



UiO • Universitetet i Oslo

# Automatisering som forutsetning for leseflyt

*En kvantitativ undersøkelse av et nytt mål på automatisert  
ordgjenkjenning*

Caroline Svendsen Nordli

Masteroppgave i Spesialpedagogikk  
Institutt for Spesialpedagogikk

Universitetet i Oslo

Vår 2020

## Sammendrag

*Leseflyt* har blitt identifisert til å være en kritisk komponent i utviklingen av funksjonelle leseferdigheter, spesielt på grunn av den nære sammenhengen leseflyt ser ut til å med leseforståelse (Kuhn, Schwanenflugel & Meisinger, 2010). Utviklingen fra langsom ordavkodning til en umiddelbar gjenkjenning av ord – forstått som *automatisert ordgjenkjenning*, menes å være helt avgjørende for å kunne utvikle leseflyt (Ehri, 2005). Hastighet har i denne sammenheng vært sett på som hovedindikator på automatisering. Foreløpig ser automatisert ordgjenkjenning kun ut til å ha blitt målt gjennom ord som presenteres isolert, noe som avviker fra en normal lesesituasjon. Avanserte, effektive lesere ser derimot ut til å være i stand til å lese tekst der flere ord presenteres samtidig, hvilket indikerer at disse leserne er i stand til å lese ord hurtig og korrekt uten at tilstøtende ord forstyrrer prosessen. Det er slik sett foreslått at automatisering også involverer en form for sekvensiell prosessering av ord (Altani, Protopapas, Katopodi & Georgiou, 2020). Formålet med gjeldende studie er å bruke et nytt mål på automatisert ordgjenkjenning, for å undersøke hvorvidt ord som blir presentert isolert og ord som blir presentert med tilstøtende ord differensierer i å predikere leseflyt. Dette med et ønske om at resultatene kan bidra til mer inngående kunnskap om automatiseringens rolle i lesing, spesielt i lys av leseflytutvikling.

### Metode

Deltakere besto av 62 voksne, hovedsakelig universitetsstudenter mellom 19-30 år, uten kjente lærevansker. For å kartlegge hastighet på automatisert ordgjenkjenning ble det brukt et eksperiment basert på øyesporing og et «visual world paradigm». Dette involverte bakvendt maskering av ord, hvorpå disse ble presentert isolert, med tilstøtende ord, og med tilstøtende symboler, for så å dekkes av en maske etter 75 ms. Hvert målord ble presentert med fire bilder, hvorpå ett av bildene korresponderte med ordet. Hastigheten ble registrert gjennom deltakeres museklikk på valgt bilde. Eksperimentet besto av totalt 120 oppgaver av denne typen, fordelt på tre blokker med 40 oppgaver i hver blokk. Leseflyt ble kartlagt gjennom norsk versjon av TOWRE.

## **Resultater**

Resultatene fra den bivariate korrelasjonsanalysen viser at leseflyt korrelerer på .01 eller .05 med tre av fire forklaringsvariabler, og antyder en tydelig sammenheng mellom ordgjenkjenningshastighet og leseflyt. Gjennomført regresjonsanalyse kunne imidlertid ikke vise til en signifikant forskjell i hvorvidt ord som presenteres isolert og ord som presenteres med tilstøtende ord predikerer leseflyt. utfordringer knyttet til multikollinearitet, størrelse på utvalg og antall forklaringsvariabler antas å påvirke resultatet.

## **Konklusjon**

Gjeldende studie kan ikke gi noe endelig svar på hvorvidt de to presentasjonsformene differensierer i å predikere leseflyt. Da utvalget antas å bestå av hovedsakelig gode, funksjonelle lesere, argumenterer det for å undersøke om resultatet blir tilsvarende med et utvalg som i større grad består av yngre lesere som er i prosessen med å utvikle leseferdigheter.

