsstørrelser* (målt med norsktilpassede utgaver av BPVS for henholdsvis norsk og urdu/panjabi) undersøkes med en kvantitativ tilnærming og et deskriptivt design. Tospråksutvalget i studien består av seks år gamle barn med pakistansk bakgrunn (N = 61), og et enspråklig sammenlikningsutvalg består av seks år gamle barn med norsk bakgrunn (N = 190). *Språksituasjonen* til barn med pakistansk bakgrunn beskrives gjennom sentrale trekk ved tospråklige og minoritetsspråklige kontekster for språkutvikling, og ytterligere ved hjelp av kunnskap om språkrelaterte forhold for gruppen med pakistansk bakgrunn i Norge.

Studien viser at mange barn med pakistansk bakgrunn har lavere norskordforråd enn norsk-enspråklige barn. I tillegg viser studien at barn med pakistansk bakgrunn har mindre urdu/panjabiordforråd enn norskordforråd. En vurdering av disse resultatene i lys av flere sentrale kontekstuelle aspekter ved barnas språksituasjon, viser at resultatene i stor grad er forståelige og delvis forventet. Barna har utviklet, og bruker, tospråklige ferdigheter i en tospråklig kontekst. Resultatene ses som et naturlig uttrykk for at tospråklig kompetanse utvikles under andre kontekstuelle betingelser enn enspråklig kompetanse, men lave ordforrådsstørrelser må allikevel tas på alvor som en risikofaktor med hensyn til skolefaglig utvikling.

Masteroppgaven er skrevet i tilknytning til forskergruppen *Child Language and Learning* og det longitudinelle forskningsprosjektet *Language development in minority language children* ved Institutt for spesialpedagogikk (Universitetet i Oslo).

Forord

Denne masteroppgaven er et resultat av godt samarbeid med mange andre. Jeg vil først og fremst rette en takk til forskergruppen *Child Language and Learning* for anledning til å benytte verdifullt datamateriale i dette masterprosjektet. Takk også til lærekrefter og administrasjon ved Institutt for spesialpedagogikk for mange lærerike og godt organiserte studieår, spesielt seniorkonsulent Hilde Fjeldheim som har tilrettelagt gjennomføring av masteroppgaveemnet på en særdeles god måte. En særlig stor takk til min veileder ph.d. Jannicke Karlsen ved Institutt for spesialpedagogikk (UiO) for uvurderlig støtte gjennom hele prosjektet.

Kjetil og Belinda: Hjertelig takk for all støtte og inspirasjon i arbeidet med denne masteroppgaven. En hjertelig takk også til mine foreldre som har støttet meg på alle måter gjennom studietiden. Og kjære alle dere etter hvert svært gode venner og studiekamerater som jeg har lært med, lært av og levd livet med – på Blindern: Så lite vi visste, så mye vi har lært, men du verden så mye som ennå er ulært! Takk for tiden og kunnskapen dere har gitt meg! En spesielt stor takk må tilslutt gå til min mann, og nærmeste studiekamerat gjennom alle disse årene, Lars. Takk for særdeles godt samarbeid gjennom alle prosjekter, tross ulike fagfelt – din kunnskap blir min, og min kunnskap blir din.

Det er med glede, men også vemodighet, jeg setter sluttstrek for arbeidet med denne masteroppgaven. Det har vært en inspirerende og lærerik prosess, og jeg håper at resultatet kan bli til glede og interesse for flere enn meg selv.

Liv Therese Norberg Haga,

Kristiansand, 28. november 2012

Innholdsfortegnelse

SAMMENDRAG	V
FORORD	VII
INNHOLDSFORTEGNELSE	IX
1 INNLEDNING	1
2 RESEPTIVT BREDDEORDFORRÅD	5
3 TO SPRÅK – FLERE ASPEKTER	8
3.1 TOSPRÅKLIGHET	8
3.2 TOSPRÅKLIGHET I EN MINORITETSSPRÅKLIG KONTEKST	14
3.3 MINORITETSSPRÅKLIGE MED PAKISTANSK BAKGRUNN I NORGE	19
4 METODE	27
4.1 DESIGN	27
4.2 UTVALG OG POPULASJONER	28
4.3 INSTRUMENTER	32
4.4 FREMGANGSMÅTE	33
4.5 STATISTISKE ANALYSER	34
4.6 VALIDITET OG RELIABILITET	35
4.7 ETISKE VURDERINGER	36
5 RESULTATER	38
5.1 FORDELINGER	38
5.2 SIGNIFIKANSTESTER	43
5.3 EFFEKTSTØRRELSER	44
6 VALIDITETSVURDERING	47
7 TOSPRÅKLIG ORDFORRÅD I TOSPRÅKLIG KONTEKST	55
7.1 KJENNETEGN PÅ RESEPTIVT BREDDEORDFORRÅD HOS BARN MED PAKISTANSK BAKGRUNN	55
7.2 RESEPTIVT BREDDEORDFORRÅD I LYS AV ASPEKTER VED BARNAS SPRÅKSITUASJON	57
7.3 OPPSUMMERING	66
7.4 NOEN GENERELLE OG PEDAGOGISKE IMPLIKASJONER	67
8 AVSLUTTENDE KOMMENTARER	70
REGISTER	72
LITTERATURLISTE	73
VEDLEGG	79

1 Innledning

Med en stadig større andel barn i Norge med tospråklig bakgrunn er det viktig å vite hva som kjennetegner *tospråklig* kompetanse i motsetning til den *enspråklige* kompetansen til majoriteten av barn i Norge. Kunnskap om dette kan gi viktig informasjon om ulike gruppers betingelser for å lykkes i utdanningssystemet. Også det at tospråklige barn er overrepresentert i spesialundervisning (Baca & Cervantes, 2004; Espenakk et al., 2007; Kunnskapsdepartementet, 2010), tilsier at vi trenger bedre kunnskap om hva som kjennetegner normal tospråklig kompetanse.

Det å lære et språk kan betegnes som en svært komplisert prosess som påvirkes av kognitive og individuelle faktorer, men også i høy grad av de ulike sosiale og kontekstuelle omstendighetene språklæringen foregår i (Romaine, 1995; Bialystok, 2001; Baker, 2011). Dette innebærer at en *tospråklig* utviklingskontekst der flere språk brukes i varierende grad og kanskje til ulike formål, kan gi andre ferdigheter enn en *enspråklig* utviklingskontekst, der ett språk brukes i alle anledninger. Studier viser for eksempel ofte at tospråklige har mindre ordforråd i hvert av sine språk enn enspråklige (Bialystok, Luk, Peets & Yang, 2010). I tillegg kan ulike tospråklige kontekster påvirke hvilke ferdigheter tospråklige utvikler i hvert av sine språk. Dersom hjemmespråket er et minoritetsspråk, kan for eksempel tospråklige barn utvikle høyere kompetanse i samfunnsspråket enn i hjemmespråket (Verhallen & Schoonen, 1998), selv om det motsatte også kan være tilfelle. En nylig oversettelse av den internasjonalt mye anvendte ordforrådstesten *British Picture Vocabulary Scale – BPVS* – (Lyster, Horn & Rygvold, 2010; Monsrud, Thurmann-Moe & Bjerkan, 2010), muliggjør en nærmere undersøkelse av tospråkliges ordforrådskompetanse i en *norsk* sammenheng. Med BPVS kartlegges reseptivt breddeordforråd, som er et delområde av et barns totale semantiske kompetanse.

Barn med pakistansk bakgrunn utgjør den største gruppen av minoritetsspråklige skoleelever i Norge (Statistisk sentralbyrå, 2002). Som tospråklige med minoritetsbakgrunn og morsmål med røtter i en differensiert språkregion har denne gruppen en relativt kompleks språksituasjon. Hva kjennetegner den spesifikke konteksten for disse barnas tospråklige ordforrådsutvikling? Hvilken ordforrådskompetanse har barna i hvert av sine språk og i forhold til enspråklig norskspråklige barn? For å nærme meg dette komplekset har jeg valgt følgende problemstilling:

Hva kjennetegner det reseptive breddeordforrådet til barn med pakistansk bakgrunn, og hvordan kan dette ordforrådet forstås i lys av aspekter ved barnas språksituasjon?

Avgrensninger og presiseringer

Jeg ønsker altså å undersøke hva som kjennetegner det reseptive breddeordforrådet til tospråklige barn med pakistansk bakgrunn, og hvordan denne ordforrådskompetansen kan forstås i lys av aspekter ved barnas språksituasjon. Noen avgrensninger og presiseringer til problemstillingen er nødvendige: Når det gjelder *kjennetegn på reseptivt breddeordforråd*, velger jeg å undersøke dette på en kvantitativ måte og avgrenser til en undersøkelse av størrelsen på ordforråd i henholdsvis norsk og urdu/panjabi, samt størrelsen på totalordforråd. I tillegg undersøkes størrelsen på reseptivt breddeordforråd hos et norskspråklig sammenlikningsutvalg.

Barn med pakistansk bakgrunn skal i problemstillingen forstås som seks år gamle barn med urdu og/eller panjabi som hjemmespråk. Utvalg og populasjon spesifiseres ytterligere i metodedelens av oppgaven.

Aspekter ved barnas språksituasjon. For å avgrense velger jeg å ta utgangspunkt i aspekter som i vid forstand kan regnes som sosiale, eller som sosialt betingede – altså som kontekstuelle. Dermed utelukkes flere perspektiver som kunne være like relevante og interessante å fortolke resultatene i – for eksempel kognitive eller mer individfokuserede innfalsvinkler til fortolkning. Språkkompetanse behandles således i denne studien som et resultat av sosial interaksjon, i henhold til en funksjonalistisk (Bialystok, 2001; Goldenberg, Reese & Rezaei, 2011) tilnærming til språkutvikling. Dette innebærer at jeg legger til grunn at den språkerfaring barn får gjennom ulike sosiale kontekster, gir et sentralt bidrag til deres språkkompetanse.

Med basis i en slik funksjonalistisk tilnærming velger jeg å beskrive barnas språksituasjon ved hjelp av aspekter som kjennetegner tospråklige- og minoritetsspråklige kontekster for språkutvikling generelt, men også ved hjelp av aspekter som gjelder språkgruppen med pakistansk bakgrunn i Norge spesielt. Disse aspektene velger jeg i hovedsak å redegjøre for gjennom teori og tidligere empiri. Dette gir et grunnlag som gjør det mulig å se til et mangfold av aspekter ved barnas språksituasjon når resultatene – barnas reseptive breddeordforråd – skal forstås.

Begrepsavklaringer

Spørsmål om hva som skal regnes som personers eller gruppers morsmål er generelt komplisert, og særlig vanskelig når det tospråklige utvalget i studien delvis består av barn som har vokst opp med både urdu, panjabi og norsk. I sammenhenger der det ikke er naturlig å referere til ett spesifikt språk velger jeg derfor for enkelhetsskyld å bruke språk 1 (S1) eller hjemmespråk når jeg refererer til *urdu og panjabi*, og språk 2 (S2) eller samfunnspråk når jeg refererer til *norsk*. Dette betyr at S1 refererer til to språk, og at S2 også kan være hjemmespråk for noen av barna i studien. S1 brukes altså synonymt med urdu/panjabi, og S2 brukes synonymt med norsk.

Videre velger jeg, i tråd med praksis på feltet, å benytte betegnelsen tospråklig på en slik måte at betegnelsen også inkluderer flerspråklighet (se for eksempel Romaine, 1995). Utvalgene i studien betegnes dermed som henholdsvis *tospråklig utvalg* og *enspråklig utvalg*. *Tospråklig utvalg* viser således til utvalget som består av barn som snakker urdu, panjabi og norsk, og *enspråklig utvalg* viser til utvalget som består av barn som snakker norsk.

Det tospråklige utvalget i studien kan regnes som *minoritetsspråklig* i henhold til slik minoritetsspråklige defineres etter *språk- og kulturbakgrunn* i St.meld. nr. 16 (2006–2007) (Kunnskapsdepartementet, 2006): ”Med minoritetsspråklig bakgrunn menes her barn med en annen språk- og kulturbakgrunn enn norsk, samisk, svensk, dansk og engelsk” (s. 43). ’Minoritetsspråklig’ er en godt etablert betegnelse på personer eller grupper i Norge som kommer fra språklige og kulturelle minoriteter. Jeg velger derfor å bruke denne betegnelsen der det er naturlig å fremheve gruppens minoritetsposisjon i samfunnet. For øvrig betegnes minoritetsspråklige som *tospråklige*.

Disposisjon

I *kapittel to* gir jeg en kort introduksjon til en forståelse av reseptivt breddeordforråd – altså den typen ordforråd som jeg kartlegger i denne studien.

Kapittel tre inneholder en redegjørelse for den språksituasjonen jeg vil se resultatene i lys av: Vi skal se på hva som generelt kjennetegner tospråklige og minoritetsspråklige kontekster, og deretter på forhold ved språkgruppen med pakistansk bakgrunn i Norge spesielt.

I *kapittel fire* skal jeg redegjøre kort for valget av et kvantitativt og rent *deskriptivt* design i analyser, og ellers for ulike sider ved gjennomføringen av prosjektet. Det er altså *to* utvalg som inngår i studien: Et tospråklig utvalg som skal regnes som hovedutvalg, samt et enspråklig utvalg som gir et sammenlikningsgrunnlag for ordforrådsstørrelser. For å få noe mer utfyllende bakgrunnsinformasjon om utvalgene presenterer jeg også noen få bakgrunnsvariabler som kan regnes som sosioøkonomiske.

I *kapittel fem* presenteres resultatene, og i *kapittel seks* gjør jeg en validitetsvurdering av slutninger som kan trekkes av resultatene, og øvrige slutninger i studien.

I *kapittel sju* drøfter jeg oppgavens problemstilling. Jeg starter med en oppsummering og kort drøfting av resultatene – hva kjennetegner det reseptive breddeordforrådet til barn med pakistansk bakgrunn? Deretter drøfter jeg hvordan disse kjennetegnene på ordforråd kan forstås i lys av aspekter ved barnas språksituasjon. Til slutt skal vi se kort på noen generelle og pedagogiske implikasjoner av det oppgaven viser.

2 Reseptivt breddeordforråd

Meningsbærende ord

Ord kan brukes til mye: Til å hente frem kunnskap vi allerede har, til å uttrykke det vi mener, og samtidig til å lære nye ord og dermed få ny kunnskap (Stahl & Nagy, 2006). Kunnskap om ords meningsinnhold, semantisk kompetanse, kan derfor ses som en betydningsfull kompetanse – på et vis er ord både kunnskap i seg selv, middel for kommunikasjon og redskap til å erverve stadig ny kunnskap. Men hva er et ord, og hva er et ords mening? I en viss forstand kan ord beskrives som en språklig enhet som har både en *formside* – selve det fonologiske eller ortografiske uttrykket, og en *innholdsside*, eller mening (Golden, 2009); ord kan altså beskrives som ”merkelapper” på et kjent meningsinnhold. På den andre siden kan vi også si at ord i seg selv, som språklige ytringer, ikke har en umiddelbar gitt betydning, men heller får sin mening fra den konteksten de ytres i (Sveen, 1996). Ords faktiske mening kan derfor ikke oppfattes uten forankring i ytring og ytringskontekst (Sveen, 2005). Gitt en slik forståelsesramme er det først når ordet er i *bruk* at vi kan få en forståelse av hva det refererer til, og slik gripe meningsinnholdet.

Det finnes flere ulike typer kontekster som til sammen gir ord sin mening: For eksempel ytres et ord med en *form* (fonologisk og morfologisk kontekst) og som en del av en frase (syntaktisk kontekst). Slike kontekster bidrar til å gi ordet sin betydning i den aktuelle frasen. Videre, og ikke minst, får ord sin mening fra den aktuelle samtalen det inngår i og fra den øvrige situasjonelle og kulturelle konteksten (Rommetveit, 1972). Vi ser at dette utfordrer et ”merkelappsyn” på ord. Rommetveit (1972) beskriver godt hvordan et ord ikke kan karakteriseres som en enkel merkelapp med ett avgrenset meningsinnhold. Han hevder at ord heller må forstås som *meningspotensialer* med semantisk assosiative nettverk av meninger knyttet til seg – et ord har med andre ord ikke én bestemt mening, men mange *potensielle* meninger. *Innholdsaspekter* kan ikke utskilles fra *formaspekter*, og i praktiske kommunikasjonssammenhenger påvirker også den aktuelle situasjonen, samt den videre konteksten, ords mening. Kort sagt: Et ord kan ha ulike betydninger i ulike kontekster.

Semantisk kompetanse

En konsekvens av dette er blant annet at det ikke kan fastsettes et ”innslagspunkt” for når et ord kan regnes som behersket eller lært. Ifølge Rommetveit (1972) er dette snarere en læringsprosess som best kan beskrives som et *kontinuum* av beherskelse – fra spedbarnets meget personlige og konkrete situasjonsbeherskelse av ords meningsinnhold, gjennom gradvis innholdsutbygging og abstrahering av ords mening til en stadig rikere kunnskap om ordets potensielle innhold og bruksmåter i ulike kontekster; dets meningspotensial.

Dersom vi legger til grunn et slikt syn på hva ord er, blir det tydelig at *mange* aspekter er involvert i det å kunne et ord – innholdsaspekter, formaspekter og bruksaspekter virker sammen, og kunnskap om alle tre trengs for å kunne bruke og forstå et ord. Nation (2001) viser på en tydelig måte hvordan semantisk kompetanse består av ferdigheter på flere områder og nivåer. Overordnet deler han ordforrådskunnskap inn i de samme tre aspektene som vi har sett til nå: Kunnskap om ords *form*, om ords *mening*, og om ords *bruk*. Hver av disse komponentene deler han imidlertid igjen inn i flere kompetanseområder. Kunnskap om ords *mening* involverer blant annet kunnskap om hvilket meningsinnhold et ord har, hvilken ordform som kan uttrykke dette meningsinnholdet, hvilke meninger som kan inkluderes i konseptet i vid forstand, hvilke assosiasjoner meningen gir, og hvilke andre ord som kan brukes i stedet (Nation, 2001). På liknende måte deler Nation (2001) inn de to andre hovedkomponentene i flere underkunnskaper. Poenget, som det er viktig å understreke, er at semantisk kompetanse kan beskrives som en svært sammensatt kompetanse. Videre er det tydelig at semantisk kompetanse heller ikke er en statisk kompetanse, men heller en kompetanse i stadig endring og utvikling. Dette dynamiske ved semantisk kompetanse legger også Rommetveit (1972, s. 63) vekt på: ”Ordforståing [...] er difor slett ikkje noko me har *alt* eller *ingen ting* av, men noko som stendig endrar seg og aukar i nyanse og kompetanserikdom.”

Kartlegging av reseptivt breddeordforråd – et lite innblikk i en større helhet

I vokabularkartlegging brukes ofte en mer grovkornet inndeling av muntlig ordforrådskompetanse, og noen klare skiller virker etablert (Milton, 2009): Det skiller for det første mellom *reseptive*, eller passive ferdigheter, og *ekspressive*, eller produktive ferdigheter. Reseptive ferdigheter forstås vanligvis som den evnen en har til å kjenne ords mening når de brukes av andre, mens ekspressive ferdigheter anses som evnen til selv å bruke ord som en

kjenner meningsinnholdet til (Milton, 2009). Vanligvis utvikles reseptive ferdigheter før ekspressive – det vil si at evnen til å forstå hva ord betyr utvikles før evnen til å bruke ordene produktivt selv. Reseptivt ordforråd vil dermed som oftest være større enn det ekspressive ordforrådet (Milton, 2009). Videre skilles det mellom *breddeordforråd* og *dybdeordforråd*. Milton (2009) poengterer at dette skillet i praksis kan være vanskelig å sette, men den vanligste oppfatningen beskriver han slik: ”Breadth of knowledge refers to the number of words a learner knows and depth of knowledge refers to what the learner knows about these words” (s. 13). Et slikt skille mellom bredde- og dybdeordforråd bekreftes langt på vei gjennom observasjoner av tidlig språkutvikling: Et barn kan ha en ordform lagret i sitt leksikon som en del av sitt *breddeordforråd*, uten en fullstendig forståelse av dets mening knyttet til seg Oulette (2006). Dette korresponderer som vi ser godt med Rommetveits (1972) forståelse av ordforrådsutvikling som et kontinuum.

Ut fra denne grovinndelingen kan reseptivt breddeordforråd forstås som *mengden ord en person behersker (tilstrekkelig til å identifisere et adekvat meningsinnhold) når ordene brukes av andre*. Kartlegging av barns reseptive breddeordforråd gir dermed kunnskap om en liten, og det kan hevdes kunstig isolert, del av barns helhetlige semantiske kompetanse. Dette kan synes noe uheldig. Milton (2009) minner imidlertid om at semantisk kompetanse er en så sammensatt kompetanse at det åpenbart ikke vil finnes én test som alene vil kunne fange opp alle (eller i det hele tatt flere) typer ordforråds kunnskap. Derfor brukes ofte flere tester, som *hver for seg* kartlegger ett eller flere elementer av ordforråds kunnskap. En test av for eksempel reseptivt breddeordforråd representerer nettopp dette; en mulighet til å måle én type ordforråds kunnskap. Kartlegging av reseptivt ordforråd gir, tross (åpenbare) begrensninger, altså verdifull informasjon om en bestemt type språklig kompetanse: Altså et lite, men viktig innblikk i en større helhet.

3 To språk – flere aspekter

Formålet med denne delen av oppgaven er å beskrive ulike aspekter ved språksituasjonen til tospråksutvalget i studien som kan være relevante å se til i en drøfting av resultatene. Jeg begynner på et overordnet nivå med å beskrive forhold ved *tospråklighet* og smalner etter hvert inn til å beskrive generelle forhold ved en *minoritetsspråks*situasjon. Fremstillingen av disse mer generelle aspektene velger jeg å ordne ved hjelp av fire dimensjoner som regnes som sentrale for å beskrive tospråklighet som overordnet fenomen: En ferdighetsdimensjon, en funksjonsdimensjon, en aldersdimensjon og en kontekstdimensjon. I tillegg til disse generelle aspektene skal vi til slutt se nærmere på hva som kjennetegner språkgruppen med pakistansk bakgrunn i Norge.

Hensikten er at forhold jeg presenterer her, skal utfylle hverandre – ikke til en fullstendig fortolkningsramme som uttømmer relevant forståelseskontekst, men til et meningsfullt *utvalg* av aspekter som det kan være relevant å se resultatene i denne studien i lys av – en teoretisk og empirisk referanseramme som gjør det mulig å besvare problemstillingens andre ledd. I hele denne bakgrunnsdelen vil jeg redegjøre for sentrale teoretiske forhold og trekke inn relevante studier som belyser eller eksemplifiserer disse.

3.1 Tospråklighet

Ferdigheter

Aller først bør vi kanskje spørre: Hva er tospråklighet, eller hvem er tospråklige? Det synes åpenbart at tospråklighet må innebære *ferdigheter* i to språk på en eller annen måte, men i hvilken grad må to språk beherskes for at en person skal kunne regnes som tospråklig? Må en kunne snakke flytende, og både lese og skrive på begge språk, eller er det nok å kunne ytre noen fraser? Grosjean (1997) hevder at tospråklighet ofte oppfattes som det å beherske to språk like godt som en enspråklig behersker ett språk, men at dette ikke passer med den faktiske variasjonen man finner blant tospråklige i praksis. I kontrast til denne vanlige oppfatningen beskriver derfor de fleste nyere teoretiske fremstillinger *ferdighetsdimensjonen* i tospråklighet som et kontinuum: Tospråklighet kan eksistere i mange varianter – fra full beherskelse av alle språkferdigheter, til for eksempel bare å kunne forstå eller bruke noen få ytringer på et annet språk (Romaine, 1995; Bialystok, 2001; Baker, 2011). Forskjellige

språkferdigheter (lytte, lese, tale og skrive) kan også være til stede i *ulik* grad (Baker, 2011). Videre kan også ferdigheter innenfor disse områdene variere – for eksempel kan tospråklige ha bedre ferdigheter i å uttrykke semantisk innhold i det ene språket enn i det andre (Romaine, 1995). Tospråklighet innebærer, som vi ser, åpenbart ulike typer ferdigheter i to språk, men ikke nødvendigvis ferdigheter på *alle* språkområder, og heller ikke nødvendigvis like *gode* ferdigheter på alle språkområder. Det anslås at mellom halvparten og to tredjedeler av verdens befolkning kan regnes som tospråklig (Baker, 2011). Bare et fåtall av disse har ifølge Grosjean (1997) like gode ferdigheter i begge språk. På grunn av denne variasjonen er det ikke enkelt å finne en godt svar på hvem som er tospråklige, eller i det hele tatt en dekkende beskrivelse av hva tospråklighet er, med utgangspunkt i en slik ferdighetsdimensjon (Baker, 2011; Romaine, 1995). En ofte sitert beskrivelse av hva tospråklighet er og hvem som kan regnes som tospråklige, og som nettopp ikke setter absolutte grenser for ferdighet i to språk, finner vi hos Grosjean (2008, s. 10): ”Bilingualism is the regular use of two or more languages (or dialects), and bilinguals are those people who use two or more languages (or dialects) in their everyday lives”. En slik forståelse vektlegger heller selve *bruken* av to språk i det daglige – personer som bruker to språk i hverdagslivet, er i en slik forstand å regne som tospråklige.

Funksjon

En slik forståelse av tospråklighet leder naturlig over til en bruksdimensjon, eller *funksjonsdimensjon*, som tospråklighet som fenomen ofte forstås i lys av (Romaine, 1995; Baker, 2011): Til hva og i hvilke sammenhenger brukes de to ulike språkene? Som oftest bruker ikke tospråklige begge språk i alle sammenhenger, men heller forskjellige språk til ulik bruk eller i ulike situasjoner. Man sier gjerne at språkene brukes i ulike *domener* (Baker, 2011). Eksempler på slike domener kan være familie, skole og arbeidsplass. Ett språk kan brukes hjemme, og et *annet* språk kanskje på skolen.

En naturlig konsekvens av det at forskjellige språk ofte brukes i hver sine domener, er at tospråklige sjelden har lik språkkompetanse i de forskjellige språkene (Baker, 2011). Et språk som primært brukes og utvikles i en uformell hjemmesituasjon, vil ikke nødvendigvis fungere til å snakke om forhold som krever språkkompetanse som er tilegnet i andre domener og på et annet språk (Romaine, 1995). Hvilken *funksjon* et språk har, henger, som vi ser, dermed tett sammen med hvilke *ferdigheter* som utvikles på det språket.

