

# Betydningen av ordkunnskap for utviklingen av leseforståelse

Ellen Irén Brinchmann



Avhandling for graden PhD  
Institutt for spesialpedagogikk  
Det utdanningsvitenskapelige fakultetet

UNIVERSITETET I OSLO

2016

© Ellen Irén Brinchmann, 2016

*Doktoravhandlinger forsvart ved  
Det utdanningsvitenskapelige fakultet, Universitetet i Oslo.  
Nr. 245*

ISSN 1501-8962

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

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

# Sammendrag

I denne avhandlingen utforskes betydningen av ordkunnskap for utviklingen av leseforståelse. Temaet belyses gjennom tre empiriske studier:

I den første studien, en metaanalyse, ble relasjonen mellom vokabular i førskolealder og senere leseforståelses- og ordavkodingsferdigheter undersøkt. Resultatene fra studien viste en moderat sammenheng mellom tidlige vokabularferdigheter og leseforståelse i skolealder, mens den prediktive relasjonen mellom vokabular og ordavkodning var noe svakere. Samlet sett framstår imidlertid vokabular som en robust prediktor av leseforståelse. Dette funnet har flere praktiske implikasjoner. For det første vil man med utgangspunkt i førskolebarns vokabularferdigheter kunne identifisere individer i risiko for å utvikle lesevansker. Resultatene antyder dessuten at det kan være nyttig å arbeide med barns ordkunnskap for å forhindre at slike vansker oppstår.

Den andre studien i avhandlingen baseres også på longitudinelle observasjonsdata, men er i motsetning til metaanalysen en primærstudie. I studien utforsket vi selve utviklingen av ordkunnskap ved å undersøke forholdet mellom grammatisk og leksikalsk vekst i en kohort av barn i førskolealder. I studien ble det identifisert gjensidige forbindelser mellom barns grammatiske og leksikalske utvikling, og funnene understreker således at ordkunnskapens bidrag til leseforståelse må ses i en større språklig sammenheng. Resultatene indikerer for øvrig at ulike områder av språket kan samhandle og forsterke hverandre, noe som kan være hensiktsmessig å dra nytte av i arbeidet med barns ordkunnskap.

I likhet med studie I og II har den siste studien i avhandlingen en kvantitativ metodisk tilnærming. Fra et metodologisk perspektiv skiller likevel studien seg fra de to andre ved å bygge på data fra et kvasi-eksperiment. Mer spesifikt ble effekten av en intervensjon designet for å stimulere svake leseres ordkunnskap undersøkt. Intervensjonen hadde en bred tilnærming til undervisning i ordkunnskap, hvor forbindelsen mellom ordkunnskap og andre språkområder ble vektlagt. Etter intervensjonens slutt ble det funnet statistisk signifikante effekter av undervisningen på flere ulike språklige mål samt et standardisert mål på leseforståelse. Resultatene tyder med andre ord at arbeid med ordkunnskap kan være en god metode for å fremme elevers leseforståelsesferdigheter.

# Forord

Aller først vil jeg takke alle elevene, lærerne og skolelederne som har vært involvert i arbeidet mitt. Uten deres velvilje hadde ikke prosjektet kommet særlig langt, og jeg er dypt takknemlig for tilliten dere viste meg.

Jeg vil også rette en helt spesiell takk til min hovedveileder, Sol Lyster, for at du har holdt meg i hånda, dyttet meg i ryggen, åpnet dører og av og til kastet meg ut i bølgene. Du har alltid hatt en vanvittig tro på meg og arbeidet mitt, og det setter jeg umåtelig stor pris på. Jeg er enormt glad for at akkurat du har vært veilederen min.

Takk til alle i forskergruppen Child Language and Learning. Det har vært både inspirerende og lærerikt å være med i en gruppe bestående av så engasjerte og dyktige forskere som dere. En særskilt takk går til Bente Hagtvet som leder gruppa på en helt fabelaktig måte, til Hanne Hjetland for et ypperlig samarbeid og til Janne von Koss Torkildsen og Monica Melby-Lervåg for gode råd helt på tampen.

Jeg ønsker også å takke Johan Braeken, min metodiske læremester og medforfatter. Det har vært en glede å jobbe med deg. Takk til Astrid Gillespie og Kay Ove Brinchmann for språklige innspill. Og så er jeg utrolig takknemlig for mine kollegaer ved Institutt for Spesialpedagogikk (ISP), spesielt gjengen på stipendiatrekka. Dere er gull.

Mange takk til min arbeidsgiver, ISP, for å ha gitt meg muligheten til å gjennomføre dette arbeidet samt National Graduate School in Educational Research og Fulbright Foundation for å ha utvidet mine horisonter. Rick Wagner ved Florida State University fortjener også en kjempetakk for å vise en gjestfrihet større enn det noen kan forvente. Jeg vil også nevne Kjersti Larsen, seminarleder på UTVIT 1000, som ga meg en flyvende start inn i academia. Du er et forbilde for alle som jobber med å utdanne framtidige forskere og pedagoger.

Og så vil jeg rette en stor takk til mine venner og familie som har sørget for at jeg har et liv utenfor academia. Det har vært så utrolig viktig. Og til Magnus, verdens beste medisin mot stress. Du holder hodet mitt kaldt og hjertet varmt.

Oslo, januar 2016

Ellen Irén Brinchmann

# Innholdsfortegnelse

1	Avhandlingens formål .....	1
1.1	Begrepsavklaringer .....	2
1.2	Avhandlingens oppbygning .....	2
2	Teoretiske perspektiver .....	3
2.1	Ordkunnskap i teorier om lesing .....	3
2.1.1	Ordkunnskap og leseforståelse .....	3
2.1.2	Ordkunnskap og avkoding .....	5
2.2	Ordkunnskap i teorier om språk .....	7
2.2.1	Språk – del eller helhet? .....	7
2.3	Ordkunnskap som fokus for undervisning .....	11
2.3.1	Hvordan kan man fremme elevers ordkunnskap? .....	11
2.3.2	Undervisning for hvem? .....	14
2.4	Presentasjon av studier .....	16
2.4.1	Artikkel I .....	16
2.4.2	Artikkel II .....	17
2.4.3	Artikkel III .....	17
3	Metodologiske refleksjoner .....	19
3.1	Artikkel I: Generalisering av sammenheng .....	19
3.1.1	En forstudie .....	19
3.1.2	Validitet i metaanalyser .....	20
3.1.3	Estimering av populasjonseffekter i univariate analyser .....	21
3.1.4	Utvalgsfeil på metanivå .....	24
3.2	Artikkel II: Måling av endring .....	26
3.2.1	Tradisjonelle tilnærminger .....	27
3.2.2	Valg av analysemetode .....	29
3.2.3	Konsekvenser for validitet .....	31
3.3	Artikkel III: Evaluering av effekt .....	35
3.3.1	Mill's method of difference .....	36
3.3.2	Matching .....	37
3.3.3	Statistisk justering av gruppeforskjeller .....	39
3.4	Etiske betraktninger .....	41

4	Diskusjon av hovedfunn.....	44
4.1	Språklig dynamikk.....	44
4.1.1	Effekten av undervisning i ordkunnskap på elevers språklige kompetanse.....	46
4.2	Betydningen av ordkunnskap for utviklingen av leseforståelse .....	47
4.2.1	Prediktive relasjoner.....	48
4.2.2	Effekten av undervisning i ordkunnskap på elevers leseforståelse .....	51
	Litteraturliste .....	54
	Liste over vedlegg .....	64
	Artikler .....	70

# 1 Avhandlingens formål

I denne avhandlingen utforskes hypotesen om at kunnskap om ord er instrumentelt for utviklingen av leseforståelse. For enkelte framstår antakelig denne hypotesen mer som et selvinnyttende faktum enn som et velegnet tema for vitenskapelig gransking. Å vite hva ordene man leser betyr, er tross alt en logisk nødvendighet for å kunne lese med forståelse. Jeg vil likevel påstå at det finnes svært gode grunner til å vie hypotesen oppmerksomhet – spesielt om man vurderer forholdet fra et praktisk pedagogisk ståsted. Selv om forbindelsen mellom ordkunnskap og leseforståelse er selvfølgelig, er det langt fra selvfølgelig at ordkunnskap fortjener et undervisningsmessig fokus, ei heller er det åpenbart hvem som eventuelt behøver en slik undervisning eller hvordan undervisningen best kan gjennomføres. Slike pedagogiske beslutninger må nødvendigvis hvile på mer enn det åpenbare; beslutningene krever dypere innsikt i de språklige og skriftspråklige mekanismene som styrer forholdet mellom ordkunnskap og leseforståelse. Målet med mitt doktorgradsarbeid har vært å bidra til å utvikle nettopp denne innsikten.

Mer konkret bidrar avhandlingen til å belyse tre problemområder knyttet til hypotesen som er nevnt over. Det første problemområdet i mitt doktorgradsarbeid handler om styrken på relasjonen mellom ordkunnskap og leseforståelse, nærmere bestemt i hvilken grad ordkunnskap kan predikere individuelle forskjeller i leseforståelse. Dette er et viktig område å belyse. Hvis det viser seg at ordkunnskap ikke forklarer hvorfor noen elever utvikler god leseforståelse mens andre ender opp med å streve, vil det framstå som lite hensiktsmessig å fokusere på ordkunnskap i leseopplæringen. Det andre problemområdet er mer indirekte knyttet til avhandlingens overordnede hypotese. Mer spesifikt dreier det seg om tilegnelsen av ordkunnskap og hvordan denne tilegnelsen kan relateres til øvrig språklig utvikling. Å lære mer om slike utviklingsmessige forhold er viktig av flere grunner. For det første kan det være med på å plassere ordkunnskapens bidrag til leseforståelse i en større språklig sammenheng og belyse hvor meningsfullt det faktisk er å fokusere på det unike ved denne språklige variabelen. Ved å utforske selve utviklingen av ordkunnskap vil en dessuten kunne lære mer om hvordan denne utviklingen kan fremmes, noe som er svært relevant for pedagogiske formål. Det pedagogiske perspektivet adresseres imidlertid mer direkte i det tredje av avhandlingens problemområder, som dreier seg om hvorvidt undervisning i ordkunnskap er en effektiv metode for å øke elevens leseforståelsesferdigheter.

## 1.1 Begrepsavklaringer

Gjennom avhandlingen brukes hovedsakelig tre til dels overlappende begreper for å omtale individers kunnskap om ord; ordkunnskap, vokabular og leksikalsk kunnskap. I det meste av litteraturen som det refereres til i denne avhandlingen, er «vokabular» det av disse begrepene som brukes hyppigst. Som oftest beskriver dette begrepet en størrelse, nærmere bestemt antallet ord et individ bruker og forstår innenfor et språk. Slik vil også jeg bruke dette begrepet. Når jeg i tillegg velger å bruke begrepet «ordkunnskap», er det for å understreke at et individs kunnskap om ord handler om mer enn en størrelse. Riktignok er det i litteraturen vanlig å skille mellom «dybde- og breddevokabular», mellom mengden ord man kan og fylde i ens kunnskap om disse ordene, men når alt kommer til alt refererer både dybde- og bredde til forståelsen av enkeltords betydning (Kieffer & Lesaux, 2012). Med ordkunnskap mener imidlertid jeg kunnskap om ord i videst mulig forstand, ikke bare kunnskap om hva enkeltord betyr, men også mer generell kunnskap om ords form og funksjon, semantiske forbindelser mellom ulike ord, konkrete og abstrakte meningsinnhold, hvordan ord er bygd opp av mindre meningsbærende enheter og så videre. Forhåpentligvis vil noe av denne rikdommen i individers kunnskap om ord komme til syne i mitt doktorgradsarbeid. Det siste begrepet som brukes i avhandlingen, leksikalsk kunnskap, anser jeg ikke som kvalitativt annerledes fra de øvrige hovedbegrepene. Dette begrepet brukes simpelthen for å anerkjenne at ordkunnskap er del av individers kognitive system og har således en mental side. Begrepet «leksikalsk kunnskap» vil derfor hovedsakelig brukes i diskusjoner av ordkunnskap som psykologisk fenomen.

## 1.2 Avhandlingens oppbygning

Innledningsvis har jeg skissert tre problemområder som til sammen er med på å belyse avhandlingens overordnede tema. I neste kapittel vil jeg utdype disse problemområdene og gjøre rede for de teoretiske perspektivene som har dannet grunnlaget for de empiriske studiene i mitt doktorgradsarbeid. I kapittel tre diskuteres metodiske forhold knyttet til disse studiene før studienes hovedfunn drøftes i kapittel fire. Studiene presenteres i sin helhet til slutt i denne avhandlingen.



## 2 Teoretiske perspektiver

### 2.1 Ordkunnskap i teorier om lesing

I en teoretisk drøfting av leseforståelse er det i dag vanskelig å komme utenom Gough og Tunmers (1986) kjente lesemodell «the simple view of reading». Modellen har hatt stor innflytelse i leseforskningsmiljøet og etter tre tiår framstår den som stadig like aktuell. I følge modellen er leseforståelse et produkt av to distinkte komponenter; avkodning og språkforståelse. Leseforståelse oppstår i et naturlig samspill mellom komponentene; gjennom avkodning omsettes skrift til fonologiske enheter, som deretter blir bearbeidet og gitt mening ved hjelp av språklige forståelsesmekanismer. Hvilke konkrete mekanismer som driver de to komponentene utdypes imidlertid ikke.

Spesielt vanskelig kan det være å tyde hva språkforståelseskomponenten i «the simple view of reading» egentlig inneholder. I følge Kirby og Savage (2008) kan språkforståelse, som definert av Gough og Tunmer (1986), mer eller mindre forstås som «all form for verbal evne» og Kirby og Savage etterlyser derfor mer spesifikke teorier om relasjonen mellom språk- og leseforståelse. Det særskilte fokuset på ordkunnskap i denne avhandlingen kan ses som et forsøk på imøtekomme en slik anmodning. Jeg er imidlertid ikke alene om å være opptatt av denne tematikken, og i det følgende vil jeg se nærmere på noen av de teoretiske perspektivene som fokuserer på ordkunnskapens unike bidrag til leseforståelse. Målet er ikke å gi en uttømmende beskrivelse av disse perspektivene, men å gi en kort redegjørelse av hva vi vet, og ikke vet, om betydningen av ordkunnskap for utviklingen av leseforståelse.

#### 2.1.1 Ordkunnskap og leseforståelse

Som nevnt innledningsvis er det nærmest så opplagt at ordkunnskap har betydning for leseforståelse at det virker overflødig å påpeke det. Til tross for dette fremhever mange nettopp ordkunnskap som en variabel av særskilt betydning for leseforståelse (se f.eks. Oulette, 2006; Perfetti & Stafura, 2014; Quinn, Wagner, Petscher, & Lopez, 2015; Tunmer & Chapman, 2012). Noe av årsaken til ordkunnskap har mottatt en slik oppmerksomhet, kan muligens ses i sammenheng med forskjellene som eksisterer mellom muntlig og skriftlig språk. I en studie av den relative distribusjonen av ord i ulike språklige kontekster, viste for

eksempel Hayes og Ahrens (1988) at selv barnebøker har en høyere tetthet av sjeldne ord<sup>1</sup> enn dialoger mellom voksne personer. Denne forskjellen mellom muntlig og skriftlig språk øker dersom en tar utgangspunkt i akademisk tekstmateriale, som har en langt høyere frekvens av vanskelige, abstrakte og fagspesifikke ord sammenlignet med narrativ litteratur (Chall, 1996). Det er med andre ord grunn til å tro at skriftlig materiale byr på betraktelig større leksikalske utfordringer enn muntlige språksituasjoner, noe som kan forklare hvorfor mange elever strever med leseforståelse uten å ha bemerkelsesverdige vansker med talespråklig kommunikasjon. En slik tolkning støttes av studier som har vist at forenkling av vokabularet i en tekst, sammen med noe reorganisering av tekstens struktur, fremmer leseforståelse (Freebody & Anderson, 1983). Man ser også at undervisning i ords betydning øker forståelsen av tekst som inneholder disse ordene (Elleman, Lindo, Morphy, & Compton, 2009).

Betydningen av ordkunnskap i teorier om leseforståelse har antakelig også blitt forsterket av at forskjeller i elevers vokabularferdigheter til dels kan forklare det store gapet i akademiske prestasjoner som man finner mellom elever med høy og lav sosioøkonomisk status (Chall, Jacobs, & Baldwin, 1990). Hart og Risley (1995) demonstrerte for eksempel markante forskjeller i den leksikalske utrustingen til barn fra amerikanske familier med ulikt inntektsnivå, og mange har siden hevdet at barn fra lavere sosiale lag ofte mangler den avanserte ordkunnskapen som skolefaglige tekster krever (Sanacore & Palumbo, 2008). Enkelte vil imidlertid mene at en slik forklaring er ufullstendig. I følge Hirsch (2007) er manglende vokabularferdigheter bare del av et større problem knyttet til elevers bakgrunnskunnskap. Ord vil i et slikt perspektiv kun representere små deler av større konseptuelle kunnskapsstrukturer, og en tenker at det er disse kunnskapsstrukturene som er ansvarlig for forskjellene i elevers faglige prestasjoner og ikke vokabularferdigheter per se (Anderson & Freebody, 1981).

Hvorvidt manglende ordkunnskap forårsaker vansker med leseforståelse diskuteres ikke bare i relasjon til elever med lav sosioøkonomisk status; det har også blitt drøftet i forbindelse med mer kliniske grupper av elever, som for eksempel elever med såkalte spesifikke forståelsesvansker. Det vil si elever som, til tross for gode ordavkodningsferdigheter, strever med leseforståelse. Sammenlignet med andre elevgrupper har disse elevene ofte svake språklige ferdigheter, inkludert relative svekkelser i ordkunnskap (Catts, Hogan, & Fey, 2003; Nation, Clarke, Marshall, & Durand, 2004). Andre studier har imidlertid demonstrert at elever

---

<sup>1</sup> Sjeldne ord ble definert som ord med frekvens under 10 000 basert på et stort korpus av engelsk tekstmateriale.

med spesifikke forståelsesvansker også strever med å generere inferenser og har manglende evne til å overvåke egen forståelse (Cain & Oakhill, 1999, 2006). I tråd med dette har enkelte foreslått at svekkede vokabularferdigheter ikke er en proksimal årsak til leseforståelsesvansker, men en indikator for et mer generelt problem knyttet til verbal kognisjon (Anderson & Freebody, 1981).

For å oppsummere diskusjonen så langt, råder det liten tvil om at ordkunnskap er en nødvendig betingelse for leseforståelse. Om man ikke kan ordene i en tekst, vil man heller ikke forstå teksten som helhet. Samtidig vet man at gode vokabularferdigheter ikke er tilstrekkelig; ordkunnskap alene sikrer ikke leseforståelse. Det store stridsspørsmålet handler dermed ikke om hvorvidt ordkunnskap har betydning for leseforståelse eller ikke, men om *hvor stor* betydningen egentlig er. Så langt har jeg vurdert dette spørsmålet med utgangspunkt i at ordkunnskap fremmer det Gough og Tunmer (1986) betegner som språkforståelse. Det er imidlertid også mulig å tenke seg at ordkunnskap bidrar indirekte til leseforståelse, via ordavkoding. Denne muligheten vil jeg diskutere i det kommende delkapittelet.

### **2.1.2 Ordkunnskap og avkoding**

Som tidligere nevnt postulerer «the simple view of reading» at avkoding og språkforståelse er distinkte komponenter av leseforståelse; en påstand som støttes av flere ulike forskningsfunn. Man ser for eksempel at avkodings- og språkforståelseskomponenten ofte forklarer unike deler av variasjonen i leseforståelse (Vellutino, Tunmer, Jaccard, & Chen, 2007). Videre pekes det på at det finnes individer som har vansker i én av komponentene uten å ha vansker i den andre (Catts et al., 2003) og at de to komponentene er fundert på ulike ferdigheter (Storch & Whitehurst, 2002). Akkurat dette siste punktet, at ulike ferdigheter underbygger lesekomponentene, utfordres av Perfetti og Stafura (2014) i det de hevder at det mentale leksikon, hvor representasjoner av ord er lagret, er et knutepunkt hvor ordavkoding og forståelsesprosesser møtes. Mer spesifikt hevder de at dersom ord er lagret på en slik måte at ordenes form og mening er godt representert, vil de raskt og effektivt kunne identifiseres og integreres i en mental modell av teksten som leses (Perfetti & Stafura, 2014). Gitt at dette stemmer vil altså ordkunnskap ikke bare ha betydning for forståelsesprosesser, men også for ordavkoding.

At ordkunnskap har betydning for avkoding er ingen ny tanke; flere ulike teoretiske perspektiver har viet plass til leksikalsk kunnskap i sine hypoteser om ordlesing (se f.eks.

Seidenberg & McClelland, 1989; Share, 1995; Stanovich, 1980). Selv om det er visse forskjeller i hvordan disse teoriene beskriver forholdet mellom ordavkoding og leksikalske prosesser, inneholder de fleste tilnærmingene en forestilling om såkalte top-down strategier. Enkelt sagt handler bruk av top-down strategier om hvordan avkodingen av et ord fasiliteres av den språklige konteksten som ordet befinner seg i samt leserens kunnskap om ordets betydning. Dette kan eksemplifiseres ved hjelp av følgende setning: «Stjernekipperen fikk øye på en komet i teleskopet». Gitt setningskonteksten, vil leseren som vet hva et «teleskop» er antakelig kunne identifisere ordet raskere enn leseren som ikke er kjent med ordets betydning. At top-down strategier er av reell betydning for ordavkoding, har blitt demonstrert i eksperimentelle studier som viser at både hurtighet og nøyaktighet i avkodingen av ord øker når ordene presenteres i kontekst (Nation & Snowling, 1998; Stanovich, West, & Feeman, 1981).

Funnene fra de eksperimentelle studiene jeg nevner over bekrefter at top-down strategier er virksomme under onlinebearbeiding av ord. Share (1995) argumenterer imidlertid for at top-down strategier også spiller en rolle for selve utviklingen av ordavkoding. I sin teori om ortografisk utvikling hevder han at en ortografisk representasjon av et bestemt ord dannes gjennom gjentatte vellykkede fonologiske rekodinger<sup>2</sup> av ordet. Fordi ords skrivemåte ofte ikke samsvarer med ordets fonologiske struktur vil imidlertid en ren fonologisk rekoding i mange tilfeller føre til at ord avkodes på gal måte, noe som i verste fall kan føre til feillæring av ords ortografi. Det er i denne sammenhengen at top-down strategier er av betydning. I følge Share (1995) vil nemlig de fleste ikke-lydrette ord, når de forekommer i sammenhengende tekst, ha nok regularitet i sine fonem-grafem forbindelser til at leseren kan greie å identifisere riktig målord – gitt at leseren har ordet lagret i sitt mentale leksikon. Over tid vil top-down strategier dermed hjelpe leseren til å danne ordspesifikke ortografiske representasjoner, hvilket betyr at semantiske prosesser ikke begrenser seg til onlinebearbeiding, men også fremmer leserens utvikling av automatiserte ordavkodingsferdigheter.