## **Forord**

Først og fremst vil jeg rette en takk til Universitetet i Oslo med Athanassios Protopapas og Laoura Ziaka for muligheten til å bli med på å pilotere deres forskningsprosjekt. Dere har gitt meg masse nyttige erfaringer som jeg vil ta med meg videre. Det setter jeg pris på.

En ekstra takk til min veileder, Athanassios Protopapas, som alltid har vært tilgjengelig og kommet med konstruktive tilbakemeldinger fra start til slutt i denne prosessen. Det inspirerer.

Takk til min medstudent Dzan Zelihic for godt samarbeid.

Takk til alle mine medstudenter for to fine år på Universitetet i Oslo.

Helt til slutt vil jeg takke min familie og min samboer, som alltid er støttende og motiverende når jeg trenger det.

# Innholdsfortegnelse

<b>SAMMENDRAG</b> .....	<b>II</b>
<b>FORORD</b> .....	<b>IV</b>
<b>INNHOLDSFORTEGNELSE</b> .....	<b>V</b>
<b>1. INNLEDNING</b> .....	<b>9</b>
1.1 PROBLEMSTILLING .....	10
1.2 OPPGAVENS OPPBYGGING .....	11
<b>2. TEORETISK BAKGRUNN</b> .....	<b>12</b>
2.1 LESING = AVKODING X FORSTÅELSE? .....	12
2.2 LESEFLYT .....	13
2.3 UTVIKLINGEN AV LESEFLYT - FRA AVKODING TIL ORDGJENKJENNING .....	14
2.4 LABERGE OG SAMUELS AUTOMATISERINGSTEORI – ET VIKTIG BIDRAG .....	15
2.5 LOGANS KRITERIER FOR AUTOMATISERTE FERDIGHETER .....	17
2.6 LESEFLYT = AUTOMATISERING? .....	18
2.7 AUTOMATISERING – NØDVENDIG MED ANDRE PERSPEKTIVER? .....	19
2.8 «INTERFERENCE» SOM RETTESNOR? .....	20
2.9 BRUK AV «MASKING» I SELVSTENDIG MÅL PÅ AUTOMATISERT ORDGJENKJENNING? .....	21
2.10 AUTOMATISERING SOM «MULTI-ITEM PROCESSING» .....	22
2.11 Å BRUKE «VISUAL WORLD PARADIGM» I AVDEKKING AV AUTOMATISERT ORDGJENKJENNING	23
2.12 TEORETISK OPPSUMMERING OG IMPLIKASJONER FOR OPPGAVENS FORSKNINGSSPØRSMÅL	24
<b>3. METODE</b> .....	<b>26</b>
3.1 UTVALG .....	26
3.2 INNSAMLING AV DATA .....	26

3.3	EKSPERIMENTELL OPPGAVE .....	27
3.4	GJENNOMFØRING.....	32
3.5	VALIDITET OG RELIABILITET .....	32
3.5.1	<i>Indre validitet</i> .....	33
3.5.2	<i>Ytre validitet</i> .....	33
3.5.3	<i>Statistisk validitet</i> .....	34
3.5.4	<i>Begrepsvaliditet</i> .....	34
3.6	FORSKNINGSETISKE HENSYN .....	35
<b>4.</b>	<b>RESULTATER</b> .....	<b>36</b>
4.1	DESKRIPTIVE ANALYSER .....	36
4.1.1	<i>Vurdering av variabelen leseflyt</i> .....	39
4.1.2	<i>Vurdering av variablene som representerer automatisk ordgjenkjenning</i> .....	40
4.1.3	<i>Vurdering av variabelen ordgjenkjenningshastighet på isolert høyfrekvente ord</i> .....	40
4.1.4	<i>Vurdering av variabelen ordgjenkjenningshastighet på isolerte lavfrekvente ord</i> .....	42
4.1.5	<i>Vurdering av variabelen ordgjenkjenningshastighet på høyfrekvente ord med ord-flanker</i>	43
4.2	SAMLET VURDERING AV VARIABLENES FORDELING .....	45
4.3	BIVARIAT KORRELASJONSANALYSE.....	45
4.3.1	<i>Resultater bivariat korrelasjonsanalyse</i> .....	46
4.4	REGRESJONSANALYSE .....	49
4.4.1	<i>Transformering av data</i> .....	50
4.4.2	<i>Statistiske forutsetninger for regresjonsanalyse</i> .....	52
4.4.3	<i>Resultater av regresjonsanalyser</i> .....	55