Grosjean (2008) teoretiserer dette samspillet mellom ferdigheter og funksjon ved hjelp av det han kaller *the complementary principle*. Komplementærprinsippet beskrives som det at tospråkliges ferdigheter naturlig *kompletterer* hverandre til en total funksjon på alle livsområder, eller domener. Bak dette prinsippet ligger også oppfatningen om at det bare er de språkferdigheter som *faktisk* trengs til kommunikasjon som utvikles: Hvis et språk aldri brukes til en bestemt funksjon (for eksempel skolearbeid), vil ikke ferdigheter utvikles for denne funksjonen på det språket (Grosjean, 2008). Og dersom et språk brukes i et mindre antall domener, vil det ikke utvikles like omfattende ferdigheter i dette språket som i et språk som brukes i mange domener (Grosjean, 2008). I lys av en slik komplementær forståelse av tospråklige ferdigheter og tospråklig funksjon, ser vi at det vil være naturlig at selv personer som bruker to språk på en daglig basis ikke nødvendigvis har den samme kompetansen i hvert av sine språk som enspråklige, som bruker det samme språket til alle funksjoner og i alle domener.

Vi skal se litt nærmere på en studie som eksemplifiserer denne forståelsen av tospråklige ferdigheter som noe annet enn enspråklige ferdigheter med hensyn til *ordforråds*kompetanse spesielt. Bialystok, Luk, Peets og Yang (2010) ønsket å undersøke hvorvidt enkeltstudier som viste forskjellige ordforrådsstørrelser mellom tospråklige og enspråklige, reflekterte et generelt mønster, og sammenstilte derfor resultater fra flere mindre studier av reseptivt ordforråd hos enspråklige og tospråklige. Det sammenslåtte utvalget besto av til sammen 1738 barn (772 enspråklige og 966 tospråklige). Alle de tospråklige barna snakket i tillegg til engelsk et annet språk hjemme (hjemmespråket varierte mellom svært mange ulike språk). Foreldrene rapporterte at barna hadde gode ferdigheter i begge språk, og at de brukte begge i hverdagen – ”on a daily basis” (Bialystok et al., 2010, s. 526). Bialystok et al. (2010) undersøkte reseptivt ordforråd for hele dette utvalget (tospråklige og enspråklige), og fant at tospråklige skåret signifikant lavere enn enspråklige. Denne forskjellen fant de videre at gjaldt for alle alderstrinn de undersøkte; fra 3 til 10 år. Det var ganske jevne utvalgsstørrelser av tospråklige og enspråklige på alle alderstrinn. Det at Bialystok et al. (2010) hadde et så stort utvalg, og at hjemmespråkene varierte mellom mange ulike språk, viser at det er god grunn til å tro at det er noe ved en tospråklig språksituasjon i seg selv som gir slike ulikheter i ferdigheter mellom enspråklige og tospråklige.

I følge Thordardottir (2011) er det en generell og tydelig tendens til slike funn, selv om like ordforrådsstørrelser også forekommer. Dette fenomenet, altså den tydelige forskjellen i

ordforrådsstørrelser mellom tospråklige og enspråklige, refereres ofte til som *the vocabulary gap* (Thordardottir, 2011) – en *vokabularavstand* kan vi kanskje kalle dette på norsk. Forklaringen som vanligvis gis for denne vokabularavstanden mellom tospråklige og enspråklige, er ifølge Thordardottir (2011, s. 428) nettopp det at tospråklige barn må dele læringstiden mellom ulike språkmiljøer: ”Amount of language input is known to influence strongly the rate of language growth in both monolingual and bilingual children.” *Semantisk kompetanse* vil være spesielt utsatt for dette input/erfarings-elementet, siden vokabular må læres ”ord for ord” og derfor i enda større grad er avhengig av direkte erfaring over lengre tid – i motsetning til grammatiske regler og for eksempel avkodingsferdigheter som kan tilegnes som et begrenset system over noe kortere tid (Bialystok & Feng, 2011; Thordardottir, 2011).

Oller og Eilers (2002) har sett nærmere på akkurat denne *vokabularavstanden*, basert på en omfattende studie av spansk- og engelsktalende barn i Miami. Denne studien inkluderte flere tester av ulike språkferdigheter, blant annet den amerikanske versjonen av BPVS (PPVT; Peabody Picture Vocabulary Test) og den spanske versjonen (TVIP), og hadde som primærmål å undersøke sammenhengen mellom tospråklig opplæring og tospråklige ferdigheter. Datamaterialet viste imidlertid også interessante mønstre med hensyn til ordforråds kunnskap – det viste seg kort sagt at det var større forskjeller mellom tospråklige og enspråklige med hensyn til ordforråd enn ved enkelte andre språkferdigheter. På bakgrunn av dette hypotiserer Oller og Eilers (2002) at vokabularavstanden mellom tospråklige og enspråklige kan forstås ut fra nettopp det at tospråklige gjerne bruker ulike språk i ulike domener, og at dette derfor rimeligvis gir seg utslag i at ulike ord læres på ulike språk: ”Some words are learned in one language but not the other, because those words tend to be used in situations where one language, but not the other is the medium of exchange”, skriver Oller og Eilers (2002, s. 10). Oller og Eilers (2002) betegner dette som tospråklig ordforråds *distribuerte* karakter: Domenespesifikke ord læres bare på ett av språkene (singelord), mens mindre domenespesifikke ord vil finnes på begge språk (dobbelord). Tospråklige vil altså ha mange dobbelord – ord som brukes over flere domener, men kan *også* ha flere singelord som bare brukes i det ene språket. Vi ser at dette korresponderer godt med Grosjeans (2008) komplementærprinsipp – distribusjonshypotesen kan kanskje ses som en *spesifisering* av dette: Mens komplementærprinsippet er ment å beskrive en totalitet av tospråklige ferdigheter, er distribusjonshypotesen mer spesielt innrettet mot å forklare forskjeller i *ordforråds*kompetanse mellom en- og tospråklige.

Antakelsen om at ordforrådsforskjeller mellom enspråklige og tospråklige kan skyldes en slik distribusjon, støttes av funn i Oller, Pearson og Cobo-Lewis (2007). Denne studien reanalyserte data fra Miami-studien til Oller og Eilers (2002) og da *spesielt* med tanke på å teste distribusjonshypotesen om tospråklig ordforråd. Oller et al. (2007) fant en klar tendens til at ordforrådet til tospråklige spansk- og engelsktalende barn var distribuert mellom de to språkene på den måten at en signifikant andel ord forekom bare i ett av språkene. Også Bialystok et al. (2010), som vi så over, fant at vokabularavstanden mellom enspråklige og tospråklige kunne nyanseres i lys av at ulike ord brukes i ulike domener. De klassifiserte ord i PPVT som enten hjemmeord eller skoleord, og fant at vokabularavstanden mellom enspråklige og tospråklige i *engelsk* i stor grad kunne tilskrives forskjeller i ord som brukes i *hjemmedomener*: Tospråklige hadde ikke like mange hjemmeord på *engelsk* som de enspråklige hadde. Dette ser Bialystok et al. (2010) som en naturlig følge av at engelsk ikke brukes så utstrakt i tospråklige hjem som i enspråklige, noe som korresponderer godt med distribusjonsprinsippet.

Alder

Fenomenet tospråklighet bør imidlertid forstås i lys av flere dimensjoner. For eksempel trekker Baker (2011) frem *alder for introduksjon* til et annet språk som sentralt i forståelsen av tospråklig kompetanse. Korresponderende med dette gjøres det et skille i faglitteraturen mellom det å lære to språk parallelt fra fødselen av, som ofte refereres til som *simultan tospråklighet*, og det å lære ett språk først for så å introduseres for et annet på et senere tidspunkt; noe som betegnes som *suksessiv tospråklighet* eller andrespråklæring (Øzerk, 2003; Baker, 2011). Vanligvis brukes betegnelsen suksessiv tospråklighet eller andrespråklæring dersom barnet introduseres for andrespråket etter omtrent treårsalder (Baker, 2011). Det er imidlertid ingen fullstendig etablert enighet om dette innslagspunktet, og Hammer, Miccio & Rodriguez (2004) foreslår en kontinuumsforståelse også av denne dimensjonen i tospråklighet. For å sette forskjeller på spissen kan vi allikevel si at simultantospråklig utvikling i teorien som oftest refererer til tilfeller der barnet lærer to språk helt parallelt i hjemmet, og at suksessiv tospråklighet/andrespråklæring brukes om tilfeller der et annet språk introduseres først *etter* 3–4-årsalder, og da i en kontekst utenfor hjemmet.

Denne distinksjonen, eller disse ytterpunktene på alderskontinuumet, vil kanskje spesielt påvirke ordforrådsferdigheter, siden ordforrådsutvikling som vi så over innebærer at nye ord

må høres og brukes for å læres: Ordforrådsutvikling tar tid. I tråd med dette hevder for eksempel Verhallen og Schoonen (1998) at en av de største utfordringene for andrespråklærere er å utvikle nok ordforråd på andrespråket. Flere studier har også vist at det tar lang tid å utvikle ferdigheter på et andrespråk – kanskje så mye som fra fem til ti år eller mer (Cummins, 1994; Thomas & Collier, 1997). I tillegg er ordforrådsutvikling tett knyttet til erfaringer på det aktuelle språket (Bialystok, 2001; Lindberg 2008; Wold, 2008). Andrespråklærere kan for eksempel bare utvikle ordforrådskompetanse i de situasjoner hvor de kommer i direkte kontakt med andrespråket, for eksempel i barnehage, skole eller i butikken (Wold, 2008).

Aldersdimensjonen belyser derfor helt åpenbart noen forskjeller som kan ha betydning for *graden* av ferdigheter som utvikles på de ulike språkene – en tidligere læringsstart vil normalt føre til et høyere ferdighetsnivå enn et senere læringsstart (Bialystok, 2001). I tillegg ligger det implisitt i aldersdimensjonen at også *ulike* typer ferdigheter kan følge av en tidligere eller senere introduksjon, ettersom en senere introduksjon som oftest også innebærer introduksjon i andre domener enn hjemmet. I henhold til komplementærprinsippet (Grosjean, 2008) vil for eksempel et barn i simultantospråklig utvikling i større grad bruke to språk *innenfor* det samme domenet – som oftest hjemmedomenet, og dermed utvikle kompetanse på begge språk i ferdigheter som kreves der. Andrespråklærere vil i motsetning til dette kanskje i større grad få *domenespesifikk* kunnskap – først i hjemmet på ett språk, og deretter utenfor hjemmet, for eksempel i barnehage eller skole, på et annet språk.

For å oppsummere så langt, kan vi si at tospråklige *ferdigheter* ser ut til å utvikles i et tett samspill med hvilken *funksjon* de ulike språkene har, og hvilken *alder* andrespråket introduseres. Dette gir en god indikasjon på at tospråklig kompetanse (nærmest med nødvendighet) kan regnes som noe kvalitativt annerledes enn enspråklig kompetanse: Dersom tospråklige ferdigheter utvikles i ulike domener, og gitt et underliggende premiss om at ferdigheter bare utvikles i de domenene der de faktisk trengs, følger en naturlig kompetanseforskjell mellom enspråklige og tospråklige.

3.2 Tospråklighet i en minoritetsspråklig kontekst

I tillegg til ferdighetsdimensjonen, funksjonsdimensjonen og aldersdimensjonen bør tospråklighet også forstås i lys av at språkkompetanse utvikles i samspill med ulike kulturelle og samfunnsmessige omstendigheter (Baker, 2011; Bialystok, 2001) – en *kontekstdimensjon*. Vi skal her se nærmere på noen forhold som kjennetegner *minoritetsspråklige* kontekster, men først: Hva ligger i begrepet minoritetsspråklig?

Romaine (1995) kategoriserer språkkontekster for tospråklig utvikling blant annet etter grad av støtte for hjemmespråket i samfunnet for øvrig. Med utgangspunkt i denne systematiseringen kan det settes et grovt skille mellom to grunnleggende ulike kontekster for tospråklig utvikling hos barn: I noen tilfeller bruker en eller begge av foreldrene et språk som i større eller mindre grad også anvendes i samfunnet – da vil barnet møte dette hjemmespråket også i flere sammenhenger. Et eksempel kunne være barn i Norge som vokser opp med engelskspråklige foreldre. I motsetning til dette vokser mange opp med et hjemmespråk som ikke anvendes i særlig grad i samfunnet for øvrig: Språket er således et *minoritetsspråk* i den forstand at det snakkes av en mindre gruppe i samfunnet, og ikke brukes i samfunnskonteksten utover de arenaene der mennesker som bruker det samme minoritetsspråket møtes. Også Øzerk (2003) trekker opp nettopp denne overordnede distinksjonen, og betegner de to kategoriene av tospråklighet som henholdsvis *majoritetstospråklige* og *minoritetstospråklige*. Hva innebærer en minoritetsspråklig kontekst for utvikling av tospråklige ferdigheter? Vi skal i det følgende se litt nærmere på *hjemmespråklige* forhold og *samfunnsforhold* i en minoritetsspråkkontekst. Til slutt skal vi se kort på en type kontekstuelle forhold som også ser ut til å virke inn på særlig ordforrådsutvikling, nemlig sosioøkonomiske.

Minoritetsspråk i hjemmet

Hjemmespråkmiljøet i minoritetsspråklige kontekster vil naturlig nok stå i et kontinuerlig brytningsforhold mellom bruk av minoritetsspråket og bruk av samfunnspråket. Derfor ser man ofte at språkmiljøet i minoritetsspråksfamilier preges av en mer skiftende bruk av språk. Mancilla-Martinez og Kieffer (2010) fant for eksempel i en stor longitudinell undersøkelse av 1538 minoritetsspråklige barn med varierende minoritets-S1 og engelsk som S2, at språkbruk i familiene deres endret seg over tid. Mancilla-Martinez og Kieffer (2010) hevder derfor at språkbruk i minoritetsspråklige familier må betraktes som et *dynamisk* fenomen. I bare

halvparten av familiene holdt nivået for bruk av minoritetsspråket seg stabilt, mens i mer enn en tredjedel av utvalget endret mønsteret seg i den forstand at minoritetsspråket ble brukt mindre med tiden (Mancilla-Martinez & Kieffer, 2010). Disse endringene viste seg gjeldende for både barn og foreldre, og knyttet særlig til det at barna etter hvert lærer majoritetsspråket utenfor hjemmet (Mancilla-Martinez & Kieffer, 2010). Det at barna starter å lære og bruke majoritetsspråket kan altså føre til at også *foreldrene* går over til å bruke majoritetsspråket hjemme, noe også Sirén (1995) viste i en omfattende tverrsnittstudie av ulike minoritetsspråksgruppers hjemmespråksbruk i Sverige.

Det at foreldrene begynner å bruke majoritetsspråket kan selvsagt ha flere årsaker. Blant annet er det ofte en naturlig konsekvens av lengre botid i landet. Det er i tillegg vist spesielt at foreldrenes språkbruksmønster også i stor grad påvirker utviklingen av *minoritetsspråket hos barna*. De Houwer (2007) viser for eksempel at dersom *begge* foreldre snakker minoritetsspråket, og bare én forelder snakker majoritetsspråket i tillegg til minoritetsspråket, økes sjansene for at barnet skal bli tospråklig. Dersom begge foreldre snakker *begge* språk, minsker imidlertid sannsynligheten for tospråklig utvikling hos barnet (De Houwer, 2007). De Houwer (2007) fant for øvrig at barn som vokser opp i en minoritetsspråklig kontekst uten unntak lærer *majoritetsspråket*, det vil si uavhengig av foreldrenes språkbruksmønster, men samtidig altså at minoritetsspråket sto i fare for ikke å bli brukt. Liknende resultater fant også Sirén (1995), som rapporterer at foreldrenes kommunikasjon med barnet har betydning for barnets språkbruk: Dersom mor eller far bruker *begge* språk, fører dette til at også barnet bruker minoritetsspråket mindre i kommunikasjon med den forelderen. Denne effekten viser seg sterkest i familier hvor bare én av foreldrene bruker et minoritetsspråk – her viser Siréns (1995) resultater at minoritetsspråksbruken ble overført til barnet i 80–90 % av tilfellene der minoritetsspråksforelderen utelukkende brukte *minoritetsspråket* hjemme, mens denne andelen falt der minoritetsspråksforelderen brukte begge språk (avhengig av om det var far eller mor som var minoritetsspråklig, til henholdsvis 26 og 41 %). Siréns (1995) studie var for øvrig basert på en rekke ulike minoritetsspråksgrupper i det svenske samfunnet.

Slike studier peker mot at graden av bruk av minoritetsspråket i familien kan ha en betydning for språkutviklingen til minoritetsspråklige barn – noe som kan ses som naturlig i og med at foreldrene er barnas viktigste kilde til utvikling av minoritetsspråket i tidlig alder. Dette synes også sannsynlig i lys av generell kunnskap om tospråklighet som vi så over: Dersom domenene for minoritetsspråket i utgangspunktet er få, og majoritetsspråket brukes også der,

blir tiden til og erfaringene i hjemmespråket mindre. Sirén (1995) bemerker i tillegg at barn riktignok har kapasitet til å tilpasse seg og bruke ulike språk overfor ulike personer, men at dersom en forelder ikke aktivt signaliserer et behov for adressering på minoritetsspråket og heller selv veksler mellom språkene, er det sannsynlig at barnet vil følge ”the principle of least resistance” (s. 80). Sirén (1995) spesifiserer ikke dette nærmere, men vi forstår at det kan innebære at det språket som brukes mest i barnets liv for øvrig kanskje faller naturligst for barnet å anvende, særlig ettersom det også kan være sannsynlig at kompetansen etter hvert også er høyest i dette språket.

Vi skal straks se på minoritetsspråk i lys av samfunnsaspekter, men først må det også poengteres at denne innflytelsen fra majoritetsspråket kan føre til at utviklingen av S1 ikke bare går saktere, men også stagner ved at barn i språklige minoritetsfamilier selv i ganske ung alder går over til å bruke majoritetsspråket mest (Wong-Fillmore, 1991; Øzerk, 2003; Baker, 2011). Dette fenomenet, at samfunnspråket utvikles på bekostning av hjemmespråket, betegnes som *subtraktiv* tospråklig utvikling (Baker, 2011; Øzerk, 2003). Dette står i motsetning til en *additiv* tospråklig utvikling, hvor tilegnelse av samfunnspråket ikke går på bekostning av kompetanse i hjemmespråket (Baker, 2011; Øzerk, 2003).

Minoritetsspråk i samfunn

Det er ikke bare hjemmespråksforhold som kan påvirke en minoritetsspråklig utvikling, men også forhold som kan lokaliseres i den videre *samfunnskonteksten*. Vi kan for det første anta at graden av bruk av minoritetsspråket i samfunnet kan påvirke ferdigheter i minoritetsspråket: Dersom barn fra minoritetsspråksfamilier ikke møter sitt hjemmespråk i samfunnet forøvrig, vil dette redusere domenene for utvikling av kompetanse i minoritetsspråket. Verhallen og Schoonen (1998) viste for eksempel at barn med tyrkisk bakgrunn i Nederland etter noen år med opplæring i majoritetsspråket nederlandsk hadde bedre dybdeordforrådskompetanse i nederlandsk enn i tyrkisk – men fortsatt ikke like gode ferdigheter i nederlandsk som enspråklige. Studien inkluderte 40 ni til elleve år gamle andrespråkslærere med tyrkisk som S1, som i hovedsak hadde brukt bare tyrkisk hjemme (noen svært få av barna snakket nederlandsk med søsken). I motsetning til dette er *tospråklige* opplæringsmodeller vist å styrke hjemmespråksferdigheter. Den tidligere nevnte Miami-studien til Oller og Eilers (2002) undersøkte språkferdigheter hos enspråklige (engelsk) og tospråklige (spansk-engelsk), og hadde som nevnt som *primærmål* å undersøke effekten av

tospråklige undervisningsopplegg som vektla bruk av hjemmespråk. Det viste seg at barn som hadde fått 60 % av undervisningen fra barnehage og oppover på hjemmespråket og 40 % av undervisningen på majoritetsspråket, hadde mye sterkere S1-ferdigheter i femte klasse enn tospråklige barn som hadde vært undervist i bare engelsk. Studien viste samtidig at ferdighetene i S2 ikke var dårligere for barna som hadde fått tospråklig opplæring. Denne sammenhengen gjaldt også når sosioøkonomiske forhold ble kontrollert for (Oller & Eilers, 2002). Slike studier peker mot at graden av bruk av minoritetsspråk i samfunnet for øvrig påvirker språkutviklingen for tospråklige barn – selv om minoritetsspråket brukes hjemme, er altså ikke dette nødvendigvis tilstrekkelig for en god utvikling av S1.

For det andre er det også sannsynlig at samfunnsforhold kan påvirke ferdigheter i minoritetsspråk på en mer *indirekte* måte. Minoritetsspråk har vanligvis ikke en spesielt høy status i samfunnet for øvrig – der det er *majoritetsspråket* som gir prestisje. Baker (2011, s. 127) skriver for eksempel: “With a higher status majority language ever-present on the screen, in the street, at school and in shops, children quickly learn which language has prestige, power and preference.” Dersom språkene ikke har likeverdig status i samfunnet, kan dette kanskje føre til at barn heller velger å bruke majoritetsspråket, også i situasjoner der det hadde vært mulig å bruke minoritetsspråket. Et eksempel på at samfunnet sender indirekte signaler om hvilket språk som har mest verdi, kan være nettopp valg av opplæringsform – som vi så over. Dersom samfunnet gjennom ensidig vektlegging av majoritetsspråket sender signaler om at minoritetsspråket ikke er viktig, kan dette være en indirekte årsak til forsinket eller stagnerende hjemmespråkutvikling (Monsrud et al., 2010). I denne forbindelse er det interessant at Thordardottir (2011) fant ganske like reseptive ordforrådsstørrelser mellom tospråklige og enspråklige i en samfunnskontekst av jevnbyrdighet mellom språkene engelsk og fransk (Montréal, Canada), der de tospråklige barna i tillegg også hadde lært begge språk simultant i oppveksten. Vi ser av dette at forholdet mellom et minoritetsspråk og et majoritetsspråk også impliserer et ujevnt maktforhold som ikke kan undervurderes som indirekte påvirkningsfaktor i en minoritetsspråklig kontekst.

I noen tilfeller kan konsekvensen av et sterkt majoritetsspråk over lengre tid (over generasjoner) bli et språkskifte for hele minoritetsspråksgrupper. Romaine (1995) beskriver mønstret for slike språkskifter slik: “The classic pattern is that a community which was once monolingual becomes transitionally bilingual as a stage on the way to the eventual extinction of its original language” (s. 39f). I et slikt generasjonsspråksskifte er det altså en midtfase,

hvor minoritetsspråksgruppen bruker både minoritetsspråket og majoritetsspråket, før overgangen til majoritetsspråket inntreffer – vanligvis etter tre generasjoner (Alba, Logan, Lutz & Stults, 2002; Hurtado & Vega, 2004). Hurtado og Vega (2004) hevder at det er velkjent at slike språkskifter skjer, ofte også raskere enn tre generasjoner, men at det også er eksempler på at minoritetsspråket bevares selv etter at majoritetsspråket er blitt det dominante i bruk. En bevaring av minoritetsspråksferdigheter på tross av et sterkt dominerende majoritetsspråk kan for eksempel understøttes av kontinuerlig immigrasjon fra språkets opprinnelsesland. Romaine (1995) påpeker også at en fase av tospråklighet for minoritetsgrupper ikke *nødvendigvis* følges av et språkskifte; her viser hun til den spanske minoriteten i USA som fornyer sitt minoritetsspråk blant annet nettopp gjennom slike immigrasjonsmønstre.

Minoritetsspråk og sosioøkonomisk status

Vi skal se på et siste aspekt som kan ha konsekvenser for språkutvikling; sosioøkonomiske og sosiokulturelle forhold. Dette er en sekkebetegnelse på en lang rekke forhold som kan sies å være økonomiske eller kulturelle trekk ved grupper. Vanligvis regnes for eksempel *inntektsnivå* som en sosioøkonomisk faktor, mens for eksempel utdanningsnivå og tilgang på datamaskiner og bøker i hjemmet regnes som sosiokulturelle faktorer (Bakken, 2003). I det videre refererer jeg for korthetsskyld til slike forhold som sosioøkonomiske forhold, eller sosioøkonomisk status (SØS). Det som bringer opp dette temaområdet som et selvstendig vurderingsaspekt i denne sammenhengen, er at en minoritetsspråkssituasjon ofte (men selvsagt ikke alltid) følges av lav SØS (Snow & Kim, 2007; Bakken, 2003; Øzerk, 2003). Bakken (2003) har for eksempel vist at minoritetsspråksstatus i stor grad kan forbindes med lav SØS i *norsk* sammenheng. Samtidig er sosioøkonomiske forhold vist å korrelere høyt med nettopp ordforrådskunnskap – ordforrådskunnskap er kanskje det språkområdet som i størst grad påvirkes av slike forhold (Aukrust, 2005). En av de mest kjente studiene som har vist en samvariasjon mellom sosioøkonomisk status og språkferdigheter, er Hart og Risley (1995) som fant at forskjeller i barns ordforrådsutvikling korrelerte høyt med foreldrenes sosioøkonomiske status. Videre viste en studie (Dollaghan et al., 1999) av 241 engelsk-enspråklige treåringer at *reseptive ordforrådsstørrelser* (målt med PPVT-R) økte signifikant ($< .01$) over grupper inndelt etter mødrenes utdanningsnivå, spesifisert av Dollaghan et al. (1999) til *under high school*, *fullført high school* og *fullført college*.

Årsaken til en slik *samvariasjon* er komplisert (Bialystok 2001), og uten å gå inn i dette bakteppet av mulige forklaringer (se for eksempel også Edwards, 2009; Øzerk, 2003; Bakken, 2003) nøyer vi oss med å konstatere at mye tyder på at sosioøkonomiske forhold er et relevant aspekt å ta med i betraktning når ordforrådskompetanse skal vurderes, i tillegg til de aspektene vi allerede har sett på av språklig karakter. Her kan det bemerkes at den nære sammenhengen mellom en minoritetssituasjon og lav SØS gjør at det kan være vanskelig å skille ut hva som påvirker ordforrådsstørrelser *mest* for minoritetsspråklige: Som vi så over, er det sannsynlig at en minoritetsspråkssituasjon i seg selv representerer en utfordring med hensyn til å utvikle ordforråd på to språk, men det er tydelig at også sosioøkonomiske forhold kan virke inn på denne ordforrådsutviklingen uavhengig av språksituasjon for øvrig. Oller & Eilers (2002) fant imidlertid sterke *selvstendige* effekter fra sosioøkonomisk status på språkferdigheter i en minoritetsspråklig kontekst: Minoritetsspråklige (spansk- og engelsktalende) barn med høy SØS presterte bedre enn minoritetsspråklige (spansk- og engelsktalende) barn med lav SØS. Slike funn indikerer at sosioøkonomisk status gir et selvstendig bidrag *innenfor* konteksten av minoritetsspråklig utvikling.