Selv om rasjonalen for å inkludere leksikalske prosesser i teorier om ordavkoding er god, er det empiriske grunnlaget for å anta at det eksisterer en relasjon mellom ordkunnskap og utviklingen av avkoding ikke entydig. Korrelasjoner mellom ordavkoding og vokabular er vanligvis relativt små, og i longitudinelle studier av forholdet står vokabularferdigheter noen

---

<sup>2</sup> Omdanning fra skrift til lyd.

ganger fram som en signifikant prediktor av ordavkodning og andre ganger ikke (Muter, Hulme, Snowling, & Stevenson, 2004; NICHD<sup>3</sup>, 2005; Storch & Whitehurst, 2002). Det må imidlertid merkes at i studiene det vises til her har avkodning typisk blitt målt ved hjelp av standardiserte tester av enkeltordslesing, mens det i følge rasjonalen over vil forventes en tydeligere sammenheng mellom ordkunnskap og lesing av ord i kontekst og/eller lesing av ikke-lydrette ord. Det kan se ut som om studier som fokuserer på akkurat disse distinksjonene rapporterer noe høyere korrelasjoner, men også her er resultatene varierende (Ouellette & Beers, 2010; Ouellette, 2006; Ricketts, Nation, & Bishop, 2007). Det er med andre ord mye vi fremdeles ikke vet om betydningen av ordkunnskap for avkodning.

## 2.2 Ordkunnskap i teorier om språk

Som jeg allerede har vært inne på har flere antydnet at betydningen av ordkunnskap for leseforståelse kan være noe overvurdert. Anderson og Freebody (1981) foreslår som tidligere nevnt at vokabularferdigheter muligens kun er en indikator for verbal kognisjon og at vokabularet i seg selv ikke er instrumentelt for leseforståelse. På samme måte kan man hevde at såkalte top-down strategier i avkodning er knyttet til et bredt spekter av semantiske ferdigheter, hvorav ordkunnskap bare representerer én mulig dimensjon. Ikke nok med det. Flere har i senere tid fremmet mer globale syn på språkforståelse, hvor «språk» mer eller mindre forstås som et udelt hele (Foorman, Koon, Petscher, Mitchell, & Truckenmiller, 2015; Klem, Gustafsson, & Hagtvet, 2015). Dette er et perspektiv som er viktig å vurdere i tilknytning til avhandlingens overordnede tema. Spørsmålet om hvilken betydning ordkunnskap spiller for leseforståelse, hviler nemlig på en forutsetning om at ordkunnskapens bidrag til leseforståelse er noe eget, noe unikt som kan ses adskilt fra andre språklige variabler. Hvorvidt dette er en meningsfull distinksjon kan likevel diskuteres.

### 2.2.1 Språk – del eller helhet?

#### *Empiriske tilnærminger*

Oakhill og Cain (2011) er tydelige på at man i studier av leseforståelse må måle og analysere bidraget fra ulike språkforståelseskompontener hver for seg, slik at det er mulig å kvantifisere ulike språkvariablers relative betydning for leseforståelse. Analyse av variablers relative

---

<sup>3</sup> National Institute of Child Health and Human Development (NICHD)

betydning gjøres typisk i korrelasjonelle studier hvor en ved hjelp av regresjon kan undersøke hvor stor del av variasjonen i leseforståelse som kan forklares av en variabel etter at variasjon som forklares av andre språklige variabler er fjernet. Det er imidlertid flere problemer ved en slik analytisk tilnærming. Dersom man for eksempel tar utgangspunkt i rådata når analysene gjennomføres, kan ulik grad av målefeil i variablene føre til skjevestimering av effektstørrelser og variablenes relative betydning kan dermed feiltolkes (Nunnally & Bernstein, 1994). Et større problem kan likevel sies å ligge på et mer konseptuelt plan. Når det unike ved enkeltvariabler vektlegges, ignoreres samtidig variablenes felles forklaringsvarians. Lager man således kunstige skiller mellom variabler som i realiteten ikke kan ses uavhengig av hverandre?

Dette er et spørsmål som enkelte har forsøkt å besvare ved hjelp av faktoranalytiske tilnærminger. Faktoranalyser tar utgangspunkt i ulike variablers fellesvarians og brukes til å utforske dimensjonaliteten i bestemte fenomener, som for eksempel språkforståelse. Ved hjelp av faktoranalyse kan man dermed undersøke om et fenomen best burde forstås som et udelte hele eller om det er meningsfullt å dele det inn i separate komponenter (for mer om faktoranalyse, se s. 33). Selv om det finnes relativt få empiriske studier som eksplisitt tar for seg dimensjonalitet i språkforståelse, finnes noen unntak. Klem, Gustafsson og Hagtvatn (2015) fant for eksempel at en stor del av variasjonen i syntaktiske, morfologiske og leksikalske ferdigheter hos en gruppe barn i førskolealder, kunne forklares av en unidimensjonal faktorstruktur. I tråd med dette fant Tomblin og Zhang (2006) at en rekke leksikalske og grammatiske variabler målt i skolealder ladet på én enkelt underliggende språkfaktor, selv om en teoretisk modell bestående av to høyt korrelerte dimensjoner («grammatikk» og «vokabular») også hadde en god tilpasning til data. Fra et faktoranalytisk ståsted kan det altså se ut som om ulike språkvariabler har mer til felles enn det som separerer dem, og i lys av slike funn er det naturlig å være kritisk til det å trekke skillelinjer mellom slike variabler. I forlengelsen av dette hevder Foorman et al. (2015) at en må vise stor varsomhet i å analysere og tolke bidraget fra ulike språklige variabler til leseforståelse hver for seg. I en tverrsnittstudie av 1792 elever fra fjerde til tiende klasse viste de at en generell språkfaktor basert på leksikalske og syntaktiske mål forklarte fra 72 % til hele 99 % av variasjonen i leseforståelse på de ulike klassetrinnene. Når et sett variablers fellesvarians har så sterk prediksjonskraft, etterlater det lite rom til å fokusere på de enkelte variablenes unike bidrag.

Til tross for at analyser som inneholder faktorer typisk forklarer mer av variasjonen i leseforståelse enn analyse av enkeltvariabler hver for seg, er heller ikke faktoranalytiske tilnærminger uproblematiske. Den psykologiske statusen til underliggende faktorstrukturer er for eksempel høyst uklar. Faktorer omtales typisk som «teoretiske begreper» fordi de ikke er direkte observerbare – en til dels misledende terminologi med tanke på at faktorer er rene empiriske modeller basert på korrelasjoner mellom enkeltvariabler. Selv om et sett språklige variabler lader på samme faktor, betyr det altså ikke at disse variablene per definisjon er ett og samme psykologiske fenomen. For å kunne hevde dette, må man vite hvorfor variablene korrelerer, noe som er vanskelig å bedømme ved hjelp av empiri alene.

### *Teoretiske perspektiver*

Hvorvidt språk består av et sett separate komponenter eller best forstås som et udelt hele, har vært diskutert innenfor språkvitenskapelige tradisjoner i en årrekke. Spesielt har distinksjonen mellom vokabular og grammatikk blitt mye debattert, og atskillig spalteplass har vært viet til å diskutere relasjonen mellom disse to variablene. Selv om denne debatten på langt nær omfavner hele kompleksiteten i språkforståelsesbegrepet, er den interessant fordi den i bunn og grunn handler om mekanismene som skjuler seg bak språklige variabelers samvariasjon.

Innenfor et leksikalistiske tradisjoner vil man for eksempel hevde at leksikalske og grammatiske mål korrelerer høyt fordi vokabularet er en drivende kraft i barns grammatiske utvikling (Bates & Goodman, 2001). I følge dette perspektivet involverer læring av ord en samtidig læring av ordenes grammatiske egenskaper; først etter at en viss mengde ord er lagret i leksikon vil grammatiske strukturer kunne abstraheres og generaliseres til nye ordsammenhenger (Marchman & Bates, 1994). Tilegnelsen av grammatikk er med andre ord mer eller mindre en direkte funksjon av leksikalsk utvikling, og grammatikk vil dermed ikke kunne karakteriseres som en selvstendig språkkomponent. Et slikt syn har flere følger for tolkningen av de empiriske tilnærmingene som er diskutert over. For det første betyr det at de generelle språkfaktorene som Klem et al. (2015), Tomblin og Zhang (2006) og Foorman et al. (2015) identifiserte, vil være grunnleggende leksikalske av natur. For det andre betyr det at bruk av grammatikk og vokabular som uavhengige variabler i samme regresjonsanalyse, vil kunne føre til en underestimert av vokabularets unike bidrag. Med bakgrunn i leksikalistiske perspektiver vil altså større vekt tillegges ordkunnskap, på bekostning av grammatikk, i analyse av individuelle forskjeller i leseforståelse.



Få andre teoretiske perspektiver tilegner imidlertid vokabularet en så overlegen rolle i språkutviklingen som det leksikalistiske teorier gjør. I følge Chomsky (1965) er for eksempel deler av menneskers grammatiske system medfødt, i motsetning til vokabular som er helt og fullt tillært. Dette innebærer blant annet at grammatiske og leksikalske representasjoner tilegnes gjennom separate språkmekanismer og lagres på forskjellig måte i hjernen. I praksis vil en dermed kunne forvente en større grad av uavhengighet mellom vokabular og grammatikk – de er ulike språklige domener i ordets rette forstand. Selv om Chomsky (1995) i senere tid har tonet ned skillet mellom vokabular og grammatikk, vil man med bakgrunn i nativistiske syn på språkutvikling kunne argumentere for at disse variablene bidrar med noe eget, og at det kanskje til og med er det unike ved variablene som er mest interessant å fokusere på. Baser på et slikt teoretisk grunnlag vil det i det minste kunne anses som hensiktsmessig å behandle komponentene separat i empiriske tilnærminger til leseforståelse.

I en tredje teoretisk tilnærming legges det vekt på at vokabular og grammatikk deler varians fordi den leksikalske og grammatiske utviklingen til dels drives av felles læringsmekanismer. Spesielt har statistisk læring blitt framhevet som en viktig mekanisme i tilegnelsen av både grammatikk og vokabular, så vel som for utvikling innenfor andre språklige og ikke-språklige domener (Saffran & Thiessen, 2007). Enkelt sagt handler statistisk læring om evnen til å spore regulariteter i ens omgivelser, eller mer konkret å oppdage samvariasjon mellom ulike elementer i input. Denne typen læring er implisitt og forekommer fra første stund. Et nyfødt barn møter for eksempel tale som en sammenhengende strøm av lyder. Lydstrømmen består imidlertid ikke av tilfeldige lydsekvenser, og flere studier har vist at barn bruker statistiske forbindelser mellom stavelser til å oppdage hvor ett ord stopper og et annet ord begynner (Saffran, 2001; Saffran, Newport, & Aslin, 1996). Det som er interessant i vår forbindelse er at med denne oppdagelsen endrer også input karakter – fra lydstrøm til en sekvens av ord – og mulighetene ligger dermed til rette for at barnet kan lære seg ords betydning. Ikke nok med det, identifikasjonen av enheter på ordnivå betyr også at mønstre i distribusjonen av ord kan spores, barnet oppdager hvilke type ord som predikerer andre ord, og syntaktiske «regler» læres (Gomez & Gerken, 1999; Saffran & Wilson, 2003). Måten input oppfattes på endrer seg altså når barnet lærer noe nytt, noe som igjen åpner opp for sporing av ytterligere statistiske sammenhenger. Denne naturlige forbindelsen mellom gammel og ny kunnskap forsetter gjennom utviklingen, og slik males bildet av språket som et dynamisk system med konstante vekselvirkninger mellom ulike former for språklig læring (Gómez & Gerken, 2000).



Selv om ulike områder av språket bygger på hverandre, er likevel en viss form for domenespesifisitet mulig å forestille seg – også innenfor perspektiver som fokuserer på generelle læringsmekanismer. Saffran og Thiessen (2007) påpeker eksempelvis at én og samme læringsmekanisme kan produsere svært ulike former for kunnskap avhengig av hvilke informasjonskilder som er tilgjengelig for den som lærer. Et ord refererer for eksempel som oftest til noe som eksisterer i den fysiske eller sosiale virkeligheten, og det er dermed mulig å bygge på ens egen erfaring med objekter og sosiale situasjoner i konstruksjon av leksikalske representasjoner. Grammatikk, og kanskje spesielt syntaks, er imidlertid ikke like eksplisitt representert i den materielle verden og tilegnelsen av grammatiske representasjoner vil i større grad baseres på ren lingvistisk informasjon. Dessuten er statistisk læring bare én av mange læringsmekanismer og det er ikke utenkelig at forskjellige språklige representasjoner bygger på ulike kombinasjoner av læringsmekanismer. Eksplisitte sosiale læringsmekanismer er for eksempel svært viktig for ordlæring (etter hvert kan for eksempel barn spørre direkte hva ord betyr), mens det er mer uvisst hvilken rolle slike mekanismer spiller for tilegnelsen av grammatikk (Saxton, 2010). Når alt tas i betraktning er med andre ord ikke all språklig læring nødvendigvis ett og samme psykologiske fenomen, og det gir dermed en viss mening i å se språket som sammensatt av ulike komponenter.

Som gjennomgangen over viser, finnes det flere ulike syn på spørsmålet om hvorvidt språkforståelse består av et sett unike bestanddeler eller ikke, og innenfor de språkvitenskapelige disipliner eksisterer det fremdeles ingen faglig konsensus rundt hvilket syn som best beskriver virkeligheten. Så langt har jeg vurdert denne tematikken fra et teoretisk og empirisk perspektiv, med fokus på språklig læring innenfor et naturlig utviklingsforløp. Noen av disse perspektivene vil jeg likevel ta med meg videre i det jeg nå forsetter med en mer praktisk pedagogisk diskusjon av avhandlingens overordnede tema. Mer spesifikt vil jeg drøfte ordkunnskapens plass i leseopplæringen.

## **2.3 Ordkunnskap som fokus for undervisning**

### **2.3.1 Hvordan kan man fremme elevers ordkunnskap?**

Som tidligere nevnt hevder Oakhill og Cain (2011) at det i studier av leseforståelse er viktig å måle og analysere bidraget fra ulike språkforståelseskomponenter hver for seg. Begrunnelsen henter de fra et undervisningsmessig ståsted: «... it is more likely to be feasible to train

specific skills rather than to train ‘language ability’» (s. 115). Med dette sitatet peker Oakhill og Cain (2012) indirekte på nok et problem ved å fokusere på språklige variabelers fellesvarians gjennom faktoranalytiske tilnærminger. Fordi de psykologiske mekanismene som ligger bak generelle språkfaktorer er uklare, er slike faktorer vanskelige utgangspunkt når man designer undervisningsopplegg. For hvordan kan en «generell språkfaktor» trenes? Det er, som Oakhill og Cain (2012) påpeker, betydelig enklere å fokusere på spesifikke ferdigheter i leseopplæringen.

Uavhengig av om ulike språklige komponenter er selvstendige eller ikke på et psykologisk plan, har mangfoldige intervensjonsstudier vist at enkeltferdigheter kan trenes gjennom direkte undervisning (f.eks. fonologisk bevissthet: Bus & van Ijzendoorn, 1999; vokabular: Elleman et al., 2009; morfologiske ferdigheter: Goodwin & Ahn, 2010). Selv om det er mulig å fremme spesifikke språkferdigheter, har det imidlertid vært vanskelig å demonstrere at denne treningen har overføringsverdi til standardiserte tester av leseforståelse. Dette gjelder også for intervensjoner som har fokusert på elevers ordkunnskap, typisk ved å gi opplæring i enkeltords betydning. I slike studier ser man ofte store undervisningseffekter på ordene som har blitt trent, mens effekten av undervisning på elevenes leseforståelse er svært begrenset (Elleman et al., 2009). Dette betyr ikke nødvendigvis at ordkunnskap ikke er instrumentelt for leseforståelse. Nagy og Herman (1984) hevder at mengden ord som behøves for at endringer i elevers vokabular skal bre seg til elevenes leseforståelsesferdigheter, simpelthen er for stor til at undervisning i enkeltord er en effektiv pedagogisk tilnærming. Det aller meste av vokabularlæring skjer tross alt ikke gjennom direkte undervisning, men ved «tilfeldig» og ofte ubevisst læring av ord fra kontekst (Sternberg, 1987). Nagy (2005) hevder derfor at fokus på enkeltord alene ikke er tilstrekkelig for å fremme leseforståelse, en må også stimulere elevenes selvstendige ordlæring.

Men hvordan kan man støtte opp under elevers selvstendige ordlæring? Det er mulig å forestille seg flere ulike tilnærminger. Først og fremst er det viktig å sørge for at elevene møter et ordrikt input, slik at de får mulighet til å utvide sin ordkunnskap ved å møte ord de ikke allerede kan. Mye tyder for øvrig på at ikke bare ordrikdom, men også variabilitet i input kan være med på å fremme den implisitte læringen av nye ord. Studier av statistisk læring har for eksempel demonstrert at variabilitet i hvilke referenter som opptrer sammen med et ord, er av betydning for evnen til å kunne koble ordet til riktig referent (Yu & Smith, 2007). Et enkelt eksempel kan brukes for å illustrere dette. Hvis et barn alltid møter ordet «kniv» i

spisesituasjoner, kan ordet forveksles med andre referenter som er tilstede i slike situasjoner (f.eks. gaffel, glass, osv.). Hvis derimot barnet også møter ordet «kniv» i situasjoner som ikke inkluderer andre spiserelaterte gjenstander, som for eksempel når det skal spikkes en trefigur, vil imidlertid denne tvetydigheten kunne løses. Variabilitet er dessuten viktig for at sammenhengen mellom et ord og en spesifikk referent skal kunne generaliseres til andre referenter av samme art (Perry, Samuelson, Malloy, & Schiffer, 2010). Jo flere eksemplarer av en kniv et barn møter, jo lettere vil det som er felles og konstant ved knivene tre fram, og barnet vil i framtiden kunne generalisere ordet til nye og annerledes kniver. Det samme vil for øvrig kunne tenkes å gjelde ord av mer abstrakt karakter. Når man møter ord som for eksempel «konsekvens» eller «absurd» i flere forskjellige sammenhenger, vil ordenes definerende trekk bli tydeligere og meningsinnholdet vil lettere kunne abstraheres (Beck, McKeown, & Kucan, 2002).

I tillegg til å fremme elevers automatiske ordlæring ved å sørge for et rikt og variert input, kan man også lære elever mer eksplisitte strategier for selvstendig ordlæring. Et eksempel på en slik strategi er det Anglin (1993) omtaler som morfologisk problemløsning<sup>4</sup>. Anglin (1993) demonstrerte i sin studie av elevers vokabularkunnskap at enkelte elever kunne forklare ord de ikke hadde hørt tidligere ved å trekke analogier til andre ord med lik morfemisk struktur. Én elev kunne for eksempel forklare hva ordet «treelet» betyr (et lite tre) ved å ta utgangspunkt i ordet «riverlet» (en liten elv), mens en annen elev kom fram til samme løsning ved å trekke analogi til ordet «piglet» (grisunge). Selv om morfologisk problemløsning for mange antakelig er en ubevisst mekanisme, har flere foreslått at strategien også kan læres. Kieffer og Lesaux (2012) hevder for eksempel at dersom elever lærer å gjenkjenne ords morfemiske bestanddeler, og de samtidig får opplæring i hva de ulike bestanddelene betyr, vil de bli bedre i stand til å utlede ukjente ords betydning ved hjelp av ordenes morfemer.

Som allerede nevnt hevder Nagy (1989) at undervisning i enkeltord er en lite fruktbar metode for å fremme elevers ordkunnskap, fordi mengden ord som må læres for å skape ringvirkninger til eksempelvis leseforståelse simpelthen er for stor. I dette henseende kan et fokus på morfemer i undervisningen muligens være mer effektivt. Morfemkunnskap er nemlig generativt, og over tid vil morfemkunnskap dermed kunne føre til kumulativ vokabularvekst. Der kan likevel argumenteres for at også direkte opplæring i enkeltords betydning har sin

---

<sup>4</sup> Morfologi handler om hvordan ord bygges opp ved hjelp av morfemer, språkets minste meningsbærende enheter. Ordet «uvennlig» består for eksempel av tre morfemer: avledningsprefikset *u-*, rotmorfemet *venn* og avledningssuffikset *-lig*. Kombinasjonen av disse morfemene utgjør til sammen ordets mening.

plass i undervisningen. Graves (2000) hevder for eksempel at et jevnlig fokus på ukjente ord, kan stimulere elevenes bevissthet over og nysgjerrighet på ords betydning. Direkte vokabularundervisning kan dermed være med på å motivere elevene til selv å oppsøke ords betydning når de møter på ord de ikke forstår, og slik vil fokus på enkeltord kunne bidra til å stimulere elevenes selvstendige ordlæring.

### 2.3.2 Undervisning for hvem?

En naturlig del av diskusjonen om ordkunnskapens plass i leseopplæringen, er spørsmålet om hvem slik undervisning er relevant for. Dette spørsmålet kan ses som en forlengelse av tematikken som ble drøftet i kapittel 2.1.2, fordi svaret henger sammen med hvorvidt ordkunnskap har betydning for avkodingsprosesser eller ikke. For å illustrere dette kan spørsmålet gjerne omformuleres: Vil elever med avkodingsrelaterte leseforståelsesvansker ha utbytte av vokabularundervisning, eller burde slik undervisning forbeholdes elever med semantisk relaterte vansker?