4.5	OPPSUMMERING AV RESULTATER .....	60
<b>5.</b>	<b>DISKUSJON .....</b>	<b>61</b>
5.1	DRØFTING AV RESULTATENE I LYS AV TEORI.....	61
5.2	DRØFTING AV RESULTATENE I LYS AV STUDIENS VALIDITET OG RELIABILITET .....	63
5.2.1	<i>Statistisk validitet</i> .....	63
5.2.2	<i>Ytre validitet</i> .....	65
5.2.3	<i>Indre validitet</i> .....	66
5.2.4	<i>Begrepsvaliditet</i> .....	67
<b>6.</b>	<b>KONKLUSJON .....</b>	<b>69</b>
	<b>LITTERATURLISTE .....</b>	<b>70</b>
	<b>APPENDIKS 1: SAMTYKKESKJEMA.....</b>	<b>75</b>
	<b>APPENDIKS 2: NORSK VERSJON AV TOWRE.....</b>	<b>78</b>
	<b>APPENDIKS 3: HØYFREKVENTE ORD .....</b>	<b>79</b>
	<b>APPENDIKS 4: LAVFREKVENTE ORD.....</b>	<b>81</b>
	<b>APPENDIKS 5: LISTE OVER ALLE ELEMENTER I EKSPERIMENTELL OPPGAVE.....</b>	<b>83</b>
	<b>APPENDIKS 6: ILLUSTRASJONER FRA EKSPERIMENTELL OPPGAVE.....</b>	<b>87</b>
	<b>APPENDIKS 7: KOLINEARITETSSTATISTIKK .....</b>	<b>90</b>

## Liste over figurer:

Figur 1: Histogram av variabelen leseflyt .....	39
Figur 2: Q-Q-diagram av variabelen leseflyt.....	39
Figur 3:Histogram av variabelen ordgjenkjenningshastighet på isolerte høyfrekvente ord .....	41
Figur 4:Q-Q-diagram av variabelen ordgjenkjenningshastighet på isolerte høyfrekvente ord.....	41
Figur 5: Histogram av variabelen ordgjenkjenningshastighet på isolerte lavfrekvente ord .....	42
Figur 6: Q-Q-diagram av variabelen ordgjenkjenningshastighet på isolerte lavfrekvente ord.....	42
Figur 7:Histogram av variabelen ordgjenkjenningshastighet på høyfrekvent ord med ord-flanker .....	43
Figur 8:Q-Q-diagram av variabelen ordgjenkjenningshastighet på høyfrekvent ord med ord-flanker .....	43
Figur 9:Histogram av variabelen ordgjenkjenningshastighet på lavfrekvente ord med ord-flanker.....	44
Figur 10:Q-Q-diagram av variabelen ordgjenkjenningshastighet på lavfrekvente ord med ord-flanker .....	44
Figur 11:Q-Q-diagram, første sekvens.....	52
Figur 12: Q-Q-diagram, andre sekvens .....	52
Figur 13: Q-Q-diagram, tredje sekvens.....	52
Figur 14: Punktdiagram over residualer, første sekvens.....	53
Figur 15: Punktdiagram over residualer, andre sekvens.....	53
Figur 16: Punktdiagram over residualer, tredje sekvens.....	53