For å oppsummere kort: Vi ser at en *minoritetsspråkssituasjon* innebærer et kontinuerlig brytningsforhold mellom hjemmespråk og samfunnsspråk, og at kompetansen til minoritetsspråklige barn kan påvirkes direkte av hvor mye tid og erfaring de får i de ulike språkene – både i hjemmet og i samfunnet. Kompetansen kan også påvirkes indirekte gjennom en ulik maktbalanse mellom minoritetsspråket og majoritetsspråket, og av sosioøkonomiske bakgrunnsforhold. Skillet mellom faktorer i hjemmet og faktorer i samfunnet representerer for øvrig selvsagt en forenkling. I praksis virker faktorer som dette i et samspill: Hjemmespråksforhold utvikles for eksempel i en vekselvirkning med hvilke verdier samfunnet generelt har til tospråklighet (Hammer et al., 2004).

3.3 Minoritetsspråklige med pakistansk bakgrunn i Norge

Vi skal nå snevre inn konteksten ytterligere og se spesielt på språklige aspekter for gruppen av personer med pakistansk bakgrunn i Norge. Men først: Hvem skal regnes som personer med pakistansk bakgrunn? Andreassen og Dzamarija (2008) skiller mellom *innvandrere* og *norskfødte med innvandrerforeldre*, og regner både innvandrere fra Pakistan og norskfødte med foreldre innvandret fra Pakistan som personer med pakistansk bakgrunn. Tall fra

Statistisk sentralbyrå (Andreassen & Dzamarija, 2008) viser at det var ca. 16 000 innvandrere med pakistansk bakgrunn i Norge i 2008 og ca. 13 000 norskfødte med foreldre innvandret fra Pakistan. Hvis vi slår sammen innvandrere og norskfødte med foreldre innvandret fra Pakistan til én gruppe av personer med pakistansk bakgrunn, utgjør disse den nest største innvandrergruppen i Norge med ca. 30 000 innbyggere (Andreassen & Dzamarija, 2008). Dette betyr også at det er en relativt stor andel (45 %) av personer med pakistansk bakgrunn som består av *norskfødte med innvandrerforeldre*. Dette er den desidert største andelen av norskfødte med innvandrerbakgrunn sammenliknet med andre innvandrergrupper i Norge (Andreassen & Dzamarija, 2008). En slik stor andel av norskfødte med innvandrerforeldre kan ses som en naturlig følge av at denne gruppen har en lang innvandringstid til Norge: Arbeidsinnvandring fra Pakistan startet allerede i 60- og 70-årene (Engen & Kulbrandstad, 2004). Til tross for at gruppen har lang innvandringshistorie og en høy andel norskfødte med innvandrerforeldre, kan gruppen sett under ett allikevel sies å ha en nær tilknytning til pakistansk språk og kultur (Karlsen & Lykkenborg, 2012). Gruppen har blant annet en lang og sterk tradisjon for familieegjenforening og familieetablering (Andreassen & Dzamarija, 2008). Andelen innvandrere som kommer fra Pakistan for familieetablering, altså for å stifte familie i Norge, er relativt høy sammenliknet med andre innvandrergrupper (Andreassen & Dzamarija, 2008). Som vi så over, kan slik tilflytting utgjøre en sentral språk- og kulturkonserverende faktor.

Vi skal kort se på gruppens *pakistanske språkbakgrunn*. Denne må regnes som kompleks; det tales en rekke ulike språk i områdene mellom India og den vestlige delen av Asia (Thiesen, 2003b), og pakistanske områder er således språkkulturelt svært mangfoldige. Urdu er Pakistans nasjonalspråk, og det dominerende språket for eksempel i pakistansk presse. Det er imidlertid bare et fåtall som har dette som morsmål; det finnes heller ikke et eget område i Pakistan hvor urdu er det dominerende talespråket (Thiesen, 2003a). Det språket som snakkes av flest personer i Pakistan (omkring to tredjedeler) er *panjabi* (Thiesen, 2003b). Til tross for at panjabi altså snakkes av en majoritet har det ingen offisiell status. Det læres heller ikke i skolene, og brukes sjelden som skriftspråk (Thiesen, 2003b). Panjabi er altså et typisk lavstatusspråk, og forbindes med mangel på utdanning og med den pakistanske landsbygda (Karlsen & Lykkenborg, 2012). Mange oppgir derfor *urdu* som sitt førstespråk selv om de snakker panjabi med hverandre (Thiesen, 2003a). En årsak til at det ikke er et større krav om å gjøre panjabi til et offisielt språk, gitt den store utbredelsen, kan ifølge Thiesen (2003b) være at språket står så nært urdu at det ikke er et veldig stort problem for de som snakker panjabi å

tilegne seg og bruke urdu som formelt språk. Urdu står i tillegg mye nærmere panjabi språklig sett enn de *øvrige* språkene som benyttes i regionen (Thiesen, 2003b). Thiesen (2003b, s. 32) beskriver den språklige forskjellen mellom urdu og panjabi på en folkelig måte slik: ”Forskjellen mellom panjabi og urdu er i samme størrelsesorden som mellom dansk og skikkelig norsk landmål, kanskje litt større.” Innenfor panjabi er det imidlertid store dialektale forskjeller (Thiesen, 2003b) – uten at vi skal gå nærmere inn på det her. Det kan altså se ut som at det både er både språklige, pragmatiske og sosiokulturelle årsaker til at urdu har den høye statusen det har – på bekostning av panjabi.

Språksituasjonen i Norge

De fleste i Norge med pakistansk bakgrunn kommer fra panjabitalende områder i Pakistan og snakker følgelig panjabi (Thiesen, 2003a). Også mange av disse oppfatter imidlertid *urdu* som sitt førstespråk (Thiesen, 2003a). Det er dessuten urdu som læres i norsk-pakistanske moskeer, og NRKs programmer for pakistanere er på urdu (Thiesen, 2003a). Statusforskjellen mellom urdu og panjabi regnes som en viktig grunn til at personer med pakistansk bakgrunn i Norge ofte oppgir urdu som sitt morsmål til tross for at de snakker mest panjabi (Kristiansen, 1980). I tillegg til denne statusforskjellen mellom urdu og panjabi er det heller ingen tvil om at pakistanske språk har lav status i det *norske* samfunnet sett under ett: Pakistanske språk brukes, i likhet med de fleste minoritetsspråk i Norge, ikke systematisk i det øvrige samfunnet. Det er for eksempel norsk som er enerådende i næringsliv og i det offentlige generelle kommunikasjon. Norsk har altså en sterkt dominerende stilling, noe som er naturlig i og med at det brukes av om lag 95 % av befolkningen (Engen & Kulbrandstad, 2004). Dette fører igjen til at det blir naturlig å bruke norsk i kommunikasjon med det overveldende flertallet av norsktalende (Engen & Kulbrandstad, 2004). I tillegg fungerer norsk som *lingua franca* (språk som brukes i kommunikasjon mellom mennesker som ikke har noe felles språk) mellom språklige minoriteter i Norge (Engen & Kulbrandstad, 2004).

En konsekvens av dette (både bakgrunnen i Pakistan og situasjonen i Norge) er at majoriteten av personer med pakistansk bakgrunn har minst tre språk å forholde seg til (Karlsen & Lykkenborg, 2012): norsk, urdu og panjabi. Studier har også vist at barn med pakistansk bakgrunn i stor grad bruker både norsk og urdu/panjabi i daglig interaksjon, med bruk av kodeveksling mellom disse språkene (Aarsæther, 2002).

Når det gjelder spørsmålet om i hvilken alder barn med pakistansk bakgrunn introduseres for norsk, hevder Engen og Kulbrandstad (2004) at disse ofte har en suksessiv tospråklig utvikling. Selj (2008) nevner ikke barn med pakistansk bakgrunn spesielt, men hevder at minoritetsspråklige i Norge ofte vil ha en typisk suksessiv tospråklig utvikling, til tross for at man ofte ser at norsk introduseres tidlig, i og med at primærsosialiseringen først og fremst foregår på det språket som brukes hjemme. Dette er også i tråd med Monsrud et al. (2010) sin forståelse: De fleste minoritetsspråklige barn i Norge følger en suksessiv tospråklig utvikling ved at de lærer hjemmespråket først, og deretter norsk i barnehage eller skole.

Vi må også se kort på kontrastive aspekter for barn med pakistansk bakgrunn – altså på forskjellen mellom språkene. Urdu og panjabi er beslektet, men forskjellen mellom dem kan allikevel regnes som ganske store, som vi så over. Forskjellene mellom urdu/panjabi og *norsk* må imidlertid helt klart regnes som betydelige: Urdu/panjabi og norsk beskrives som strukturelt svært ulike språk (Kristiansen, 1980). På alle områder (for eksempel setningsstruktur, verbbygging, preposisjoner og pronomen) beskriver Kristiansen (1980) store forskjeller. Det området som imidlertid innebærer aller størst vanskeligheter for personer med pakistansk bakgrunn i Norge, er ifølge Kristiansen (1980) ordforrådet: Grunnleggende begreper er ofte svært forskjellige – for eksempel kan ikke begrepet *å være* oversettes direkte, og verbet *å ha* finnes ikke. Dette er bare noen få eksempler, men viser at oppgaven å lære seg norsk kan by på ekstra utfordringer når forskjellene er så store. Før vi ser på studier som har undersøkt ordforrådet til barn med pakistansk bakgrunn i Norge, skal vi derfor se på en studie som belyser hvilke konsekvenser det kan ha for ordforrådsutvikling at språkene er så ulike.

Lervåg og Melby-Lervåg (2009) gjennomførte en metastudie hvor de undersøkte blant annet forskjeller i muntlige språkferdigheter mellom enspråklige og tospråklige. Muntlige språkferdigheter brukes her som en samlebetegnelse for ulike typer vokabular- eller begrepsforståelse, og blant annet lytteforståelse. Disse ferdighetene er målt på forskjellig måte i de studiene som inngår i metaanalysen, blant annet med ulike språkversjoner av BPVS. Lervåg og Melby-Lervåg (2009) analyserte data på muntlig ordforråd fra flere mindre studier hvor gjennomsnittlig alder for både enspråklige og tospråklige var 6–7 år. Metaanalysen viste for det første at tospråklige hadde langt svakere muntlige språkferdigheter på sitt andrespråk enn det de enspråklige hadde på det samme språket – i tråd med tendensen til en vokabularavstand som vi så svært tydelig over. Studien viser imidlertid også at *språklige forskjeller* mellom de to språkene kan påvirke denne vokabularavstanden. For å undersøke

dette gjorde Lervåg og Melby-Lervåg (2009) et skille mellom de tilfellene der S1 var et europeisk språk, og de tilfellene der S1 var et ikke-europeisk språk, og fant at barn med et asiatisk eller afrikansk førstespråk og europeisk/engelsk andrespråk hadde svakere muntlige andrespråksferdigheter i forhold til enspråklige barn, enn tospråklige barn med et europeisk førstespråk og et europeisk/engelsk andrespråk (Lervåg & Melby-Lervåg, 2009, s. 274). Lervåg og Melby-Lervåg (2009) foreslår flere mulige forklaringer på dette funnet. Den første kan være at strukturelt ulike språk inneholder færre språklig-strukturelle likheter (som lånord, kognater og liknende), og at det derfor er vanskeligere å overføre språkferdigheter fra det ene til det andre språket. En annen mulig forklaring er ifølge Lervåg og Melby-Lervåg (2009) at det kan være større sosiokulturelle og/eller sociolingvistiske forskjeller mellom asiatiske/afrikanske kulturer og europeiske kulturer. Dette kan gjøre det vanskeligere å integreres i kulturen og det tilhørende språket enn dersom de to språkkulturene er mer beslektet. En tredje forklaring mener Lervåg og Melby-Lervåg (2009) kan være av mer sosioøkonomisk art: Det kan være større forskjeller i arbeidstilknytning og utdanningsnivå i grupper som har mer avstand til den nye kulturen, noe som igjen kan virke inn på både mengden og kvaliteten på barnas andrespråk (Lervåg & Melby-Lervåg, 2009).

Reseptivt breddeordforråd hos barn med pakistansk bakgrunn

Vi skal se på tre studier som er gjort av det reseptive breddeordforrådet til barn med pakistansk bakgrunn i Norge; Monsrud et al. (2010), Lervåg og Aukrust (2010) og Meling (2011).

Monsrud et al. (2010) gjennomførte en normeringsstudie av BPVS for seks ulike minoritetsspråkgrupper (albansk, pakistansk, somalisk, tamil, tyrkisk og vietnamesisk) i Norge. BPVS var i forkant av denne studien oversatt til disse minoritetsspråkene, og da til urdu og panjabi for barn med pakistansk bakgrunn. Det totale normeringsutvalget bestod av 979 barn i aldersgruppen seks til tolv år, uten mistanke om spesialpedagogiske behov. Det oppgis at det var mellom 100 og 200 barn fra hver språkgruppe, men ikke nøyaktig hvor mange som hadde urdu og/eller panjabi som hjemmespråk. Det oppgis heller ikke hvor store utvalg det var på hvert alderstrinn, men gitt at det var 100-200 barn for hver språkgruppe, kan vi anta at antall barn som er testet på urdu/panjabi for hvert alderstrinn (6–12 år) *i gjennomsnitt* ligger et sted mellom 16 og 33.

Resultatene for barn med pakistansk bakgrunn skiller seg i denne studien ut som de svakeste hjemmespråksresultatene på dette alderstrinnet i sammenlikning med de andre minoritetsspråksgruppene resultater. Barn med pakistansk bakgrunn i aldersgruppen 6–7 år oppnådde i gjennomsnitt 44 på urdu/panjabi denne studien (BPVS toppskår på begge språk for alle aldre er 144). Resultatene for de andre språkene viser: albansk ($M = 66$), somalisk ($M = 59$), tyrkisk ($M = 82$) vietnamesisk ($M = 61$). Dette viser at barn med pakistansk bakgrunn skårer ganske mye lavere enn disse andre minoritetsspråksgruppene. Dersom resultatene for disse fire andre språkgruppene slås sammen, altså *ses under ett*, skårer disse i gjennomsnitt 67 i hjemmespråket.

Barna ble samtidig testet med den norske oversettelsen av BPVS, og her fikk barn med pakistansk bakgrunn i gjennomsnitt 53 riktige. Norskresultatet til barn med pakistansk bakgrunn skiller seg ikke særlig fra norskresultatene til de andre minoritetsspråksgruppene – det er verken det laveste eller det høyeste: De andre fem minoritetsspråksgruppene (albansk ($M = 58$), somalisk ($M = 42$), tamil ($M = 66$), tyrkisk ($M = 55$), vietnamesisk ($M = 63$)) skårer sett under ett i gjennomsnitt 56.8 på norsk. (Merk: Gjennomsnittsresultatene for disse andre minoritetsspråksgruppene *sett under ett* er min utregning fra opplysninger om de respektive språkernes gjennomsnitt i tabell i Monsrud et al., 2010, s. 52–53.) Standardavvik er ikke oppgitt i denne normeringsstudien, og Monsrud et al. (2010) bemerker for øvrig at aldersgruppene er forskjøvet i testingen, slik at data ikke kan brukes for å sammenlikne hjemmespråks- og norskresultater for *samme gruppe barn*, men at resultatene per aldersgruppe på henholdsvis norsk og hjemmespråk er representative for utvalget som helhet.

Monsrud et al. (2010) fant også *totalordforråd* for minoritetsspråksgruppene – som regnes ut ved å summere riktige skårer uavhengig av språk. Her er det ikke oppgitt egne totalskåre for de ulike alderstrinnene eller for de ulike språkgruppene, men resultatene viser at samlet ordforråd for hele utvalget ($N = 979$) er vesentlig høyere enn dersom språkene ses hver for seg (Monsrud et al., 2010).

Lervåg og Aukrust (2010) undersøkte blant annet reseptive ordforrådsstørrelser hos 288 andreklassinger, som en del av en studie av utvikling i leseforståelse. Utvalget bestod av 198 enspråklig norske og 90 tospråklige med urdu som hjemmespråk (det er ikke gjort noe skille mellom urdu og panjabi i denne studien). Ingen av barna hadde lærevansker ved denne studiens start, og foreldrene oppga at barna snakket både urdu og norsk hjemme, med en liten

overvekt av urdu. Barna snakket imidlertid mer norsk enn urdu både med søsken og venner. De fleste barna i studien var andre- eller tredjegerasjonsinnvandrere.

Denne studien viste at det tospråklige utvalget skårer signifikant lavere på norsk enn det enspråklige norskutvalget: Det norskspråklige utvalgets gjennomsnittresultat på en oversatt norskversjon av PPVT var 99.16 ($SD = 14.25$), mens det tospråklige utvalget oppnådde i gjennomsnitt 62.16 ($SD = 15.23$) riktige. Lervåg og Aukrust (2010) undersøkte i tillegg urduordforrådet til tospråksutvalget og fant her et gjennomsnitt på 37.77 ($SD = 18.25$). Urduresultatet for disse *andreklassingene* (90 mnd./7;5 år) hos Lervåg og Aukrust (2010) er som vi ser lavere enn det Monsrud et al. (2010) fant for barn med pakistansk bakgrunn i seks-sjuårsalder ($M = 44$). Dette kan skyldes mange forhold, for eksempel forskjeller i utvalg og utvalgsstørrelser (delvis ukjent for normeringsstudien), og/eller det at det er brukt forskjellige oversettelser av ordforrådstesten, eller det kan være et resultat av tilfeldighet.

Meling (2011) studerte i sin masteroppgave ordforrådet til fem år gamle barn med pakistansk bakgrunn ($N = 63$). Hun fant at barna skåret i gjennomsnitt 32.08 på BPVS-norsk ($SD = 11.587$), og i gjennomsnitt fikk 21.43 riktig på BPVS-urdu/panjabi ($SD = 10.077$). Her er oversettelsen den samme som i normeringsstudien, og barna er testet på enten urdu eller panjabi avhengig av hvilket språk de oppgav å bruke mest hjemme (Meling, 2011).

Som vi ser, viser alle disse studiene at det er norskordforrådet som er størst hos barn med pakistansk bakgrunn. Melings (2011) studie er for øvrig gjort på det samme utvalget som jeg bruker i min studie. Det er også den siste studien vi nå skal se på; Karlsen og Lykkenborgs (2012) studie av *hjemmespråkmiljøet* til barn med pakistansk bakgrunn.

Språkbruksmønstre i norsk-pakistanske familier

Som vi så over kan det bety mye for språkutvikling hvilke språkbruksmønstre familien har. Karlsen og Lykkenborgs (2012) studie er en detaljrik undersøkelse av språkbakgrunn og språkbruk i familier med pakistansk bakgrunn, og jeg skal bare kort referere noen hovedfunn her.

Karlsen og Lykkenborg (2012) undersøkte språkbruk i 67 familier med pakistansk bakgrunn. Flesteparten av foreldrene var førstegerasjonsinnvandrere, men noen (15 % av mødrene og 7–8 % av fedrene) var også født i Norge (Karlsen & Lykkenborg, 2012). Dette betyr at barna i familiene hovedsakelig er andregenerasjonsinnvandrere, men også at noen er

tredjegerasjonsinnvandrere. De aller fleste av disse foreldrene vurderte et eller begge av de pakistanske språkene som sitt morsmål.

Karlsen og Lykkenborg (2012) identifiserte tre hovedgrupper med hensyn til selve *bruken* av norsk, urdu og panjabi i familiene: Familier som snakket mest urdu, familier som snakket både urdu og norsk, og familier som snakket både panjabi, urdu og norsk. Dette viser at gruppen totalt sett har en sammensatt bruk av flere språk i hjemmet: Pakistanske språk brukes i høy grad, men samtidig brukes også norsk av mange. Panjabi rapporteres imidlertid totalt sett brukt av en noe mindre gruppe (Karlsen & Lykkenborg, 2012).

Karlsen og Lykkenborg (2012) undersøkte også hvilke språkaktiviteter som gjøres med barna på enten norsk, urdu eller panjabi. Her viser det seg for det første at de fleste foreldre gjør ulike språklige aktiviteter daglig med barna – dette kan for eksempel være at de leser med barna, forteller eventyr eller leker. De fleste foreldrene rapporterer bruk av enten urdu eller norsk i disse aktivitetene med barna, mens panjabi bare rapporteres brukt av en mindre gruppe (to til sju foreldre avhengig av aktivitet). Når det gjelder fordelingen av hvilke språk som brukes i de ulike aktivitetene, fant Karlsen og Lykkenborg (2012) at 66 % av foreldrene leser daglig eller ukentlig for barna på norsk, mens 37 % leser daglig eller ukentlig for barna på urdu. Det er videre omtrent like mange som oppgir å bruke norsk som urdu når de forteller eventyr, mens over halvparten av foreldrene leker daglig med barna på norsk, og en enda større andel leker med barna daglig på urdu.

Dette viser at urdu og norsk brukes i omtrent like stor grad i språklige aktiviteter med barna i familier med pakistansk bakgrunn. Det viser også at selv om de fleste foreldrene anser et eller begge av de pakistanske språkene som sitt morsmål, brukes allikevel like ofte norsk i språklig samhandling med barna. Videre, gitt at de fleste norskpakistanere kommer fra panjabitalende områder i Pakistan og at *panjabi* derfor skulle kunne regnes som de fleste av foreldrenes morsmål (Thiesen, 2003b), kan det også se ut til at det bare er et fåtall som oppgir å bruke sitt faktiske morsmål mest i språklige aktiviteter med barna, men heller velger urdu eller norsk. Karlsen og Lykkenborg (2012) påpeker imidlertid at gruppens lange botid i Norge kan påvirke norsk- og urdukunnskapen til foreldrene, slik at det er uklart hvilket språk som er deres sterkeste – det er ikke nødvendigvis noe sammenfall mellom hvilket språk foreldrene oppfatter som sitt morsmål og hvilket språk som er deres sterkeste. Man kan heller ikke se bort ifra at det også kan være en underrapportering av bruk av panjabi, gitt språkets lavere status (Karlsen & Lykkenborg, 2012).

4 Metode

I arbeidet med denne masterstudien har jeg som forskningsassistent i forskergruppen *Child Language and Learning* (CLL) ved Institutt for spesialpedagogikk, Universitetet i Oslo (Institutt for spesialpedagogikk, 2010a) bidratt til å samle inn data for to av denne forskergruppens utvalg: CLLs kohort og utvalget til *Language development in minority language children* (Institutt for spesialpedagogikk, 2010b). Jeg har gjennom dette også fått anledning til å bruke data fra disse to store utvalgene i min egen masteroppgave og samtidig verdifull metodisk erfaring med innsamling og registrering av data. Jeg har stått fritt til å velge problemstilling og design for min egen masterstudie, og skal nå først begrunne de valgene jeg har gjort med hensyn til design. Jeg skal videre beskrive de to utvalgene, instrumentene som er brukt (BPVS for kartlegging av reseptivt breddeordforråd og spørreskjema for enkelte bakgrunnsvariabler), samt fremgangsmåten for datainnsamlingen og hvilke statistiske analysemetoder jeg velger. Til slutt skal jeg drøfte prosjektets reliabilitet, og vurdere sentrale etiske problemer.

4.1 Design

Jeg ønsker å finne informasjon om store gruppers egenskaper på få variabler og deretter referere resultatene på en måte som synliggjør hovedtendenser, og velger derfor en *kvantitativ* og deskriptiv tilnærming til analyser av *reseptivt breddeordforråd*. En kvantitativ tilnærming er også en naturlig følge av at *CLL* og *Language development in minority language children* samler inn data for reseptivt breddeordforråd som egner seg for en slik tilnærming.

Videre velger jeg å søke en forståelse for disse resultatene i barnas språksituasjon på en ”ikke-statistisk” måte. Dette er fordi jeg ønsker å belyse hvordan mange ulike og samvirkende aspekter *til sammen* kan gi en forståelse av tospråklig kompetanse på en måte som vanskelig lar seg vise gjennom statistisk analyser. Flere forhold som jeg ønsker å se til for en forståelse av resultatene, lar seg heller ikke enkelt tallfeste. Et eksempel kan være samfunnsaspekter som kan påvirke språkutvikling på en indirekte måte. Et naturlig valg for å forstå resultatene er derfor å se til en bred kunnskapsbasis om barnas språksituasjon – redegjort for over ved hjelp av teori og tidligere empiri.

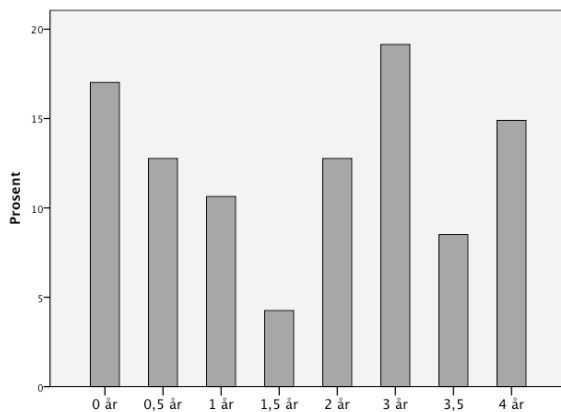
4.2 Utvalg og populasjoner

Tospråklig utvalg

Det tospråklige utvalget i dette masterprosjektet er altså hentet fra *Language development in minority language children*. Dette er en pågående longitudinell studie ved Institutt for spesialpedagogikk ved Universitetet i Oslo som undersøker språkutvikling hos barn med urdu og/eller panjabi som hjemmespråk. Utvalgskriteriene for deltakelse i den longitudinelle studien var at barnet var født i 2005, hadde gått i barnehage det siste året før skolestart, hadde urdu og/eller panjabi som hjemmespråk, og ikke hadde kjente språkvansker eller utviklingsforstyrrelser. Utvalget ble gjort ved at et stort antall barnehager i ulike bydeler i Oslo ble kontaktet. I de barnehagene som ønsket å bidra til studien, ble foreldre til barn fra den aktuelle språkgruppen kontaktet via barnehagen med forespørsel om deltakelse. Utvalget er altså ikke gjort tilfeldig, men basert på bekvemmelighet ved at de som meldte sin interesse og samtidig oppfylte utvalgskriteriene, ble valgt ut til å delta. Det kan selvsagt ikke utelukkes at det kan være systematiske forskjeller mellom barnehager som valgte å delta og barnehager som ikke valgte å delta, og heller ikke at det kan være forskjeller mellom hvilke foreldre som valgte og ikke valgte å delta. Slike forskjeller kan imidlertid samtidig ikke unngås, selv ved tilfeldig utvalg, ettersom deltakelse er basert på frivillighet. Når det er sagt kan vi anta at det likevel er en viss "ikke-statistisk tilfeldighet" i flere ledd av utvelgelsen; for eksempel er ikke barna funnet gjennom relaterte miljøer eller bestemte lag i populasjonen. Barna som nå deltar er også spredt over flere områder i Oslo by og omegn.