Med bakgrunn i «the simple view of reading» argumenterer Snowling og Hulme (2012) for at ulike typer lesevansker krever forskjellige pedagogiske tilnærminger. Mer spesifikt hevder de at undervisning av elever med avkodingsvansker må konsentreres om arbeid med bokstav-lyd forbindelser og annen koderelatert trening, mens elever med spesifikke forståelsesvansker behøver undervisning som fremmer språkforståelse generelt og ordkunnskap spesielt. Snowling og Hulmes (2012) argumentasjon bygger på ideen om at den mest effektive undervisningen av elever med leseforståelsesvansker, er den som tar utgangspunkt i de proksimale årsakene til elevenes vansker. Det finnes også empirisk støtte for denne ideen. Elleman et al. (2009) gjennomførte en metaanalyse der de blant annet oppsummerte effekten av å trene på ords betydning for forståelsen av tekst som inneholder disse ordene. Analysen viste at studier av elever med semantisk relaterte lesevansker<sup>5</sup> rapporterte tre ganger så stor effekt av undervisningen sammenlignet med studier av elever med typisk utvikling. Spesielt interessant er det å merke at ingenting tydet på at elevene med vansker lærte seg flere ord enn andre elever. Resultatene indikerte altså at det var selve overføringsverdien av vokabularundervisningen til leseforståelse som var særskilt stor for elevene som hadde svake språkferdigheter til å begynne med. I tråd med Snowling og Hulmes (2012) resonnering, ser

---

<sup>5</sup> Elleman et al. (2009) spesifiserer ikke selv at elevene i disse studiene hadde semantisk relaterte lesevansker. En gjennomgang av primærstudiene de refererer til viste imidlertid at de aktuelle elevutvalgene enten hadde generelle lærevansker eller spesifikke vansker med semantiske aspekter av språket.

det med andre ord ut som vokabularundervisning er spesielt nyttig for elever som har semantiske vansker.

Selv om elever med avkodingsrelaterte leseforståelsvansker helt klart behøver undervisning som adresserer de primære årsakene bak deres vansker (Snowling & Hulme, 2011), kan det likevel tenkes at også denne elevgruppen vil ha godt utbytte av vokabularundervisning. Som jeg allerede har vært inne på legges det i flere teorier om ordavkodning vekt på betydningen av semantiske ferdigheter, deriblant ordkunnskap. Til tross for dette har det ikke vært vanlig å fokusere på vokabularferdigheter i intervensjoner for elever som strever med avkodning, noe som ikke er rart med tanke på at vokabular sjelden framstår som en sterk prediktor av ordavkodning. Interessant nok finnes det likevel studier som antyder at enkle korrelasjoner kan være misvisende, fordi relasjonen mellom ordkunnskap og avkodning ikke er generell; semantiske ferdigheter er først viktige når avkodingen ikke strekker til. Flere eksperimentelle studier har for eksempel vist at effekten av kontekstuell fasilitering varierer avhengig av hvilke ordavkodningsferdigheter man har i utgangspunktet. West, Stanovich, Feeman og Cunningham (1983) fant for eksempel at yngre elever i større grad lente seg på kontekst under ordavkodning enn eldre elever, og Nation og Snowling (1998) viste at elever med dysleksi hadde tydeligere utbytte av kontekstuell støtte enn både elever med normale leseferdigheter og elever med spesifikke forståelsvansker.

Siden såkalte top-down strategier ser ut til å være mest virksomme når avkodingen ikke fungerer som den skal, har enkelte foreslått at elever kan kompensere for manglende ordavkodningsferdigheter ved å anvende sin semantiske språkkompetanse (Snowling, Gallagher, & Frith, 2003). Denne hypotesen støttes blant annet av longitudinelle studier av elever med familiær risiko for dysleksi. I en metaanalyse av slike studier fant Snowling og Melby-Lervåg (2016) at elever som ikke utviklet problemer i møtet med skriftspråket, til tross for sin genetiske disposisjon, hadde markant bedre semantiske språkferdigheter i førskolealder sammenlignet med andre risikoelever som senere utviklet lesevansker. Fra et utviklingsmessig perspektiv ser det dermed ut som om gode språkferdigheter er med på å moderere virkningene av en genetisk risiko for avkodingsvansker. At elevers språkkompetanse fungerer som en viktig moderatorvariabel har også blitt demonstrert i enkelte intervensjonsstudier. Al Otaiba og Fuchs (2006) viste for eksempel at semantiske språkferdigheter predikerte effekten av en koderelatert intervensjon, noe som indikerer at gode språkferdigheter er en forutsetning for å profittere på avkodingsrettet undervisning.

For å oppsummere, finnes det altså flere grunner til å hevde at elever med svekkede ordavkodingsferdigheter kan dra nytte av undervisning som stimulerer deres ordkunnskap. Få studier har imidlertid undersøkt denne hypotesen direkte, og nøyaktig hvor effektiv slik undervisning er for denne elevgruppen gjenstår dermed å se.

## 2.4 Presentasjon av studier

Så langt har jeg i denne avhandlingen diskutert tre generelle problemområder knyttet til avhandlingens overordnede tema. Disse problemområdene har også dannet bakteppe for de tre studiene som til sammen utgjør mitt doktorgradsarbeid.

### 2.4.1 Artikkel I

**Tittel:** *The Role of Vocabulary Knowledge in the Development of Reading Comprehension*

**Forfattere:** Hjetland, H.N, Brinchmann, E.I., & Lyster, S.A.H.

**Sammendrag:** Den første studien som presenteres i denne avhandlingen utforsker styrken i relasjonen mellom ordkunnskap og leseforståelse gjennom en metaanalytisk tilnærming. Mer konkret, undersøkte vi i studien to relaterte problemstillinger: (a) I hvilken grad kan vokabular i førskolealder predikere leseforståelse i skolealder, og (b) i hvilken grad kan tidlige vokabularferdigheter predikere senere ordavkodning? I tillegg til å undersøke den direkte relasjonen mellom vokabular og utviklingen av leseforståelse, undersøkte vi altså hvorvidt det kan eksistere en indirekte relasjon mellom vokabular og leseforståelse som går via ordavkodning.

Basert på resultatene fra 41 primærstudier fant vi en gjennomsnittlig korrelasjon mellom tidlige vokabularferdigheter og senere leseforståelse på  $r = .40$ . En sammenfatning av resultatene fra 30 studier som rapporterte korrelasjoner mellom vokabular og ordavkodning avdekket på sin side en effektstørrelse på  $r = .28$ . Selv om ingen av disse korrelasjonene er særskilt store, er det likevel tankevekkende at vokabularferdigheter målt i førskolen kan forklare en så substansiell del av variasjonen i leseforståelse langt inn i skolealder.

## 2.4.2 Artikkel II

**Tittel:** *The Relationship between Grammar and the Lexicon Revisited*

**Forfattere:** Brinchmann, E.I., Braeken, J., & Lyster, S.A.H

**Sammendrag:** I artikkel II rapporteres en studie som undersøker på hvilken måte tilegnelsen av ordkunnskap kan relateres til øvrig språklig utvikling. Nærmere bestemt, undersøkte vi i studien fire alternative hypoteser om mekanismene som styrer samvariasjonen mellom leksikalsk og grammatisk vekst: (a) vokabular driver utviklingen av grammatikk, (b) grammatikk driver utviklingen av vokabular, (c) det er gjensidige relasjoner mellom grammatisk og leksikalsk utvikling, og (d) det er ingen utviklingsmessige relasjoner mellom de to språkområdene.

For å teste disse hypotesene gjennomførte vi årlige kartlegginger av grammatikk og vokabular i en i en kohort av barn med typisk utvikling ( $N = 217$ ), fra barna i utvalget var fire år gamle til fylte seks år. Ved hjelp av latente forandringsskåreanalyser fant vi at individuelle forskjeller i vokabular ved ett måletidspunkt kunne forklare påfølgende vekst i grammatikk, og motsatt – grammatisk kompetanse kunne forklare variasjon i utviklingen av vokabular. Med bakgrunn i disse resultatene er det god grunn til å tro at det eksisterer en gjensidig og dynamisk relasjon mellom grammatisk og leksikalsk vekst.

## 2.4.3 Artikkel III

**Tittel:** *Lexical Quality Matters: Effects of Word Knowledge Instruction on the Language and Literacy Skills of Third-and Fourth-Grade Poor Readers*

**Forfattere:** Brinchmann, E.I., Hjetland, H.N., & Lyster, S.A.H

**Sammendrag:** I den tredje og siste studien i denne avhandlingen ønsket vi å utforske hvorvidt undervisning i ordkunnskap kan være en hensiktsmessig metode for å bedre elevers leseforståelsesferdigheter. Som utgangspunkt hadde vi en hypotese om at undervisning som stimulerer elevenes kunnskap om ords form og mening kan støtte fundamentet som underbygger elevenes leseforståelsesferdigheter, nærmere bestemt elevenes ordavkodning og språklige forståelse.

Denne hypotesen undersøkte vi ved å teste effekten av en intervensjon designet for å fremme elevers ordkunnskap. Intervensjonen ble evaluert ved hjelp av et pretest-posttest design med én intervensjons- og én kontrollgruppe matchet på leseforståelse. Totalt besto utvalget av 118 svake lesere fra grunnskolens tredje og fjerde klassetrinn. Etter å ha mottatt 60 minutters undervisning tre ganger i uka i 10 uker, viste intervensjonsgruppen signifikant bedre framgang enn kontrollgruppen på flere ulike mål; inkludert forskerutviklede mål på elevenes leksikalske og morfologiske kunnskap samt standardiserte mål på elevenes grammatiske kompetanse og leseforståelse. Vi kunne derimot ikke se noen forskjell på intervensjons- og kontrollgruppens framgang på ordavkodning. Til tross for manglende effekt på elevenes ordavkodingsferdigheter konkluderte vi likevel med at undervisning i ordkunnskap kan være en nyttig framgangsmåte for å fremme elevenes leseforståelse.



## 3 Metodologiske refleksjoner

Som det kommer fram i forrige kapittel er det tre relaterte, dog svært forskjellige studier som ligger til grunn for denne avhandlingen. Den følgende diskusjonen vil av den grunn være organisert i tre deler som hver tar for seg metodologiske vurderinger av særskilt relevans for den enkelte studie. I en fjerde og siste del vil både studiespesifikke og overordnede refleksjoner knyttet til forskningsetiske forhold drøftes.

### 3.1 Artikkel I: Generalisering av sammenheng

#### 3.1.1 En forstudie

I den første artikkelen i denne avhandlingen utforsket vi to relaterte problemstillinger gjennom en metaanalytisk tilnærming. Først ble forholdet mellom tidlig vokabular og senere leseforståelsesferdigheter undersøkt ved å sammenfatte resultatene fra 41 ulike primærstudier. Data fra 30 av disse studiene ble også brukt til å se nærmere på forholdet mellom vokabularferdigheter og ordavkodning.

Datamaterialet som utgjør grunnlaget for artikkel I er hentet fra et prosjekt som i skrivende stund ikke er avsluttet (Hjetland, Brinchmann, Lyster, Hagtvvet, & Melby-Lervåg, 2015). I utgangspunktet var dette hovedprosjektet ment å inngå som et arbeid i avhandlingen, men fordi tidsrammen for prosjektet ble forskjøvet, lot det seg ikke gjøre. I stedet ble det vurdert som hensiktsmessig å utarbeide en mindre forskningsrapport til bruk i avhandlingen, hvor deler av materialet fra hovedprosjektet inngår. Metaanalysen i artikkel I kan dermed ses på som en forstudie som ved et senere tidspunkt vil videreutvikles, og enkelte metodologiske valg må ses i lys av at studien er del av et større arbeid. I det følgende vil jeg derfor kort beskrive hovedprosjektet som studien har sitt utspring fra.

Prosjektet som artikkel I er knyttet til består av en metaanalyse som går under navnet «Preschool predictors of later reading comprehension ability: A systematic review» (Hjetland et al., 2015). Prosjektet har fokus på forebygging av lesevansker og målet for metaanalysen er å oppsummere forskning på språklige og kognitive forløpere til leseforståelse. Innsikt i hvilke

ferdigheter som utgjør grunnlaget for leseforståelse er sentralt for å kunne identifisere barn i risiko for senere vansker og for utarbeidelse av gode forebyggende tiltak.

Mer konkret har det i prosjektet blitt søkt etter og lokalisert studier som rapporterer korrelasjoner mellom leseforståelse og tre brede kategorier av prediktorvariabler; avkodingsrelaterte ferdigheter (f.eks. fonologisk bevissthet og hurtig benevning), språkforståelsesferdigheter (f.eks. grammatikk og vokabular) og domenegenerelle kognitive ferdigheter (f.eks. nonverbal intelligens og arbeidsminne). Det er dette datamaterialet som artikkel I bygger på, og med bakgrunn i avhandlingens overordnede tema ble det i denne artikkelen fokusert på én av de aktuelle prediktorvariablene, nemlig vokabular. At søketermer var konstruert med et annet studieformål i tanke, har hatt få konsekvenser for metaanalysen i artikkel I. Naturlig nok var søket mye bredere enn det ville vært om vi bare var på utkikk etter vokabulardata – en hel rekke andre prediktorvariabler var tross alt inkludert som søketermer. Studier som ikke rapporterte data som var relevant for problemstillingen i artikkel I kunne imidlertid lett ekskluderes i ettertid, og at noen overflødige termer var med i det opprinnelige søket hadde derfor liten betydning.

Det som derimot har hatt betydning for analysen i artikkel I, var inklusjonskriteriene som ble bestemt i hovedprosjektet. I dette prosjektet ble det avgjort at en studie som minimum måtte rapportere minst én korrelasjon mellom en av prediktorvariablene og leseforståelse for å bli inkludert i metaanalysen. Hvilke konsekvenser dette inklusjonskriteriet har hatt vil jeg komme tilbake til nedenfor. Øvrige metodologiske forhold vil imidlertid gjennomgå uten videre referanse til prosjektet som artikkel I er en del av. Om det likevel er ønskelig med mer informasjon om hovedprosjektet, vises det til prosjektets protokoll (Hjetland et al., 2015).

### **3.1.2 Validitet i metaanalyser**

Selv om det er mange fordeler ved metaanalytiske tilnærminger, er metaanalysens primære fortrinn i følge Matt og Cook (2009) muligheten til å undersøke en vitenskapelig slutningsrekkevidde. Gjennom en metaanalyse vil en med andre ord kunne vurdere gyldigheten av en slutning på tvers av personer, tider, steder og operasjonaliseringer av teoretiske begreper. I den følgende metodologiske diskusjonen vil jeg derfor, med utgangspunkt i studien vi gjennomførte, fokusere særskilt på ulike problemer knyttet til ytre validitet i metaanalytiske tilnærminger. Før vi går videre kan det derfor være nyttig å ta en kort kikk på rasjonalen bak generalisering av forskningsfunn i metaanalyser.

### *Generalisering i metaanalytiske tilnærminger*

En metaanalyses generaliserbarhet er fundamentalt knyttet til forestillingen om populasjonsverdier og betydningen av utvalgsfeil (på engelsk: sampling error). I primærstudier utføres statistiske analyser på et utvalg av individer fra en større populasjon. Fordi et utvalg ikke inkluderer absolutt alle individer i målpopulasjonen vil imidlertid et effekttestimat basert på utvalget skille seg fra populasjonens sanne effektstørrelse. Forskjellen mellom et effekttestimat basert på et utvalg av individer og effektens sanne verdi, kalles utvalgsfeil (Borenstein, Hedges, Higgins, & Rothstein, 2009). Jo mindre denne forskjellen er, mer generaliserbart vil effekttestimatet kunne sies å være.

Utvalgsfeil er en ukjent enhet og vi kan således aldri vite hvor stor den er og hvilken betydning den har for våre statistiske analyser. Det man imidlertid vet er at utvalgsfeil reduseres i takt med at utvalgets størrelse øker. Dette prinsippet er helt grunnleggende i metaanalytiske tilnærminger og er direkte årsak til at metaanalyser regnes for å ha bedre ytre validitet enn primærstudier. For det første vil et utvalg individer i en metaanalyse alltid være større enn i enkeltstudier, og således representere en større andel av målpopulasjonen. For det andre vil man i en metaanalyse sammenfatte effektestimater fra enkeltstudier ved hjelp av en algoritme som minimerer utvalgsfeil ved å vekte studiene ulikt basert på utvalgsstørrelse. Dette betyr at små studier med store utvalgsfeil teller mindre i analysen enn store studier med mindre utvalgsfeil. Alt i alt vil en dermed kunne anta at det gjennomsnittlige effekttestimatet som kalkuleres i metaanalysen er nærmere populasjonens sanne effektstørrelse sammenlignet med de enkelte primærstudiene som inngår i analysen.

### **3.1.3 Estimering av populasjonseffekter i univariate analyser**

I vår metaanalyse undersøkte vi som tidligere nevnt sammenhengen mellom vokabular og to avhengige variabler; leseforståelse og ordavkoding. Selv om metaanalysen hadde to forskjellige avhengige variabler, kom all informasjon om disse variablene fra samme utvalg av studier. Dette betyr at to ulike effektestimater (ett for leseforståelse og ett for ordavkoding) i praksis kunne stamme fra ett enkelt utvalg individer, noe som ikke er helt uproblematisk.

## *Statistisk avhengighet*

En grunnleggende forutsetning i metaanalytiske tilnærminger er at effektestimater som inngår i analysen er statistisk uavhengige, det vil si at de er basert på ulike grupper av individer (Gleser & Olkin, 2009). At samme gruppe individer opptrer flere ganger i en metaanalyse vil være uheldig av flere grunner. For det første vil metaanalysens generaliserbarhet reduseres, fordi enkelte individer overrepresenteres i analysen. For det andre vil to effektestimater basert på samme gruppe individer ha korrelerte utvalgsfeil. Hvis begge effektestimater inngår i en metaanalyse vil dette kunne føre til underestimering av den gjennomsnittlige effektstørrelsens varians og/eller feil i kvantifiseringen av heterogenitet mellom studiene i analysen (Scammacca, Roberts, & Stuebing, 2014).

Det finnes flere ulike måter å håndtere statistisk avhengighet på i en metaanalyse. I vår studie benyttet vi oss av metoden som ofte regnes som den enkleste; en såkalt «shifting-unit-of-analysis approach» (Cooper, 1998). En slik metode innebærer at man i stedet for å inkludere statistisk avhengige effektestimater i en og samme analyse, behandler estimatene hver for seg i to separate analyser. Selv om vi var interessert i én uavhengig variabels sammenheng med to avhengige variabler (en multivariat forbindelse), gjennomførte vi dermed i praksis to univariate metaanalyser. Selv om denne metoden muligens er den enkleste, er den ikke nødvendigvis den beste. På den ene siden er det en klar fordel at informasjon fra en og samme studie kan brukes flere ganger, samtidig som den statistiske uavhengigheten mellom effektestimater bevares. På den andre siden vil man ved å bruke denne metoden miste muligheten til å kunne se alle variablene i sammenheng, og man kan dermed gå glipp av verdifull informasjon. Statistisk avhengighet mellom variabler kan nemlig ta ulik form. Som allerede nevnt vil to effektestimater basert på samme gruppe individer ha korrelerte utvalgsfeil. Det er imidlertid også mulig at effektene er korrelerte på populasjonsnivå, altså at det er en avhengighet mellom de sanne effektstørrelsene (Cheung, 2015). Det kan for eksempel tenkes at studier som rapporterer en høyere sann korrelasjon mellom vokabular og ordavkodning også vil rapportere en høyere sann korrelasjon mellom vokabular og leseforståelse. Denne korrelasjonen mellom sanne effekter vil i så fall ligge i naturen til fenomenet som studeres, og vil derfor kunne bidra til å belyse problemstillingen man undersøker. For å kunne estimere en slik form for sann avhengighet mellom variabler i en metaanalyse, må imidlertid datamaterialet håndteres gjennom mer komplekse multivariate

framgangsmåter, hvor for eksempel korrelasjoner mellom avhengige variabler også inkluderes i analysen.

### *Manglende effektstørrelser*

I tillegg til å miste en kilde til potensielt interessant informasjon, finnes det også andre ulemper ved å bruke en «shifting-unit-of-analysis approach» for å håndtere statistisk avhengighet. Én av disse ulempene er relatert til hvordan manglende data håndteres i univariate metaanalyser. I vår metaanalyse fant vi 41 studier som rapporterte en korrelasjon mellom vokabular og leseforståelse. Tretti av disse studiene oppga også en korrelasjon mellom vokabular og ordavkodning. Det vil si at 11 av studiene som oppga korrelasjon mellom vokabular og leseforståelse manglet effektestimater på forholdet mellom vokabular og ordavkodning. Når ordavkodning og leseforståelse analyseres hver for seg i to univariate analyser, anvendes en metode for håndtering av manglende data som på engelsk refereres til som «pairwise deletion». I praksis betyr dette at de 11 studiene med manglende ordavkodningsdata simpelthen slettes fra analysen av forholdet mellom vokabular og ordavkodning. Sett bort i fra at en slik tilnærming kan gi tap av statistisk styrke, er «pairwise deletion» uproblematisk så lenge det er rimelig å anta at det er helt tilfeldig at de 11 aktuelle studiene mangler data på ordavkodning (Cheung, 2015). Dette er imidlertid ikke en antakelse som nødvendigvis kan tas for gitt. La oss si at forskerne som ikke inkluderer ordavkodningsdata i sine studier, har unnlatt å rapportere disse dataene fordi korrelasjonen mellom ordavkodning og vokabular ikke var signifikant. Denne typen selektiv rapportering av effektestimater er eksempel på såkalt publikasjonsskjevhet, og i et slikt scenario vil det være langt fra tilfeldig at studiene mangler data. Dersom disse studiene slettes fra analysen vil en gjennomsnittlig effektstørrelse basert på de resterende studiene kunne gi et skjevt bilde av den sanne populasjonsverdien.

Selv om data som mangler av ikke-tilfeldige årsaker representerer et problem for enhver metaanalytisk tilnærming, finnes det metoder som håndterer manglende data bedre enn univariate framgangsmåter. Ved hjelp av mer komplekse multivariate metoder vil man for eksempel kunne ta i bruk «Maximum Likelihood»-teknikker for å estimere effektstørrelser (se f.eks. Cheung, 2015). Enkelt sagt bruker slike analyseteknikker alle tilgjengelige datapunkter for å estimere aktuelle parametere – ingen data slettes (Kline, 2011). Med en slik metode ville dermed de 11 studiene med manglende ordavkodningsdata indirekte bidratt inn i estimeringen

av sammenhengen mellom vokabular og ordavkoding. Denne teknikken regnes derfor for å gi effkestimater som er nærmere populasjonsverdien sammenlignet med «pairwise deletion».