Det er noe frafall av deltakere siden starten av studien *Language development in minority language children*. Av de 67 som bekreftet deltakelse i utgangspunktet, har det på det nåværende tidspunkt (2012) vært mulig å innhente data fra 61 barn. Slikt frafall er naturlig og kan ha mange årsaker – blant annet kan det være vanskelig å "finne igjen" noen av barna når de har begynt på skolen. Utvalget består nå av 61 barn som har gjennomført BPVS på norsk – 60 av disse har også gjennomført BPVS på enten urdu (39) eller panjabi (21). Utvalgets gjennomsnittsalder på testtidspunktet er 77.8 måneder, det vil si ca. seks og et halvt år. Utvalget er kjønnsbalansert, med 48.1 % gutter og 51.9 % jenter. Populasjonen det eventuelt kan generaliseres til i denne studien, kan dermed beskrives som *populasjonen av seks år gamle barn med pakistansk bakgrunn bosatt i Oslo og omegn som har gått i barnehage (minst) ett år før skolestart, og som ikke har kjente avvik fra normalutvikling*.

Utvalgets språksituasjon er i stor grad redegjort for over. Jeg ønsker i tillegg til dette å vise utvalgets fordeling over en sentral språklig bakgrunnsvariabel. *Figur 1* gir en oversikt over når barnet ble introdusert for norsk, og viser en spredning på alle verdier fra 0 til 4 år. Den høyeste prosentandelen finner vi på 3 år. Variabelen kan også re-grupperes etter distinksjonen mellom simultan tospråklighet og andrespråklæring som redegjort for over, altså til om barnet ble introdusert for norsk *før* eller *etter* treårsalder. En slik gruppering viser at 56.6 % ble introdusert for norsk *før* de var tre år, og at 43.4 % ble introdusert for norsk *etter*.



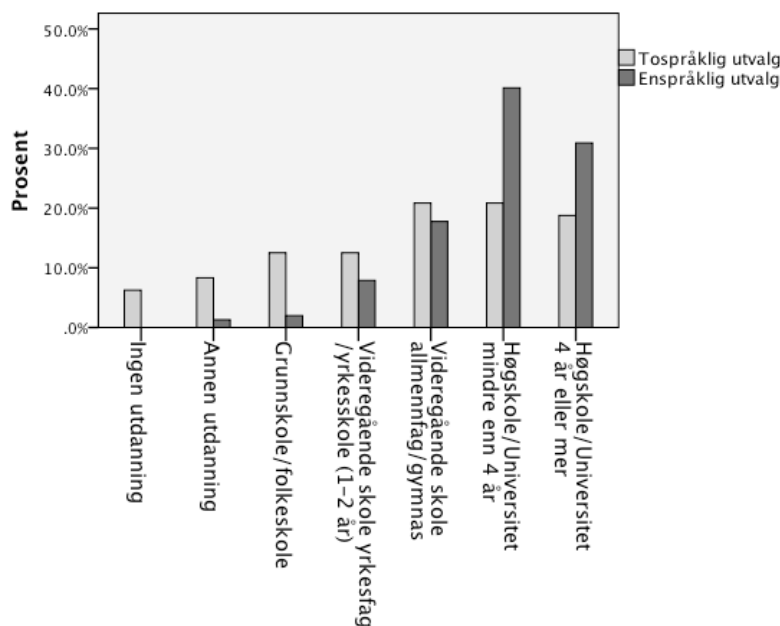
Figur 1. Hvor gammelt var barnet da det ble introdusert for norsk språk?

Enspråklig utvalg

For å kunne sammenlikne det tospråklige utvalgets gjennomsnittsskårer på BPVS-norsk med et enspråklig utvalg bruker jeg data fra kohorten til forskergruppen *Child Language and Learning*. Denne kohorten består av 190 norskspråklige barn, uten kjente avvik fra normalutvikling. Utvalget er gjort tilnærmet statistisk tilfeldig i en østlandskommune som regnes som en gjennomsnittskommune i landssammenheng med hensyn til sosioøkonomiske forhold. Data fra det enspråklige utvalget er fra 2010, altså fra to år før data fra tospråksutvalget, da kohorten var på samme alderstrinn som tospråksutvalget: seks år / første klasse.

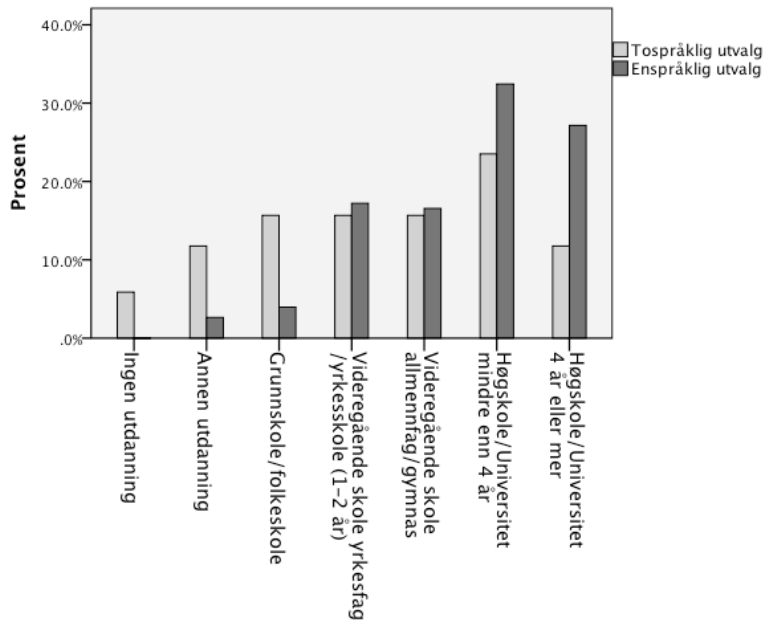
Sosioøkonomiske bakgrunnsvariabler for begge utvalg

For å kunne forstå deskriptive resultater, er det viktig å ha god kjennskap til utvalgenes bakgrunn på antatt relevante områder. I tillegg skal vi derfor også se på enkelte sosioøkonomiske bakgrunnsvariabler – nå for begge utvalg. Her er kasus som verken inneholdt resultat på BPVS norsk eller BPVS urdu/panjabi ekskludert, slik at utvalget korresponderer med mitt prosjektets utvalg (N = 61, N = 190). Variablene jeg har valgt å bruke, er *mors høyeste utdanning*, *fars høyeste utdanning* og *antall billedbøker for barn i hjemmet*.



Figur 2. Mors høyeste utdanning

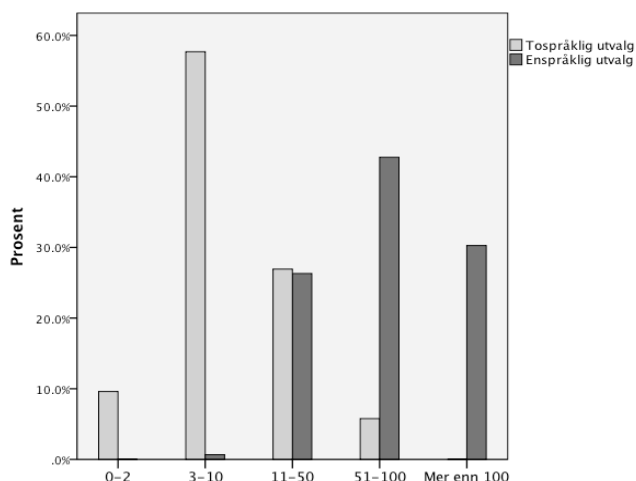
Mors og fars utdanning. Som det kommer frem av *figur 2 og 3*, er det en generell tendens til at foreldrene i det tospråklige utvalget har lavere utdanning enn foreldrene i det enspråklige utvalget. Prosentfrekvenser i *figur 2* viser at 18.8 % av mødrene i det tospråklige utvalget oppgir å ha ingen utdanning eller grunnskole/folkeskole, mot 2 % i det enspråklige utvalget. Andelen mødre i det tospråklige utvalget som har høyere utdannelse (universitet/høgskole – mindre enn 4 år eller 4 år og mer) er til sammen 39,6 %, mens den i det enspråklige utvalget er 71 %. Det er med andre ord vesentlige forskjeller mellom utvalgene på denne variabelen.



Figur 3. Fars høyeste utdanning

For *fedrene* (figur 3) ser vi den samme tendensen til noe lavere utdanningsnivå i det tospråklige utvalget, men den er ikke like sterk som for mødrene. I det tospråklige utvalget er det til sammen 21.6 % som oppgir ingen utdanning eller grunnskole/folkeskole. Den tilsvarende andelen i det enspråklige utvalget er 4 %. Det er til sammen 35.3 % av fedrene i det tospråklige utvalget som oppgir å ha høyere utdanning (universitet/høyskole – mindre enn 4 år eller 4 år og mer), mens 59.7 % av fedrene i det enspråklige utvalget oppgir det samme.

Antall billedbøker for barn i hjemmet. Her (figur 4) ser vi en sterk tendens til at foreldre i det tospråklige utvalget oppgir å ha færre billedbøker for barn i hjemmet enn foreldrene i det enspråklige utvalget. Det er 67.3 % i det *tospråklige* utvalget som oppgir å ha inntil ti bøker for barn, altså et relativt lavt antall bøker. Tilsvarende tall for det *enspråklige* utvalget er 0.7 %. Ingen i det tospråklige utvalget har oppgitt å ha mer enn 100 bøker, mens 30 % i det enspråklige utvalget oppgir dette. Forskjellene vi ser på disse sosioøkonomiske bakgrunnsvariablene må sies å være ganske tydelige, og for å undersøke dette statistisk vil forskjellene mellom dem bli signifikanstestet i resultatdelen av studien.



Figur 4. Antall billedbøker for barn i hjemmet

4.3 Instrumenter

Som hovedinstrument benytter jeg BPVS for å kartlegge barnas reseptive breddeordforråd, og skal nå kort beskrive denne testen. I tillegg skal jeg informere om spørreskjemaet som bakgrunnsvariablene er hentet fra.

BPVS-norsk og BPVS-urdu/panjabi

BPVS II (heretter BPVS) er utformet på grunnlag av den reviderte amerikanske Peabody Picture Vocabulary Test (PPVT-R) fra 1981 og er en britisk ordforrådstest som måler reseptivt breddevokabular (Dunn, Dunn, Whetton & Burley, 1997). I dette prosjektet bruker jeg en norsk oversettelse av denne (BPVS-norsk), og to oversettelser for pakistanske språk i Norge (BPVS-urdu og BPVS-panjabi).

Lyster, Horn & Rygvold (2010) har oversatt og standardisert den norske versjonen (se også Institutt for spesialpedagogikk, 2010c). Testen er organisert i flere blokker à 12 bildesett, og hver blokk har stigende vanskelighetsgrad. Ett bildesett består av fire svart/hvitt-tegninger av objekter, handlinger eller konsepter. Ett og ett slikt bildesett presenteres for barnet, testlederen sier et angitt ord, og barnet blir bedt om å peke på det av de fire bildene som det synes passer best til ordet (se vedlegg 1 og 2 for instruksjoner og et eksempel fra manualen). BPVS har for øvrig en nedre grense på 11 riktige ord innenfor en blokk à 12 bildesett. Denne nedre grensen

må etableres; det vil si at man går tilbake i oppgavesett helt til barnet klarer 11 av 12 riktige, og deretter fortsetter testingen blokk for blokk, helt til barnet får åtte feil innenfor en blokk.

BPVS er også oversatt og normert for tospråklige med minoritetsbakgrunn i Norge (Monsrud et al., 2010), deriblant som nevnt for pakistanske språk – urdu og panjabi. Testen er bygget opp og gjennomføres på samme måte som BPVS-norsk. I tilpasningen til minoritetsspråk ble det forsøkt å ha en så lik tilnærming som mulig til enten den engelske eller den norske originalversjonen (Monsrud et al., 2010). Som oftest var det mulig å finne ord som hadde likt meningsinnhold, men i noen tilfeller ble vanskelighetsgraden ulik det norske eller engelske utgangspunktet. Det er for eksempel stor variasjon i hvor høyfrekvente ord er på de ulike språkene, og dette kan bidra til at vanskelighetsgraden kan bli ulik på tvers av språkene (Monsrud et al., 2010).

Spørreskjema

For å samle inn bakgrunnsvariablene til studien er det benyttet spørreskjema til foreldrene (både for tospråks- og enspråksutvalget). Innsamlingen av disse dataene ble gjort av forskergruppene på et tidligere tidspunkt (henholdsvis 2010 og 2008), altså da barna var fem år. Spørreskjemaet til det tospråklige utvalget ble for øvrig utdelt både på norsk og urdu og foreldrene kunne selv velge hvilket av språkene de ønsket å svare på. Spørreskjemaet er omfattende og inneholder spørsmål om forskjellige områder relatert til barnets tidlige språkutvikling, barnets og familiens språkbruk, samt ulike bakgrunnsfaktorer. Det at spørreskjemaet er omfattende (elleve sider), kan bidra til at svarprosenten for enkelte av spørsmålene er noe lav (varierende mellom 77 % og 85 % for alle variablene jeg bruker). Noen spørsmål kan også være av sensitiv art, og kan være unngått besvart av den grunn. Svarprosenten må allikevel regnes som akseptabel.

4.4 Fremgangsmåte

BPVS inngår som en av flere tester i forskergruppens testbatteri, og brukes både i det enspråklige og i det tospråklige utvalget. Selve testingen ble gjennomført av forskningsledere og flere forskningsassistenter. Assistentene ble kurset i hvordan de enkelte testene skulle administreres og skåres, og hvordan en legger til rette for at barn skal få en god opplevelse av testingen. Barna fikk god muntlig informasjon om testsituasjonen og i tillegg et oversiktsark

som de kunne krysse av på når de var ferdig med en oppgave, slik at de kunne se hvor mange oppgaver som gjensto. Som belønning for innsatsen fikk de klistremerker etter hver testøkt og en premie etter siste gang.

Barna ble testet på skolen i skoletiden. BPVS ble administrert først i disse testbatteriene og tar ca. 10–15 minutter. For det tospråklige utvalget ble BPVS-norsk gjennomført på første testdag og BPVS-urdu/panjabi andre testdag. Det var minst en ukes opphold mellom de to testdagene. Barna fikk spørsmål om hvilket språk de snakker *mest* hjemme (”Hvilket språk snakker du mest hjemme?”), og kunne velge mellom urdu eller panjabi. Barnas svar ble valgt som testspråk for BPVS-urdu/panjabi. BPVS-norsk ble gjennomført ved at testleder leser opp instruksjoner og testord i henhold til testmanualen, mens BPVS-urdu/panjabi ble gjennomført ved at testinstruksjon og testord ble spilt av fra bånd – lest inn av morsmålsbrukere av de respektive språkene. I begge tilfeller registrerte testleder nøyaktig hvilke svar barna gav på hvert testord/item.

4.5 Statistiske analyser

Dataprogrammet ”Statistical Packages for the Social Sciences” (SPSS) brukes til å registrere, systematisere og analysere datamaterialet. Jeg vil gjøre deskriptive analyser som summerer opp og viser mønstre i utvalgene (De Vaus 2002), og gjennomgå sentraltendens, spredning, skjevhet og kurtosis for de ulike fordelingene. Hensikten med dette er blant annet å undersøke om gjennomsnittsmålene, som jeg planlegger å basere slutningsstatistikk og den påfølgende drøftingen av resultatene på, er gode mål for sentraltendensen i utvalgene. Jeg vil også undersøke deskriptive mål for undergrupper i tospråksutvalget basert på en persentil-inndeling som blir beskrevet nærmere i forbindelse med presentasjon av resultatene. Hensikten med dette er først og fremst å tydeliggjøre og å kunne differensiere nivåforskjeller innad i tospråksutvalget. I tillegg til deskriptive analyser vil jeg bruke *analytisk statistikk* – også kalt slutningsstatistikk. Hensikten med dette er å se om utvalgsmønstre (altså forskjeller i ordforrådsstørrelser) kan sluttes, eller generaliseres, til populasjonene i studien (De Vaus, 2002).

4.6 Validitet og reliabilitet

Validitet dreier seg på et overordnet nivå for det første om *hvilke* slutninger som kan trekkes i en studie – og for det andre om *gyldighetsgraden* av disse slutningene. Lund (2002) påpeker i denne forbindelse at vi aldri kan oppnå helt sikre slutninger – i beste fall kan vi oppnå en høy sikkerhet i slutninger vi gjør. Som et utgangspunkt for vurdering av validiteten i dette prosjektet velger jeg å benytte det velkjente validitetssystemet til Cook og Campbell slik dette presenteres i Lund (Cook & Campbell, 1979, ref. i Lund, 2002). Systemet omfatter fire ulike typer validitet; statistisk validitet, begrepsvaliditet, indre validitet og ytre validitet; og gir derfor en god struktur til å drøfte gyldigheten av ulike typer slutninger. Systemet er utarbeidet for kausal forskning, men store deler av det er også relevant for ikke-kausale studier (Lund, 2002), slik som denne studien. Jeg skal presentere disse ulike validitetstypene nærmere og drøfte resultatenes gyldighet i lys av dem etter presentasjonen av resultatene.

Reliabilitetsspørsmål omhandler pålitelighet i selve målingene vi gjør; gir instrumentene vi velger, og måten data innhentes på, pålitelige data? Her spørres det altså ikke etter *hva* som måles, men bare hvorvidt målingen i seg selv er nøyaktig og pålitelig (Kleven, 2002a). God reliabilitet blir dermed også åpenbart et viktig fundament for å vurdere gyldigheten (validiteten) for øvrig, i den forstand at upålitelige målinger svekker grunnlaget for å kunne trekke gyldige slutninger. Reliabiliteten til en måling kan styrkes blant annet ved å redusere muligheten for at tilfeldige målefeil påvirker resultatene (Kleven, 2002a). Dette kan oppnås for eksempel ved å bruke standardiserte metoder for innsamling av data med klare retningslinjer for måling og registrering av data, og ved at de som gjør denne jobben får tilstrekkelig informasjon og trening slik at ulike metoder for innsamling og skåring av data ikke forekommer.

Slike reliabilitetstrusler er i stor grad unngått i denne studien: BPVS regnes for det første som en test som både er enkel å administrere og skåre, og det er derfor regnet som usannsynlig at individuelle forskjeller i datainnsamling og registrering forekommer (Bialystok et al., 2010). Videre fikk alle som skulle teste barn for forskergruppene god informasjon om prosedyrer for gjennomføring og skåring. I tillegg styrkes reliabiliteten av at registrering av skårer er gjort på en nøyaktig måte: Skårer ble registrert som enten feil eller riktig på hvert item, samt med en manuelt utregnet sumskår for hvert barn. Dette gjør det mulig å sjekke om verdien *sumskår*, som brukes i videre analyser, stemmer med summen av enkeltitem. Verdiene ble kontrollert

av den enkelte testleder når dataene ble lagt inn, samt av forskningsleder i etterkant – tastefeil eller summeringsfeil som ble oppdaget ble kontrollert mot protokoller og rettet opp. Til sammen bidrar slike sikkerhetstiltak mot prosedyrefeil til å gi resultatene i denne studien høy reliabilitet. I tillegg viser BPVS høy inter-reliabilitet på item-nivå, slik dette måles med Cronbachs Alpha. Resultatene av denne analysen blir presentert sammen med de øvrige resultatene.

4.7 Etiske vurderinger

I et etisk perspektiv vil det med hensyn til enkeltpersoner som deltar i studien være avgjørende at det gis god informasjon om studien, slik at foreldre på et kvalifisert grunnlag kan vurdere om de ønsker at deres barn skal delta. De må ikke bare informeres om at anonymitet og konfidensialitet ivaretas, men også om retten til å *ikke* delta og retten til å kunne trekke seg uten begrunnelse på et hvilket som helst tidspunkt i prosessen (De Vaus, 2002). For denne studien er disse kravene ivaretatt av de overordnede forskergruppene: Tillatelse fra Personvernombudet for forskning (ved Norsk samfunnsvitenskapelig datatjeneste, NSD), samt informert samtykke fra foreldrene, foreligger. Når foreldre samtykker på barns vegne, hviler det imidlertid et ekstra ansvar på forskere for å sikre at også barnas interesser ivaretas i forskningen (NESH, 2009, 19. mai). Dette er forsøkt ivaretatt gjennom å gjøre barnet oppmerksom på at det er frivillig å delta, samt ved å tilrettelegge testsituasjonen på en aldersadekvat og individsensitiv måte slik at deltakelsen oppleves som positiv for det enkelte barn.

En annen sentral etisk utfordring vil være å vurdere den skaden som eventuelt kan påføres barn som deltar (De Vaus, 2002). Her må særlig et *metodisk* moment i dette prosjektet vurderes kritisk: Undersøkelsen innebærer at barn hentes ut av den normale klassesituasjonen, noe som i prinsippet innebærer en viss risiko for stigmatisering. En risiko for slik stigmatisering bør veies opp mot fordelene som kan oppnås ved å gjennomføre undersøkelsen. Her kan det bemerkes at flere ting tyder på at barna *selv* opplevde deltakelsen som positiv: Alle testledere rapporterer om en svært positiv innstilling fra barna til å delta, også andre testdag, noe som kan være et tegn på at oppgavene ble oppfattet som spennende, og at barna ikke oppfattet det som negativt å bli tatt ut av klassen. Det gjennomgående inntrykket er at barna som var med i undersøkelsen, hadde en god opplevelse – mange fortalte at de hadde vist klassen oppgaver de hadde gjort og klistremerkene de hadde fått.

Til slutt vil jeg trekke frem sentrale problemer knyttet til selve fortolkningen og formidlingen av resultatene. Først kan det stilles spørsmål ved at barn fremstilles som en *gruppe*. Dette innebærer åpenbart en generalisering som skjuler et mangfold. Slik generalisering åpner i tillegg også for potensiell stigmatisering (NESH, 2009, 20. mai). Dette er uheldige, men mulige, konsekvenser som det er viktig å forebygge. Sett i et annet perspektiv bør man imidlertid heller ikke la være å samle informasjon om ulike demografiske gruppers kompetanse: ”Hensynet til den enkeltes integritet eller til en hel gruppes selvbylde kan [...] føre til at verdifull kunnskap ikke blir formidlet”, skriver Dalen (2011, s. 22). Her er det altså motstridende hensyn som aktualiseres. Dersom forskning skal gjennomføres på gruppenivå, bør det således særlig reflekteres over hvilke innfallsvinkler som velges for fortolkning av materialet. Bevisste valg må også tas med hensyn til hvordan gruppene fremstilles, og hvordan resultater formidles: Som forsker må man være oppmerksom på at slike valg kan få konsekvenser for oppfatningen av sårbare grupper i samfunnet og i utdanningssystemet. I denne studien har jeg valgt som hovedperspektiv å tolke resultater på en måte som er sensitiv til tospråklig kompetanse som noe *kvalitativt annet* enn enspråklig kompetanse. Denne innfallsvinkelen korresponderer med det Grosjean (1985) kaller et tospråklig betraktningsspunkt på tospråklighet (*the bilingual view on bilingualism*). Dette mener jeg kan bidra til at sårbare gruppers integritet ivaretas, fordi tospråklighet gjennom dette perspektivet ses som en naturlig utviklet lingvistisk kompetanse som er kvalitativt annerledes, men ikke mindre verdifull, enn enspråklig kompetanse (Grosjean, 1985). Når et slikt grunnleggende perspektiv er etablert, er det samtidig et viktig ansvar å belyse gruppeforskjeller som kan ha konsekvenser for allmennpedagogisk tilrettelegging. Dette kan bidra til at utdanningsbetingelser for hele grupper bedres, og er derfor et ansvar som må tas på alvor innenfor pedagogisk forskning.

5 Resultater

Når jeg nå skal presentere resultatene i studien, er det viktig å innledningsvis påpeke at disse er beregnet med *råskårer* fra BPVS-norsk og BPVS-urdu/panjabi. Gjennomsnittsverdiene som oppgis er altså gjennomsnittlig antall riktige ord av totalt 144 mulige. Det er imidlertid ikke forventet at barn på seks år skal oppnå maksskåre, da testen er normert for bruk opp til 16 år.

Vi skal først se på utvalgenes ordforrådsstørrelser – først på utvalgsnivå og deretter for grupperinger av tospråksutvalget. Videre skal vi se på signifikanstester av gjennomsnittsforskjellene mellom BPVS-norsk for enspråks- og tospråksutvalget, for BPVS-norsk og BPVS-urdu/panjabi innad i tospråksutvalget, samt av forskjellen i bakgrunnsvariablene. Til slutt skal jeg vise effektstørrelser for utvalgte gjennomsnittsforskjeller. Merk at *BPVS-total* er beregnet ved å legge sammen antall riktige uavhengig av språk: Dersom ordet var kjent på ett eller begge språk, fikk barnet ett poeng. Dersom ordet ikke var kjent på noen av språkene fikk barnet null poeng. BPVS-total viser altså antall ord som er kjent i enten det ene eller begge språkene.

5.1 Fordelinger

Siden variablene er på intervallnivå, kan vi bruke mål som gjennomsnitt (mean/M) og standardavvik (standard deviation/SD) for å beskrive fordelingene (De Vaus, 2002). *Gjennomsnittsverdier* gir verdifull informasjon om utvalgenes sentraltendens, og er et av de vanligste målene for dette (De Vaus, 2002). *Tabell 1* viser fordelingene for tospråksutvalgets og enspråksutvalgets skårer på BPVS. Vi ser for det første at tospråksutvalgets BPVS-urdu/panjabi fremstår som svært lav sammenliknet med både tospråksutvalgets BPVS-norsk og enspråksutvalgets BPVS-norsk. Tospråksutvalgets BPVS-norsk ligger også betydelig under det *enspråklige* utvalgets gjennomsnitt. Det er også en forskjell i gjennomsnittsskår for de barna som har gjennomført BPVS-urdu og de som har gjennomført BPVS-panjabi, men denne er ikke vesentlig i forhold til de andre forskjellene. Vi kan merke oss at også BPVS-total ligger betydelig under BPVS-norsk for enspråklige.

Tabell 1. Resultater BPVS (råskårer – maks 144)

	N	Min	Max	M	SD	Skew	Krt	Cr. Alpha
Tospråklig utvalg								
BPVS Norsk	61	26	81	51.15	13.141	.313	-.635	.929
BPVS Urdu/Panjabi	60	10	76	28.62	12.188	1.351	3.087	.932
BPVS Urdu	39	10	76	30.13	13.407	1.345	2.673	.943
BPVS Panjabi	21	14	45	25.81	9.163	.445	-.428	.894
BPVS Total	60	33	87	55.27	12.998	.317	-.659	.937
BPVS enspråklig utvalg	190	43	104	73.39	11.945	.173	-.497	.914

Som vi ser gir disse gjennomsnittsmålene allerede en god indikasjon på forskjeller mellom utvalgene, og innad i tospråksutvalget. Som uttrykk for sentraltendens kan imidlertid et gjennomsnitt forstyrres av ekstremt høye eller lave skårer, og i tillegg kan ulike fordelinger også få det samme gjennomsnittsuttrykket (De Vaus, 2002). Sagt på en annen måte: Et gjennomsnitt sier lite om hvordan skårene *fordeler* seg rundt gjennomsnittet. Vi er derfor nødt til å se på flere andre mål for å se om mean er et godt uttrykk for fordelingenes sentraltendens.

For å få mer informasjon om resultatenes fordeling rundt gjennomsnittet, kan vi se på ulike *spredningsmål* – for eksempel standardavvik (Standard Deviation/SD). Standardavvik er et uttrykk for utvalgets gjennomsnittlige avvik fra mean (Gall, Gall & Borg, 2007). Jo høyere standardavviket er, desto dårligere uttrykker gjennomsnittsverdien fordelingen sentraltendens (De Vaus, 2002). Vi ser av tabellen at standardavvikene er noe høyere for tospråksutvalget (med unntak av BPVS-panjabi) enn for enspråksutvalget. Dette betyr at det er en noe større spredning av resultatene for tospråksutvalget. Forskjellene er imidlertid ikke betydelige, og SD gir dermed jevnt over en god indikasjon på at mean uttrykker sentraltendensen på en akseptabel måte. *Range* (uttrykt som minimums- og maksimumsverdier) gir også et inntrykk av spredningen i en skåredistribusjon, men på en litt annen måte enn SD (Gall et al., 2007). Her ser vi for eksempel at BPVS-panjabi har en mye lavere range enn de andre fordelingene, noe som også avspeiles i BPVS-panjabi sitt SD. Vi ser også at range for det enspråklige

utvalget starter og slutter betraktelig høyere enn for BPVS-total og BPVS-norsk for tospråklige, noe som også underbygger at sentraltendensen uttrykkes godt gjennom mean.

Skjevhet (skewness/Skew) er et mål for hvor symmetrisk eller jevn fordelingen er. Dersom skjevheten er 0, er resultatene symmetrisk fordelt over og under gjennomsnittet (De Vaus, 2002). Dersom skjevheten er *positiv*, vil fordelingen ha en overvekt *under* gjennomsnittet, og ved *negativ* skjevhet vil fordelingen ha en overvekt *over* gjennomsnittet (De Vaus, 2002). Alle resultatene er her positive – dette betyr at resultatene har en overvekt *under* sitt gjennomsnitt. Men i hvor stor grad? Det kan være vanskelig å avgjøre hvor skjev en fordeling er ved bare å se på skjevhetsmålet i seg selv – det finnes i prinsippet ingen øvre grense på skjevhetsskalaen, og alle fordelinger som ikke har en skjevhetsverdi på 0 er per definisjon skjeve (De Vaus, 2002). Christophersen (2012) foreslår at skjevheter på ± 2 må regnes som store avvik. En annen tommelfingerregel er at en skjevhet som er dobbelt så stor som skjevhetens *standardfeil* (standardfeil til skjevhet), må regnes som skjev (De Vaus, 2002). Med en standardfeil til skjevhet på .306 (ikke oppgitt i tabellen) ser vi at tospråksutvalgets fordeling på BPVS-norsk (Skew .313) etter denne tommelfingerregelen ikke kan regnes som spesielt skjev. Enspråksutvalgets fordeling kan heller ikke regnes som skjev (Standardfeil til skjevhet = .176), og heller ikke BPVS-total (Standardfeil til skjevhet = .309). BPVS-urdu/panjabi (Skew 1.351) må imidlertid regnes som tydelig høyreskjev med en standardfeil til skjevhet på .309. I henhold til Christophersen (2012) er imidlertid ikke avviket fra normalfordeling helt uakseptabelt.

Kurtosis (Krt) er et mål for i hvor stor grad fordelingen er enten spiss eller flat – det vil si om skårene spisser seg rundt visse verdier eller om de fordeler seg flatere utover verdiskalaen. En normalfordeling er verken spiss eller flat, og ville få en kurtosis på 0 (og i tillegg en skjevhet på 0) (De Vaus, 2002). Spisse fordelinger vil få en positiv kurtosis, mens flate vil få en negativ (De Vaus, 2002). Vi ser svakt negative, altså noe flate fordelinger for tospråksutvalgets BPVS-norsk, BPVS-panjabi, BPVS-total og enspråksutvalgets BPVS – fordelingene flater altså noe mer ut enn en normalfordeling. BPVS-urdu/panjabi og BPVS-urdu har positive kurtoser, noe som indikerer relativt spisse fordelinger. Dette innebærer at skårefordelingen her er noe mer sentrert rundt visse skårer enn de ville vært i en jevnere normalfordeling.

Oppsummerende kan vi si at fordelingenenes gjennomsnitt i stor grad ser ut til å avspeile sentraltendensene på en akseptabel måte. Vi skal nå videre se nærmere på nivåforskjeller mellom undergrupper i tospråksutvalget.

Fordelinger for grupperinger av tospråksutvalget

For å få et tydeligere bilde av nivåforskjeller i tospråksutvalget velger jeg å gruppere fordelingene med cut-points på 33.33-persentilen og 66.66-persentilen. Dette betyr at barna deles inn etter om resultatene deres ligger under, mellom eller over de verdiene som kutter de respektive fordelingene i tre tilnærmet like store deler. Disse grupperingene avspeiler dermed nivåforskjeller mellom lavtskårende, middelsskårende og høyt-skårende barn. Inndelingen gir ni grupperinger (*tabell 2*): (1) Barn som skårer under 33.33-persentilen på BPVS-norsk, (2) mellom 33.33- og 66.66-persentilen på BPVS-norsk og (3) over 66.66-persentilen på BPVS-norsk – og så videre på samme måte for BPVS-urdu/panjabi og BPVS-total. Dette betyr for eksempel at grupperingen *BPVS-norsk Lav (26–43)* utgjøres av de barna som skåret under 33.33-persentilen på BPVS-norsk.

Tabell 2. Resultater BPVS-norsk og BPVS-urdu/panjabi for grupperinger. Range oppgitt i parentes.

Grupperinger	N	M	SD	Skew	Krt
BPVS-norsk Lav (26–43)	21	37.43	4.812	-.940	.006
BPVS-norsk Middels (44–56)	21	50.52	4.082	-.253	-1.293
BPVS-norsk Høy (57–81)	19	66.89	6.903	.411	-.651
	61				
BPVS-urdu/panjabi Lav (10–24)	20	17.35	4.332	.512	-1.214
BPVS-urdu/panjabi Middels (25–30)	22	27.36	1.649	.195	-1.096
BPVS-urdu/panjabi Høy (31–76)	18	42.67	11.438	1.605	3.165
	60				
BPVS-total Lav (33–47)	21	41.57	4.467	-.594	-.672
BPVS-total Middels (49–60)	20	54.95	3.576	-.193	-1.066
BPVS-total Høy (61–87)	19	70.74	6.822	.429	.485
	60				

BPVS-norsk. Gjennomsnittsmålene forteller at det bare er den sterkeste gruppen fra tospråksutvalget som nærmer seg det enspråklige utvalgets gjennomsnitt: *BPVS-norsk Høys* gjennomsnitt på 66.89 ligger ca. et halvt standardavvik under gjennomsnittet til enspråksutvalget (73.39/*tabell 1*). De to nederste gruppene får skårer som er over halvannet standardavvik under enspråksutvalget. Vi kan også se at maksimumsskåren for *BPVS-norsk Lav* (43) tangerer minimumsskåren for enspråksutvalget (43/*tabell 1*). Denne nyanseringen bidrar til en forståelse av forholdet mellom tospråksutvalgets skårer og enspråksutvalgets skårer: Ikke bare er gjennomsnittet lavere, som vi så over – det er også bare de to øverste tredjedelene av tospråksutvalget som overhodet overlapper enspråksutvalgets skårer.

BPVS-urdu/panjabi. Det påfallende her er *BPVS-urdu/panjabi Høy* sitt høye standardavvik. Ved høye standardavvik kan mean være et dårlig uttrykk for sentraltendensen. Her er imidlertid median 40, som ikke skiller seg betydelig fra mean. Videre understreker skjevheten i grupperingene tendensen vi så i totalutvalget, men med en nyansering: Vi ser at den positive skjevheten i de to lavest skårende gruppene er ganske svak – langt under skjevheten for totalutvalget. For *BPVS-urdu/panjabi Høy* ser vi imidlertid en klar høyreskjevhet. Vi ser også at det er en betydelig større range for *BPVS-urdu/panjabi Høy* (31–76) enn for de to lavest skårende gruppene, men *BPVS-urdu/panjabi Høy* har samtidig en spiss fordeling. Dette viser at resultatene er mer sentrert rundt visse skårer, enn spredt flatere utover. Til sammen tydeliggjør denne nivåbaserte gruppeinndelingen av *BPVS-urdu/panjabi* at mean for hele utvalget ikke tilslører skjevheter i utvalget i negativ retning – gjennomsnittet er ikke ”kunstig” lavt, det vil si trukket ned av noen få svært lave skårer. Det er snarere motsatt: Tyngdepunktet/skjevheten er klart positiv, og gjennomsnittet trekkes således noe opp av noen få svært gode skårer.

BPVS-total. En beregning av *BPVS-total* ligger altså fortsatt langt under enspråkliges gjennomsnittsordforråd. I disse undergruppene ser vi i henhold til dette at mean for *BPVS-total Lav* ligger ca. 2 ½ SD under gjennomsnittet for enspråklig norske (73.39/*tabell 1*), mens *BPVS-total Middels* ligger halvannet standardavvik under gjennomsnittet for enspråklig norske. *BPVS-total Høy* skiller seg bare ubetydelig fra enspråkliges gjennomsnitt, med 70.74. Dette betyr at en beregning av totalordforråd fører til at den sterkeste tredjedelen av tospråksutvalget kommer omtrent opp på gjennomsnittet for enspråklig norske. Dette så vi imidlertid også gjaldt for *BPVS-norsk Høy*, som hadde 66.89 – altså bare et halvt SD under enspråksutvalget. De to nederste grupperingene av *BPVS-total* ligger fortsatt tydelig under

normalområdet for enspråklig norske, akkurat som de to nederste grupperingene av BPVS-norsk. En beregning av BPVS-total endrer med andre ord ikke avstanden til enspråklige i spesielt stor grad.

5.2 Signifikanstester

Forskjellen mellom tospråksutvalg og enspråksutvalg på BPVS-norsk

Med *t-tester* testes en såkalt nullhypotese, det vil si sannsynligheten for at det *ikke* er noen forskjell mellom populasjonene (Gall et al., 2007). For å undersøke dette med hensyn til BPVS-norsk for tospråks- og enspråksutvalget, gjennomfører jeg en independent-samples *t*-test (se vedlegg 3 for resultater i tabell). Her er for det første *Levenes Test* ikke signifikant, og derfor refereres resultatene som er beregnet ut fra antakelsen om at variansen (*populasjonenes standardavvik*/ σ) i de to populasjonene er lik. Resultatet viser en signifikant forskjell mellom gjennomsnittet for tospråksutvalget ($M = 51.15$, $SD = 13.141$) og enspråksutvalget ($M = 73.39$, $SD = 11.945$), $t(249) = -12.347$, $p = .000$ (two-tailed). Dette innebærer at vi med høy sikkerhet kan forkaste nullhypotesen, men hvor *stor* forskjellen er mellom populasjonene, sier dette resultatet ingenting om – dette er informasjon som strengt tatt er utilgjengelig for oss, så lenge vi ikke kan teste hele populasjonen. Vi kan imidlertid komme et godt stykke på vei ved å se på konfidensintervallet som beregnes i *t*-testen. Konfidensintervall brukes for å *estimere* populasjonens gjennomsnittsforskjell; det beregnes en øvre og nedre grense som med stor sannsynlighet (95 %) dekker *populasjonenes gjennomsnittsforskjell* (Gall et al., 2007). Sagt på en annen måte: Dersom vi hadde gjort hundre utvalg av populasjonen, så ville bare fem av dem hatt et konfidensintervall som ikke inneholdt populasjonenes faktiske forskjell. I dette tilfellet estimeres gjennomsnittsforskjellen i populasjonen til (med 95 % sannsynlighet) å ligge mellom -25.79 og -18.69. Vi ser at verdiene er negative, slik at 0 ikke dekkes av konfidensintervallet. Dette betyr at vi kan være rimelig sikre på at forskjellen går i den retningen den gjør i utvalgene (enspråklige skårer høyere enn tospråklige), og på at forskjellen er stor også mellom populasjonene. Det er imidlertid altså mulig at populasjonenes forskjell er både større og mindre enn det vi ser mellom utvalgene, men ikke at den er null, og at forskjellen med 95 % sannsynlighet ligger innenfor estimert konfidensintervall.

Forskjellen mellom BPVS-norsk og BPVS-urdu/panjabi i tospråksutvalget

En paired-samples t-test velges for å teste sannsynligheten for nullhypotesen med hensyn til differansen mellom tospråksutvalgets BPVS-urdu/panjabi og BPVS-norsk, altså sannsynligheten for at differansen i populasjonen er null (se vedlegg 4 for resultater i tabell). Her er utvalget på 60, siden det bare foreligger BPVS-urdu/panjabiresultat for 60 barn. Dette innebærer også at mean for BPVS-norsk ikke er helt identisk med det *tabell 1* viser, der $N = 61$. T-testen viser en signifikant forskjell mellom gjennomsnitt på BPVS-urdu/panjabi ($M = 28.62$, $SD = 12.188$) og gjennomsnitt på BPVS-norsk ($M = 51.32$, $SD = 13.185$), $t(59) = -11.030$, $p = .000$, (two-tailed). Dette betyr at nullhypotesen om ingen forskjell i populasjonen av barn med pakistansk bakgrunn kan forkastes også her. Videre estimeres forskjellen til å ligge mellom -18.582 og -26.818 i populasjonen med 95 % sannsynlighet.

Signifikanstesting av distribusjonene for sosioøkonomiske bakgrunnsvariabler

For å undersøke om forskjellene jeg fant med hensyn til sosioøkonomiske bakgrunnsvariabler mellom tospråksutvalget og enspråksutvalget er signifikante, gjør jeg en Mann-Whitney U-test for uavhengige utvalg. Mann-Whitney U er et non-parametrisk alternativ til t-tester som kan brukes dersom variabler er på ordinalnivå (Gall et al., 2007). Ved bruk av Mann-Whitney U sammenliknes *medianen* mellom to grupper, i motsetning til mean ved t-tester (Pallant, 2010). Testen viste at nullhypotesen kan avvises (med et signifikansnivå på .05) med hensyn til distribusjonen av *mors høyeste utdanning* ($p = .013$) og *hvor mange billedbøker har dere for barn i hjemmet* ($p = .000$). Det er altså sannsynlig at de forskjellene som vi ser i utvalgene på disse bakgrunnsvariablene avspeiler faktiske tendenser i populasjonene. Nullhypotesen må imidlertid opprettholdes med hensyn til *fars høyeste utdanning* ($p = .051$) – her er det altså noe mer sannsynlig at den forskjellen som finnes i utvalgene er tilfeldig og ikke avspeiler en tendens i populasjonene.

5.3 Effektstørrelser

Effektstørrelser for forskjellen mellom mean er interessante mål som kan gi utfyllende informasjon om de forskjellene som er signifikanstestet, fordi de gir et estimat på den relative ”magnitudo”, eller styrke, på en forskjell (Gall et al., 2007; Pallant, 2010). Effektstørrelser tilfører med andre ord ytterligere et mål på statistisk *power*/styrke til funn fra signifikansanalysene (Gall et al., 2007). Vi kan imidlertid også bruke effektstørrelser til å

vurdere forskjellen mellom andre gjennomsnitt – her er det spesielt interessant å undersøke forskjellen mellom henholdsvis BPVS-total og BPVS-urdu/panjabi i forhold til enspråkliges gjennomsnitt; dette vil være en nyansering til forskjellen mellom tospråksutvalgets BPVS-norsk og enspråklige. I tillegg kan vi se på forskjellen mellom BPVS-urdu og BPVS-panjabi; som vi så over var denne liten. Til slutt vil jeg undersøke effektstørrelser på utviklingen til barna fra de var fem år og bruker her Melings (2011) resultater som beskrevet tidligere.

Et vanlig brukt mål på effektstørrelse er *Cohens d* (Gall et al., 2007; Pallant, 2010). Cohens *d* uttrykker forskjeller mellom grupper i standardavviksenheter (*d*), og regnes ut ved hjelp av mean og SD for gruppene som skal sammenliknes (Pallant, 2010). Dette betyr enkelt sagt at dersom effektstørrelsen (*d*) er 1.0, er forskjellen mellom gruppens mean en standardavviksenhet – altså rimelig stor. Cohen (ref. i Pallant, 2010) anga tentative retningslinjer for tolkning av *d* – og foreslo at verdier opp til .2 kunne tolkes som små, verdier på .5 som middels og verdier fra .8 som store effekter.

Tabell 3. Cohens *d*, beregnet med mean og SD. Mean oppgitt.

					Cohens <i>d</i>
Forskjeller mellom tospråklige og enspråklige:					
	<i>Tospråklige</i>		<i>Enspråklige</i>		
*	BPVS-norsk	51.15	BPVS-norsk	73.39	1.77
	BPVS-total	55.27	BPVS-norsk	73.39	1.45
	BPVS-urdu/panjabi	28.62	BPVS-norsk	73.39	3.71
Forskjeller innad i tospråksutvalget:					
*	BPVS-norsk	51.15	BPVS-urdu/panjabi	28.62	1.77
	BPVS-norsk	51.15	BPVS-total	55.27	.31
	BPVS-urdu	30.13	BPVS-panjabi	25.81	.37

Tabellen fortsetter på neste side

Tabell 3 forts.

Forskjeller fra 5-årsalder til 6-årsalder:				
5 år		6 år		
BPVS-urdu/panjabi	21.43	BPVS-urdu/panjabi	28.62	.64
BPVS-norsk	32.08	BPVS-norsk	51.15	1.53

* Forskjellen er signifikanstestet.

Vi ser som forventet (*tabell 3*) svært store effektstørrelser for forskjellene mellom tospråklige og enspråklige på BPVS-norsk, og også innad i tospråksutvalget for forskjellen mellom BPVS-norsk og BPVS-urdu/panjabi (merk: begge disse beregningene viste $d = 1.77$). Effektstørrelsene tilfører dermed ekstra styrke til signifikanstestene over. Videre kan vi merke oss at BPVS-total ikke minsker effektstørrelsen mellom tospråklige og enspråklige i særlig grad. Forskjellen mellom BPVS-urdu/panjabi og enspråklige er, ikke overraskende, spesielt stor.

Innad i tospråksutvalget ser vi for øvrig at forskjellene mellom BPVS-norsk og BPVS-total, og mellom BPVS-urdu og BPVS-panjabi, må regnes som relativt små. Effektstørrelsene viser også tydelig at utviklingen fra femårsalder har vært mye mindre på urdu/panjabi enn på norsk.

6 Validitetsvurdering

Før jeg går videre til å oppsummere og drøfte disse resultatene, skal jeg gjøre en validitetsvurdering av dem og av studien for øvrig – altså en vurdering av hvilke slutninger som kan gjøres i studien. Resultatene kan for det første anses som pålitelige i seg selv: Vi så tidligere at prosedyrene for testingen bidrar til at studien har en god reliabilitet, og Cronbachs Alpha er – som vi så i resultatene (*tabell 1*, s. 39) – høy, da verdier over .8 regnes som høye (Pallant, 2010). Dette viser at også testen har god indre konsistens. Men hvilket grunnlag har vi for å komme videre fra disse konkrete, reliable resultatene for utvalgene – altså fra det vi faktisk har målt – til det vi ønsker å undersøke? I problemstillingen spør jeg etter kjennetegn ved *reseptivt breddeordforråd hos barn med pakistansk bakgrunn*, og hvordan slike kjennetegn kan forstå i lys av *barnas språksituasjon*. Fra et validitetsperspektiv må vi spørre: Har jeg målt reseptivt breddeordforråd på en god måte, og kan resultatene generaliseres til populasjonen av barn med pakistansk bakgrunn? Har jeg et godt grunnlag for å forstå resultatene i lys av aspekter ved barnas språksituasjon? Vi skal se på disse og flere spørsmål som angår validitet, med utgangspunkt i Cook og Campbells validitetssystem (1979, ref. i Lund, 2002).

Statistisk validitet

Spørsmålet som stilles når en studies statistiske validitet skal vurderes, er i hvilken grad det kan trekkes holdbare *statistiske* slutninger fra kjente utvalgsresultater til *populasjonen* utvalget er hentet fra (Lund, 2002). Slike slutninger kan gjøres med ulike former for analytisk statistikk (slutningsstatistikk): I hovedsak går skillet her mellom parametriske og non-parametriske analyser (Gall et al., 2007). I denne studien har jeg valgt *parametriske* analyser for signifikanstesting av gjennomsnittsforskjeller på BPVS. Parametriske analyser er sterke verktøy som kan gi gode indikasjoner på ukjente egenskaper ved populasjonen (parametre) (Gall et al., 2007). Slike analyser gjøres imidlertid med utgangspunkt i at flere betingelser gjelder, blant annet at data er på intervallnivå, og at populasjonene er normalfordelt. Gyldigheten av beregningene vil svekkes dersom disse betingelsene ikke er oppfylt (Gall et al., 2007). Når det gjelder den statistiske gyldigheten i denne studien, må vi derfor vurdere om det var riktig å bruke *parametriske* analyser. Variablene var riktignok på intervallnivå, men i tillegg bør altså populasjonene være tilnærmet normalfordelt dersom slike mål skal kunne gi et gyldig resultat. Som vi så over, kan ikke *BPVS-urdu/panjabi* regnes som jevnfordelt: Den

er tydelig høyreskjev. Gitt krav til normalfordeling ble det vurdert om uteliggere (i dette tilfellet noen svært få høye skårer) skulle fjernes fra fordelingen før denne testen skulle gjennomføres for å få en mer tilnærmet normalfordeling; det vil si for å gjøre beregninger om *populasjonen* uten disse ekstremskårene. Slike korrigeringer bør imidlertid bare gjøres dersom man har grunn til å tro at det har skjedd noe feil i testingen, eller at testpersonen ikke har vært utsatt for like testbetingelser som de andre, og at resultatene for uteliggerne dermed ikke kan regnes som reliable og representative (Gall et al., 2007). Det er ingen grunn til å tro at noe slikt er tilfelle her, derfor ble ingen kasus fjernet fra videre analyse. Jeg valgte allikevel den parametriske t-testen. En begrunnelse for dette er at parametriske analyser som nevnt er sterkere statistiske metoder. Dette betyr at de finner forskjeller mer effektivt enn *non-parametriske* alternativer, som riktignok ikke forutsetter noe om populasjonen, men som også er svakere – det vil si at de utnytter utvalgstillene dårligere og ofte også krever større utvalg for å vise signifikans på tendenser, og dermed gir større sannsynlighet for feil slutning (Gall et al., 2007). En annen begrunnelse for anvendelse av t-testen er at flere hevder at den er rimelig robust mot selv ganske store avvik fra normalfordeling – særlig dersom utvalget er av en viss størrelse (Gall et al., 2007; Pallant, 2010). Pallant (2010) antyder at en utvalgsstørrelse på over 30 vil være tilstrekkelig til at brudd på forutsetningen om normalfordeling ikke er et så stort problem. Det er derfor grunn til å tro at resultatet fra t-testen er gyldig, siden utvalget er rimelig stort ($N = 60$), og siden avviket fra normalfordeling ikke var helt uakseptabelt. For t-testen for *uavhengige utvalg*, mellom tospråksutvalgets og enspråksutvalgets BPVS-norsk, så vi at betingelsen om jevnfordeling var oppfylt.

Selv om analysemetodene slik kan forsvares, bør det videre gjøres en vurdering av muligheten for å trekke *feil* slutning av signifikansanalysene: Såkalte *type I-feil* innebærer at man feilaktig slutter at det er en forskjell mellom populasjonene eller i populasjonen når det i virkeligheten ikke er det (altså at man feilaktig forkaster nullhypotesen) (Gall et al., 2007). I tillegg bør forskjellene være rimelig sterke for at slutning til populasjonen kan gjøres (Lund, 2002). Resultatene i denne studien viser sterk signifikans ($p = .000$), og derfor er sannsynligheten for feilaktig forkasting av nullhypotesen svært liten. Vi så videre at konfidensintervallene gav en tydelig indikasjon på en relativt stor forskjell også i populasjonen. Det at effektstørrelsene i tillegg var høye ($d = 1.77$ for begge forskjellene) viser at tendensen er sterk, noe som tilfører styrke til signifikanstesten og dermed ytterligere styrker muligheten for å generalisere funn på statistisk grunnlag.

Med hensyn til signifikanstesting av forskjellene mellom utvalgenes *sosioøkonomiske bakgrunnsvariabler*, kan bruken av Mann Whitney U forsvares ettersom variablene kan rangeres og dermed er på *ordinalnivå* (Gall et al., 2007; Pallant, 2010). Her kan vi i tillegg si at muligheten for å begå type 1-feil, det vil si å feilaktig slutte at det er en forskjell i populasjonene, er betydelig redusert på grunn av svært lave signifikansverdier for *mors høyeste utdanning* ($p = .013$) og *antall billedbøker for barn i hjemmet* ($p = .000$). Med hensyn til *fedrenes utdanningsnivå*, som viste en lav men ikke signifikant p ($= .051$), er faren imidlertid tilstede for å begå en *type 2-feil* – altså det å feilaktig anta nullhypotesen når det i virkeligheten er en forskjell.