I retrospekt er det ikke vanskelig å erkjenne at en multivariat analysemetode ville vært et bedre alternativ enn den univariate «shifting-unit-of-analysis»-tilnærmingen vi brukte i vår studie. Når det er sagt vil en univariat metode – gitt at forutsetninger om manglende data ikke brytes – gi presise estimater av hovedeffekter (Cheung, 2015). Så lenge man har nok statistisk styrke til å identifisere disse effektene, vil altså bruk av univariat analyse i seg selv ikke representere en betydelig trussel mot metaanalysens validitet.

### **3.1.4 Utvalgsfeil på metanivå**

Som allerede nevnt vil en metaanalyse typisk ha bedre ytre validitet enn primærstudier fordi utvalgsfeil minimeres, noe som gir effkestimater nærmere populasjonsverdien. En forutsetning for at en metaanalyse skal gi presise effkestimater, er imidlertid at alle relevante studier inkluderes i analysen, eller i det minste, at studiene i analysen er et representativt utvalg av relevante studier. Det er dermed ikke bare manglende effektstørrelser som kan representere et problem for estimering av populasjonseffekter; hvis hele studier mangler fra analysen kan det utgjøre en stor trussel mot metaanalysers ytre validitet, og denne trusselen er det verdt å ta på alvor. Mange studier vil av ulike grunner aldri publiseres og i praksis er det dermed umulig å lokalisere samtlige kandidater i en populasjon av relevante studier. Hvis det er systematiske forskjeller mellom studiene som lokaliseres og de som forblir ukjente, vil dette representere utvalgsfeil på metanivå. I motsetning til utvalgsfeil i primærstudier, er det imidlertid ikke like selvsagt hvordan feilkilder knyttet til sampling av studier skal håndteres innenfor rammen av en metaanalyse.

#### ***Publikasjonsskjevhet***

Den kanskje mest omtalte kilden til utvalgsfeil på metanivå oppstår når studier med bestemte resultater publiseres sjeldnere enn andre, slik at studier som er lett tilgjengelige ikke er representative for alle studier som har undersøkt en bestemt problemstilling. Fordi informasjon om upubliserte studier er vanskelig å oppdrive, er det ofte umulig å vite nøyaktig hvor stor innvirkning en slik publikasjonsskjevhet har hatt for resultatene i en metaanalyse. Det man imidlertid vet er at enkelte faktorer øker risikoen for publikasjonsskjevhet. Det er for eksempel mye som tyder på at studier med eksperimentelle design er mer utsatt for

publikasjonsskjevhet enn observasjonsstudier, spesielt innenfor medisinske fagområder og klinisk psykologi (Schmidt & Hunter, 2015). At vi i all hovedsak inkluderte longitudinelle observasjonsstudier fra en pedagogisk/utviklingspsykologisk tradisjon i vår studie kan dermed anses som en fordel. En annen og mer spesifikk risikofaktor for publikasjonsskjevhet, er mangel på funn av statistisk signifikante sammenhenger. Også her kan vår metaanalyse ha hatt en fordel. I longitudinelle studier er det svært vanlig at flere variabler undersøkes samtidig, og det vil dermed være av mindre betydning for eventuell publikasjonsstatus om enkelte sammenhenger ikke er statistisk signifikante. I den forbindelse kan det også nevnes at både vokabular, ordavkoding og leseforståelse ofte inngår i studier hvor helt andre variabler er i fokus. I slike tilfeller er det rimelig å anta at hvorvidt sammenhenger som involverer disse variablene er statistisk signifikante eller ei, ikke vil være avgjørende for om studien publiseres eller ikke.

Selv om faktorene over taler til vår fordel, er det ikke mulig fullstendig å utelukke at det eksisterer publikasjonsskjevhet i populasjonen av studier som vi ville undersøke. Vi anså det derfor som svært viktig å følge gode prosedyrer for å sikre at vi lokaliserte et representativt utvalg studier. Vi gjorde for eksempel et svært omfattende litteratursøk i syv forskjellige databaser, hvorav flere var databaser for såkalt «grå litteratur» (studier som ikke er publiserte i vitenskapelige tidsskrift). I følge Borenstein et al. (2009) er et grundig litteratursøk kanskje det viktigste virkemidlet en har for å minske følgene av publikasjonsskjevhet. Når dette er sagt, kan det likevel hevdes at vi ikke utnyttet det brede litteratursøket til fulle. I søket fikk vi for eksempel treff på flere upubliserte doktorgradsavhandlinger. Basert på informasjonen som ble oppgitt i avhandlingenes sammendrag var det imidlertid ofte usikkert hvorvidt studien inneholdt relevante data eller ikke. I følge våre planlagte screeningprosedyrer skulle vi ved slike tvilstilfeller lese dokumentet i fulltekst for å avgjøre om studien skulle inkluderes eller ekskluderes. Mange av de aktuelle avhandlingene var imidlertid ikke gratis tilgjengelige i fulltekstformat, og av økonomiske hensyn valgte vi kun å kjøpe inn doktorgradsavhandlinger hvor det kunne sannsynliggjøres ut fra sammendraget at relevante data ble rapportert. Denne beslutningen kan ha ført til skjevheter i vårt utvalg av studier, selv om det er lite trolig at dette forholdet har representert en alvorlig feilkilde.

### ***Eksklusjon av studier***

Som nevnt innledningsvis kan metaanalysen i artikkel I karakteriseres som en forstudie innenfor et større prosjekt. I dette prosjektet ble det på et tidlig tidspunkt avgjort at studier



som manglet informasjon på leseforståelse ikke skulle inkluderes i datagrunnlaget. Dette betyr at alle studier som rapporterte data på sammenhengen mellom vokabular og ordavkoding i fraværet av leseforståelsesdata ble ekskludert. En substansiell del av studiepopulasjonen kan derfor ha manglet fra analysen av vokabular og ordavkoding, og det er ikke en tilfeldig mekanisme som har ført til at noen studier mangler. En kan med andre ord ikke å utelukke at de manglende studiene er systematisk forskjellige fra de som per nå inngår i metaanalysen. Blant de ekskluderte studiene kan det for eksempel eksistere undersøkelser som fokuserer særskilt på sammenhengen mellom ordavkoding og vokabular. Undersøkelser av denne typen inneholder i enkelte tilfeller eksperimentelle mål som er spesielt sensitive med tanke på å fange opp sammenhengen mellom vokabular og ordavkoding, noe som kan tenkes å produsere høyere effektestimater (se diskusjon s. 6-7). Jeg ønsker likevel å påpeke at det i forbindelse med skrivingen av artikkel I ble foretatt et mer kvalitativt søk etter nettopp slike studier. Ingen av studiene som dukket opp i dette søket oppfylte imidlertid de øvrige inklusjonskriteriene i metaanalysen (f.eks. longitudinelt studiedesign) og det er dermed lite sannsynlig at vår eksklusjon av studier uten leseforståelsesmål utgjør en svært alvorlig trussel mot metaanalysens generaliserbarhet.

## 3.2 Artikkel II: Måling av endring

I studien som rapporteres i artikkel II ønsket vi å undersøke utviklingsmessige relasjoner mellom grammatikk og vokabular, og vi gjennomførte derfor tre årlige målinger av disse konstruktene i en kohort av barn med typisk utvikling. Studien representerer dermed et relativt karakteristisk eksempel på et såkalt paneldesign, som innebærer at observasjoner av to eller flere ulike fenomener har blitt foretatt gjentatte ganger hos en bestemt gruppe individer. I planleggingen og gjennomføringen av longitudinelle observasjonsstudier av denne typen, er det uten tvil mange metodologiske forhold som er viktig å vurdere. Fordi jeg personlig ikke var med i selve oppstarten av studien i artikkel II<sup>6</sup>, var imidlertid den aller største beslutningen jeg var involvert i knyttet til valg av analysemetode. Siden dette valget sto så sentralt, ønsker jeg i all hovedsak å vie den metodologiske drøftingen av studien i artikkel II til analysemessige forhold. Jeg vil starte med å diskutere noen av de alternative analysemetodene vi overveide, før jeg går videre til de konkrete vurderingene som lå til grunn

---

<sup>6</sup> Artikkel II er basert på data fra en longitudinell studie av barns språkutvikling ledet av forskergruppen «Child Language and Learning» ved Institutt for Spesialpedagogikk, Universitetet i Oslo. Studien hadde oppstart i 2007 og ble avsluttet i 2012.



for vårt endelige valg av analytisk tilnærming. Til slutt vil jeg drøfte hvilke konsekvenser dette valget kan sies å ha hatt for studiens validitet.

### 3.2.1 Tradisjonelle tilnærminger

Bakgrunnen for å velge longitudinelle paneldesign er som oftest knyttet til et ønske om å undersøke *endring over tid*, eller rettere sagt, å studere forskjeller i individuelle endrings- eller utviklingsmønstre (McArdle, 2009). Som allerede nevnt var vi i vår longitudinelle studie interessert i å se nærmere på utviklingen av grammatikk og vokabular. Helt konkret ønsket vi å se på hvorvidt individuelle forskjeller i vokabular kunne forklare variasjon i utviklingen av grammatikk og motsatt – om individuelle forskjeller i grammatikk kunne forklare variasjon i utviklingen av vokabular. Slike utviklingsmessige relasjoner mellom variabler kan analyseres på flere ulike måter. Selv om det fremdeles ikke råder klar konsensus om hvilken metode som er best egnet, har spesielt to analysemetoder blitt mye brukt. Lenge var såkalte autoregressive modeller ansett for å være selve gullstandarden av longitudinell analyse, mens i de senere år har latente vekstkurvmodeller vunnet stadig mer terreng (Geiser, 2012). Selv om vi ikke brukte disse metodene i vår studie, lå en grundig vurdering av metodenes egenskaper til grunn for vårt endelige valg av analyse. Det er derfor relevant å se litt nærmere på hva metodene innebærer.

#### *Autoregressive modeller*

Autoregressive modeller (Jöreskog & Sörbom, 1979) bygger på antakelsen om at tidligere atferd er den beste prediktoren for senere atferd av samme art. En univariat autoregressiv modell består dermed av en serie målinger av et fenomen, hvor man ved hjelp av regresjon kontrollerer for effekten av samme variabel ved foregående tidspunkt. Hvert repeterte mål blir slik sett en funksjon av det samme målet på et tidligere stadium (autoregressoren) samt en tidsspesifikk residual (varians som ikke kan forklares av autoregressoren). Om man ønsker å undersøke relasjoner mellom to ulike mål, som for eksempel grammatikk og vokabular, kan dette gjøres ved å utvide en univariat autoregressiv modell til en bivariat modell. En slik modell består av to sett av univariate målinger med sine respektive autoregresjoner. Ved å inkludere såkalte kryssladninger i modellen, kan man undersøke hvorvidt den ene variabelen ved ett måletidspunkt predikerer individuelle forskjeller i den andre variabelen ved neste tidspunkt. Når man kontrollerer for autoregressive effekter, vil signifikante kryssladninger i

følge Schlueter, Davidov og Schmidt (2007) indikere den «rene» effekten av en variabel på en annen.

Det er flere fordeler ved autoregressive modeller om man ønsker å studere relasjonen mellom to variabler over tid – slik vi ville i vår studie. I tillegg til å representere en sterk grad av statistisk kontroll, er det for eksempel en fordel at utviklingsperioden man undersøker deles inn i distinkte tidsintervaller ( $t_1$ ,  $t_2$ ,  $t_3$  osv.). Man vil dermed kunne etablere temporal presedens, altså fastslå at en tenkt årsak går forut for dens effekt, noe som anses som et viktig kriterium for slutninger om retningsforhold mellom variabler. En annen stor fordel ved denne analytiske tilnærmingen er muligheten til å kunne identifisere tidsspesifikke sammenhenger mellom variabler. Dette vil være spesielt viktig i tilfeller hvor relasjonen mellom to variabler endrer seg over tid.

Til tross for klare fordeler, har imidlertid autoregressive modeller flere begrensninger. I analyse av utviklingsmessige relasjoner, er det for eksempel en ulempe at effektestimater kun baseres på kovariansstrukturer. Dette betyr at «endring over tid» i en autoregressiv modell ikke representerer endringer i et individs absolutte skårer, men endring i relativ rangering mellom individer i et utvalg. Så lenge den relative rangeringen mellom individene forblir den samme, spiller det altså ingen rolle om hvert enkelt individ har økt, redusert eller uendret skåre fra ett tidspunkt til et annet. Identiske effektparametere vil bli estimert og faktisk endring blir med andre ord ignorert. Siden vi i vår studie ville undersøke hvorvidt individuelle forskjeller i vokabular kunne forklare variasjon i utviklingen av grammatikk og motsatt, må dette regnes som en stor mangel.

### *Latente vekstkurvmodeller*

I likhet med autoregressive modeller brukes latente vekstkurvmodeller til å analysere gjentatte målinger (minimum tre) av en bestemt atferd i et utvalg individer. I motsetning til autoregressive modeller baseres imidlertid vekstkurvmodeller på gjennomsnittsverdier i tillegg til kovariansstrukturer. Dette gjøres ved å modellere to primære latente vekstfaktorer: et intercept, som representerer den gjennomsnittlige startverdien ved første måletidspunkt, og en slopefaktor som reflekterer gjennomsnittlig endring i absolutte skårer over tid. Disse to vekstfaktorene har hver sin varians som uttrykker individuell variasjon rundt gjennomsnittet, samt hver sin residual som representerer tilfeldige feil (Bollen & Curran, 2006).

Fordi vekstfaktorene i en latent vekstkurvemodell har et gjennomsnitt og varians, vil det i en slik tilnærming være mulig både å identifisere individuelle utviklingsmønstre og analysere relativ rangering mellom individer. Denne kombinasjonen av analyse på individ- og gruppenivå, gjør vekstkurvemodeller svært godt egnet til å undersøke problemstillinger knyttet til individuelle forskjeller i faktisk utvikling. Man kan for eksempel undersøke hvorvidt det er sammenheng mellom relativ rangering ved startpunktet og hvor mye man endrer seg over tid, om utviklingens «form» er den samme for alle individer eller hvorvidt eksterne variabler kan forklare forskjeller i individuelle vekstkurver (Geiser, 2012). Dette siste eksempelet er spesielt interessant i forbindelse med vår studie. Ved hjelp av latent vekstkurveanalyse vil man altså kunne undersøke hvorvidt variasjon i en variabel kan forklare utvikling i en annen – ikke rent ulikt målet for vår studie.

Selv om latente vekstkurvemodeller i utgangspunktet virker godt egnet til å belyse vårt forskningsspørsmål, har slike modeller flere begrensninger i analyse av utviklingsmessige relasjoner mellom variabler. Disse begrensningene er hovedsakelig knyttet til hvordan den latente slopefaktoren er estimert. I en vekstkurvemodell representeres ikke de enkelte måletidspunktene i en studie eksplisitt, men uttrykkes gjennom en underliggende vekstkurve som er generalisert over alle tidspunkt (Bollen & Curran, 2006). Det vil dermed ikke være mulig å vurdere tidsspesifikke påvirkninger fra en variabel til en annen, og modellen vil følgelig være lite sensitiv i tilfeller hvor en relasjon er begrenset til deler av den utviklingsmessige perioden som studeres. Av samme grunn vil man heller ikke kunne etablere temporal presedens, noe som begrenser slutninger om retningsforhold mellom variablene. Når alt kommer til alt vil man altså ikke kunne undersøke dynamiske relasjoner mellom variabler.

### **3.2.2 Valg av analysemetode**

I studien som rapporteres i artikkel I ble både en autoregressiv tilnærming og latente vekstkurvemodeller vurdert som mulige analysemetoder. I stedet valgte vi en analysemodell som med norske termer kan beskrives som en latent forandrings-skåremodell (på engelsk: latent change score model). Når jeg likevel har valgt å inkludere relativt inngående beskrivelser av prinsippene bak autoregressive og latente vekstkurvemodeller, er det fordi latente forandrings-skåremodeller kombinerer flere av egenskapene til både autoregressive metoder og vekstkurvemodeller.

### *Latente forandrings-skåremodeller*

Latente forandrings-skåremodeller er basert på samme prinsipp som enkle forandrings-skårer. Forandrings-skårer ( $D$ ) blir ofte brukt til å indikere endring mellom to måletidspunkt og lages ved å regne ut differansen mellom et individs skårer på samme variabel målt ved to ulike tidspunkt ( $D = Y[2] - Y[1]$ ). Skåren gir dermed en direkte indikasjon på hvor mye et individ har forandret seg fra ett tidspunkt til et annet. Dette er i utgangspunktet et godt prinsipp om man ønsker å studere endring, men siden reliabiliteten til differansen mellom to skårer ofte er lavere enn for skårene hver for seg, har mange argumentert for at forandrings-skårer burde unngås (Kessler, 1977).

I likhet med enkle forandrings-skårer ser også latente forandrings-skåremodeller på endring fra ett måletidspunkt til et annet. Modellen tar imidlertid utgangspunkt i variabler på faktornivå ( $f[1]$  og  $f[2]$ ) i stedet for observerte skårer, og problematikken rundt akkumulerte målefeil vil dermed unngås. En annen viktig forskjell mellom enkle forandrings-skårer og latente forandrings-skårer er at i stedet for simpelthen å subtrahere en skåre fra en annen, defineres  $f[2]$  som en funksjon av  $f[1]$  (ganget med konstanten 1) pluss differansen mellom  $f[1]$  og  $f[2]$  (ganget med konstanten 1). Dette kan uttrykkes gjennom følgende likning:  $f[2] = 1 * f[1] + 1 * (f[2] - f[1])$ . Om man legger til enda en måling ( $f[3]$ ), vil denne på samme måte være en funksjon av dens foregående måletidspunkt pluss endring. Sagt på en annen måte,  $f[1]$  påvirker  $f[2]$  direkte og er samtidig en indirekte del av alle framtidige måletidspunkt (McArdle, 2009). Modellen tar altså effekten av samme variabel ved tidligere tidspunkt med i betraktning. Dette kan minne om prinsippet bak en autoregressiv modell, selv om statistikken som ligger til grunn for de to ulike tilnærmingene er svært forskjellig.

En annen sentral forskjell mellom autoregressive tilnærminger og latente forandrings-skåremodeller, er at endring i individers absolutte skårer er eksplisitt representert som et eget parameter i modellen (den latente forandrings-skåren). Dette gjør latente forandrings-skåremodeller godt egnet til å undersøke problemstillinger knyttet til utvikling. Den latente forandrings-skåren har for eksempel både et gjennomsnitt og varians, og man vil dermed kunne analysere interindividuelle forskjeller i intraindividuelle endringer. Som tidligere nevnt er dette også en viktig egenskap ved latente vekstkurvemodeller. Når utviklingsmessige relasjoner mellom to variabler skal undersøkes, har likevel latente forandrings-skåremodeller enkelte fortrinn sammenlignet med vekstkurvemodeller. Ved å koble sammen to sett av repeterte målinger i en bivariat latent forandrings-skåremodell, kan

man ved hjelp av regresjon modellere såkalte krysskoblinger mellom en variabel ved ett måletidspunkt og en annen variabels latente forandrings-skåre. Slik vil man kunne vurdere om variasjon i en variabel kan forklare *endringen* som skjer i den andre variabelen mellom to målinger. De ulike målingene behandles her som distinkte tidsintervaller, og i motsetning til latente vekstkurvemodeller vil man således kunne etablere både temporal presedens og vurdere tidsspesifikke sammenhenger.

Med bakgrunn i drøftingene over, valgte vi som allerede nevnt å bruke en latent forandrings-skåremodell for å analysere datamaterialet i vår longitudinelle studie. Selv om vi vurderte denne tilnærmingen som bedre egnet enn autoregressive modeller og latente vekstkurvemodeller for studiens formål, er heller ikke denne analysemetoden helt uproblematisk. I neste avsnitt vil jeg derfor diskutere hvordan våre modellspesifikasjoner kan ha begrenset studiens slutningsvaliditet.

### **3.2.3 Konsekvenser for validitet**

Før jeg går videre med diskusjonen, kan det være nyttig å repetere noen sentrale trekk ved studiens metode samt studiens hovedfunn. Som tidligere nevnt, var målet for studien å undersøke hvorvidt individuelle forskjeller i vokabular kunne forklare variasjon i utviklingen av grammatikk og motsatt. Disse utviklingsmessige relasjonene ble undersøkt årlig i en kohort av barn med typisk utvikling fra barna i utvalget var fire år gamle til fylte seks år. Ved hjelp av en latent forandrings-skåremodell fant vi statistisk signifikante bivariate krysskoblinger mellom variablene, og basert på disse funnene konkluderte vi med at det er en dynamisk relasjon mellom utviklingen av grammatikk og vokabular. I det følgende vil holdbarheten i denne slutningen vurderes.

#### ***Slutninger om statistisk sammenheng***

Sjansen for å slutte feilaktig at det er en sammenheng mellom to variabler forbindes ofte med statistisk signifikans. Mer spesifikt kan man si at sjansen for å forkaste en korrekt nullhypotese (type I-feil) er lik et effektparameters signifikansnivå. Med en  $p$ -verdi på .05 er det for eksempel 5 % sannsynlighet for at en observert sammenheng kan være forårsaket av tilfeldigheter. De statistisk signifikante bivariate krysskoblingene som vi identifiserte i vår studie, hadde relativt lave  $p$ -verdier (verdiene rangerte fra  $p = .001$  til  $p = .05$ ) og om man kun

baserer seg på disse verdiene må sannsynligheten for at vi har begått en type I-feil anses som liten.

Når det kommer til beregning av statistisk signifikans framheves som oftest utvalgsstørrelse som sentralt; store utvalg gir sikrere analyser og mindre  $p$ -verdier. I tillegg til utvalgsstørrelse, påvirkes imidlertid  $p$ -verdier også av størrelsen på en aktuell sammenheng. Det er av den grunn svært viktig at effektestimater er korrekte, og at eventuelle irrelevante faktorer som kunstig forsterker statistiske sammenhenger tas hånd om. Tidligere har jeg hevdet at latente forandrings-skåremodeller tilbyr en sterk grad av statistisk kontroll, men dette er imidlertid en sannhet med visse modifikasjoner. Den latente forandrings-skåremodellen som vi tok i bruk i vår studie er basert på arbeid fremmet av McArdle og Nesselroade (1994). I senere tid har imidlertid Ferrer og McArdle (2003) tatt til orde for å utvide den opprinnelige tilnærmingen ved å kontrollere for startverdier og total vekst når man undersøker tidsspesifikke krysskoblinger mellom variabler. Bakgrunnen for dette er knyttet til følgende fenomen: utviklingen av ulike ferdigheter har en tendens til å korrelere, uavhengig av om det er dynamiske relasjoner mellom disse ferdighetene eller ikke. Når man ser på utviklingen av to variabler, vil det dermed ofte være en generell sammenheng mellom både variablenes startverdi (intercept) og hvor fort man vokser (slope). Altså, skårer man høyt på det ene målet til å begynne med, er det sannsynlig at man skårer høyt på det andre. Og på samme måte – om man vokser raskt på det ene målet, vokser man antakelig også raskt på det andre. Om man imidlertid ønsker å undersøke dynamiske relasjoner mellom variablene som er uavhengige av denne generelle tendensen, er det viktig å ta høyde for samvariasjonen som eksisterer mellom variablenes intercept og slope. Hvis denne sammenhengen ikke tas hånd om, kan man stå i fare for overestimere størrelsen på eventuelle tidsspesifikke krysskoplinger.

I vår studie kontrollerte vi ikke for sammenhengen mellom variablenes startverdi og totale vekst. For å gjøre dette må man ta i bruk en såkalt dual forandrings-skåremodell hvor man både modellerer vekst fra måletidspunkt til måletidspunkt ved hjelp av latente forandrings-skårer samt en overordnet vekstkurve som er generalisert over de latente forandrings-skårene (tilsvarende en slopefaktor i en latent vekstkurvemodell). For å kunne estimere denne vekstkurven behøver man imidlertid minst tre forandrings-skårer og følgelig minimum fire måletidspunkt.

Modellen som vises i artikkel II består av tre måletidspunkt og to latente forandrings-skårer. Slik modellen framstår nå, uten kontroll for samvariasjon i generell vekst, vil det dermed være

en viss risiko for at de bivariate effektene mellom grammatisk og leksikalsk utvikling er noe kunstig forhøyet. Dette øker risikoen for at vi kan ha begått en type I-feil i vår studie. Det må likevel nevnes at vi i utgangspunktet hadde tilgang til et fjerde måletidspunkt, og i den innledende fasen av analyseprosessen forsøkte vi derfor å analysere datamaterialet ved hjelp av en dual forandringsskåremodell. Modelleringen av variablenes overordnede vekstfaktorer resulterte imidlertid i en svært dårlig modelltilpasning, noe som indikerer at modellen ikke stemte godt overens med data. I tillegg til dette hadde vi også store betenkeligheter knyttet til å bruke det fjerde og siste målet av grammatikk, fordi veksten allerede ved det tredje tidspunktet var i ferd med å flate ut (målet nærmet seg takeffekt). Denne tendensen var enda mer markant ved t4, og vi bestemte oss derfor for å fjerne det siste måletidspunktet i de endelige analysene.

### ***Målefeil***

Et annet forhold som påvirker størrelsen på effektestimater i en modell, er tilstedeværelsen av målefeil. I våre analyser ble målefeil håndtert gjennom en faktoranalytisk tilnærming. Hensikten med faktoranalyse er å identifisere latente variabler (faktorer) basert på fellesvariansen til et sett observerte variabler (indikatorer) som er ment å måle samme teoretiske begrep (Brown, 2006). I praksis gjøres dette ved å dele variansen til den enkelte indikator i to komponenter: (1) varians som er felles med andre indikatorer i analysen, og (2) en residual som består av tilfeldige målefeil samt sann unik varians. Den samlede fellesvariansen til de enkelte indikatorene utgjør en egen latent variabel – en teoretisk variabel som ikke er direkte observerbar. Basert på denne fremgangsmåten estimerte vi to primære faktorer per måletidspunkt i studien vår, ett for grammatikk og ett for vokabular.

Fordi grammatikk og vokabular ble representert som latente variabler i analysene, vil man kunne argumentere for at trusselen fra eventuelle målefeil er liten. Latente variabler er, som vist over, fri for tilfeldige målefeil, og variablene uttrykker dermed sann skårevariens (Brown, 2006). Selve metoden vi tok i bruk for å definere våre latente variabler, har imidlertid sterke begrensninger. Instrumentene som målte grammatikk og vokabular ble simpelthen delt inn i fire deler, slik at hvert individ hadde fire observerte skårer per instrument ved hvert måletidspunkt. Disse skårene ble så brukt som indikatorer i en faktoranalyse. Selv om denne tilnærmingen riktignok tar hånd om tilfeldige målefeil, vil de latente variablene være like sårbare for systematiske målefeil som sine observerte sidestykker. Siden kun ett instrument representerte henholdsvis grammatikk og vokabular er det for eksempel en særskilt risiko for



at viktige sider ved de teoretiske begrepene kan ha vært noe underrepresenterte – at fyllden i begrepene ikke har blitt fanget opp – og at vi dermed kan ha skjevestimert konstruktens utviklingsmessige relasjoner.

Selv om analysene i vår studie inneholdt to primære variabler, grammatikk og vokabular, er også forandringsskårene operasjonaliseringer av et teoretisk begrep, nemlig begrepet endring. Det er dermed viktig ikke å glemme disse variablene når tilstedeværelsen av eventuelle målefeil vurderes. I likhet med de andre variablene i analysene er forandringsskårene latente variabler og således fri for tilfeldige målefeil. Disse endringsparameterne er imidlertid uløselig knyttet til sine respektive primærfaktorer (se s. 30), og eventuelle systematiske feil i målingen av grammatikk og vokabular vil dermed forplante seg til forandringsskårene. Det er for eksempel vanskelig å argumentere for at de latente forandringsskårene virkelig representerer sann endring grammatikk og vokabular, hvis grammatikk og vokabular ikke er målt på en god måte.

### ***Slutninger om kausalitet***

Siden vi tok i bruk et longitudinelt paneldesign, kan det i utgangspunktet virke lite relevant å drøfte kausalitet i tilknytning til vår studie. Av grunner som jeg vil utdype videre i neste kapittel, regnes slike design typisk som uegnede til å si noe om årsakssammenhenger. Til tross for dette er kausale relasjoner ofte fokus for longitudinelle observasjonsstudier, selv når dette ikke uttrykkes eksplisitt. Det er vanskelig å bestribe at vi også er skyldige i dette henseende. Den viktigste grunnen til at vi valgte å ta i bruk en latent forandringsskåremodell i stedet for en vekstkurvemodell, var tross alt fordi vi ønsket å utforske dynamiske relasjoner mellom grammatikk og vokabular – altså, vi ville undersøke hvorvidt de to variablene påvirker hverandres utvikling.

En av grunnene til at observasjonsstudier ikke egner seg til å foreta kausale slutninger, er at slike studier mangler manipulasjon av uavhengig variabel. Gitt at det er kausalitet i forholdet mellom to variabler, er det dermed i praksis vanskelig å fastslå hvilket retningsforhold som eksisterer mellom variablene. For å bøte på dette legges det i observasjonsstudier ofte stor vekt på temporal presedens, altså at den tenkte årsaken må observeres før dens effekt. I tillegg til å ta hensyn til dette tidsaspektet brukte vi i vår studie ulike statistiske teknikker for å undersøke hvilke hypoteser om retningsforhold som kunne støttes av datamaterialet fra vårt utvalg. Enkelt sagt gjøres dette ved å sammenligne alternative teoretiske modeller som



uttrykker ulike sett av statistiske påstander om relasjonene mellom variablene i studien. En modell som postulerer at en tenkt påvirkning går fra den ene variabelen til den andre, kan for eksempel sammenlignes med en modell hvor dette retningsforholdet er snudd.

Sammenligningen gjøres ved hjelp av forskjellige statistiske tilpasningsindekser (på engelsk: fit indices) som tester hvor godt de gitte statistiske påstandene samsvarer med det observerte datamaterialet (Brown, 2006). Hvis det er en signifikant forskjell i tilpasningen til to modeller forkastes den teoretiske modellen med dårligst statistisk tilpasning.

Gjennom en serie slike modellsammenligninger fant vi som tidligere nevnt støtte til en hypotese om at grammatikk og vokabular påvirker hverandre gjensidig gjennom utviklingen. Bruk av statistiske tilpasningsindekser er imidlertid ingen eksakt vitenskap. For det første finnes det flere mulige indekser man kan bruke, som hver har sine svakheter. Én test vil for eksempel avhenge av studiens utvalgsstørrelse, en annen av modellens kompleksitet mens en tredje påvirkes av statistisk estimeringsmetode (Brown, 2006). Videre finnes det ulike retningslinjer, ikke bare for hvilke tilpasningsindekser som bør brukes, men også for hvilke grenseverdier som faktisk indikerer god modelltilpasning (Hu & Bentler, 1999; Marsh, Hau, & Wen, 2004; West, Taylor, & Wu, 2012). Viktigere enn alt dette er det imidlertid å påpeke at selv om en teoretisk modell har en ypperlig modelltilpasning, betyr det ikke nødvendigvis at modellen er sann. Det er fullt mulig at flere ulike teoretiske modeller samsvarer godt med data, og i beste fall vil man dermed kunne slutte at en bestemt modellspesifisering representerer én plausibel forklaring på de statistiske relasjonene i datamaterialet. Alt i alt er det med andre ord all grunn til å stille seg kritisk til kausale tolkninger av sammenhengene vi fant mellom vokabular og grammatikk i vår studie.

### **3.3 Artikkel III: Evaluering av effekt**

Som det kommer frem i diskusjonen av artikkel II har man i dag relativt sofistikerte metoder for å teste ulike kausale modeller basert på observasjonsdata. Selv ved å ta i bruk disse metodene vil imidlertid en kausaltolkning i slike modeller i beste fall bygge på en forestilling om såkalt «constant conjunction»: observasjonen av at én hendelse inntreffer i forbindelse med annen (Roth, 2006). At to hendelser opptrer sammen er imidlertid ingen garanti for at de er kausalt relaterte, og kanskje verre – et kausalitetsbegrep basert på en assosiasjonistisk sammenheng inneholder ikke noe element av produksjon; ingen informasjon om *hvordan* en

bestemt effekt kan frambringes. Data fra observasjonelle studier danner dermed et magert grunnlag for pedagogisk praksis (Kvernbekk, 1997).

I studien som presenteres i artikkel III ønsket vi å undersøke effekten av en intervensjon som var designet for å bedre svake leseres språklige- og skriftspråklige ferdigheter. For å gjennomføre denne studien tok vi i bruk et kvasi-eksperimentelt design. I motsetning til observasjonelle studier, er kausalbegrepet i slike studiedesign basert på prinsippet om manipulasjon: hvis manipulasjon av variabel *X* assosieres med endring i variabel *Y*, kan man sannsynliggjøre at det er kausalitet i forholdet mellom variablene (Cook & Campbell, 1979). Siden manipulasjon består av en aktiv handling, gir en slik tilnærming direkte informasjon om hva man kan gjøre for å bedre elevers læringsutbytte – noe som har tydelig relevans for pedagogisk praksis.

Manipulasjon er ikke det eneste prinsippet som karakteriserer kausalbegrepet i studier med eksperimentelle design. Sentralt er også kausalbegrepets kontrafaktiske natur. Sammenlignet med prinsippet om manipulasjon er imidlertid den kontrafaktiske kausalitetstenkningen langt mindre intuitiv og, ikke minst, vanskeligere å realisere i praksis. I det følgende vil jeg derfor løfte fram noen metodologiske utfordringer vi møtte i vår intervensjonsstudie som kan ha truet grunnlaget for å kunne trekke kontrafaktiske kausalitetsslutninger basert på studiens resultater.

### **3.3.1 Mill's method of difference**

Før jeg går videre med den metodologiske diskusjonen av artikkel III er det hensiktsmessig å se litt på nærmere på logikken bak kausale slutninger i intervensjonsstudier.

Målet med en intervensjonsstudie er å utgjøre en forskjell. Det vil si, å skape en effekt som ikke ville forekommet uten intervensjonen. Dette målet gir en åpenbar utfordring knyttet til evaluering av intervensjonens effekt fordi effekten man ønsker å studere er kontrafaktisk. Hvis man for eksempel har gitt et tiltak til en gruppe elever, er det umulig å vite hva som ville skjedd dersom elevene ikke hadde fått tiltaket. For å imøtekomme denne logiske selvmotsigelsen bygger eksperimentelle studiedesign på en såkalt «method of difference», en kausalitetstenkning som først ble beskrevet av filosofen John Stuart Mill og som senere har blitt videreutviklet av vitenskapsteoretikere som eksempelvis Pearl (2009) og Menzies og Price (1993). Denne metoden tar utgangspunkt i to kausale systemer som har lik distribusjon

av kausale faktorer som er relevante for verdien av en variabel  $Y$ . Så lenge disse faktorene er likt distribuert i de kausale systemene, er det rimelig å anta at også systemenes  $Y$ -verdier er identiske. Om man innfører en intervensjon i ett av de to kausale systemene, samtidig som alle de andre kausale faktorene holdes konstante, vil en eventuell forskjell i systemenes  $Y$ -verdier således kunne tolkes som intervensjonens effekt.

Studien som presenteres i artikkel III er på mange måter en typisk «method of difference»-studie. To grupper av elever representerte parallelle kausale systemer. Én av gruppene mottok en planlagt intervensjon, og forskjellen mellom gruppenes framgang i løpet av studieperioden ble tolket som intervensjonens effekt. Selv om prinsippet er enkelt, opplevde vi at det i praksis var vanskelig å sørge for at metodens viktigste forutsetning ble innfridd, nemlig at distribusjonen av relevante  $X$ -variabler var identisk i de to gruppene. Den eneste måten man kan nærme seg en innfrielse av dette kriteriet, er ved tilfeldig fordeling av individer over forsøksbetingelser. For at skolene i prosjektet vårt skulle greie å administrere intervensjonen var imidlertid en slik randomisering ikke mulig å gjennomføre, og vi måtte dermed ta i bruk alternative metoder for å sikre at forsøksgruppene var ekvivalente på relevante variabler. I vår studie brukte vi to ulike metoder som alternativ til randomisering; matching og statistisk justering av gruppeforskjeller.

### **3.3.2 Matching**

Ved matching i studier med kvasi-eksperimentelle design, tar man typisk utgangspunkt i en gruppe individer som har blitt valgt ut til å motta en intervensjon. Målet med matchingen er å danne en sammenlignbar kontrollgruppe ved at man for hvert individ i intervensjonsgruppen finner minimum ett annet individ med så like karakteristikk som mulig på relevante variabler. Jo flere variabler man greier å matche på jo bedre er det, men i praksis er det ofte svært vanskelig å matche på mer enn noen få variabler samtidig (Stuart, 2010). En av de viktigste beslutningene man tar ved bruk av matching, er dermed hvilke variabler matchingen skal baseres på.

I tilfeller hvor man må prioritere mellom flere mulige matchingvariabler, er det viktig å velge mål som er sterkt relatert til utfallsvariabelen ( $Y$ ) av interesse (Stuart, 2010). Fordi intervensjonen i vår studie først og fremst var rettet mot å bedre elevers leseforståelsesferdigheter, valgte vi nettopp leseforståelse (målt før intervensjonens start) som primær matchingvariabel. Leseforståelse ved ett tidspunkt er tross alt den sterkeste

prediktoren for leseforståelse ved et senere tidspunkt, og slik sett var det naturlig å matche elever på akkurat denne variabelen. Dette valget var imidlertid ikke uproblematisk.

Leseforståelse er en svært sammensatt ferdighet, hvilket betyr at to elever som har lik skåre på en leseforståelsestest kan ha relativt ulike kognitive profiler. En elev som strever med ordavkodning kan for eksempel ende opp med samme skåre som en elev med manglende begrepsforståelse, til tross for at det er helt ulike årsaker som ligger til grunn for elevenes lesevaner. Selv med identiske skårer vil ikke disse to elevene utgjøre en god match.

Et annet problem ved å matche elever på en bestemt testskåre er at testskårer ofte er feil. Forstyrrelser i testsituasjonen, dagsformen til individet som testes, unøyaktigheter i skåring og mange andre ytre og indre faktorer kan føre til at et individs skåre på en test ikke representerer individets sanne skårer. Fordi betydningen av målefeil for den enkeltes skåre som oftest er ukjent, er det dermed vanskelig å slå fast hvorvidt to individer med like testskårer også matcher i sanne skårer.

På grunn av problemene ved matching som er beskrevet over, valgte vi å matche elever i to omganger. I begynnelsen av studien ble det gjennomført en screening av elevene på alle aktuelle klassetrinn i kommunen som prosjektet foregikk i. På bakgrunn av screeningen ønsket vi å gjøre en innledende matching av intervensjonselever og kontroll elever, og vi gjennomførte derfor gruppeadministrerte tester av elevenes leseforståelse, vokabularkunnskap samt morfologiske og ortografiske ferdigheter. Fordi man ved gruppeadministrerte tester har liten kontroll på elevenes testatferd, var imidlertid ikke målet i denne runden å matche elever på eksakte skårer, men å lokalisere proksimale matcher som kunne kartlegges videre med individuelt administrerte tester. Helt konkret prøvde vi å matche elever med tilnærmet like leseforståelsesskårer som hadde så lik distribusjon av skårer som mulig på de øvrige screeningvariablene. Fordi det ikke alltid var åpenbart hvilke potensielle kandidater som utgjorde den beste matchen, oversamlet vi kontrollgruppen slik at alle som utgjorde en omtrentlig match til noen i intervensjonsgruppen ble tatt med videre i prosessen. Disse elevenes leseforståelsesferdigheter ble kartlagt igjen i forbindelse med pretest, hvorpå en endelig en-til-en matching av elever ble gjennomført.

Selv om vi tok i bruk en relativt omfattende matchingprosedyre, er det noe usikkert hvor vellykket prosedyren kan sies å ha vært. På den ene siden viste analyser av såkalte propensityskårer at disse skårene var godt balanserte mellom gruppene. Dette betyr ikke nødvendigvis at vi lyktes med å finne en perfekt match til hver elev i intervensjonsgruppen.

Tvert i mot, de enkelte elevparene kan ha vært svært ulike. Propensityskårer har sitt utspring i teori om matching, men bygger på en litt annen logikk enn matching som baseres på konkrete testskårer. En propensityskåre representerer et helt sett av bakgrunnsvariabler (her: screeningvariablene) uttrykt gjennom én enkelt tallstørrelse; sannsynligheten til å bli valgt ut til å motta en intervensjon (Rosenbaum & Rubin, 1983). Når to grupper har lik distribusjon av propensityskårer betyr det dermed at individene i gruppene har, gitt bakgrunnsvariablene, like stor teoretisk mulighet til å motta intervensjonen. I praksis betyr det at selv om de enkelte parene av matchede individer ikke har identiske skårer på konkrete variabler, vil den samlede distribusjonen av bakgrunnsvariabler som definerer propensityskåren være lik i de to gruppene.

Til tross for at propensityskårer basert på screeningdataene var balanserte på tvers av intervensjons- og kontrollgruppen, viste analyser av pretestdata en annen tendens. På dette måletidspunktet skåret kontrollgruppen noe bedre enn intervensjonsgruppen på de fleste av variablene som ble målt – selv på leseforståelse, som var variabelen gruppene hadde blitt matchet på. På to av målene var denne forskjellen signifikant (ordavkoding og regning), men siden statistisk signifikans avhenger av utvalgsstørrelse kan det tenkes at vi ville funnet flere signifikante forskjeller mellom gruppene dersom utvalget hadde vært større. Det var med andre ord en del som tydet på at det eksisterte enkelte forskjeller mellom gruppene før intervensjonen ble satt i gang, noe som kan ha truet troverdigheten til kausale tolkninger av intervensjonens effekt.

### **3.3.3 Statistisk justering av gruppeforskjeller**

Når det er preeksisterende forskjeller mellom grupper i en intervensjonsstudie, er det svært vanlig å justere disse forskjellene statistisk i en såkalt kovariansanalyse. Prinsippet bak en slik analyse er relativt enkelt. Minimum tre variabler inngår i analysen; avhengig variabel ( $Y$ ), uavhengig variabel (*Gruppe*) og kovariat ( $Cov$ ). Gruppeforskjeller på avhengig variabel analyseres så ved å holde kovariaten konstant på tvers av individer. Mer spesifikt fjernes varians i  $Y$  som kan forklares av  $Cov$  før forskjellen mellom to gruppers verdi på  $Y$  tolkes. Individuelle forskjeller på kovariaten vil med andre ord «elimineres», og betydningen av en ujevn distribusjon av kovariaten i gruppene som studeres, vil i teorien reduseres.

Selv om vi valgte å bruke kovariansanalyse i vår studie, er bruk av denne analysemetoden i studier med kvasi-eksperimentelle design omdiskutert. Det er flere grunner til dette. For det

første er det en viktig statistisk forutsetning i kovariansanalyse at kovariat og uavhengig variabel (*Gruppe*) er ukorrelerte (Miller & Chapman, 2001). Dette skaper en paradoksal situasjon for kvasi-eksperimentøren. Hvis det er substansielle forskjeller mellom to grupper på en kovariat, vil sannsynligvis kovariaten og gruppevariabelen korrelere. Ved å bruke kovariansanalyse for å jevne ut gruppeforskjeller, bryter man dermed automatisk en av de mest sentrale statistiske forutsetningene for analysen. Et slikt brudd kan ha ulike konsekvenser for analysens resultater, men ofte framheves risikoen for å begå type-II feil (Owen & Froman, 1998). Når kovariat og gruppevariabel korrelerer, vil nemlig disse variablene dele varians. Det betyr at all fellesvariens mellom kovariaten og gruppevariabelen fjernes fra den avhengige variabelen når kovariaten føres inn i analysen, og gruppevariabelens mulighet til å forklare varians i den avhengige variabelen reduseres følgelig. I våre data fant vi som ventet statistisk signifikante korrelasjoner mellom gruppevariabelen og de to variablene hvor det var signifikante forskjeller mellom intervensjons- og kontrollgruppen ved pretest (ordavkodning:  $r = -.20$ ; regning:  $r = -.30$ ). I analysene hvor disse variablene ble brukt som kovariater, kan dermed intervensjonseffektene ha vært noe skjevestimerte.

Korrelasjonene mellom gruppevariabelen og de øvrige kovariatene var noe lavere, og ikke statistisk signifikante (rangerte mellom  $r = -.18$  og  $.09$ ). Fordi de fleste korrelasjonene var høyere enn null er det likevel vanskelig å utelukke at de kan ha påvirket effektestimaterne.