Begrepsvaliditet

En god begrepsvaliditet innebærer at det er samsvar mellom begrepet man ønsker å måle, slik det er definert teoretisk, og begrepet slik vi har lyktes å operasjonalisere det med de instrumenter som velges til målingen (Kleven, 2002a). Et sentralt teoretisk begrep i dette prosjektet er *reseptivt breddeordforråd*, og instrumentet, eller testen som skal måle dette, er BPVS. En trussel mot begrepsvaliditet kan blant annet være at testen som er valgt ikke dekker hele det teoretiske begrepet (Kleven, 2002a) – i dette tilfellet at testen ikke fanger opp barnas faktiske reseptive breddeordforråd. Hva kan vi si om dette med hensyn til BPVS? I forskning som sammenlikner språkferdigheter, er det svært ofte BPVS som brukes når reseptivt breddevokabular skal undersøkes (Bialystok et al., 2010). Testen er således svært anerkjent og regnes som en valid operasjonalisering av reseptivt breddeordforråd. Det er ingen grunn til å tro at den norske oversettelsen (BPVS-norsk) skal være mindre gyldig (Lyster et al., 2010). Validitet ble også vurdert i forbindelse med normeringen av BPVS for ulike minoritetsspråk i Norge ved å undersøke samvariasjon mellom BPVS og testen *Begreper fra språk 6–16* (dybdevokabular) (Monsrud et al., 2010). Korrelasjonen mellom disse resultatene var svært høy for BPVS-norsk (.70), og noe lavere, men fortsatt moderat til sterk (.53) for minoritetsspråksversjonene. Dette understøtter antakelsen om at breddeordforråd slik det måles med BPVS gir et lite, men viktig, innblikk i en større helhet av semantisk kompetanse.

En trussel mot begrepsvaliditeten som allikevel bør vurderes, er at reseptivt breddeordforråd via BPVS ikke er undersøkt i faktiske kommunikasjonssammenhenger, men i en relativt dekontekstualisert ramme. Som vi så tidligere, vil selve konteksten et ord normalt ytres i gi betydelig informasjon om dets meningsinnhold (Rommetveit, 1972; Nation, 2001). Det er

derfor grunn til å anta at *reseptivt breddeordforråd slik det måles med BPVS*, ikke nødvendigvis er et direkte uttrykk for det reseptive breddeordforrådet et barn har tilgjengelig i naturlig kommunikasjon, eller for eksempel i forbindelse med lesing. I begge disse kontekstene vil det være flere ulike typer informasjon tilgjengelig om hva et ord betyr – for eksempel *morfosyntaks* (det vil si informasjon om hvilken rolle ordet har i setningen det ytres i), den *semantiske* konteksten for ytringen og samtalen, og den øvrige *sosiale og kulturelle kontekst* for ytringen (Sveen, 2005). Innenfor rammen av en naturlig samtale vil barn også ha mulighet til å oppklare misforståelser og sjekke om et ord er forstått riktig. Nation (2001) hevder at 5–10 % av ukjente ord som opptrer *i kontekst* til en viss grad vil kunne gjettes riktig og samtidig læres, og at ord som kan gjettes riktig i kontekst men som ikke læres til neste gang, antakelig er mange flere enn dette igjen. Dette tyder på at barn kan beherske et ord i praksis (i hvert fall i en slik grad at samtalen ikke bryter sammen), som de ikke ville få poeng for på BPVS. Videre så vi tidligere at ord ikke *enten* beherskes *eller* ikke beherskes, men forstås mer og mer i et gradvis kontinuum (Rommetveit, 1972; Nation, 2001). BPVS-resultater bør derfor ikke tolkes unyansert, nærmest som et svart-hvittbilde av barns kompetanse – denne kompetansen må heller ses som i stadig endring, og BPVS som en operasjonalisering av et utsnitt i tid på dette forståelseskontinuumet.

De videre vurderinger av begrepsvaliditeten til reseptivt breddeordforråd skal gjøres spesielt i lys av muligheten for å kartlegge tospråklig kompetanse. I hvilken grad måler BPVS *tospråklige* barns reseptive ordforråd på en god måte? For det første måler BPVS ordforråd på de to språkene atskilt, og fanger derfor ikke opp et samlet ordforråd. Ordforråd som er utviklet i hjemmekontekst vil ikke nødvendigvis være utviklet på norsk, og motsatt (Bialystok et al., 2010). Denne trusselen mot begrepsvaliditeten for tospråklig ordforråd kan til en viss grad veies opp ved å beregne totalordforråd, noe som også er gjort i denne studien. Oller og Eilers (2002) påpeker imidlertid at selv om man slår sammen skårer på ordforrådstester fra to språk, trenger vi andre typer tester med andre typer normer for virkelig å kunne forstå tospråklig ordforråds-kunnskap. Vi kan for eksempel tenke oss at et tospråklig ordforråd kan ha ”hull” (fordi ulike språk ofte brukes i ulike domener), som ikke et enspråklig ordforråd vil ha. Når testen er utformet for å måle ettspråkskompetanse som normalt ikke har slike ”hull”, vil tospråklig vokabularkompetanse kunne fremstå som kunstig svak. I denne studien er det derfor en styrke at BPVS faktisk er normert for flere minoritetsspråksgrupper i Norge – dette gir et visst tospråklig sammenlikningsgrunnlag, selv om testen som sådan er utformet for å teste ettspråkskompetanse. Likevel er det grunn til å ta denne trusselen mot begrepsvaliditet

på alvor: Den kan gjelde i dobbel forstand for tospråksutvalget, ettersom vi kan anta at flere bruker både urdu og panjabi hjemme (Karlsen & Lykkenborg, 2012). Når barnet testes på enten urdu eller panjabi, er det derfor en mulighet for at ikke totalS1-ordforrådet fanges opp. Det kan videre tenkes at testen ikke avspeiler reseptivt breddeordforråd på *alle* barnas *sterkeste* hjemmespråk: Mine egne observasjoner i testsammenheng viser at det ofte var vanskelig for barna å bestemme seg for *ett* språk som de snakket mest hjemme – mange svarte at de snakket både litt urdu og litt panjabi.

Til slutt kan vi spørre om det kan være kulturelle forskjeller som kan bidra til at BPVS kan være vanskelig å bruke på tvers av språk – til tross for at den er oversatt. Monsrud et al. (2010) poengterer dette i beskrivelsen av normeringsarbeidet med testen, gjennom å påpeke at ordforråd i stor grad er knyttet til erfaringer og utviklet i barnets kulturelle kontekst. Ord som er høyfrekvente på ett språk eller i én kultur, er ikke nødvendigvis det på et annet språk eller i en annen kultur (Monsrud et al., 2010). For å sette denne trusselen *på spissen*; dersom barn med pakistansk bakgrunn bruker et ordforråd som er kulturelt svært ulikt det man ser i engelskspråklige kulturer og norskspråklige kulturer, vil en BPVS-utgave som er oversatt fra disse språkenes versjoner fange opp deres faktiske reseptive ordforråd dårlig. Dette kan gi feilkilder i testen som kan svekke begrepsvaliditeten. Det er imidlertid forsøkt å ta høyde for slike feilkilder i så stor grad som mulig i normeringen, men slike faktorer vil alltid måtte tas i betraktning når resultatenes gyldighet skal vurderes (Monsrud et al., 2010).

Til sammen kan forhold som er beskrevet her, svekke begrepsvaliditeten potensielt ganske betraktelig med hensyn til måling av tospråklige barns breddeordforråd. Det er imidlertid ingen grunn til å tro at BPVS er noe dårligere enn andre tester med hensyn til disse problemene. I tillegg gir testen også utvilsomt noen fordeler: For eksempel muliggjøres sammenlikninger mellom tospråklige og enspråklige innenfor hvert språk. En klar styrke er også at testen er oversatt til både urdu og panjabi. Når forbeholdene over er tatt, kan vi derfor bruke testen som et redskap for å kartlegge også tospråklige barns kompetanse, og problemet med at testen måler hvert av språkene for seg kan til en viss grad veies opp av muligheten for å beregne totalordforråd på en ganske enkel måte. Som vi så over, er det også på generelt grunnlag god grunn til å tro at testen gir en god *indikasjon* på størrelsen av et barns breddeordforråd; dette vil også gjelde for tospråklige. Øvrige fordeler ved testen er i tillegg mange: Data kan samles inn fra store grupper relativt raskt, med mulighet for å kvantifisere data på intervallnivå som egner seg for parametrisk slutningsstatistikk. Det er også en av få

tester for reseptivt breddeordforråd som er oversatt og normert for minoritetsspråk i norsk sammenheng, noe som i seg selv gjør den til et svært verdifullt redskap for vurdering av tospråklig kompetanse.

Et annet sentralt teoretisk begrep i prosjektet, ”aspekter ved barnas språksituasjon”, er beskrevet gjennom teori og tidligere empiri over, og i tillegg med bakgrunnsvariabler fra spørreskjema. Her er det selvsagt ikke praktisk mulig å dekke hele det teoretiske begrepet – barnas faktiske språksituasjon er uendelig kompleks. Det er imidlertid svært sentrale aspekter ved tospråklighet og minoritetsspråklighet jeg har tatt utgangspunkt i når språksituasjonen skal spesifiseres (Baker, 2011; Grosjean; 2008; Bialystok, 2001; Romaine, 1995). Jeg mener derfor at beskrivelsen i tilstrekkelig grad fanger opp *noen* sentrale aspekter ved barnas språksituasjon – altså tilstrekkelig til å besvare problemstillingen meningsfullt innenfor de avgrensningene som er gjort (jf. innledningen). Dette gjør at forståelsen av resultatene får gyldighet i lys av et begrenset, men på alle måter sentralt utsnitt (en ”begrepsoperasjonalisering”) av sosiale og kontekstuelle aspekter ved barnas språksituasjon.

Når det gjelder *spørreskjemaet*, er de sosioøkonomiske variablene relativt enkle å operasjonalisere godt. I motsetning til dette er den språklige bakgrunnsvariabelen basert på vurdering av et mer sammensatt forhold – et svar på spørsmålet *hvor gammelt var barnet når det ble introdusert for norsk* gir for eksempel ikke noe informasjon om hvorvidt barnet også etter dette ble omgitt av mye norsk. Det kan også være vanskelig for foreldre å huske når barnet ble introdusert for norsk, og kanskje også å vurdere akkurat hva som skal regnes som *introduksjon for norsk*. Dette er en svakhet ved spørreskjemaets validitet, og må tas i betraktning når informasjonen skal tolkes. Det er imidlertid en styrke at foreldrene fikk så mange svaralternativer, slik at vi får et godt inntrykk av variasjonen i utvalget på denne variabelen.

Indre validitet

Indre validitet handler først og fremst om de statistiske slutningene som trekkes *innad* i en studie (Kleven, 2002b). Statistiske analyser i denne studien er ikke-kausale – altså beskrivende: Funn (deskriptive kjennetegn) forsøkes heller forstått gjennom ”ytre” fortolkninger, og ikke gjennom ”indre” kausalitets- eller samvariasjonsanalyser. Derfor aktualiseres ikke de vanlige spørsmålene som angår indre validitet i Cook og Campbells validitetssystem (Lund, 2002). Noen anliggender som i en viss forstand gjelder indre

sammenhenger er allikevel aktuelle; for eksempel spørsmålet om en eventuell retesteffekt ved andre gangs testing av reseptivt breddeordforråd for tospråksutvalget. Slike retesteffekter er ofte en feilkilde i pedagogisk forskning (Lund, 2002). I min studie kan vi ikke utelukke at kjennskap til testsituasjon gir en mer komfortabel ramme ved andregangs testing med BPVS-urdu/panjabi, og heller ikke en potensiell hukommelseffekt av å ha sett de samme bildene en uke før. I denne forbindelse er det viktig å informere om at barnet ved førstegangstesting aldri får bekreftet om de svarer riktig eller feil. Fasiten er altså ukjent for barnet, også ved andregangstesting. BPVS er videre standardisert på en slik måte at det ikke ses såkalte instrumenteringseffekter; det vil si effekter som kan oppstå dersom noe ved *selve måleinstrumentet* gir kunstige resultater i form av for eksempel en gulv- eller takeffekt (Lund, 2002).

Ytre validitet

Ytre validitet handler om hvorvidt det er mulig å gjøre ikke-statistiske generaliseringer av funn; for eksempel om hvordan en best mulig velger individer slik at funn kan generaliseres til å gjelde populasjonen (Lund, 2002). Utvalget bør med andre ord være *representativt* for populasjonen for å sikre generaliseringsmuligheter (Lund, 2002). Som vi så i utvalgsbeskrivelsen, er ikke *tospråksutvalget* i denne studien gjort statistisk tilfeldig. Dette kan svekke ytre validitet (og for så vidt statistisk validitet – siden gyldigheten av slutningsstatistikk forutsetter at utvalget er representativt (Gall et al., 2007)). Det som imidlertid *styrker* muligheten for at utvalget allikevel kan regnes som rimelig representativt, er (1) en relativt stor utvalgsstørrelse, (2) at det ikke var noen styring i utvalgsprosessen mot enkelte grupperinger i populasjonen, og (3) at resultatene viser fordelinger uten uakseptable skjevheter. Jeg vil i den følgende drøftingen derfor anse utvalget som tilfredsstillende representativt (med de forbehold som er anført) for *populasjonen av seks år gamle barn med pakistansk bakgrunn bosatt i Oslo og omegn, som har gått i barnehage (minst) ett år før skolestart, og som ikke har kjente avvik fra normalutvikling* – heretter betegnet, i samsvar med problemstillingen, 'barn med pakistansk bakgrunn'.

Ytre validitet handler også om hvorvidt den kunnskapen vi har fra andre kilder (teori og empiri) kan bidra til at resultater eller vurderinger får en ytre gyldighet (Lund, 2002). Det er altså ikke bare begrepsvaliditet og statistisk validitet som bestemmer generaliseringsmulighetene. Når det gjelder *resultatene* i denne studien og *problemstillingens*

første ledd, vil jeg gjøre en slik vurdering i starten av drøftingsdelen, der jeg oppsummerer resultatene og ser dem i lys av relevante studier. For resten av drøftingsdelen, der jeg besvarer *problemstillingens andre ledd*, må det nevnes at selv om jeg ikke gjør *statistisk* indre slutninger i denne studien, gjøres allikevel slutninger om mulige sammenhenger mellom barnas språksituasjon og deres kompetanse. Når jeg drøfter resultatene i lys av barnas språksituasjon, innebærer dette derfor å spekulere i nettopp ”indre” sammenhenger: Hvorfor er resultatene som de er? Her må det understrekes at disse slutningene gjennomgående gjøres med forsiktighet, men også at det er hevet over tvil at språksosiale, kontekstuelle forhold gir et sentralt bidrag til barns språkkompetanse (som vist i teoridelen, og ellers behandlet spesielt for eksempel i Romaine, 1995; Bialystok, 2001; Hammer et al., 2004; Eilers, Pearson & Cobo-Lewis, 2004; Baker, 2011; Goldenberg, Reese & Rezaei, 2011). *Ytre*, teoretisk-emipriske, begrunnelser kan slik bidra til at også slutninger om presumptivt *indre* sammenhenger mellom språksituasjon og ordforråd kan få en akseptabel gyldighet.

Validitet oppsummert

Denne validitetsvurderingen viser at slutninger som gjøres i studien alt i alt må regnes som av akseptabel kvalitet. Med forbehold om at utvalget er tilfredsstillende representativt, noe jeg altså legger til grunn videre, kan det statistisk generaliseres fra denne studien. Slik jeg vurderer det finner vi de mest reelle truslene mot *begrepsvaliditet*, men også her må konklusjonen være at gyldigheten må anses som akseptabel. Det må imidlertid legges til grunn i den påfølgende drøftingen at BPVS ikke helt fanger opp reseptivt breddeordforråd slik det kan antas å fungere i praksis, og videre at det er en mulighet for at tospråklig ordforråd ikke fanges opp like godt som enspråklig ordforråd, gitt potensielle kulturelle forskjeller og mulighet for at ikke totalS1-ordforrådet er fanget opp.

7 Tospråklig ordforråd i tospråklig kontekst

For å få en oversikt over resultatene, som altså i stor grad kan regnes som reliable og generaliserbare, skal jeg nå først gi en kort oppsummering og vurdering av det jeg regner som hovedfunn i henhold til spørsmålet *hva kjennetegner reseptivt breddeordforråd hos barn med pakistansk bakgrunn?* For å styrke resultatenes ytre validitet, underbygges vurderingen med resultater fra andre nylige norske studier av reseptivt breddeordforråd hos barn med pakistansk bakgrunn. Deretter skal jeg drøfte resultatene i lys av aspekter ved barnas språksituasjon.

7.1 Kjennetegn på reseptivt breddeordforråd hos barn med pakistansk bakgrunn

Som vi har sett, avspeiler gjennomsnittresultater sentraltendenser på en god måte. Barn med pakistansk bakgrunn skårer i gjennomsnitt 51.15 på BPVS-norsk, noe som er på linje med det Monsrud et al. (2010) fant i normeringsstudien ($M = 53$) for pakistanske språk i aldersgruppen *6 til 7 år*. Videre skårer barn med *norsk* bakgrunn i gjennomsnitt 73.39 på BPVS-norsk. Dette er også i samsvar med resultatene fra den norske normeringsstudien av BPVS for enspråklig norske, der Lyster et al. (2010) fant at enspråklig norske barn i aldersgruppen *6;6–6;11 år* i gjennomsnitt skåret 77.89 ($N = 45$, $SD = 12.95$). Samsvar med disse normeringsstudiene underbygger resultatene i min studie, og dermed også funnet at ca. to tredjedeler av barn med pakistansk bakgrunn skårer under normalområdet for enspråklig norske på norsk. En beregning av totalordforråd for tospråksutvalget endrer ikke dette bildet i særlig grad.

BPVS-urdu/panjabi i min studie ($M = 28.62$) korresponderer imidlertid ikke like godt med skåren for pakistanske språk som ble funnet i normeringsstudien ($M = 44$) for aldersgruppen *6 til 7 år*. Som tidligere nevnt er det lite rapportering om utvalg og utvalgsstørrelser for spesifikke alderstrinn i normeringsstudien, og vi må derfor generelt legge til grunn en viss usikkerhet knyttet til denne som sammenlikningsgrunnlag. Det kan tenkes at utvalget i min masterstudie er mer representativt, dette vet vi ikke, men i alle fall er det større. Og hvis vi i stedet sammenlikner med Lervåg og Aukrust (2010), som fant et urdu-mean på 37.77 ($SD = 18.25$) for *andreklassinger*, er ikke et mean for førsteklassinger på 28.62 overraskende –

kanskje snarere forventet. Gitt stor utvalgsstørrelse i Lervåg og Aukrust (2010, $N = 90$), og en ukjent, men antakelig lav utvalgsstørrelse i normeringsstudien, kan vi anta at Lervåg og Aukrust (2010) utgjør et bedre sammenlikningsgrunnlag. Det er heller ikke, som vi tydelig har sett, grunnlag for å si at mean for BPVS-urdu/panjabi i denne masterstudien skulle avspeile sentraltendensen på en kunstig lav måte.

Basert på disse gjennomsnittsmålene viser også signifikanstesting en forskjell mellom tospråklige og enspråklige på norsk, og mellom tospråkliges BPVS-norsk og BPVS-urdu/panjabi – norskordforrådet er størst. Signifikanstestene får tydelig power gjennom p -verdier, konfidensintervall og effektstørrelser.

Beregning av *totalordforråd* for barn med pakistansk bakgrunn gir ingen vesentlig økning i forhold til BPVS-norsk ($d = .31$). Dermed skiller resultatene i denne masterstudien seg fra resultatene i normeringsstudien (Monsrud et al., 2010). Det er grunn til å feste lit til dette normeringsresultatet gitt den betydelige totale utvalgsstørrelsen ($N = 979$), selv om det er mulig at en differensiering i alderstrinn og språkgrupper ville vist et annet resultat. Det er uansett tydelig at barn med pakistansk bakgrunn skiller seg fra (mange) andre minoritetsspråksgrupper i Norge på dette punktet, da økningen rapporteres som så markant i normeringsstudien (vist grafisk s. 49, Monsrud et al., 2010) og samtidig er så liten i denne masterstudien. Jeg finner heller ikke noe grunnlag for å betvile gyldigheten av resultatet i min studie.

Resultatene, både sett for seg selv og i sammenlikning med andre studier, fremstår som akseptabelt pålitelige og gyldige, og gir dermed et godt lite innblikk i ordforrådskompetansen til barn med pakistansk bakgrunn. Hovedfunn, altså kjennetegn på *barn med pakistansk bakgrunns reseptive breddeordforråd* som tas med videre til drøfting av problemstillingens andre ledd, er:

- I. Barn med pakistansk bakgrunn har signifikant lavere norskordforråd enn enspråklig norske. En beregning av totalordforråd endrer ikke dette forholdet til enspråklige.
- II. Barn med pakistansk bakgrunn har signifikant ulik ordforrådsstørrelse på sine to språk. S2-ordforrådet er størst. Fra femårsalder har S2-ordforrådet økt betraktelig mer enn S1-ordforrådet.

7.2 Reseptivt breddeordforråd i lys av aspekter ved barnas språksituasjon

I henhold til problemstillingen skal jeg nå drøfte hvordan kjennetegn ved barn med pakistansk bakgrunns reseptive breddeordforråd kan forstås i lys av aspekter ved barnas språksituasjon. For å strukturere drøftingen vil jeg ta utgangspunkt i de ulike dimensjonene ved tospråklighet som jeg har redegjort for over: Ferdighet, funksjon, alder og kontekst. Disse tospråksdimensjonene anses altså i det følgende som overordnede aspekter (altså kjennetegn/trekk) ved barnas språksituasjon. Gjennomgående gjøres drøftingen med et spesielt blikk til kunnskap om gruppen med pakistansk bakgrunn i Norge, og til andre studier som det er relevant å se resultatene i lys av.

Ferdighets- og funksjonsaspekter ved barnas språksituasjon

Det er rimelig, kanskje selvfølgelig, å betegne barn med pakistansk bakgrunn som *tospråklige* slik Grosjean (2008) definerer dette – personer som bruker to eller flere språk (eller dialekter) i hverdagslivet. Vi kan regne med at barn med pakistansk bakgrunn i varierende grad bruker både urdu, panjabi og norsk i en kombinasjon på ulike livsdomener (Karlsen & Lykkenborg, 2012; Aarsæther, 2002). Vi så tidligere at en tospråklig situasjon svært ofte innebærer at tospråklige får en ulik kompetanse i hvert av sine språk: Kompetanse vil utvikle seg ulikt etter hvilke ferdigheter som kreves i de ulike språkene på de ulike domenene (Romaine, 1995; Grosjean, 2008; Baker, 2011). Det at resultatene viser en *ulik* ordforrådskompetanse i S1 og S2 er derfor ikke overraskende i lys av slike *ferdighets-og funksjonsaspekter* ved tospråklighet. En ulik kompetanse vil følge ganske naturlig dersom domenene for bruk av S1 og S2 ikke krever samme type eller grad av ferdigheter – korresponderende med komplementærprinsippet (Grosjean, 2008).

Fordi tospråklig vokabular ofte har vist seg å være distribuert mellom de to språkene, anbefales det å gjøre beregninger av sammenlagt ordforråd for tospråklige (Bialystok, 2001; Oller & Eilers, 2002; Oller et al., 2007). Dersom ordforrådet er distribuert – altså dersom flere ord bare kjennes i det ene språket og flere andre ord bare kjennes i det andre språket, ville man forvente at ordforrådsstørrelsen øker merkbart dersom ordforrådene legges sammen. I denne studien er det imidlertid helt tydelig funnet at totalordforrådet til de tospråklige barna ikke øker nevneverdig i forhold til deres BPVS-norsk. Dette resultatet impliserer at barna i

snitt har få singelord på urdu/panjabi. De får derfor bare en ubetydelig høyere skår når urdu/panjabi legges til (med forbehold om at BPVS-urdu/panjabi ikke har nevneverdig kulturell bias (jf. begrepsvaliditet over), slik at en eventuell faktisk distribusjon ikke fanges opp). Dette resultatet står i motsetning til flere andre store studier av tospråklig kompetanse (Oller et al., 2007; Monsrud et al., 2010; Bialystok et al., 2010), og vi skal derfor komme tilbake til dette etter hvert.

I lys av en generell tospråklig situasjon er det heller ikke overraskende at barn med pakistansk bakgrunn har lavere ordforråd på norsk enn enspråklig norske. Flere andre studier som har sammenliknet tospråklig og enspråklig vokabular, viser den samme tendensen til en slik *vokabularavstand* (Verhallen & Schoonen, 1993; Oller et al., 2007; Lervåg & Melby-Lervåg, 2009; Lervåg & Aukrust, 2010; Bialystok et al., 2010). Med den store variasjonsbredden i språk og kontekster som dekkes av disse studiene, tyder mye på at det er funksjonsaspekter (altså at tiden og domenene fordeles mellom språkene) ved en generell tospråklig språksituasjon *i seg selv* som legger føringer for utvikling av språkferdigheter. Vokabularavstanden jeg finner mellom tospråklige og enspråklige er altså forventet i lys av slike generelle trekk ved en tospråkssituasjon.

Videre kan det heller ikke utelukkes at vokabularavstanden også er påvirket av den store språklige avstanden mellom urdu/panjabi og norsk. Som vi så over kan de strukturelle forskjellene mellom språkene innebære en stor utfordring, kanskje særlig med hensyn til ervervelse av ordforråd (Kristoffersen, 1980). Funn av Lervåg og Melby-Lervåg (2009) støtter antakelsen om at vokabularavstanden kan bli større mellom strukturelt ulike språk. Lervåg og Melby-Lervåg (2009) hevder videre at en forklaring på dette ikke nødvendigvis bare er å finne i selve den språklige forskjellen, men også i antatt større sosiokulturelle og sociolingvistiske avstander mellom språkkulturene. Det er imidlertid viktig å påpeke at en vokabularavstand *også* ses der S1 og S2 er nærmere hverandre språklig og kulturelt sett (Oller & Eilers, 2002; Lervåg & Melby-Lervåg, 2009; Bialystok et al., 2010). Det er derfor mulig at vokabularavstanden mellom barn med pakistansk bakgrunn og enspråklig norske kan få tilstrekkelig forklaring i lys av generelle trekk ved en tospråklig situasjon *i seg selv*, men det er heller ikke utenkelig at den store strukturelle og kulturelle avstanden mellom språkene kan bidra til at resultatet viser en så klar forskjell, med tydelig styrke i både signifikanstest og effektstørrelse.

Aldersaspekter ved barnas språksituasjon

Vi ser at barn med pakistansk bakgrunn helt klart kan betegnes som tospråklige, og at ulik størrelse på S1- og S2-ordforrådet, samt vokabularavstanden til enspråklige, synes rimelig i lys av deres grunnleggende tospråklige språksituasjon. Hvorvidt barn med pakistansk bakgrunn følger en *simultan* eller *suksessiv* tospråklig utvikling, er imidlertid et vanskeligere spørsmål. Som vi har sett viser bakgrunnsvariabelen *når ble barnet introdusert for norsk* en rimelig jevn fordeling fra null til fire år. Et skille ved treårsalder delte utvalget omtrent i to, med en liten overvekt av barn som ble introdusert for norsk før tre år. Dette indikerer en mer typisk andrespråkutvikling for litt under halvparten av barna – det vil si at utviklingen primært har foregått på hjemmespråket inntil norsk ble introdusert senere, og da typisk i andre domener. Vi kan anta at en slik andrespråkskontekst typisk vil medføre lavere kompetanse i det språket som introduseres senere (Bialystok, 2001), og at dette kanskje særlig gjelder for ordforråd; i og med at ordforrådsutvikling forutsetter erfaring med, og bruk av, ord (Bialystok, 2001; Lindberg, 2008; Wold, 2008; Thordardottir, 2011) over lengre tid (Cummins, 1994; Thomas & Collier, 1997). Det at barn med pakistansk bakgrunn skårer gjennomsnittlig lavere på norsk enn enspråklige, er dermed heller ikke uventet i lys av dette *aldersaspektet*, ettersom litt under halvparten bare har lært norsk i 2–3 år før denne testen ble gjort.