En annen sentral ulempe ved kovariansanalyse i studier med kvasi-eksperimentelle design, er at den statistiske justeringen begrenser seg til kovariatene som føres inn i analysen. Til tross for at vi i vår studie hadde flere ulike mål som potensielt kunne blitt brukt som kovariater, valgte vi som hovedstrategi å bruke elevenes skårer ved pretest som kovariat i analyse av samme mål ved posttest. Selv om det kunne vært relevant å bruke flere andre kovariater per analyse, er det ikke alltid like selvfølgelig hvilke variabler man eventuelt burde velge. Mens det for eksempel er svært rimelig å tenke at ordavkodning ved ett tidspunkt har betydning for ordavkodning ved et senere tidspunkt, er det ikke like opplagt at ordavkodning er en kausal faktor som er relevant for andre utfallsvariabler. Ville det eksempelvis vært meningsfullt å bruke ordavkodning som kovariat i analyse av intervensjonens effekt på elevenes vokabularferdigheter? Denne problemstillingen kan være svært vanskelig å løse på en god måte. På den ene siden kan inklusjon av irrelevante kovariater føre til tap av frihetsgrader, større standardfeil og reduksjon av muligheten til å identifisere en intervensjonseffekt (Allen, 1997). På den andre siden vil utelatelse av relevante kovariater skape skjevhet i effektestimater og følgelig true eventuelle kausale slutninger basert på analysene (Shadish,

Cook, & Campbell, 2002). I vår studie prøvde vi å unngå en eventuell over- eller underspesifikasjon av modellene vi satte opp, ved å tilføye tidligere nevnte propensityskårer som kovariat i analysene. På denne måten kunne vi kontrollere for flere kovariater samtidig, uten nevneverdig reduksjon av effektiviteten i våre analyser (d'Agostino, 1998). Siden gruppene i utgangspunktet var relativt balanserte på disse skårene, er det imidlertid lite trolig at dette har gitt store utslag på analysenes resultater.

Selv om det alltid er en risiko forbundet med å utelate observerte variabler fra en kovariatanalyse, er denne risikoen på sett og vis håndterbar fordi den er kjent. For det første har man tross alt informasjon om de utelatte variabelenes distribusjon, og man kan sannsynliggjøre at det ikke eksisterer betydelige forskjeller mellom gruppene på disse variablene. Man har også mulighet til å undersøke hvorvidt et bestemt resultat kan repliseres av ulike analysemodeller og dermed få et inntrykk av eventuelle problemer knyttet til modellspekifiserte. Etter at hovedanalysene som presenteres i artikkel III var gjennomført, testet vi for eksempel ut flere alternative kovariansmodeller som bygde på teorier om ulike kausale relasjoner (ordavkodning ved pretest ble for eksempel lagt til som kovariat i analyse av leseforståelse). Alle disse modellene hadde resultater som sammenfalt med analysene i artikkel III, noe som styrker troverdigheten til resultatene. En langt mindre håndterbar trussel er imidlertid kovariatene som aldri ble målt. Alle analysene som er gjennomført i vår studie hviler på en forutsetning om at det ikke eksisterer noen skjulte forskjeller mellom gruppene, og at alle uobserverte variabler dermed trygt kan ignoreres (Stuart, 2010). Denne forutsetningen er det imidlertid umulig å vite hvorvidt man innfrir fordi man i kvasi-eksperimentelle studier ikke har noen form for kontroll over de uobserverte variablene. I siste instans er det dermed vanskelig å vurdere hvor troverdig en eventuell kausaltolkning av resultatene i vår studie kan sies å være.

### **3.4 Etske betraktninger**

I alle deler av et forskningsprosjekt vil en som forsker måtte ta stilling til en hel rekke etiske problemstillinger. Ikke sjelden vil ulike etiske hensyn motstride hverandre, og i møte med slike normkonflikter er det ikke alltid enkelt å treffe velbegrunnede valg. Likevel er det få som vil bestride at av ulike etiske hensyn, veier hensynet til forskningens deltakere som oftest tyngst. Av den grunn vil jeg i det følgende drøfte enkelte situasjoner i studiene som har blitt gjennomført hvor ivaretagelsen av deltakerne kan ha vært truet.



Som hovedregel skal det i alle forskningsprosjekter som involverer personer hentes inn fritt og informert samtykke fra prosjektets deltakere (NESH<sup>7</sup>, 2006 – pkt 9). Når prosjektets deltakere er barn, kan dette kravet være vanskelig å innfri. Rent formelt er det barnas foresatte som samtykker på deres vegne, men barnas egen aksept er like fullt nødvendig dersom barna er gamle nok til å uttrykke den (NESH, 2006 - pkt. 12). Hvor fritt og informert et barns samtykke kan sies å være, er det imidlertid ikke enkelt å vurdere. Befring (2007) peker på at barn ofte synes det er vanskelig å protestere, de innretter seg lettere etter voksnes ønsker og kan mangle oversikt over konsekvensene ved å delta i et forskningsprosjekt. Det er derfor svært viktig å være sensitiv overfor barnas reaksjoner og respektere deres vilje dersom de på ulike måter motstrider deltakelse. Både i gjennomføringen av studien i artikkel II og i screeningen av barn til intervensjonen i artikkel III opplevde vi at enkelte deltakere ikke ville fullføre kartleggingen. Selv om disse ønskene ble tatt hensyn til, er det likevel en viss risiko for at flere kan ha hatt lyst til å avslutte uten at dette ble fanget opp. Denne utfordringen kan også ha gjort seg gjeldene i gjennomføringen av selve intervensjonen i artikkel III.

Intervensjonen ble gjennomført på barnas skoler av lærere tilsatt ved disse skolene, og for elevene framsto antakelig intervensjonen slik sett som en reell skolesituasjon. Siden undervisning i Norge er obligatorisk kan valgfriheten til å delta i intervensjonsstudien ha blitt opplevd som lite reell. Lærerne som implementerte intervensjonen ble derfor oppfordret til å utvise faglig skjønn og ta hensyn til elevene dersom de ga uttrykk for ubehag ved deltakelse.

Selv om det er en viss risiko for at elevene som mottok intervensjonen har opplevd deltakelsen som obligatorisk, er det likevel mange fordeler ved at intervensjonen ble implementert av barnas lærere. At omgivelsene er kjente bidrar til å skape trygghet i forskningssituasjonen, noe som kan ha redusert eventuelle psykososiale påkjenninger ved å delta i studien. I så måte kan testsituasjonene muligens ha utløst en større grad av usikkerhet hos elevene. Personene som administrerte testene var tross alt fremmede for barna og ved alle individuelle tester var barna stort sett alene med vedkommende. Av den grunn ble det lagt stor vekt på å skape trygghet i testsituasjonen og sørge for at barna opplevde mestring i arbeid med kartleggingsoppgavene. Å sørge for at barna opplevde trivsel og mestring var likevel mest utfordrende i forbindelse med de gruppeadministrerte testene som ble gjennomført i opptakten til intervensjonsstudien. Ved testing i gruppe er barna i større grad overlatt til seg selv og det er dermed vanskeligere å følge opp hver enkelt med ros og oppmuntring

---

<sup>7</sup> Den nasjonale forskningsetiske komité for samfunnsvitenskap og humaniora (NESH)



underveis. Dessverre opplevde vi i enkelte tilfeller at elever ble opprørte da de ikke mestret oppgavene. Selv om det ikke gjaldt mange, er dette en situasjon som helt klart burde vært unngått.

Et annet etisk forhold som var spesielt knyttet til intervensjonsstudien vi gjennomførte, handlet om elevenes forventede utbytte av deltakelsen. Elevene som var valgt ut til å motta intervensjonen var alle elever som av ulike grunner strevde med lesing. Siden elevene ble tatt ut av ordinær undervisning i timene hvor intervensjonen pågikk, var det relevant å vurdere hvorvidt en eventuell gevinst av å delta i studien ville oppveie tapet av annen verdifull undervisningstid. Selv om intervensjonen var designet for å bedre elevenes leseferdigheter, var det imidlertid i praksis umulig å si noe om dette på forhånd – det var aldri noen garanti for at deltakelse ville medføre et reelt utbytte. I etterkant av studien fant vi det som ble tolket som gunstige effekter av intervensjonen og man kan dermed argumentere for at det var etisk forsvarlig å ta elevene ut av ordinær undervisning. Det er likevel verdt å merke at intervensjonseffektene er målt på gruppenivå, hvilket betyr at noen elever kan ha hatt et stort utbytte av deltakelsen, andre kan ha hatt et lite utbytte og det kan til og med ha vært enkelte elever som ville ha hatt større gevinst av ikke å delta i det hele tatt. Relatert til denne problemstillingen er også spørsmålet om hvorvidt det var riktig å ta i bruk en kontrollgruppe som ikke fikk et tiltak som de potensielt kunne hatt nytte av. Av hensyn til disse elevene gjorde vi derfor alt intervensjonsmateriell tilgjengelig for elevenes skoler etter at studien var over, og lærerne ved kontrollskolene fikk også tilbud om opplæring i intervensjonens metoder. Hvorvidt dette kom elevene i kontrollgruppen til gode, er imidlertid uvisst.

## 4 Diskusjon av hovedfunn

Utgangspunktet for denne avhandlingen er hypotesen om at ordkunnskap er instrumentelt for utviklingen av leseforståelse; en hypotese som innledningsvis ble aktualisert ved hjelp av tre ulike problemområder. Disse problemområdene kan oppsummeres kort ved hjelp av følgende spørsmål:

1. I hvilken grad kan ordkunnskap predikere individuelle forskjeller i leseforståelse?
2. Hvordan kan tilegnelsen av ordkunnskap relateres til den øvrige språklige utviklingen?
3. Er undervisning i ordkunnskap en effektiv metode for å fremme elevers leseforståelsesferdigheter?

I dette siste kapittelet av avhandlingen vil jeg diskutere på hvilken måte disse problemområdene belyses gjennom avhandlingens empiriske arbeider.

### 4.1 Språklig dynamikk

Ett av områdene som problematiseres innledningsvis er tilegnelsen av ordkunnskap, og hvordan denne tilegnelsen kan relateres til øvrig språklig utvikling. Selv om dette området ikke er direkte knyttet til avhandlingens overordnede tema, er det likevel svært relevant. For det første vil man ved å se på ordkunnskap i en større språklig sammenheng, kunne belyse hvor meningsfullt det er å fremheve ordkunnskapens unike bidrag til leseforståelse. Ved å få innsikt i utviklingen av ordkunnskap vil en også kunne lære noe om hvordan denne utviklingen best kan fremmes. Samlet sett adresseres denne tematikken i to av de empiriske arbeidene i denne avhandlingen, nærmere bestemt av studiene i artikkel II og artikkel III.

Studien som presenteres i artikkel II er den som mest konkret tar for seg ordkunnskapens rolle som språklig komponent ved å undersøke samvariasjonen som kan observeres mellom leksikalsk og grammatisk utvikling. Målet med studien var å utforske mekanismene som ligger bak denne samvariasjonen med et ønske om å si noe om ordkunnskapens rolle i språkutviklingen. Er for eksempel tilegnelsen av nye ord i språket en drivende kraft i den grammatiske utviklingen, eller er leksikalsk og grammatisk utvikling drevet av parallelle prosesser som har lite med hverandre å gjøre? Resultatene fra studien støttet ingen av disse to

mulige hypotesene. Med bakgrunn i analysene som ble gjennomført fant vi derimot støtte til en hypotese om språklig dynamikk, eller mer spesifikt at det eksisterer vekselvirkninger mellom grammatisk og leksikalsk læring gjennom utviklingen. Denne tolkningen av resultatene i vår studie har to sentrale implikasjoner for hvordan man betrakter utviklingen av grammatikk og vokabular. For det første forutsetter tolkningen at man legger en viss grad av domenespesifisitet til grunn. Det er meningsløst å snakke om dynamiske relasjoner mellom to variabler, dersom variablene ikke kan ses adskilt fra hverandre på ett eller annet nivå.

Hvorvidt en slik domenespesifisitet signaliserer at utviklingen er drevet av separate språkssystemer (Chomsky, 1965; Pinker, 1998) eller om den snarere indikerer at grammatiske og leksikalske representasjoner baseres på ulike typer informasjon i input (Saffran & Thiessen, 2007) vites imidlertid ikke. Hovedbudskapet fra studien er likevel at verken grammatikk eller vokabular er språklige komponenter som opererer i isolasjon. I et naturlig utviklingsforløp vil læring innenfor det ene området kunne bidra til ny læring i det andre. Selv om språket ikke nødvendigvis er et udeelt hele, vil ulike språklige variabler med andre ord være langt fra uavhengige av hverandre.

Den naturlige forbindelsen mellom ulike former for språklig læring, var noe vi ønsket å utnytte i intervensjonen som presenteres i artikkel III. Selv om hovedfokuset for intervensjonen var på å styrke elevers ordkunnskap, tok vi i utarbeidelsen av undervisningsopplegget derfor utgangspunkt i en teoretisk antakelse om at ulike områder av språket kan samhandle og forsterke hverandre, og således føre til høyere nivåer av språklig kompetanse. Mer konkret tok vi utgangspunkt i Perfettis (2007) hypotese om leksikalsk kvalitet (se også Perfetti & Hart, 2002). Som nevnt i kapittel 2.1.2 hevder Perfetti og Stafura (2014) at dersom ord er lagret i det mentale leksikon på en slik måte at ordenes form og mening er godt representert, det vil si holder høy leksikalsk kvalitet, vil ordene raskt og enkelt kunne identifiseres og bearbeides ved hjelp av høyere ordens forståelsesprosesser. Det som er interessant i denne forbindelse, er at et ords leksikalske kvalitet avgjøres av ens kunnskap om ordets ortografiske, fonologiske, semantiske og grammatiske egenskaper (Perfetti, 2007). I Perfettis (2007) hypotese er med andre ord ulike områder av språket bundet sammen på ordnivå, og ord kan dermed representere en inngangsport til å arbeide dynamisk med elevers språklige kompetanse. Dette er en antakelse vi ønsket å teste i vår intervensjon. I tråd med hypotesen om leksikalsk kvalitet designet vi dermed et undervisningsopplegg som stimulerte elevenes kunnskap om ords ulike egenskaper, og da særskilt ords semantiske, syntaktiske og morfologiske egenskaper. For eventuelt å bevise at denne tilnærmingen fremmer språklig

dynamikk, måtte to sentrale forventinger innfris: (a) vi måtte observere undervisningseffekter på flere ulike språkområder og (b) disse effektene måtte signalisere generalisert kunnskap, altså ikke begrense seg til de konkrete ordene som ble gjennomgått i intervensjonen.

#### **4.1.1 Effekten av undervisning i ordkunnskap på elevers språklige kompetanse**

For å undersøke effekten av intervensjonen kartla vi elevene med et testbatteri sammensatt av ulike språklige tester, inkludert både leksikalske, syntaktiske og morfologiske mål. Når det kommer til de leksikalske målene, var to vokabulartester inkludert i testbatteriet; én test som inneholdt ord som det ble undervist i under intervensjonen og én standardisert vokabulartest med et helt annet utvalg av ord. Etter intervensjonens slutt fant vi store undervisningseffekter på ordene som inngikk i intervensjonen, mens det ikke var noen statistisk signifikant effekt av undervisningen på den standardiserte vokabulartesten. Vi kunne med andre ord ikke påvise at undervisningseffektene lot seg generalisere til elevenes øvrige vokabularferdigheter. Dette funnet er imidlertid ikke overraskende. Som jeg har vært inne på tidligere kan det tenkes at direkte undervisning i ords betydning har effekt på elevenes øvrige vokabular dersom denne undervisningen er med på øke elevenes nysgjerrighet og motivasjon til selv å utforske ords betydning (Graves, 2000). Hvis denne nysgjerrigheten opprettholdes, vil direkte vokabularundervisning, over tid, kunne føre til generell vokabularvekst. Å se en slik overføringseffekt i løpet av den relativt korte perioden intervensjonen foregikk på, ville vært svært overraskende.

Når undervisningstiden tas i betraktning, var det kanskje noe overraskende at vi observerte en såkalt overføringseffekt på et standardisert mål på syntaktisk kompetanse, nærmere bestemt evnen til å formulere grammatisk korrekte setninger. Samtidig er det en viss logikk i at vi fant overføringseffekt på nettopp et syntaktisk mål. Selv om enkeltord kan sies å ha bestemte syntaktiske egenskaper, er syntaks i seg selv abstrakt og generativt (Saxton, 2010). Gitt at undervisningen hadde effekt på elevenes syntaktiske kompetanse, er det dermed naturlig at effekten lot seg generalisere utover de spesifikke ordkontekstene som ble brukt i intervensjonen.

Det samme kan delvis sies om den morfologiske ordkunnskapen som vi trente. I likhet med syntaks er morfemer generative, det vil si at ett og samme morfem kan brukes på tvers av mange ulike ord. Morfemer har derimot ikke den samme abstrakte karakteren som syntaks; de

er konkrete språklige enheter og kunnskap om ett morfem vil ikke nødvendigvis kunne overføres til andre morfemer. Vi forventet derfor at elevene som mottok intervensjonen ville kunne generalisere kunnskap om morfemer de hadde lært til nye ordkontekster, men ikke at de uttrykte økt kunnskap om morfemer som det ikke hadde blitt fokusert på i undervisningen. Disse forventningene ble innfridd. På en test som inneholdt fremmede morfologiske strukturer, viste ikke intervensjonsgruppen større framgang enn kontrollgruppen. Det gjorde de derimot på en test som var utviklet for å avdekke elevenes morfologiske problemløsningsevner, altså evnen til å utlede et ukjent ords betydning ved hjelp av ordets morfemiske bestanddeler. Alle morfemene i denne testen var blitt gjennomgått i intervensjonen, men i tilknytning til helt andre ord. Den observerte effekten indikerte dermed at elevene kunne overføre morfemkunnskapen de hadde ervervet til nye ordkontekster.

Når alt kommer til alt, var det med andre ord flere tegn som tydet på at effektene av intervensjonen ikke begrenset seg til de konkrete ordene vi fokuserte på i undervisningen, men signaliserte språklig læring i bredere forstand. Dette funnet kan representere et svar på noe av problematikken knyttet til språkrelatert undervisning som jeg har diskutert tidligere. Som nevnt har flere i senere tid forfektet et globalt syn på språk og samtidig advart mot å fokusere på det unike ved språklige enkeltvariabler (Foorman et al., 2015). Et slikt teoretisk utgangspunkt er imidlertid lite retningsgivende for praksis; når man designer undervisningsopplegg er «språk» et vanskelig begrep å ta utgangspunkt i. Resultatene fra intervensjonsstudien indikerer imidlertid at arbeid med ordkunnskap kan være en inngangsport til å fremme mer generelle språkferdigheter, og intervensjonen representerer således en mer håndfast og anvendelig tilnærming til «global» språkopplæring.

## **4.2 Betydningen av ordkunnskap for utviklingen av leseforståelse**

Selv om forskningsfunnene jeg har diskutert over er med på å belyse avhandlingens overordnede hypotese, adresseres hypotesen imidlertid mer direkte av de to øvrige problemområdene jeg introduserte innledningsvis. Et eksempel på dette kan vi finne i artikkel I hvor den prediktive relasjonen mellom ordkunnskap og leseforståelse ble undersøkt gjennom en metaanalytisk tilnærming.

## 4.2.1 Prediktive relasjoner

### *Ordkunnskap og leseforståelse*

Da vi oppsummerte resultatene fra 41 studier fant vi som tidligere nevnt en gjennomsnittlig korrelasjon mellom vokabular i førskolealder og leseforståelse i skolealder på  $r = .40$ . Basert på konvensjonelle kriterier betegner denne korrelasjonen et moderat styrkeforhold (Cohen, 1988). I og med at vi undersøkte sammenhenger som strakk seg over tid, kan det likevel argumenteres for at korrelasjonen er rimelig sterk og at vokabular er en robust prediktor av leseforståelse.<sup>8</sup> Dette funnet har flere praktiske implikasjoner. For det første vil man med utgangspunkt i førskolebarns vokabularferdigheter kunne identifisere individer i risiko for senere leseforståelsesvansker. Dette er et viktig poeng helt uavhengig av hvilken betydning ordkunnskap har for selve utviklingen av leseforståelse. Hvis man kan identifisere barn i risiko, vil man også kunne iverksette gode forebyggende tiltak før eventuelle vansker manifesteres. Resultatene antyder likevel at det kan være hensiktsmessig å inkludere arbeid med ordkunnskap som del av en slik forebygging. Siden førskolebarns vokabular kan predikere hvem som flere år senere utvikler gode og dårlige leseforståelsesferdigheter, er det nærliggende å tro at ordkunnskap er av instrumentell betydning for utviklingen av leseforståelse. Siden resultatene fra metaanalysen er basert på enkle korrelasjoner, må imidlertid en slik slutning tas med store forbehold.

Metaanalysen har ikke bare praktiske implikasjoner, den kan også bidra til videre teoriutvikling på feltet. Som nevnt vil det gjennomsnittlige effekttestimatet som kalkuleres i en metaanalyse være nærmere populasjonens sanne effektstørrelse sammenlignet med de enkelte primærstudiene som inngår i analysen. Metaanalysen kan således bidra til å øke presisjonsnivået i våre antakelser om den prediktive relasjonen mellom vokabular og leseforståelse, noe som er en klar fordel fra et teoretisk perspektiv. Vel så viktig er det imidlertid av vi gjennom arbeidet med metaanalysen var i stand til å avdekke enkelte hull i forskningsfeltet. Dette betyr at metaanalysen kan være med på å synliggjøre forhold som bør tas i betraktning i framtidige studier på området.

For eksempel, studiene som vi inkluderte i metaanalysen undersøkte i all hovedsak forholdet mellom leseforståelse og det jeg innledningsvis betegner som «breddevokabular», altså

---

<sup>8</sup> Til sammenligning kan man i artikkel II se at korrelasjonen mellom ett og samme vokabularmål administrert med to års mellomrom er på  $r = .45$ .

størrelsen på individers vokabular. Som tidligere nevnt er imidlertid ordkunnskap mer enn en størrelse, noe som ikke nødvendigvis reflekteres av forskningen på feltet. Slik vi diskuterer i artikkel I, er det for eksempel en sterk tendens blant leseforskere til å foretrekke reseptive vokabularmål som «Peabody Picture Vocabulary Test» (Dunn & Dunn, 2007) og «British Picture Vocabulary Scale» (Dunn, Dunn, Whetton, & Burley, 1997). Flere grunner til dette kan tenkes. I tillegg til at testene er kjente, enkle å administrere og har gode psykometriske egenskaper, anser mange forskere reseptive vokabulartester som «renere» mål på ordkunnskap enn for eksempel ekspressive vokabulartester som til en viss grad også måler individers formuleringsevne. Problemet med denne ensidige bruken av vokabulartester, er imidlertid at man risikerer å overse viktige sider av elevers ordkunnskap, og de teoretiske perspektivene som beskriver forholdet mellom ordkunnskap og leseforståelse blir således lite ivaretatt. Det kan for eksempel tenkes at vokabulartester som krever muntlig respons i større grad avdekker fylde i elevers kunnskap om ord og dermed representerer et mer sensitivt mål på ordkunnskap (Tilstra, McMaster, Van den Broek, Kendeou, & Rapp, 2009). Det er også mulig å ønske seg flere mer teoridrevne mål, som for eksempel tester som kartlegger elevers semantiske nettverk (se f. eks. Carlo et al., 2004), evne til å utlede ords betydning (se f.eks. Kieffer & Lesaux, 2012) og så videre. Samlet sett vil en slik testrikdom muligens kunne representere et mer solid grunnlag for teoretisering om forholdet mellom ordkunnskap og leseforståelse.

En annen begrensning i feltet dreier seg om mer overordnede metodologiske forhold. Selv om vi i vår metaanalyse kodet enkle korrelasjoner, fikk vi likevel god oversikt over hvilke metodiske framgangsmåter som regjerer i feltet. Det som var svært tydelig, var at det store flertallet av studier undersøkte sine problemstillinger ved hjelp av metoder som baserer seg på kovariansstrukturer. Når det kommer til spørsmålet om ordkunnskapens betydning for utviklingen av leseforståelse, er imidlertid slike metodiske framgangsmåter begrenset. Som jeg har vært inne på tidligere, sier en longitudinell korrelasjon ingenting om endringer i et individs absolutte skårer (se s. 28), og slike korrelasjoner vil i beste fall kunne karakteriseres som et indirekte mål på utvikling. Jeg vil dermed argumentere for at framtidige undersøkelser av forholdet mellom ordkunnskap og leseforståelse, i større grad bør benytte seg av metodiske framgangsmåter som også adresserer individers vekst.



## *Ordkunnskap og avkoding*

I tillegg til å undersøke den direkte relasjonen mellom vokabular og leseforståelse, utforsket vi i artikkel I også hvorvidt tidlig vokabular er assosiert med senere ordavkodingsferdigheter. Sammenlignet med den gjennomsnittlige korrelasjonen mellom vokabular og leseforståelse, var sammenhengen mellom vokabular og ordavkoding noe mindre ( $r = .28$ ). Selv om vi i artikkel I argumenterer for at også denne korrelasjonen må anses som substansiell når tiden mellom målinger tas i betraktning, antyder vi også at metaanalysen muligens ikke er presis nok i sin undersøkelse av den teoretiske relasjonen mellom vokabular og ordavkoding. Det er flere årsaker til dette. For det første handler mange av teoriene som beskriver forholdet mellom vokabular og ordavkoding om hvordan top-down strategier fasiliterer avkoding av ord ved lesing av sammenhengende tekst (Perfetti & Stafura, 2014; Stanovich, 1980). Denne bestemte antakelsen var det imidlertid ikke mulig å teste i metaanalysen, fordi for få studier rapporterte mål på lesing av ord i kontekst.

Det vi derimot kunne undersøke var Shares (1995) hypotese om at kontekstuell fasilitering ikke bare har betydning for onlinebearbeiding av tekst, men at top-down strategier også fremmer tilegnelsen av spesifikke ortografiske representasjoner. Gitt at dette stemmer, vil man kunne forvente å se en sammenheng mellom semantiske variabler og avkoding av ord som presenteres i isolasjon. Til tross for at det er akkurat denne hypotesen som ble adressert da vi oppsummerte den prediktive korrelasjonen mellom vokabular og enkeltordslesing, kan det hevdes at metaanalysen heller ikke er treffsikker nok i dette henseende. Som nevnt hevder Share (1995) at top-down strategier er av særskilt betydning i tilfeller hvor en fonologisk avkodingsstrategi ikke er tilstrekkelig for korrekt ordidentifikasjon, som ved lesing av ikke-lydrette ord. Igjen forelå det imidlertid ikke tilstrekkelig data fra primærstudiene som tillot analyse av en mer spesifikk sammenheng mellom vokabular og avkoding av ord med ulik ortografi.

En tredje grunn til at metaanalysen ikke var treffsikker nok i sin undersøkelse av relasjonen mellom vokabular og ordavkoding, er knyttet til funn fra eksperimentelle studier av top-down strategier i lesing. Som nevnt i kapittel 2.3.2 har flere studier foreslått at effekten av kontekstuell fasilitering ikke er generell, men at semantiske ferdigheter først og fremst er viktige når avkodingen ikke strekker til (Nation & Snowling, 1998). Hvis denne hypotesen stemmer vil en kunne forvente at styrken på sammenhengen mellom vokabular og



ordavkodning varierer langs ulike deler av distribusjonen i ordavkodingsferdigheter, noe som i så fall ikke vil fanges opp av enkle korrelasjoner.

Selv om forholdene jeg har skissert over representerer begrensninger ved metaanalysen vi gjennomførte, er metaanalysen på mange måter kun en refleksjon av forskningen som foreligger på feltet. Riktignok kan ulike metodologiske beslutninger ha ført til at sentrale studier falt utenfor analysen (se s. 26), men basert på vårt møte med et bredt sett av forskningslitteratur vil jeg likevel hevde at det er en mangel på teoridrevne studier av forholdet mellom ordavkodning og leseforståelse – i alle fall når det kommer til studier som fokuserer på de prediktive eller utviklingsmessige relasjonene mellom de to konstruktene. Dette er en begrensning ved feltet som jeg håper fremtidige studier vil imøtekomme.

#### **4.2.2 Effekten av undervisning i ordkunnskap på elevers leseforståelse**

Instrumentaliteten i forholdet mellom ordkunnskap og leseforståelse adresseres mest direkte i artikkel III hvor effektene av undervisning i ordkunnskap på elevers leseforståelsesferdigheter ble undersøkt. Mer konkret undersøkte vi en hypotese om at undervisning som stimulerer elevenes kunnskap om ords form og mening kan støtte fundamentet som underbygger elevenes leseforståelsesferdigheter, nærmere bestemt elevenes ordavkodning og språklige forståelse. Som jeg allerede har drøftet, var det mye som tydet på at intervensjonen vi gjennomførte var med på å øke elevenes språklige kompetanse. Vi kunne imidlertid ikke identifisere noen effekt av intervensjonen på elevenes avkodingsferdigheter. Til tross for at vi bare fant delvis støtte til hypotesen vi jobbet ut i fra, kunne vi observere en liten, men statistisk signifikant effekt av intervensjonen på elevenes leseforståelse ( $d = .30$ ).

I og med at elevene ikke viste en markant fremgang på ordavkodning, er det nærliggende å tolke effekten vi fant på elevenes leseforståelse som et resultat av at undervisningen i ordkunnskap styrket deres språklige kompetanse. Med bakgrunn i de samme resultatene kan man også hevde at studien er med på å avkrefte hypotesen om at ordkunnskap har betydning for avkodning. En slik slutning kan imidlertid være noe forhastet. For å evaluere effekten av intervensjonen på elevenes ordavkodning tok vi i bruk tester som involverte lesing av isolerte ord og nonord. Som det har kommet fram tidligere i avhandlingen, er det imidlertid ikke sikkert at slike mål er sensitive nok til å fange opp sammenhengen mellom ordkunnskap og

avkodning. Det er dermed en liten mulighet for at vi kan ha oversett eventuelle undervisningseffekter på elevenes ordavkodning, og i likhet med mange andre empiriske undersøkelser på feltet kunne også vår studie med fordel ha inneholdt mer teorisensitive ordavkodningsmål.

Når det er sagt, var det ikke nødvendigvis selve undervisningen i ordkunnskap som i realiteten førte til at elevene bedret sine leseforståelsesferdigheter. Intervensjonen vi gjennomførte var relativt kompleks og ikke alle komponentene av undervisningen kan karakteriseres som rent arbeid med ord. For å sette undervisningen i ordkunnskap inn i en meningsfull kontekst, ble for eksempel ordene vi fokuserte på presentert gjennom små fagtekster som elevene var med på å lese og diskutere. Dette kan ha vært med på å gi elevene en mer meningsdrevet tilnærming til tekst slik at de etter intervensjonen leste med høyere grad av kognitivt engasjement. Dessuten inneholdt intervensjonen en relativt høy grad av språklige problemløsningsoppgaver, noe som også kan ha stimulert elevenes driv mot forståelse. Det kan altså tenkes at det var andre faktorer enn ordkunnskap som var instrumentelle for elevenes økte leseforståelse.

Fordi kausale mekanismer ikke kan observeres, er det i prinsippet umulig å fastslå hvilke av komponentene i intervensjonen som utgjorde nødvendige, tilstrekkelige eller ubetydelige betingelser for den observerte effekten på elevenes leseforståelse. For å si noe om virkningen av bestemte enkeltkomponenter, måtte en eventuelt ha gjennomført flere avgrensede studier eller utført eksperimenter med mange sett av ulike forskningsbetingelser. Selv om en slik isolasjon av undervisningskomponenter muligens kunne avdekket hvilke komponenter som var mest «virkningsfulle», er det ikke alltid like meningsfullt å separere ulike undervisningskomponenter fra hverandre. Tidligere studier har for eksempel vist at elever lærer nye ord bedre dersom ordene presenteres i en meningsfull kontekst (Stahl & Fairbanks, 1986). Selv om elevenes lesing og diskusjon av tekst kan ha representert en konfunderende faktor i vår intervensjon, kan denne undervisningskomponenten også ha fasilisert tilegnelsen av ny ordkunnskap. I siste instans er det med andre ord ikke sikkert at det er mulig å peke ut enkeltkomponenter som er ansvarlige for en effekt. Det kan like gjerne tenkes at det er kombinasjonen av ulike komponenter som har vært virkningsfull.

Det var forøvrig ikke bare selve undervisningsopplegget som var komplekst. Også elevene som mottok intervensjonen utgjorde en relativt heterogen gruppe, noe som gjør det vanskelig å trekke slutninger om effekten av undervisningen for spesifikke elevpopulasjoner. Som

tidligere nevnt er en del av diskusjonen om ordkunnskapens plass i undervisningen, spørsmålet om hvem slik undervisning er relevant for. Vil for eksempel elever med avkodningsrelaterte leseforståelsesvansker ha utbytte av undervisning i ordkunnskap, eller burde undervisningen forbeholdes elever med semantisk relaterte vansker? For å undersøke dette spørsmålet burde vi ideelt hatt separate forsøksgrupper bestående av elever med ulike typer lesevansker, eller et utvalg som var stort nok til å kunne utføre subgruppeanalyser. Det fantes likevel noe vi kunne gjøre for å adressere tematikken. Da vi analyserte datamaterialet fra studien undersøkte vi for eksempel om det fantes interaksjoner mellom variablene vi målte ved pretest og den observerte effekten på elevenes leseforståelse. Analysene avdekket imidlertid ingen systematiske sammenhenger mellom bestemte individkarakteristikker og elevenes utbytte av undervisningen. Selv om disse resultatene ikke representerer et direkte svar på spørsmålet om hvem undervisning i ordkunnskap er relevant for, kunne vi i alle fall ikke påvise at effekten av intervensjonen var avhengig av hvilke ferdigheter elevene hadde til å begynne med.

Til tross for begrensningene jeg har påpekt, vil jeg hevde at intervensjonsstudien vi gjennomførte har stor betydning for praksisfeltet. Riktignok ville en isolasjon av undervisningskomponenter og bruk av mer homogene forsøksgrupper gjort resultatene fra studien lettere å tolke, men ofte vil en slik form for eksperimentell kontroll resultere i kunstige forsøkssituasjoner (Cartwright & Munro, 2010). For at en intervensjonsstudie skal være relevant for praksis, er det viktig at undervisningsopplegget som undersøkes kan overføres fra de metodiske rammene som definerer en studie til læreres og elevers skolehverdag. I vår studie prøvde vi derfor å skape en så virkelighetsnær forsøkssituasjon som mulig, for å undersøke om intervensjonen var i stand til å produsere effekter under forhold som er typiske for reelle skolesettinger. Dette innebar blant annet å inkludere elever med til dels ulike læringsforutsetninger i stedet for et klinisk utvalg av elever. Det var også viktig at undervisningsopplegget inneholdt oppgaver og aktiviteter som var varierte og meningsfylte for lærerne og elevene, slik at skoler enkelt skulle kunne ta opplegget i bruk dersom det viste seg å produsere gunstige effekter. Studien representerer med andre ord et verdifullt eksempel på hvordan undervisning i ordkunnskap kan fungere i praksis.

# Litteraturliste

- Al Otaiba, S., & Fuchs, D. (2006). Who Are the Young Children for Whom Best Practices in Reading Are Ineffective? An Experimental and Longitudinal Study. *Journal of Learning Disabilities, 39*(5), 414-431.
- Allen, M. (1997). *Understanding regression analysis*. New York: Plenum Press.
- Anderson, R., & Freebody, P. (1981). Vocabulary Knowledge. In J. Guthrie (Ed.), *Comprehension and Teaching: Research Reviews* (pp. 77–117). Newark, DE: International Reading Association.
- Anglin, J. (1993). Vocabulary Development: A Morphological Analysis. *Monographs of the Society for Research in Child Development, 58*(10), 1-166.
- Bates, E., & Goodman, J. (2001). On the inseparability of grammar and the lexicon: Evidence from acquisition. In M. Tomasello & E. Bates (Eds.), *Language Development. The Essential Readings* (pp. 134-168). Oxford: Blackwell Publishing.
- Beck, I. L., McKeown, M. G., & Kucan, L. (2002). *Bringing words to life: Robust vocabulary instruction*. New York: Guilford Press.
- Befring, E. (2007). *Forskningsmetode med etikk og statistikk*. Oslo: Samlaget.
- Bollen, K., & Curran, P. (2006). *Latent Curve Models: A Structural Equation Perspective*. Hoboken, NJ: John Wiley & Sons.
- Borenstein, M., Hedges, L., Higgins, J., & Rothstein, H. (2009). *Introduction to Meta-Analysis*. West Sussex, UK: John Wiley & Sons.
- Brown, T. (2006). *Confirmatory Factor Analysis for Applied Research*. New York: The Guilford Press.
- Bus, A., & van Ijzendoorn, M. (1999). Phonological awareness and early reading: A meta-analysis of experimental training studies. *Journal of Educational Psychology, 91*(3), 403-414.

- Cain, K., & Oakhill, J. (1999). Inference making ability and its relation to comprehension failure in young children. *Reading and writing: An Interdisciplinary Journal*, 11(5), 489-503.
- Cain, K., & Oakhill, J. (2006). Profiles of children with specific reading comprehension difficulties. *British Journal of Educational Psychology*, 76(4), 683-696.
- Carlo, M., August, D., McLaughlin, B., Snow, C., Dressler, C., Lippman, D., Lively, T. J., White, C. (2004). Closing the gap: Addressing the vocabulary needs of English-language learners in bilingual and mainstream classrooms. *Reading Research Quarterly*, 39(2), 188-215.
- Cartwright, N., & Munro, E. (2010). The limitations of randomized controlled trials in predicting effectiveness. *Journal of Evaluation in Clinical Practice*, 16, 260-266.
- Catts, H., Hogan, T., & Fey, M. (2003). Subgrouping Poor Readers on the Basis of Individual Differences in Reading-Related Abilities. *Journal of Learning Disabilities*, 36(2), 151-164.
- Chall, J. (1996). *Stages of Reading Development* (2 ed.). Fort Worth: Harcourt Brace College Publishers.
- Chall, J., Jacobs, V., & Baldwin, L. (1990). *The Reading Crisis. Why Poor Children Fall Behind*. Cambridge, MA: Harvard University Press.
- Cheung, M. (2015). *Meta-Analysis: A Structural Equation Modeling Approach*. Chichester, UK: John Wiley & Sons.
- Chomsky, N. (1965). *Aspects of the Theory of Syntax*. Cambridge, MA: MIT Press.
- Chomsky, N. (1995). *The Minimalist Program*. Cambridge, MA: MIT Press.
- Cohen, J. (1988). *Statistical Power Analysis for the Behavioural Sciences* (2 ed.). Hillsdale, NJ: Lawrence Erlbaum.
- Cook, T. D., & Campbell, D. T. (1979). *Quasi-Experimentation. Design & Analysis Issues for Field Settings*. Boston: Houghton Mifflin Company.

- Cooper, H. (1998). *Synthesizing Research: A Guide for Literature Reviews* (3 ed.). Thousand Oaks, CA: Sage Publications.
- d'Agostino, R. (1998). Tutorial in biostatistics. Propensity score methods for bias reduction in the comparison of a treatment to a non-randomized control group. *Statistics in Medicine*, *17*(19), 2265-2281.
- Dunn, L., & Dunn, L. (2007). *Peabody Picture Vocabulary Test* (4 ed.). Minneapolis, MN: Pearson.
- Dunn, L., Dunn, L., Whetton, C., & Burley, J. (1997). *The British Picture Vocabulary Scale II*. Windsor: GL Assessment.
- Elleman, A., Lindo, E., Morphy, P., & Compton, D. (2009). The Impact of Vocabulary Instruction on Passage-Level Comprehension of School-Age Children: A Meta-Analysis. *Journal of Research on Educational Effectiveness*, *2*(1), 1-44.
- Ferrer, E., & McArdle, J. (2003). Alternative Structural Models for Multivariate Longitudinal Data Analysis. *Structural Equation Modeling: A Multidisciplinary Journal*, *10*(4), 493-524.
- Foorman, B., Koon, S., Petscher, Y., Mitchell, A., & Truckenmiller, A. (2015). Examining General and Specific Factors in the Dimensionality of Oral Language and Reading in 4th–10th Grades. *Journal of Educational Psychology*, *107*(3), 884-899.
- Freebody, P., & Anderson, R. (1983). Effects of Vocabulary Difficulty, Text Cohesion, and Schema Availability on Reading Comprehension. *Reading Research Quarterly*, *18*(3), 277-294.
- Geiser, C. (2012). *Data Analysis with Mplus*. New York: The Guilford Press.
- Gleser, L., & Olkin, I. (2009). Stochastically Dependent Effect Sizes. In H. Cooper, L. Hedges, & J. Valentine (Eds.), *The Handbook of Research Synthesis and Meta-Analysis* (2nd ed., pp. 357-376). New York: Russell Sage Foundation.
- Gomez, R., & Gerken, L. (1999). Artificial grammar learning by 1-year-olds leads to specific and abstract knowledge. *Cognition*, *70*(2), 109-135.

- Gómez, R., & Gerken, L. (2000). Infant artificial language learning and language acquisition. *Trends in Cognitive Sciences*, 4(5), 178-186.
- Goodwin, A., & Ahn, S. (2010). A Meta-Analysis of Morphological Interventions: Effects on Literacy Achievement of Children with Literacy Difficulties. *Annals of Dyslexia*, 60(2), 183-208.
- Gough, P., & Tunmer, W. (1986). Decoding, reading and reading disability. *Remedial and Special Education*, 7(1), 6-10.
- Graves, M. (2000). A vocabulary program to complement and bolster a middle grade comprehension program. In B. Taylor, M. Graves, & P. van den Broek (Eds.), *Reading for meaning: Fostering comprehension in the middle grades* (pp. 116-135). Newark, DE: International Reading Association.
- Hart, B., & Risley, T. (1995). *Meaningful differences in the everyday lives of young American children*. Baltimore, MD: Paul H Brookes Publishing.
- Hayes, D., & Ahrens, M. (1988). Vocabulary simplification for children: A special case of 'motherese'? *Journal of child language*, 15(02), 395-410.
- Hirsch, E. (2007). *The knowledge deficit: Closing the shocking education gap for American children*. Boston, MA: Houghton Mifflin.
- Hjetland, H., Brinchmann, E., Lyster, S.-A., Hagtvet, B., & Melby-Lervåg, M. (2015). Preschool Predictors of Later Reading Comprehension Ability: A Systematic Review. *The Campbell Library*. Retrieved from <http://www.campbellcollaboration.org/lib/project/319/>
- Hu, L., & Bentler, P. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, 6(1), 1-55.
- Jöreskog, K., & Sörbom, D. (1979). *Advances in Factor Analysis and Structural Equation Models*. Cambridge, MA: Abt Books.
- Kessler, R. (1977). The use of change scores as criteria in longitudinal survey research. *Quality and Quantity*, 11, 43-66.



- Kieffer, M. J., & Lesaux, N. K. (2012). Knowledge of Words, Knowledge about Words: Dimensions of Vocabulary in First and Second Language Learners in Sixth Grade. *Reading and Writing: An Interdisciplinary Journal*, 25(2), 347-373.
- Kirby, J. R., & Savage, R. S. (2008). Can the simple view deal with the complexities of reading? *Literacy*, 42(2), 75-82.
- Klem, M., Gustafsson, J.-E., & Hagtvet, B. (2015). The Dimensionality of Language Ability in Four-Year-Olds: Construct Validation of a Language Screening Tool. *Scandinavian Journal of Educational Research*, 59(2).
- Kline, R. (2011). *Principles and Practice of Structural Equation Modeling* (3 ed.). New York: The Guilford Press.
- Kvernbekk, T. (1997). Kausalitet i pedagogikken? *Nordisk pedagogik*, 17(4), 226-238.
- Marchman, V., & Bates, E. (1994). Continuity in lexical and morphological development: A test of the critical mass hypothesis. *Journal of Child Language*, 21(02), 339-366.
- Marsh, H., Hau, K.-T., & Wen, Z. (2004). In Search of Golden Rules: Comment on Hypothesis-Testing Approaches to Setting Cutoff Values for Fit Indexes and Dangers in Overgeneralizing Hu and Bentler's (1999) Findings. *Structural Equation Modeling: A Multidisciplinary Journal*, 11(3), 320-341.
- McArdle, J. (2009). Latent Variable Modeling of Differences and Changes with Longitudinal Data. *Annual Review of Psychology*, 60, 577-605.
- McArdle, J., & Nesselroade, J. (1994). Using multivariate data to structure developmental change. In S. Cohen & H. Reese (Eds.), *Life-span developmental psychology: Methodological contributions* (pp. 223-267). Hillsdale, NJ: Lawrence Erlbaum Associates.
- Menzies, P., & Price, H. (1993). Causation as a Secondary Quality. *The British Journal for the Philosophy of Science*, 44, 187-203.
- Miller, G., & Chapman, J. (2001). Misunderstanding Analysis of Covariance. *Journal of Abnormal Psychology*, 110(1), 40-48.