Flere resultater fremstår imidlertid som noe overraskende i lys av de aspektene vi har sett på så langt. For det første; hvis vi skulle fortsette resonnerementet om at en lav norskskår er forventet ettersom mange av barna ble sent introdusert for norsk, er det desto mer overraskende at *urdu/panjabi-ordforrådet* er enda lavere igjen. Barn med pakistansk bakgrunn er med stor sannsynlighet fra fødselen av, og i hvert fall inntil femårsalder som bakgrunnsdataene er fra (Karlsen & Lykkenborg, 2012), eksponert for pakistanske språk hjemme – enten parallelt med norsk, eller med et varierende tidspunkt for introduksjon til norsk. Det ville derfor ikke vært helt uventet om dominansen hadde gått i den andre retningen – i det minste kunne man kanskje forvente at ferdighetene i urdu/panjabi og norsk var noe jevnere.

For det andre synes ikke S1-skåren i seg selv tilstrekkelig forklart; den er som vi ser signifikant lavere enn S2-skåren ($p = .000$, $d = 1.77$), og da selvsagt tilsvarende enda mye lavere enn gjennomsnittet for enspråklige ($d = 3.71$). For en billedliggjøring av forskjellen til enspråklige, kan vi se til tredelingen av BPVS-urdu/panjabi og sammenlikne med

aldersekvivalenter for enspråklig norske (fra normeringsstudien til Lyster et al., 2010). I grupperingene av BPVS-urdu/panjabi skårer den svakeste gruppen $M = 17$ og mellomgruppen $M = 27$. Sammenliknet med aldersekvivalenter for enspråklig norske er dette til dels betydelig lavere enn gjennomsnittet for norske 3;0–3;5-åringer ($M = 31$; Lyster et al., 2010). Den sterkeste urdu/panjabi-gruppen ($M = 42$) ligger på nivå med norske 4;0–4;5-åringer ($M = 43$; Lyster et al., 2010). Det er selvsagt svært kunstig å sammenlikne tospråklig kompetanse med enspråklig kompetanse på denne måten, men det gir et talende uttrykk for lavhetsgraden av resultatene på BPVS-urdu/panjabi. Normeringen for *minoritetsspråk* (Monsrud et al., 2010) gir også et sammenlikningsgrunnlag for S1-skåren, men da *i forhold til andre minoritetsspråklige*. Vi så over at S1-resultatet for de andre minoritetsspråksgruppene (albansk, somalisk, tyrkisk og vietnamesisk) *sett under ett* var 67 (Monsrud et al., 2010). Når dette sammenliknes med BPVS-urdu/panjabi for barn med pakistansk bakgrunn i min studie ($M = 28$), ser det ut til at barn med pakistansk bakgrunn sammenliknet med flere andre minoritetsspråksgrupper skiller seg ut med et spesielt lavt S1-ordforråd.

For det tredje: En distribusjon, og dermed øket totalvokabular, skulle vært forventet som følge av generelle funksjonsaspekter dersom språkene brukes i ulike domener. I lys av aldersaspekter kunne vi kanskje også forvente en distribusjon dersom utvalget var mer rendyrket andrespråklærere, og dermed lærte ordforråd i mer distinkt atskilte domener med S1 hjemme og S2 utenfor hjemmet. Dette gjelder antakelig også for litt under halvparten av utvalget. Allikevel viser ikke resultatene noen vesentlig økning i totalordforråd for gruppen *sett under ett*.

En akseptabel forståelse av disse resultatene synes altså ikke å finnes i generelle tospråksaspekter eller aldersaspekter ved barnas språksituasjon. Sagt på en annen måte: Dersom resultatene hadde vist at barn med pakistansk bakgrunn hadde høyest ordforråd på urdu/panjabi, men samtidig lavere ordforråd i norsk enn enspråklig norske og et jevnt distribuert ordforråd, kunne vi kanskje stoppet her – dette ville synes forståelig i lys av helt overordnet tospråklighetskunnskap om ferdighets-, funksjons- og aldersaspekter. Resultatene viser imidlertid ikke dette, og vi skal derfor i det følgende se om disse andre resultatene kan forstås i lys av kjente *minoritetsspråklige* aspekter ved barnas språksituasjon.

Kontekst: Minoritetsaspekter ved barnas språksituasjon

Hjemmespråksaspekter

Selv om pakistanske språk brukes i familier med pakistansk bakgrunn, vet vi at også norsk ble brukt i ganske stor grad i mange av familiene allerede da barna var fem år: Foreldrene brukte ikke norsk noe nevneverdig mindre enn urdu (Karlsen & Lykkenborg, 2012). Det er grunn til å tro at slike språkbruksmønstre vil kunne føre til at barn i familier med pakistansk bakgrunn kan utvikle lavere kompetanse i minoritetsspråket, fordi andre studier har vist at i familier hvor en eller begge foreldrene bruker majoritetsspråket i tillegg til minoritetsspråket er det mindre sjans for at barna utvikler en sterk kompetanse i minoritetsspråket (De Houwer, 2007; Sirén, 1995). Dersom foreldrene har kompetanse i majoritetsspråket, er det også en viss mulighet for at barna velger å bruke dette språket med foreldrene, selv om de også kunne ha brukt minoritetsspråket (Sirén, 1995). Bedre ferdighet i S2 enn S1 vil dermed kunne være en konsekvens av slike hjemmespråksforhold. Dette ville for øvrig være helt i henhold til komplementærprinsippet (Grosjean, 2008) – her spesifisert til hjemmedomenet.

Det har også gått en del tid siden dataene om språkbruk i hjemmet ble samlet inn, og situasjonen kan ha endret seg siden da. Det kan tenkes at bildet har endret seg i retning av enda mer bruk av norsk – dette vet vi ikke, men det er ikke usannsynlig i lys av Mancilla-Martinez & Kieffer (2010) og Sirén (1995) som har vist at språkbruksmønstre i minoritetsfamilier er dynamiske og ofte endrer seg i retning av øket bruk av majoritetsspråket etter hvert som barna tilegner seg dette mer og mer i sammenhenger utenfor hjemmet. Dette kan i tilfellet ytterligere forklare en høyere S2 enn S1 i seksårsalder.

Videre har panjabi, som med stor sannsynlighet er de fleste personer med pakistansk bakgrunns egentlige morsmål (Thiesen, 2003b), en lav status i pakistansk-kulturell sammenheng. Dette kan kanskje bidra til at det legges mer vekt på utvikling av *urdu* enn panjabi i hjemmet – vi så at det først og fremst er urdu som brukes i språklige aktiviteter med barna, og ikke panjabi (Karlsen & Lykkenborg, 2012). Dette kan i så fall bety at flere av foreldrene ikke bruker sitt sterkeste språk aller mest med barnet, noe som igjen kanskje kan svekke betingelsene for en sterk utvikling av S1. På den andre siden er det grunn til å anta at en del av foreldrene også har utviklet god kompetanse i norsk og urdu (Karlsen & Lykkenborg, 2012), noe som skulle tilsi at det kan være godt grunnlag for språklig interaksjon også på disse språkene.

Dersom det imidlertid *er* tilfelle at også panjabi brukes til en viss grad, kan vi anta at S1 dermed for flere barn med pakistansk bakgrunn gjennom oppveksten er blitt fordelt mellom to språk. Selv om panjabi ikke brukes av mange foreldre i språklige aktiviteter med barna (Karlsen & Lykkenborg, 2012), kan det tenkes at det brukes i *andre* sammenhenger: Kanskje det vil kunne være slik at foreldrene leser for barna på urdu, mens samtalen rundt middagsbordet foregår på panjabi eller en blanding. Det er også identifisert en undergruppe hvor en eller begge foreldre snakker med barnet på panjabi, i noen tilfeller i kombinasjon med norsk og urdu (Karlsen & Lykkenborg, 2012). Vi kan følgelig ikke utelukke at denne kompleksiteten, som med stor sannsynlighet kjennetegner den faktiske språksituasjonen til mange barn med pakistansk bakgrunn, kan føre til enda lavere S1-ordforråd enn dersom hele hjemmeordforrådet hadde gått til *ett* språk. Det er imidlertid ikke en så stor språklig avstand mellom urdu og panjabi (Kristoffersen, 1980), så det er uklart hvor store konsekvenser en blandet bruk kan ha å si for utviklingen av ordforråd i seg selv.

Til sammen kan slike hjemmespråksaspekter bidra til å gi en viss forståelse av at S1 er lavere enn S2 for barn med pakistansk bakgrunn. Dette resultatet støttes for øvrig som vi har sett av Lervåg og Aukrust (2010) og Monsrud et al. (2010) – begge studier viste lavere S1 enn S2 for barn med pakistansk bakgrunn.

Hjemmespråksforhold kan kanskje også kaste lys over at totalordforrådet knapt øker i forhold til norskordforrådet: I en hjemmespråkskontekst vil vi ikke kunne forvente en *stor* forskjell med hensyn til hvilke ord som kjennes på norsk, og hvilke ord som kjennes på urdu/panjabi, i og med at språkene antakelig brukes i en kombinasjon på mer eller mindre de samme områdene innenfor hjemmet. Sagt på en annen måte ville kanskje mer utpreget og rendyrket bruk av S1 i hjemmet med større sannsynlighet gi et mer distribuert, og dermed høyere, totalordforråd.

Samfunnsaspekter

Videre kan vi anta at barn med pakistansk bakgrunn befinner seg i en såkalt immersjon-kontekst: Hjemmespråket vektlegges med stor sannsynlighet ikke systematisk i verken barnehage eller skole. Resultatene i denne masterstudien viser sammenliknet med Melings (2011) resultater at urdu/panjabi-ordforrådet ikke har økt så mye siden barna var fem år, mens norskordforrådet har økt i betydelig større grad. Det blir naturlig å først og fremst se til samfunnsaspekter for en forståelse av disse resultatene: Siden femårsalder vet vi at alle barna har gått i barnehage og har begynt på skole. Dette innebærer at barna har vært omgitt av mye

norsk, noe som kan bidra til en forståelse av dette utviklingsmønsteret for S1 og S2 – selvsagt i samspill med hjemmespråksforhold som vist over. Vi så tidligere noen eksempler på studier som viser at ensidig vektlegging av majoritetsspråket i opplæringsammenheng fører til lavere kompetanse i hjemmespråket (Verhallen & Schoonen, 1998; Oller & Eilers, 2002). Dette underbygger antakelsen om at barn med pakistansk bakgrunn kan få en lavere S1-utvikling dersom opplæringen primært foregår på norsk.

I samfunnsforhold kan vi kanskje også finne en mer *indirekte* forklaring på et lavt S1-ordforråd. For det første har panjabi en lav status i pakistansk-kulturell sammenheng (Thiesen, 2003a): For eksempel er det urdu som brukes i norsk-pakistanske moskeer og i NRKs programmer for pakistanere (Thiesen, 2003a). For det andre vet vi at urdu/panjabi har en svak stilling i den overordnede norske samfunnskonteksten (Engen & Kulbrandstad, 2004). S1 har derfor lav status på en dobbel måte for barn med pakistansk bakgrunn; først panjabi i pakistansk-kulturell sammenheng, og deretter urdu og panjabi i norsk-kulturell sammenheng. Slik maktubalanse er viktig å ta med i betraktning når en lav S1-skår (og høyere S2-skår) skal forstås, siden dette i praksis kan svekke betingelsene for utvikling av S1 fordi bruksområdene indirekte begrenses av denne dobbelt lave statusen.

Generasjonsaspekter

Vi vet at en relativt stor andel personer med pakistansk bakgrunn i Norge er norskfødte med innvandrerforeldre – som vi så over er andelen andregenerasjonsinnvandrere desidert høyest blant personer med pakistansk bakgrunn sammenliknet med andre minoritetsspråksgrupper i Norge (Andreassen & Dzamarija, 2008). Vi vet også at barna i denne studien i hovedsak er andregenerasjonsinnvandrere, og rundt 10 % er også tredjegenerasjonsinnvandrere (Karlsen & Lykkenborg, 2012). Som vi så over er det ikke uvanlig at innvandrede minoritetsspråksgrupper, etter en overgangsperiode med bruk av to språk, over generasjoner kan skifte til å bruke majoritetsspråket mest (Romaine, 1995; Hurtado & Vega, 2004). Det er derfor nærliggende å anta at dette også kan skje for gruppen med pakistansk bakgrunn i Norge – kanskje er et svært lavt S1-ordforråd en indikasjon på nettopp dette. Her må vi imidlertid, som alltid ellers, huske at gjennomsnittresultater åpenbart skjuler en individuell variasjon. Noe av denne variasjonen er belyst gjennom grupperingen av utvalget; noen har et nærmest ikke-eksisterende S1-ordforråd slik dette måles med BPVS (*BPVS-urdu/panjabi Lav*; $M = 17$), mens andre igjen er på et høyere nivå (*BPVS-urdu/panjabi Høy*; $M = 42$). I og med at barna overveiende er andregenerasjonsinnvandrere, kan man kanskje heller ikke forvente at en slik

generasjonseffekt ennå har gitt fullt utslag, selv om et skifte kan skje relativt raskt fra den ene generasjonen til den neste (Hurtado & Vega, 2004). Imidlertid vil et skifte kanskje heller aldri komme for denne gruppen, i og med at bindingen til pakistanske språk og kultur er relativt sterk (Andreassen & Dzamarija, 2008). Vi kan anta at en sterk tradisjon for familiegjenforening (Andreassen & Dzamarija, 2008; Karlsen & Lykkenborg, 2012) bidrar til å motvirke, eller forsinke et skifte. Vi kan stille et hypotetisk spørsmål: Ville S1-ordforrådet gjennomgående vært enda lavere dersom ikke personer med pakistansk bakgrunn hadde hatt en sterk binding til pakistansk språk og kultur? Det er vanskelig å vite hvor stort utslag denne bindingen kan gi, men tilsvarende forhold er vist i andre studier å ha en viss innvirkning (Hurtado & Vega, 2004; Romaine, 1995).

Vi ser at det tegner seg opp et kompleks av samvirkende aspekter i barnas språksituasjon som kan gi en god forståelse av at barna skårer lavere på S1 enn på S2. I lys av denne totaliteten av hjemmespråksforhold, samfunnsforhold, og generasjonsforhold som preger språksituasjonen til barn med pakistansk bakgrunn, er det etter mitt skjønn grunnlag for å spørre om den lave BPVS-urdu/panjabiskåren er en indikasjon på at mange av barna er i en subtraktiv tospråklig utvikling. Det er heller ikke utenkelig at gruppen som helhet kan være på vei mot et generelt språkskifte, selv om det også er sannsynlig at overgangsfasen til samfunnsspråket vil trekkes ut, og kanskje vedvare, i og med at det er sterke bånd til pakistansk språk og kultur.

Det at totalordforrådet knapt øker, må sies å være noe oppsiktsvekkende gitt en forventning fra generell tospråksteori om distribusjon, spesifisert gjennom distribusjonshypotesen (Oller & Eilers, 2002). Det er imidlertid, som vi så, god grunn til å tro at det at både norsk og urdu/panjabi brukes *innenfor hjemmedomenet*, kan gi et slikt resultat, i og med at erfaringene i de ulike språkene derfor ikke blir så ulike. Her må det i tillegg bemerkes at Oller et al. (2007) testet distribusjonshypotesen på tospråklige spansk- og engelsktalende barn i Miami, Florida. Der er den spanske minoriteten stor, og situasjonen er heller ikke direkte sammenliknbar med norske forhold med hensyn til minoritetsspråkets status i samfunnet for øvrig: Blant annet er det gjennomført storskala tospråklige opplæringsprogrammer i denne regionen (Oller & Eilers, 2002), noe som indikerer et annet forhold til minoritetsspråk enn hva som må sies å være tilfelle i Norge. Kanskje kan dette bidra til andre hjemmespråksforhold enn vi har sett i denne studien: At foreldrene legger mer vekt på å holde minoritetsspråket i hevd hjemme enn hva som ville være tilfelle i en norsk kontekst, der foreldrene kanskje heller oppfatter nøkkelen til suksess for barna til å være god kompetanse i skole- og samfunnsspråket norsk

og dermed i større grad legger vekt på å bruke norsk hjemme. At slike eventuelle kontekstuelle forskjeller kan bidra til ulike resultater med hensyn til tospråklig kompetanse, skulle være tilstrekkelig sannsynliggjort over.

I forlengelsen av dette er det derfor spesielt interessant at vi også finner forskjeller mellom barn med pakistansk bakgrunn og flere andre minoritetsspråksgrupper i Norge på viktige ordforrådskjennetegn, blant annet nettopp på størrelsen av totalordforrådet (Monsrud et al., 2010). Vi så i tillegg at flere minoritetsspråksgrupper har bedre ordforrådskunnskap i hjemmespråket enn i norsk, men dette gjelder ikke for barn med pakistansk bakgrunn (Monsrud et al., 2010), noe resultatene i denne masterstudien viser enda mer eksplisitt. Kanskje noe av årsaken til denne forskjellen til andre minoritetsspråklige i Norge kan være å finne nettopp i et delt S1 (både urdu og panjabi), den dobbelt lave S1-statusen som sannsynliggjort over, i tillegg til en vesentlig større andel andregenerasjonsinnvandrere i befolkningen med pakistansk bakgrunn enn i minoritetsbefolkningen for øvrig.

Sosioøkonomiske aspekter

Til nå har vi utelukkende sett på det som i vid forstand kan regnes som språklig-sosiale kontekster for fortolkning. I lys av hva vi vet om sammenhenger mellom sosioøkonomisk status og ordforråd, er det naturlig å også ta sosioøkonomiske aspekter med i betraktning når vi ser at skårene ligger under et enspråksutvalg som samtidig er vist å ha høyere SØS – særlig så vi at dette gjaldt for mors utdanningsnivå og antall bøker for barn i hjemmet. Riktignok er dette få variabler, men disse er samtidig vist å samvariere med ordforrådskunnskap (Hart & Risley, 1995; Dollaghan et al., 1999). Oller & Eilers (2002) har også vist at sosioøkonomiske forhold kan påvirke språkkompetanse selvstendig innenfor en minoritetsspråklig kontekst. Det virker derfor sannsynlig at også lavere SØS kan ha bidratt til vokabularavstanden mellom tospråklige og enspråklige i min studie. I lys av drøftingen over mener jeg imidlertid at vokabularavstanden til enspråklige med stor sannsynlighet hadde vært tilstede uansett, men lav SØS kan kanskje bidra til en ytterligere forståelse av forholdet.

Det kan også tenkes at lav SØS kan ha bidratt til lavere S1-ordforråd. Imidlertid – da vi vet at minoritetsspråklige i Norge sett under ett har lavere SØS (Bakken 2003), og samtidig også vet at de andre minoritetsspråksgruppene i normeringsstudien skåret høyere enn barn med pakistansk bakgrunn på sine respektive hjemmespråk (Monsrud et al., 2010), vil jeg hevde at SØS ikke *alene* kunne gi en tilfredsstillende forklaring på lavt S1-ordforråd hos barn med

pakistansk bakgrunn. Det sosioøkonomiske aspektet kan derfor heller ses som et nødvendig aspekt i en totalvurdering fordi det kan gi en ytterligere forståelse av lave skårer, men ikke som et tilstrekkelig forklaringsaspekt i seg selv.

7.3 Oppsummering

Vi så over at kjennetegn ved barn med pakistansk bakgrunns reseptive breddeordforråd kan oppsummeres i følgende hovedfunn:

- I. Barn med pakistansk bakgrunn har signifikant lavere norskordforråd enn enspråklig norske. En beregning av totalordforråd endrer ikke dette forholdet til enspråklige.
- II. Barn med pakistansk bakgrunn har signifikant ulik ordforrådsstørrelse på sine to språk. S2-ordforrådet er størst. Fra femårsalder har S2-ordforrådet økt betraktelig mer enn S1-ordforrådet.

Formålet med den foregående drøftingen har vært å undersøke hvordan disse kjennetegnene kan forstås i lys av aspekter ved barnas språksituasjon. For å sammenfatte denne drøftingen, skal jeg nå oppsummere hovedpoengene kort.

Vokabularavstanden mellom barn med pakistansk bakgrunn og enspråklig norske ser ut til å kunne forstås godt i lys av overordnet kunnskap om ferdighets-, funksjons- og aldersaspekter ved tospråklighet – tiden og erfaringene må deles mellom to språk, og læring av norsk har også startet på et senere tidspunkt for mange. Det er også mulig at stor språkstrukturell og språkkulturell avstand mellom norsk og urdu/panjabi, samt lavere SØS for de tospråklige barna, kan bidra til denne vokabularavstanden.

For å forstå at det er *norskordforrådet* som er størst, og at norskordforrådet har økt betydelig mer enn urdu/panjabiordforrådet siden femårsalder, må vi se til et mer komplisert bakteppe av kunnskap om barnas *minoritetsspråksituasjon*. Disse resultatene synes forståelige dersom vi ser alle aspekter i barnas språksituasjon som taler for et sterkere samfunnsspråk *under ett* – for å oppsummere i korte grovtrekk: Norsk brukes mye hjemme i tillegg til urdu og panjabi, det er norsk som brukes i samfunnet for øvrig, barnas hjemmespråk kan sies å ha dobbelt lavstatus og barna er andregenerasjonsinnvandrere. Alle disse forholdene kan hver for seg til en viss grad, og til sammen i stor grad, gi en god forståelse av at S2-ordforrådet er størst og

har økt mest siden femårsalder. Disse aspektene kan også bidra til en forståelse av at S1-ordforrådet er spesielt lite for mange av barna.

Det er litt vanskeligere å finne en klar forståelse av hvorfor *totalordforrådet* ikke øker. Antakelig er noe av forklaringen å finne i de samme mekanismene som begrenser bruk og utvikling av S1. Det at norsk brukes hjemme kan som vi så kanskje føre til at barna lærer mange av de samme ordene på begge språk, og dermed får få singelord på urdu/panjabi.

7.4 Noen generelle og pedagogiske implikasjoner

Jeg finner det viktig å stille spørsmål ved betydningen av resultatene, tross at dette strengt tatt innebærer at jeg delvis går utover den teoretiske referanserammen og problemstillingen. Før jeg avslutter, skal vi derfor ta et lite blick opp fra resultatene og vurderingen av disse, og se på mulige konsekvenser av det oppgaven viser. Har det noe å si at barn med pakistansk bakgrunn har lite ordforråd i S1 og lavere S2 enn enspråklige norske? Og i tilfelle hva? Her må det igjen poengteres at BPVS-skårer ikke nødvendigvis avspeiler reseptivt breddeordforråd slik det fungerer i (tospråklig) praksis (Nation, 2001; Oller & Eilers, 2002). Det er ingen grunn til å tro at normalfungerende barn med pakistansk bakgrunn har noe dårligere kommunikativ kompetanse sammenliknet med andre normalfungerende barn (Oller & Eilers, 2002; Grosjean, 2008; Bialystok, 2011). Grosjeans (2008) komplementærprinsipp tydeliggjør dette: De ferdighetene som faktisk trengs til kommunikasjon i de ulike bruksdomenene for (og bare de), vil utvikles. Et lite S1-ordforråd kan på denne måten ses som en naturlig konsekvens av at språkferdigheter i S1 ikke trengs i like stor grad som språkferdigheter i S2. I et slikt perspektiv følger det ikke naturlig å anse tospråklig kompetanse som en svikt vis a vis enspråklige: Vurdering av resultatene i et tospråklig-sensitivt blick gir en forståelse av at tospråklig kompetanse er noe *kvalitativt annerledes* enn enspråklig kompetanse.

En slik forståelse må imidlertid følges direkte av en anerkjennelse av at en denne *annerledesheten* like fullt kan innebære utfordringer i utdanningssammenheng for grupper som har en annen språkbakgrunn enn den enspråklig norske majoriteten. Vi må ta på alvor den ordforrådssituasjonen en stor del av barn med pakistansk bakgrunn er i: Som vi har sett, kjennetegnes breddeordforrådet for ca. to tredjedeler av barna ved at det er under normalområdet sammenliknet med enspråklige. Når læringspråket i skolen er norsk, synes det åpenbart at dette kan få konsekvenser. For å underbygge dette kan vi se til studier som

viser at det er en sammenheng mellom ordforråd og skolefaglig utvikling: Cunningham og Stanovich (1997) fant for eksempel at ordforråd målt med PPVT i første klasse korrelerte signifikant med prestasjoner på flere viktige kunnskapsområder ti år senere. Forholdet mellom vokabular og *leseforståelse* har vært viet særlig oppmerksomhet – her har fokus blant annet vært nettopp på den doble funksjonen semantisk kompetanse har: Som kunnskap i seg selv, men også kilde til ny kunnskap. En selvforsterkende effekt, som tidlig hypotisert av Stanovich (1986), vil derfor kunne inntreffe i en vekselvirkning mellom vokabular og leseutvikling – barn som har lavt vokabular vil forstå mindre av tekstene de leser, og dermed miste en viktig kilde til videre vokabularutvikling. Også nyere forskning indikerer en sammenheng mellom vokabularstørrelse og særlig leseutvikling (se for eksempel Scarborough, 2001; Biemiller, 2003; Oulette, 2006; Lee, 2011). Lervåg og Aukrust (2010) viste spesielt at begrensninger hos tospråklige med hensyn til ordforråd, syntes tilstrekkelig til å forklare deres forsinkede leseforståelse senere.