- Muter, V., Hulme, C., Snowling, M., & Stevenson, J. (2004). Phonemes, Rimes, Vocabulary, and Grammatical Skills as Foundations of Early Reading Development: Evidence From a Longitudinal Study. *Developmental Psychology, 40*(5), 665-681.
- Nagy, W. (2005). Why Vocabulary Instruction Needs to Be Long-Term and Comprehensive. In E. H. Hiebert & M. L. Kamil (Eds.), *Teaching and learning vocabulary: Bringing research to practice* (pp. 27-44). Mahwah, N.J.: Lawrence Erlbaum Associates.
- Nagy, W., & Herman, P. (1984). Limitations of vocabulary instruction (Tech. Rep. No. 326). Urbana: Center for the Study of Reading, University of Illinois
- Nation, K., Clarke, P., Marshall, C., & Durand, M. (2004). Hidden Language Impairments in Children: Parallels Between Poor Reading Comprehension and Specific Language Impairment? *Journal of Speech, Language, and Hearing Research, 47*(1), 199-211.
- Nation, K., & Snowling, M. (1998). Individual Differences in Contextual Facilitation: Evidence from Dyslexia and Poor Reading Comprehension. *Child development, 69*(4), 996-1011.
- Den nasjonale forskningsetiske komité for samfunnsvitenskap og humaniora (NESH). (2006). Forskningsetiske retningslinjer for samfunnsvitenskap, humaniora, jus og teologi. Retrieved from <https://www.etikkom.no/forskningsetiske-retningslinjer/Samfunnsvitenskap-jus-og-humaniora/>
- National Institute of Child Health and Human Development (NICHD). (2005). Pathways to Reading: The Role of Oral Language in the Transition to Reading. *Developmental Psychology, 41*(2), 428-442.
- Nunnally, J., & Bernstein, I. (1994). *Psychometric Theory* (3 ed.). New York: McGraw-Hill.
- Oakhill, J. V., & Cain, K. (2011). The Precursors of Reading Ability in Young Readers: Evidence From a Four-Year Longitudinal Study. *Scientific Studies of Reading, 16*(2), 91-121.
- Ouellette, G., & Beers, A. (2010). A not-so-simple view of reading: how oral vocabulary and visual-word recognition complicate the story. *Reading and Writing: An Interdisciplinary Journal, 23*(2), 189-208.

- 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.
- Owen, S., & Froman, R. (1998). Uses and Abuses of the Analysis of Covariance. *Research in Nursing & Health*, 21, 557-562.
- Pearl, J. (2009). *Causality: Models, Reasoning, and Inference* (2 ed.). Cambridge: Cambridge University Press.
- Perfetti, C. (2007). Reading Ability: Lexical Quality to Comprehension. *Scientific Studies of Reading*, 11(4), 357-383.
- Perfetti, C., & Hart, L. (2002). The lexical quality hypothesis. In L. Verhoeven, C. Elbro, & P. Reitsma (Eds.), *Precursors of functional literacy* (pp. 189–213). Amsterdam: John Benjamins.
- Perfetti, C., & Stafura, J. (2014). Word knowledge in a theory of reading comprehension. *Scientific Studies of Reading*, 18(1), 22-37.
- Perry, L., Samuelson, L., Malloy, L., & Schiffer, R. (2010). Learn Locally, Think Globally. Exemplar Variability Supports Higher-Order Generalization and Word Learning. *Psychological Science*, 21(12), 1894-1902.
- Pinker, S. (1998). Words and rules. *Lingua*, 106(1-4), 219-242.
- Quinn, J., Wagner, R., Petscher, Y., & Lopez, D. (2015). Developmental Relations Between Vocabulary Knowledge and Reading Comprehension: A Latent Change Score Modeling Study. *Child development*, 86(1), 159-175.
- Ricketts, J., Nation, K., & Bishop, D. (2007). Vocabulary Is Important for Some, but Not All Reading Skills. *Scientific Studies of Reading*, 11(3), 235-257.
- Rosenbaum, P. R., & Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika*, 70(1), 41-55.
- Roth, A. (2006). Causation. In S. Traiger (Ed.), *The Blackwell Guide to Hume's Treatise* (pp. 95-113). Oxford: Blackwell Publishing.

- Saffran, J. (2001). Words in a sea of sounds: The output of infant statistical learning. *Cognition*, 81(2), 149-169.
- Saffran, J., Newport, E., & Aslin, R. (1996). Word Segmentation: The Role of Distributional Cues *Journal of memory and language*, 35(4), 606-621.
- Saffran, J., & Thiessen, E. (2007). Domain-General Learning Capacities. In E. Hoff & M. Shatz (Eds.), *Blackwell handbook of language development* (pp. 68-86). Oxford, UK: Blackwell Publishing.
- Saffran, J., & Wilson, D. (2003). From Syllables to Syntax: Multilevel Statistical Learning by 12-Month-Old Infants. *Infancy*, 4(2), 273-284.
- Sanacore, J., & Palumbo, A. (2008). Understanding the Fourth-Grade Slump: Our Point of View. *The Educational Forum*, 73(1), 67-74.
- Saxton, M. (2010). *Child Language: Acquisition and Development*. London, UK: Sage.
- Scammacca, N., Roberts, G., & Stuebing, K. K. (2014). Meta-Analysis With Complex Research Designs: Dealing With Dependence From Multiple Measures and Multiple Group Comparisons. *Review of Educational Research*, 84(3), 328-364.
- Schlueter, E., Davidov, E., & Schmidt, P. (2007). Applying Autoregressive Cross-Lagged and Latent Growth Curve Models to a Three-Wave Panel Study. In K. Montfort, H. Oud, & A. Satorra (Eds.), *Longitudinal Models in the Behavioral and Related Sciences* (pp. 315-336). Mahwah, NJ: Lawrence Erlbaum.
- Schmidt, F., & Hunter, J. (2015). *Methods of Meta-Analysis: Correcting Error and Bias in Research Findings* (3 ed.). Thousand Oaks, CA: Sage.
- Seidenberg, M., & McClelland, J. (1989). A distributed, developmental model of word recognition and naming. *Psychological Review*, 96(4), 523-568.
- Shadish, W., Cook, T., & Campbell, D. (2002). *Experimental and Quasi-Experimental Designs for Generalized Causal Inference*. Boston: Houghton-Mifflin.
- Share, D. L. (1995). Phonological recoding and self-teaching: Sine qua non of reading acquisition. *Cognition*, 55(2), 151-218.

- Snowling, M., Gallagher, A., & Frith, U. (2003). Family Risk of Dyslexia Is Continuous: Individual Differences in the Precursors of Reading Skill. *Child Development, 74*(2), 358-373.
- Snowling, M., & Hulme, C. (2011). Evidence-based interventions for reading and language difficulties: Creating a virtuous circle. *British Journal of Educational Psychology, 81*(1), 1-23.
- Snowling, M., & Hulme, C. (2012). Interventions for children's language and literacy difficulties. *International Journal of Language & Communication Disorders, 47*(1), 27-34.
- Snowling, M., & Melby-Lervåg, M. (2016). Oral Language Deficits in Familial Dyslexia: A Meta-Analysis and Review. *Psychological Bulletin*, (online a-head-of-print). doi:10.1037/bul0000037
- Stahl, S., & Fairbanks, M. (1986). The Effects of Vocabulary Instruction: A Model-Based Meta-Analysis. *Education & Educational Research, 56*(1), 72-110.
- Stanovich, K. (1980). Toward an interactive-compensatory model of individual differences in the development of reading fluency. *Reading Research Quarterly, 16*(1), 32-71.
- Stanovich, K., West, R., & Feeman, D. (1981). A longitudinal study of sentence context effects in second-grade children: Tests of an interactive-compensatory model. *Journal of Experimental Child Psychology, 32*(2), 185-199.
- Sternberg, R. (1987). Most vocabulary is learned from context. In M. McKeown & M. Curtis (Eds.), *The Nature of Vocabulary Acquisition* (pp. 89-105). Hillsdale, NJ: Erlbaum.
- Storch, S., & Whitehurst, G. (2002). Oral language and code-related precursors to reading: Evidence from a longitudinal structural model. *Developmental Psychology, 38*(6), 934-947.
- Stuart, E. (2010). Matching methods for causal inference: A review and a look forward. *Statistical Science, 25*(1), 1-21.

- Tilstra, J., McMaster, K., Van den Broek, P., Kendeou, P., & Rapp, D. (2009). Simple but complex: Components of the simple view of reading across grade levels. *Journal of Research in Reading, 32*(4), 383-401.
- Tomblin, J., & Zhang, X. (2006). The Dimensionality of Language Ability in School-Age Children. *Journal of Speech, Language, and Hearing Research, 49*, 1193-1208.
- Tunmer, W., & Chapman, J. (2012). The Simple View of Reading Redux: Vocabulary Knowledge and the Independent Components Hypothesis. *Journal of Learning Disabilities, 45*(5), 453-466.
- Tyler, A., & Nagy, W. (1989). The acquisition of English derivational morphology. *Journal of memory and language, 28*(6), 649-667.
- Vellutino, F. R., Tunmer, W. E., Jaccard, J. J., & Chen, R. (2007). Components of reading ability: Multivariate evidence for a convergent skills model of reading development. *Scientific Studies of Reading, 11*(1), 3-32.
- West, R., Stanovich, K., Feeman, D., & Cunningham, A. (1983). The Effect of Sentence Context on Word Recognition in Second- and Sixth-Grade Children *Reading Research Quarterly, 19*(1), 6-15.
- West, S., Taylor, A. B., & Wu, W. (2012). Model fit and model selection in structural equation modeling. In R. Hoyle (Ed.), *Handbook of structural equation modeling* (pp. 209-231). New York: The Guilford Press.
- Yu, C., & Smith, L. (2007). Rapid Word Learning Under Uncertainty via Cross-Situational Statistics. *Psychological Science, 18*(5), 414-420.

# Liste over vedlegg

Vedlegg 1: Tilråding fra Norsk samfunnsvitenskapelig datatjeneste

Vedlegg 2: Informasjonsskriv kartleggingsstudie

Vedlegg 3: Informasjonsskriv intervensjonsstudie



Harald Hårfagres gate 29  
N-5007 Bergen  
Norway  
Tel: +47-55 58 21 17  
Fax: +47-55 58 96 50  
nsd@nsd.uib.no  
www.nsd.uib.no  
Org nr. 985 321 884

Ellen Brinchmann  
Institutt for spesialpedagogikk  
Universitetet i Oslo  
Postboks 1140 Blindern  
0318 OSLO

Vår dato: 22.08.2012

Vår ref:31178 / 3 / B

Deres dato:

Deres ref:

## TILBAKEMELDING PÅ MELDING OM BEHANDLING AV PERSONOPPLYSNINGER

Vi viser til melding om behandling av personopplysninger, mottatt 17.08.2012. Meldingen gjelder prosjektet:

31178	<i>Morfologisk bevissthet og skriftspråktilegnelse</i>
Behandlingsansvarlig	<i>Universitetet i Oslo, ved institusjonens øverste leder</i>
Daglig ansvarlig	<i>Ellen Brinchmann</i>

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

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

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

Personvernombudet har lagt ut opplysninger om prosjektet i en offentlig database, <http://www.nsd.uib.no/personvern/prosjektoversikt.jsp>.

Personvernombudet vil ved prosjektets avslutning, 01.08.2017, rette en henvendelse angående status for behandlingen av personopplysninger.

Vennlig hilsen

  
Bjørn Henrichsen

  
Inga Brautaset

Kontaktperson: Inga Brautaset tlf: 55 58 26 35  
Vedlegg: Prosjektvurdering





**UNIVERSITETET I OSLO**  
DET UTDANNINGSVITENSKAPELIGE FAKULTET

**Ønsker dere å bidra til økt kunnskap om barns  
leseutvikling?**

*Forespørsel om å delta i studie*

Ved Det utdanningsvitenskapelige fakultet, Universitetet i Oslo, har vi en bred forskningssatsing på barns språk- og leseutvikling. Vi vet godt at språk er viktig for læring generelt og for utviklingen av lese- og skriveferdigheter spesielt. Vi ønsker imidlertid å øke vår kunnskap om barns talespråklige og skriftspråklige utvikling, og dermed bidra til å styrke språk- og leseopplæringen i skolen.

Vi henvender oss til dere for å invitere til prosjektsamarbeid fordi dere har et barn som er født i 2004 eller 2005. I alt henvender vi oss til barn fra 13 skoler i [redacted] kommune. Det vi ber om er tillatelse til å følge deres barn over en periode på to skoleår. En eventuell deltakelse i prosjektet vil innebære at barnets talespråklige og skriftspråklige ferdigheter blir kartlagt én gang per halvår. Konkret dreier det seg om kartlegging av ordforråd, grammatikk, lese- og skriveferdigheter samt problemløsningsferdigheter. Skriftlige prøver vil bli administrert klassevis mens den muntlige kartleggingen vil foregå individuelt. Den muntlige kartleggingen vil da bli tatt opp på lydband. Kartleggingen vil finne sted på barnets skole og gjennomføres av ansatte fra Det utdanningsvitenskapelige fakultetet som har erfaring med barn. Vår erfaring er at dette er noe barna synes er morsomt og lystbetont å være med på.

Ved første kartlegging ønsker vi også å hente inn informasjon fra foresatte om barnets språkbakgrunn og miljø (eks. hvilket språk som snakkes hjemme, om det har vært uregelmessigheter knyttet til barnets språkutvikling etc.). Dette vil samles inn ved hjelp av et kort spørreskjema og vil gi oss viktig utfyllende informasjon til undersøkelsen.

Det er selvsagt frivillig å delta i undersøkelsen og det er mulig å trekke seg fra deltakelse på et hvilket som helst tidspunkt uten å gi begrunnelse. Alle personopplysninger vil da bli anonymisert. Vi vil også nevne at det er skolen som formidler denne henvendelsen og vi er derfor ikke kjent med barnets navn før dere som foresatte eventuelt samtykker til deltakelse.

Datamaterialet vil bli behandlet konfidensielt og vi som arbeider med prosjektet har taushetsplikt. Opplysningene som samles inn vil bli brukt i studier av barns talespråklige og skriftspråklige utvikling i forskergruppen Child Language & Learning ved det Utdanningsvitenskapelige fakultetet i Oslo. Resultatene presenteres i en doktorgradsavhandling, eventuelt også flere publikasjoner, uten at det enkelte barn kan





**UNIVERSITETET I OSLO**  
DET UTDANNINGSVITENSKAPELIGE FAKULTET

gjenkjennes. Alle personopplysninger og lydopptak slettes når doktorgradsprosjektet er avsluttet, senest innen 31.08.2017. Det kan bli aktuelt med en oppfølgingsstudie. Du vil i så fall motta ny informasjon og forespørsel om å delta før hovedprosjektet avsluttes. Prosjektet er meldt til Personvernombudet for forskning, Norsk samfunnsvitenskapelig datatjeneste (NSD).

Dersom dere har anledning til å delta i undersøkelsen, ber vi dere fylle ut vedlagte samtykkeerklæring og returnere det i svarkonvolutt snarest mulig. Vi vil samarbeide med skolen om når kartleggingen kan finne sted. Første møte med barnet blir i løpet av høsten 2012.

Ta gjerne kontakt med Ellen Brinchmann dersom dere har spørsmål. På forhånd takk for verdifull hjelp.

Med vennlig hilsen

---

Solveig A. H. Lyster  
Professor

---

Ellen Brinchmann  
Stipendiat  
22854877/98459174  
e.i.brinchmann@isp.uio.no

**UiO**  **Det utdanningsvitenskapelige fakultet**  
Universitetet i Oslo

## ***Kjære foresatte!***

### **Invitasjon til videre prosjektsamarbeid**

Universitetet i Oslo arbeider med et forskningsprosjekt i [redacted] kommune, hvor vi undersøker barns talespråklige og skriftspråklige utvikling. Vi henvender oss til dere fordi dere har takket ja til at barnet deres kan delta i dette prosjektet, og vi ønsker å invitere til et videre prosjektsamarbeid i en oppfølgingsstudie.

I tilknytning til det nevnte forskningsprosjektet har det blitt utarbeidet et pedagogisk opplegg som inneholder arbeid med tale- og skriftspråklige ferdigheter. Vi ønsker å undersøke om dette opplegget kan ha positiv innvirkning på barns skoleferdigheter. Barnets skole formidler denne henvendelsen fra oss fordi de har vurdert at deres barn kan dra nytte av opplegget. I samarbeid med skolen ønsker vi derfor å tilby opplegget til deres barn.

Det pedagogiske opplegget vil kunne beskrives som et intensivt lesekurs, hvor det med utgangspunkt i korte fagtekster vil arbeides med barnas leseferdigheter og ordkunnskap. Fagtekstene er valgt ut i samarbeid med Skoletjenesten ved Deichmanske bibliotek i Oslo og involverer temaer knyttet til fagene norsk, naturfag, historie og samfunnsfag. Opplegget vil foregå i en liten gruppe på 4-6 elever og gjennomføres av pedagogisk personale ved barnets skole. Det strekker seg over 10 uker og består av 3 økter i uken som hver varer i ca 45 minutter. Oppstart vil være i januar 2014. Dersom dere har planlagt flytting i løpet av prosjektperioden, kan deres barn ikke delta i studien.

Ferdighetene som det arbeides med i opplegget vil kartlegges før og etter gjennomføringen av prosjektet, slik at man kan undersøke hvilken innvirkning opplegget har hatt på disse ferdighetene. Denne kartleggingen vil gjennomføres individuelt og vil ta 2 x 45 minutter for hvert barn. Det vil i enkelte tilfeller bli brukt lydopptak under kartleggingen samt ved gjennomføring av det pedagogiske opplegget. Dette materialet vil bli lagret på eksterne sikrede harddisker.

I tillegg til å følge barna, ønsker vi å hente inn informasjon om barnets språkbakgrunn og miljø (eks. om det har vært uregelmessigheter knyttet til barnets språkutvikling, foreldres utdanningsbakgrunn etc.). Dette vil samles inn ved hjelp av et kort spørreskjema til foresatte og vil gi oss viktig utfyllende informasjon til undersøkelsen.

Datamaterialet vil bli behandlet konfidensielt og vi som arbeider med prosjektet har taushetsplikt. Resultatene presenteres i en doktorgradsavhandling, eventuelt også flere publikasjoner, uten at det enkelte barn kan gjenkjennes. Alle personopplysninger og lydopptak slettes når doktorgradsprosjektet er avsluttet, senest innen 31.8.2017. Prosjektet er meldt til Personvernombudet for forskning, Norsk samfunnsvitenskapelig datatjeneste (NSD).

Det er selvsagt frivillig å delta i undersøkelsen og barnet kan trekke seg underveis dersom det ønsker det. Dersom man trekker seg vil alle innsamlede opplysninger bli anonymisert (navnelister slettes). Det vil ikke få konsekvenser for barnets forhold til skole dersom man ikke ønsker at barnet skal delta i prosjektet eller trekker seg fra prosjektet.

Dersom dere har anledning til å delta i oppfølgingsstudien, ber vi dere fylle ut vedlagte samtykkeerklæring og levere det til barnets skole så snart som mulig.

Ta gjerne kontakt med Ellen Brinchmann dersom dere har spørsmål. På forhånd takk for verdifull hjelp.

Med vennlig hilsen

Solveig A. H. Lyster  
Professor

Ellen Brinchmann  
Stipendiat  
22854877/98459174  
e.i.brinchmann@isp.uio.no



# Artikler

**Artikkel I:** Hjetland, H.N., Brinchmann, E.I., & Lyster, S.A.H. (2015). The Role of Vocabulary Knowledge in the Development of Reading Comprehension. *Manuskript*.

**Artikkel II:** Brinchmann, E.I., Braeken, J., & Lyster, S.A.H. (2015). The Relationship between Grammar and the Lexicon Revisited. *Manuskript*.

**Artikkel III:** Brinchmann, E.I., Hjetland, H.N., & Lyster, S.A.H. (2015). Lexical Quality Matters: Effects of Word Knowledge Instruction on the Language and Literacy Skills of Third-and Fourth-Grade Poor Readers. *Reading Research Quarterly* (online a-head-of-print).  
doi:10.1002/rrq.128

# Errataliste

**Doktorkandidat:** Ellen Irén Brinchmann

**Avhandlingens tittel:** Betydningen av ordkunnskap for utviklingen av leseforståelse

Sted	Original tekst	Korrigert tekst	Type rettelse
Kappa, vedlegg 2 s. 66, linje 7		I alt henvender vi oss til barn fra 13 skoler i [redacted] kommune.	Navn på kommune har blitt strøket av hensyn til deltakernes anonymitet.
Kappa, vedlegg 3 s. 68, linje 1		Universitetet i Oslo arbeider med et forskningsprosjekt i [redacted] kommune, hvor vi undersøker barns talespråklige og skriftspråklige utvikling.	Navn på kommune har blitt strøket av hensyn til deltakernes anonymitet.
Artikkel 1, s. 6, linje 8	... National Early Literacy Panel (NELP; 2008) concluded that vocabulary knowledge measured before the onset of formal reading instruction is a fairly weak predictor of later reading comprehension.	... National Early Literacy Panel (NELP; 2008) concluded that receptive vocabulary knowledge measured before the onset of formal reading instruction is a fairly weak predictor of later reading comprehension.	Korrektur: ordet «receptive» har blitt lagt til for å presisere funnene i studien det refereres til.
Artikkel 1, s. 23, linje 8	This means that although it is important to provide empirical validation for vocabulary instruction, it mayt be a mistake to dismiss vocabulary as a curricular objective simply because the effects of short term vocabulary interventions are modest.	This means that although it is important to provide empirical validation for vocabulary instruction, it may be a mistake to dismiss vocabulary as a curricular objective simply because the effects of short term vocabulary interventions are modest.	Korrektur: «mayt» har blitt endret til «may»
Artikkel 2, s. 9, linje 12	Yet, as of today, we know surprisingly little about the relationship between grammar and the lexicon beyond the first two or three years of life, and it is therefore difficult to make strong predictions regarding the subsequent development.	Yet, today, we know surprisingly little about the relationship between grammar and the lexicon beyond the first two or three years of life, and it is therefore difficult to make strong predictions regarding the subsequent development.	Korrektur: «as of today» har blitt endret til «today».