Her skal vi imidlertid være svært forsiktig med å ”svartmale” situasjonen. For det første kjenner vi ikke helt mekanismene bak slike sammenhenger mellom leseforståelse og ordforråd, og forholdet er heller ikke nødvendigvis *direkte* (Oulette, 2006). For det andre bygger skriftspråklige ferdigheter også på kompetanse som tospråklige har vist fordeler i sammenliknet med enspråklige, for eksempel fonologisk bevissthet (Bialystok, 2001). I en gjennomgang av studier som viser sammenhenger mellom vokabularstørrelser og leseutvikling, poengterer Bialystok og Feng (2011) derfor at svakere vokabularprestasjoner hos tospråklige i forhold til enspråklige ikke nødvendigvis har så negative konsekvenser som man ellers kan få inntrykk av. For eksempel fant Bialystok og Feng at dersom man kontrollerer for *vokabularstørrelse* (altså sammenlikner tospråklige og enspråklige som har like vokabularstørrelser), går leseutviklingen ofte like bra for tospråklige som for enspråklige (Bialystok & Feng, 2011). Når dette er sagt bør ikke betydningen av lave ordforrådsstørrelser i *praksis* på noen måte undervurderes, og dette er også et hovedpoeng hos Bialystok og Feng (2011, s. 127): ”The real experience of the child in the classroom, however, requires dealing with that vocabulary deficit; children do not enjoy the privilege of the hypothetical situation that is offered by controlled statistical analyses.” Dette som Bialystok og Feng her påpeker, må sies å være et poeng av kritisk betydning: For barnet som sitter i klasserommet og skal lese for å lære, hjelper det lite at studier viser at barn med sammenliknbare ordforrådsstørrelser ikke presterer noe bedre enn dem. På samme måte hjelper det lite at dette barnets ordforrådskompetanse kan ses som et naturlig resultat av en tospråklig eller

minoritetsspråklig kontekst for språkutvikling. Resultatene i denne studien viser derfor at mange barn med pakistansk bakgrunn kan ha stor nytte av målrettede, allmennpedagogiske tiltak for ordforrådsutvikling.

8 Avsluttende kommentarer

I lys av alle aspekter som er redegjort for, fra den ytterste generelle tospråklighetskontekst til den innerste kunnskap om akkurat disse tospråklige barnas språksituasjon, kan vi si at resultatene i denne studien synes forståelige. Det å søke forståelse for ordforrådsferdigheter i barns språksituasjon blir imidlertid, dersom vi løfter blikket, som å se på to sider av samme sak – nettopp fordi språkferdigheter aldri utvikles isolert fra kontekst. Slik sett blir det naturlig å finne samsvar mellom språkferdigheter og språksituasjon – ”uansett” hvilket resultat man finner, vil det være aspekter i språksituasjonen som kan sannsynliggjøre nettopp *det* resultatet: Som man roper i skogen får man svar, kan vi kanskje si. Allikevel mener jeg at drøftingen viser at en god og informert forståelse av tospråklig ordforrådskompetanse bare er mulig dersom man tar *flere* tospråksdimensjoner i betraktning. Ferdighets- og funksjonsdimensjonen, aldersdimensjonen og kontekstdimensjonen kan til sammen vise at sosiale kontekster gir et sentralt bidrag til barns språkkompetanse, og at språkkompetansen derfor kan variere når kontekstene varierer – fra barn til barn, fra gruppe til gruppe. Når disse tospråksdimensjonene i tillegg utfylles av kunnskap om den aktuelle tospråksgruppen, gir dette et grunnlag for en mer helhetlig forståelse av status: Barna har naturlig tilegnet seg, og tilegner seg fortsatt, ordforråds kunnskap i en grunnleggende sett tospråklig kontekst; *tospråklig ordforråd i tospråklig kontekst*.

På den andre siden har vi samtidig sett at ordforrådsstørrelsene for mange barn med pakistansk bakgrunn er vesentlig lavere enn gjennomsnittet for enspråklige norske, og at dette må tas på alvor som en risikofaktor med hensyn til skolefaglig utvikling. Et bilde skisseres dermed opp av barn i ”ordforråds-limbo”: Mellom en unik, og i seg selv verdifull, *tospråklig kompetanse* og samfunnets generelle krav til kompetanse i *majoritetsspråket*. Det blir viktig å bemerke avslutningsvis at den konteksten disse barna skal utdanne seg i, i et overordnet blikk ikke kan betraktes som tospråklig. Tittelen på oppgaven kunne i dette perspektivet kanskje derfor like gjerne vært ’*tospråklig ordforråd i enspråklig kontekst*’. I lys av et slikt perspektiv bør resultatene i denne studien føre til allmennpedagogiske tiltak som kan bedre betingelsene for barn med pakistansk bakgrunn *også* i slike enspråklige kontekster.

I et mer klinisk, spesialpedagogisk perspektiv, kan det til slutt bemerkes at den overrepresentasjonen man ser av tospråklige barn i spesialundervisning, kan tyde på at den kunnskapen vi faktisk har om hva som kjennetegner kontekster for tospråklig utvikling, ikke

brukes i tilstrekkelig grad når tospråklig kompetanse skal vurderes. Når det gjelder vurdering av tospråklig kompetanse må det imidlertid understrekes at all teori og empiri må brukes, men brukes med skjønn – ingen språkgrupper eller enkeltindivider har akkurat den samme språksituasjonen. Alle typer forhold rundt den aktuelle språkgruppen og enkeltindividets språkmiljø må derfor tas i nøye betraktning, slik at vurderinger gjøres til det beste for den aktuelle språkgruppen og i siste, eller kanskje *første og viktigste* instans, for det enkelte tospråklige barn.

Videre kunnskapsutvikling

Denne studien har gitt et lite, men like fullt informativt og viktig, innblikk i språkkompetansen til barn med pakistansk bakgrunn. Videre undersøkelser av mer *sammensatt* tospråklig ordforrådskompetanse kan bidra til å belyse tospråklig barns kompetanse på en mer nyansert måte. Det ville også være svært gunstig å undersøke sammenhenger mellom språksosial situasjon og utvikling av ferdigheter i en sammensatt semantisk indeks *over tid*, for å få et tydeligere bilde av på hvilken måte og i hvor stor grad vokabularutvikling påvirkes av ulike kontekstuelle faktorer for tospråklige barn i Norge. Dette vil tilføre verdifull informasjon om tospråklig kompetanse i en norsk kontekst, og dermed samtidig gi et grunnlag for mer kvalifiserte vurderinger av hvordan pedagogiske tiltak kan rettes inn på en mest mulig hensiktsmessig måte i et samfunn hvor enspråklige normer utvilsomt setter standard i en vid skolefaglig og samfunnsmessig kontekst.

Register

FIGUR 1:	Hvor gammelt var barnet da det ble introdusert for norsk språk	s. 29
FIGUR 3:	Mors høyeste utdanning	s. 30
FIGUR 4:	Fars høyeste utdanning	s. 31
FIGUR 5:	Antall billedbøker for barn i hjemmet	s. 32
TABELL 1:	Resultater BPVS	s. 39
TABELL 2:	Resultater BPVS for grupperinger av tospråksutvalget	s. 41
TABELL 3:	Effektstørrelser	s. 45
VEDLEGG 1:	BPVS-instruksjon	s. 79
VEDLEGG 2:	BPVS-manual og de tre første oppgavesettene	s. 80
VEDLEGG 3:	Independent-samples t-test i tabell	s. 81
VEDLEGG 4:	Paired-samples t-test i tabell	s. 82

Litteraturliste

- Alba, R., Logan, J., Lutz, A., & Stults, B. (2002). Only English by the third generation? Loss and preservation of the mother tongue among the grandchildren of contemporary immigrants. *Demography*, 39(3), 467-484.
- Andreassen, K. K & Dzamarija, M. T. (2008) Innvandrere og norskfødte med innvandrereforeldre. I G. Daugstad (Red.), *Innvandring og innvandrere* (s. 13-54). Oslo: Statistisk sentralbyrå. Hentet fra http://www.ssb.no/02/sa_innvand/sa103/sa103.pdf
- Aukrust, V. G. (2005). *Tidlig språkstimulering og livslang læring. En kunnskapsoversikt*. Oslo: Utdannings- og forskningsdepartementet. Hentet fra http://www.regjeringen.no/upload/kilde/ufd/rap/2005/0032/ddd/pdfv/267903-aukrust_sprakstimulering.pdf
- Baca, L. M. & Cervantes, T. H. (2004). *The bilingual special education interface*. New Jersey: Pearson.
- Baker, C. (2011). *Foundations of bilingual education and bilingualism*. Bristol: Multilingual Matters.
- Bakken, A. (2003). *Minoritetsspråklig ungdom i skolen. Reproduksjon eller sosial mobilitet?* Norsk institutt for forskning om oppvekst, velferd og aldring (NOVA), Utdannings- og forskningsdepartementet. Rapport 2003(15). Hentet fra http://www.nova.no/asset/2836/1/2836_1.pdf
- Bialystok, E. (2001). *Bilingualism in development. Language, literacy and cognition*. Cambridge: Cambridge University Press.
- Bialystok, E. & Feng, X. (2011). Language proficiency and its implications for monolingual and bilingual children. I A. Y. Durgunoglu & C. Goldenberg (Red.), *Language and literacy development in bilingual settings* (s. 121-138). New York: The Guilford Press
- Bialystok, E., Luk, G., Peets, K. F. & Yang, S. (2010). Receptive vocabulary differences in monolingual and bilingual children. *Language and Cognition*, 13(4), 525-531.
- Biemiller, A. (2003). Vocabulary: Needed if more children are to read well. *Reading Psychology*, 24(3-4), 323-335.
- Christophersen, K.-A. (2012). *IBM/AMOS. Databehandling og statistisk analyse*. Oslo: Akademika forlag.

- Cummins, J. (1994). The acquisition of English as a second language. I K. S. Spangenberg-Urbschat & R. Pritchard (Red.), *Kids come in all languages: Reading instruction for ESL students* (s. 36-62). Delaware, US: International Reading Association.
- Cunningham A. E, & Stanovich, K. E. (1997). Early reading acquisition and its relation to reading experience and ability 10 years later. *Developmental Psychology*, 33(6), 934-945.
- Dalen, M. (2011). *Intervju som forskningsmetode. En kvalitativ tilnærming*. Oslo: Universitetsforlaget.
- De Houwer, A. (2007). Parental language input patterns and children's bilingual use. *Applied Psycholinguistics*, 28(3), 411-424.
- De Vaus, D. A. (2002). *Surveys in social research. An introduction*. London: Routledge.
- Dollaghan, C. H, Campbell, T. F., Paradis, J. L, Feldman, H. M. et al. (1999). Maternal education and measures of early speech and language. *Journal of Speech, Language, and Hearing Research*, 42(6), 1432-1443.
- Dunn, L. M., Dunn, L. M., Whetton, C., & Burley, J. (1997). *The British picture vocabulary scale. Second edition*. London: NFER-Nelson Publishing Company.
- Edwards, V. (2009). Vocabulary, education and diversity. I: B. Richards, M. H. Daller, D. D. Malvern, P. Meara, J. Milton & J. Treffers-Daller (Red.), *Vocabulary studies in first and second language acquisition* (s. 59-73). Hampshire, UK: Palgrave.
- Eilers, R. E., Pearson B. Z. & Cobo-Lewis, A. B. (2006). Social factors in bilingual development: The Miami experience. I P. McCardle (Red.), *Childhood bilingualism: Research on infancy through school age* (s. 68-90). Clevedon, UK: Multilingual Matters Ltd.
- Engen, T. O., & Kulbrandstad, L. A. (2004). *Tospråkklighet, minoritetspråk og minoritetsundervisning*. Oslo: Gyldendal Akademisk.
- Espenakk, U., Klem, M., Rygvold, A. L., Ottem, E. & Saltveit, V. (Red.) (2007). *Språkveilederen*. Oslo: Bredtvet kompetansesenter.
- Gall, M. D., Gall, J. P. & Borg, W. R. (2007). *Educational research. An introduction*. Boston: Pearson Education.
- Golden, A. (2009). *Ordforråd, ordbruk og ordlæring*. Oslo: Gyldendal Akademisk.
- Goldenberg, C., Reese, L. & Rezaei, A. (2011). Contexts for language and literacy development among dual-language learners. I A. Y. Durgunoglu & C. Goldenberg (Red.), *Language and literacy development in bilingual settings* (s. 3-25). New York: The Guilford Press.

- Grosjean, F. (1985). The bilingual as a competent but specific speaker-hearer. *Journal of Multilingual and Multicultural Development*, 6(6), 467-477.
- Grosjean, F. (1997). The bilingual individual. *Interpreting*, 2(1-2), 163-187.
- Grosjean, F. (2008). *Studying bilinguals*. Oxford: Oxford University Press.
- Hammer, C. S., Miccio, A. W. & Rodriguez, B. L. (2004). Bilingual language acquisition and the child socialization process. I B. A. Goldstein (Red.), *Bilingual language development and disorders in Spanish-English speakers* (s. 22-50). Baltimore: Paul Brooks Publishing.
- Hart, B. & Risley, T. R. (1995). *Meaningful differences in the everyday experience of young American children*. Baltimore: Paul Brooks Publishing.
- Hurtado, A. & Vega, L. A. (2004). Shift happens: Spanish and English transmission between parents and their children. *Journal of Social Issues*, 60(1), 137-155.
- Institutt for spesialpedagogikk. (2010a). *Child Language and Learning* (nettside). Hentet fra <http://www.uv.uio.no/english/research/groups/cll/index.html> .
- Institutt for spesialpedagogikk. (2010b). *Language Development in Minority Language Children* (nettside). Hentet fra <http://www.uv.uio.no/isp/english/research/projects/minority1/index.html>
- Institutt for spesialpedagogikk. (2010c). *BPVS II* (nettside). Hentet fra <http://www.hf.uio.no/iln/tjenester/kunnskap/sprak/sprakvansker/kartlegging/bpvs/>
- Karlsen, J. & Lykkenborg, M. (2012). Språkbruksmønstre i norsk-pakistanske familier. *Norsk Tidsskrift for Andrespråksforskning*, 7(1), 53-83.
- Kleven, T. A. (2002a). Begrepsoperasjonalisering. I T. Lund (Red.), *Innføring i forskningsmetodologi* (s. 141-183). Oslo: Unipub forlag.
- Kleven, T. A. (2002b). Ikke-eksperimentelle design. I T. Lund (Red.), *Innføring i forskningsmetodologi* (s. 265-286). Oslo: Unipub forlag.
- Kristiansen, K. (1980). Hindi - urdu - og pandsjabi - norsk. Kontrastiv grammatikk. I: A. Hvenekilde (Red.), *Mellom to språk: 4 kontrastive språkstudier for lærere* (s. 81-113). Oslo: Cappelen.
- Kunnskapsdepartementet. (2006). ... og ingen sto igjen. *Tidlig innsats for livslang læring*. St.meld. nr 16 (2006–2007). Oslo: Kunnskapsdepartementet. Hentet fra <http://www.regjeringen.no/Rpub/STM/20062007/016/PDFS/STM200620070016000DDDPDFS.pdf>

- Kunnskapsdepartementet. (2010). *Mangfold og mestring. Flerspråklige barn, unge og voksne i opplæringsystemet*. NOU 2010:7. Oslo: Kunnskapsdepartementet. Hentet fra <http://www.regjeringen.no/pages/10797590/PDFS/NOU201020100007000DDDPDFS.pdf>
- Lee, J. (2011). Size matters: Early vocabulary as a predictor of language and literacy competence. *Applied Psycholinguistics*, 32(1), 69-92.
- 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.
- Lervåg, A. & Melby-Lervåg, M. (2009). Muntlig språk, ordavkoding og leseforståelse hos tospråklige: En sammenfatning av empiriske studier. *Norsk Pedagogisk Tidsskrift*, 2009(4), 264-279.
- Lindberg, I. (2008). Med andre ord i bagasjen. I L. Bjar (Red.), *Det er språket som bestemmer! Læring og språkutvikling i grunnskolen* (s. 48-71). Bergen: Fagbokforlaget
- Lund, T. (2002). Metodologiske prinsipper og referanserammer. I T. Lund (Red.), *Innføring i forskningsmetodologi*. Oslo: Unipub forlag.
- Lyster, S. A. H., Horn, E. & Rygvold, A. L. (2010). Ordforråd og ordforrådsutvikling hos norske barn og unge. Resultater fra en utprøving av British Picture Vocabulary Scale II, Second Edition (BPVS II). *Spesialpedagogikk*, 2010(9), 35-43.
- Mancilla-Martinez, J. & Kieffer, M. J. (2010). Language minority learners' home language is dynamic. *Educational Researcher*, 39(7), 545-546.
- Meling, M. (2011). *Tospråklighet og vokabular. En kvantitativ undersøkelse av språkmiljøets betydning for vokabular hos tospråklige førskolebarn* (Masteroppgave). Oslo: Universitetet i Oslo.
- Milton, J. (2009). *Measuring second language vocabulary acquisition*. Bristol, UK: Multilingual Matters.
- Monsrud, M. B., Thurmann-Moe, A.-E. & Bjerkan, K. M. (2010). Minoritetsspråklige barns ordforråd og ordforrådsutvikling. *Spesialpedagogikk*, 2010(9), 44-53.
- Nation, I. S. P. (2001). *Learning vocabulary in another language*. Cambridge, UK: Cambridge University Press.
- NESH. (2009, 19. mai). *Barns krav på beskyttelse*. Hentet fra <http://www.etikkom.no/no/Forskningsetikk/Etiske-retningslinjer/Samfunnsvitenskap-jus-og-humaniora/B-Hensyn-til-personer-5---19/12-Barns-krav-pa-beskyttelse/>

- NESH. (2009, 20. mai) *Hensynet til utsatte grupper*. Hentet fra <http://www.etikkom.no/no/Forskningsetikk/Etiske-retningslinjer/Samfunnsvitenskap-jus-og-humaniora/C-Hensyn-til-grupper-og-institusjoner-20---26/22-Hensynet-til-utsatte-grupper/>
- Oller, D. K. & Eilers, R. E. (2002). *Language and literacy in bilingual children*. Clevedon, UK: Multilingual Matters.
- Oller, D. K., Pearson, B. Z., & Cobo-Lewis, A. B. (2007) Profile effects in early bilingual language and literacy. *Applied Psycholinguistics*, 28(2), 191-230.
- Oulette, 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), 554-566.
- Pallant, J. (2010). *SPSS survival manual. A step by step guide to data analysis using SPSS*. Berkshire: Open University Press.
- Romaine, S. (1995). *Bilingualism*. Oxford, UK: Blackwell.
- Rommetveit, R. (1972). *Språk, tanke og kommunikasjon. Ei innføring i språkpsykologi og psykolingvistikk*. Oslo: Universitetsforlaget.
- Scarborough, H. S. (2001). Connecting early language and literacy to later reading (dis)abilities. Evidence, theory and practice. I S. B. Neuman & D. K. Dickinson (Red.), *Handbook of early literacy research* (s. 97-110). New York: Guilford Press.
- Selj, E. (2008). Minoritetslevene, språket og skolen. I E. Selj & E. Ryen (Red.), *Med språklige minoriteter i klassen* (s. 14-46). Oslo: Cappelen Akademiske.
- Sirén, U. (1995). Minority language transmission in early childhood. *International Journal of Early Years Education*, 3(2), 75-84.
- Snow, C. E., & Kim, Y-S. (2007). Large problem spaces. The challenge of vocabulary for English language learners. I R. K. Wagner, A. E. Muse & K. R. Tannenbaum (Red.), *Vocabulary acquisition. Implications for reading comprehension*. (s.123-139) London: The Guilford Press.
- Stahl, S. A. & Nagy, W. E. (2006). *Teaching word meanings*. New Jersey: LEA publishers.
- Stanovich, K. E. (1986). Matthew effects in reading: Some consequences of individual differences in the acquisition of literacy. *Reading Research Quarterly*, 21(4), 360-407.
- Statistisk sentralbyrå. (2002). *Elevar frå språklege minoritetar, etter morsmål*. Hentet fra <http://www.ssb.no/utgrs/arkiv/tab-2002-06-04-11.html>
- Sveen, A. (1996). Semantikk. I R. T. Endresen, H. G. Simonsen & A. Sveen (Red.), *Innføring i lingvistikk* (s. 43-64). Oslo: Universitetsforlaget.

- Sveen, A. (2005). Semantikk. I K. E. Kristoffersen, H. G. Simonsen & A. Sveen (Red.), *Språk. En grunnbok* (s. 64-94). Oslo: Universitetsforlaget.
- Thiesen, F. (2003a). Om språkforholdene i Pakistan I. Urdu, Pakistans nasjonalspråk. *Språknytt*, 31(3-4), s. 27-30. Hentet fra <http://www.sprakrad.no/upload/Snytt0334.pdf>
- Thiesen, F. (2003b). Om språkforholdene i Pakistan II. Pakistans regional- og lokalspråk. *Språknytt*, 31(3-4), s. 31-34. Hentet fra <http://www.sprakrad.no/upload/Snytt0334.pdf>
- Thomas, W. P. & Collier, V. (1997). School effectiveness for language minority students. *NCBE Resource Collection Series, Vol. 9*. Hentet fra http://www.thomasandcollier.com/Downloads/1997_Thomas-Collier97.pdf
- Thordardottir, E. (2011). The relationship between bilingual exposure and vocabulary input. *International Journal of Bilingualism*, 15(4), 426-445.
- Verhallen, M. & Schoonen, R. (1998). Lexical knowledge in L1 and L2 of third and fifth graders. *Applied Linguistics*, 19(4), 452-470.
- Wold, A. H. (2008). Utvikling av ordforråd med fokus på norsk som andrespråk. I E. Selj & E. Ryen (Red.), *Med språklige minoriteter i klassen* (s. 195-221). Oslo: Cappelen Akademiske.
- Wong-Fillmore, L. (1991). When learning a second language means losing the first. *Early Childhood Research Quarterly*, 6(3), 323-346.
- Øzerk, K. (2003). *Sampedagogikk. En studie av de norskspråklige og minoritetsspråklige elevenes læringsutbytte på småskoletrinnet i L97-skolen*. Vallset: Oplandske Bokforlag.
- Aarsæther, F. (2002). To språk i en tekst. Kodeveksling i samtaler mellom pakistansk-norske tiåringer (Doktoravhandling). Oslo: Universitetet i Oslo.

Vedlegg

Vedlegg 1: BPVS-instruksjon

A.

"Jeg vil at du skal se på noen bilder med meg. Se på alle bildene på denne siden (pek). Nå skal jeg si deg noe, og du skal peke på det bildet. Vi prøver."

"Pek på ball"

"Pek på hund"

B

"Pek på gråte"

"Pek på sover"

"Nå skal jeg vise deg flere bilder. Når jeg sier noe vil jeg at du skal peke på det bildet som du synes passer best. Etter hvert kan det hende at du blir usikker på hva du skal peke på. Da vil jeg at du skal se nøye på alle bildene og velge den du tror er riktig / passer best."

Vedlegg 2: BPVS-manual og de tre første oppgavesettene

ADMINISTRERING AV THE BRITISH PICTURE VOCABULARY SCALE

Det er viktig at du starter med øvingsoppgavene før du gjennomfører selve testen. Instruksjonen til øvingsoppgaven finner du i testboken.

Startalder: For et normalt fungerende barn begynner testingen med oppgavesettet som tilsvarer barnets alder. Start med det første leddet og gjennomfør alltid alle leddene i hvert oppgavesett.

Neдре grense: tilsvarer oppgavesettet der barnet har mer enn 1 feil. Dersom barnet har mer enn

1 feil gå bakover i oppgavesettene til nedre grense er fattsatt.

Øvre grense: tilsvarer oppgavesettet der barnet har 8 eller flere feil.

Registrering av svar: Sikriv ned nummeret på den illustrasjonen barnet peker på under "Svar". Hvis barnet svarer feil, sett i tillegg en strek gjennom sirkelen.

HUSK DISSE REGLENE:

- Gjennomfør alltid de 12 testleddene i hvert oppgavesett
- Nedre grense 1 eller ingen feil i oppgavesettet
- Øvre grense 8 eller flere feil i oppgavesettene

Oppgavesett 1 Alder 2 ½ - 3		Svar	
1. hund	(1) _____	_____	_____
2. baby	(2) _____	_____	_____
3. kant	(2) _____	_____	_____
4. hoppe	(4) _____	_____	_____
5. buss	(4) _____	_____	_____
6. drikk	(3) _____	_____	_____
7. traktor	(4) _____	_____	_____
8. løpspringe	(1) _____	_____	_____
9. port	(3) _____	_____	_____
10. lese	(2) _____	_____	_____
11. ku	(1) _____	_____	_____
12. trefmann	(3) _____	_____	_____
		Antall feil _____	

Oppgavesett 2 Alder 4 - 5		Svar	
13. stige	(2) _____	_____	_____
14. plante	(1) _____	_____	_____
15. sirkel	(4) _____	_____	_____
16. stearnllys	(2) _____	_____	_____
17. gjennomålig	(4) _____	_____	_____
18. rede/reir	(4) _____	_____	_____
19. dans	(1) _____	_____	_____
20. skilpadde	(4) _____	_____	_____
21. bonde	(3) _____	_____	_____
22. spindelhev	(3) _____	_____	_____
23. hals	(3) _____	_____	_____
24. piasevin	(1) _____	_____	_____
		Antall feil _____	

Oppgavesett 3 Alder 6 - 7		Svar	
25. pakke inn	(4) _____	_____	_____
26. frukt	(1) _____	_____	_____
27. lukte på	(3) _____	_____	_____
28. pil	(1) _____	_____	_____
29. læreren	(2) _____	_____	_____
30. fullt	(3) _____	_____	_____
31. punda	(4) _____	_____	_____
32. trimmeleiene	(4) _____	_____	_____
33. mynt	(2) _____	_____	_____
34. kbb	(1) _____	_____	_____
35. måle	(2) _____	_____	_____
36. skrelle	(3) _____	_____	_____
		Antall feil _____	

Vedlegg 3: Independent-samples t-test

Deskriptiv oversikt: BPVS-norsk

	N	Mean	Std. Deviation	Std. Error Mean
Tospråklig utvalg	61	51.15	13.141	1.683
Enspråklig utvalg	190	73.39	11.945	.867

Independent-samples T-test

	Levene's Test for Equality of Variances		t-test for Equality of Means						
	F	Sig.	t	df	Sig. (2-tailed)	Mean Difference	Std. Error Difference	95% Confidence Interval of the Difference	
								Lower	Upper
Equal variances assumed	.607	.437	-12.347	249	.000	-22.247	1.802	-25.796	-18.698
Equal variances not assumed			-11.755	93.952	.000	-22.247	1.893	-26.005	-18.489

Vedlegg 4: Paired-samples t-test

Deskriptiv oversikt: BPVS-norsk og BPVS-urdu/Panjabi

	Mean	N	Std. Deviation	Std. Error Mean
BPVS-norsk	51.32	60	13.185	1.702
BPVS-urdu/panjabi	28.62	60	12.188	1.573

Paired-samples T-test

	Paired Differences					t	df	Sig. (2-tailed)
	Mean	Std. Deviation	Std. Error Mean	95 % Confidence Interval of the Difference				
				Lower	Upper			
BPVS-norsk – BPVS-urdu/panjabi	22.700	15.941	2.058	18.582	26.818	11.030	59	.000