## Liste over tabeller:

Tabell 1: Deskriptiv statistikk for eksperimentell oppgave .....	30
Tabell 2: Deskriptiv statistikk.....	38
Tabell 3: Korrelasjonsmatriks.....	47
Tabell 4 :Deskriptiv statistikk over transformerte variabler med LOG10 .....	51
Tabell 5: Koeffisienter fra regresjonsanalyse, høyfrekvent ordgjenkjenningshastighet .....	55
Tabell 6: Oppsummerte resultater regresjonsanalyse, høyfrekvent ordgjenkjenningshastighet.....	55
Tabell 7: Modellsammenligning, høyfrekvent ordgjenkjenningshastighet.....	56
Tabell 8: Koeffisienter fra regresjonsanalyse, lavfrekvent ordgjenkjenningshastighet .....	57
Tabell 9: Oppsummerte resultater fra regresjonsanalyse, lavfrekvent ordgjenkjenningshastighet.....	57
Tabell 10: Modellsammenligning, lavfrekvent ordgjenkjenningshastighet.....	58
Tabell 11: Koeffisienter fra regresjonsanalyse, alle forklaringsvariabler .....	59
Tabell 12: Oppsummerte resultater fra regresjonsanalyse, alle forklaringsvariabler .....	59



# 1. Innledning

Et begrep som løftes frem i St.meld.nr.20 (Kunnskapsdepartementet, 2013) er *kunnskapssamfunnet*, hvor det argumenteres for at dagens samfunn stiller større krav til kompetanse enn noen gang før. Utviklingen av funksjonelle leseferdigheter synes i denne sammenheng å være helt avgjørende for å imøtekomme disse kravene. Til tross for at lesing ser ut til å ha en helt sentral rolle i overordnet del av læreplanverket, er det likevel slik at nesten 20 prosent av elevene som går ut av grunnskolen har så svake leseferdigheter at de er i risiko for å få utfordringer med videre skolegang og arbeidsliv (Kunnskapsdepartementet, 2020; Kjærnsli & Olsen, 2013). Svake leseferdigheter virker derfor å være hemmende for dem det angår, og det kan videre tenkes at det vil kunne påvirke flere aspekter ved et individs liv, også utenfor skole- og jobbsammenheng. Som spesialpedagog er derfor kunnskap om komponenter som inngår i lesing avgjørende for å kunne identifisere og følge opp de barna som strever med å oppnå tilstrekkelige leseferdigheter.

Innenfor leseforskning har det i stor grad å ha vært fokusert på to kjerneelementer som forklares som sentrale i utviklingen av gode leseferdigheter, henholdsvis *leseforståelse* og *leseflyt*. Leseforståelse blir gjerne forstått som det overordnede målet med lesing, mens leseflyt forstås som ferdigheten som muliggjør dette (Wolf & Katzir-Cohen, 2009). En utfordring som gjelder leseflyt, er imidlertid at konseptet ser ut til å være preget av manglende konsensus, og det oppleves ved nærmere ettersyn å være nokså uklart hvilke komponenter som inngår i det (Wolf & Katzir-Cohen, 2009; Kuhn et al., 2010, Hudson et al., 2009). Det er likevel en enighet om at det å kunne gjenkjenne ord hurtig og korrekt, forstått som *automatisert ordgjenkjenning*, har en betydning for flytutvikling (Ehri, 2005; Rasinski, Reutzel, Chard & Linan-Thompson, 2011).

Til tross for dette, er det få studier som virker å ha undersøkt hvorvidt automatisert ordgjenkjenning kan måles uavhengig av subjektive kjennetegn. Hastighet har vanligvis blitt brukt som hovedindikator, og den dominerende oppfatningen er at høyfrekvente ord gjenkjennes raskere enn lavfrekvente ord (Brysbart, Mandere og Keuleers, 2018). Nylig ble det gjennomført en studie som utviklet et nytt mål på automatisert ordgjenkjenning. Målet baserte seg på individuelle ord som ble presentert isolert, hvorpå dette viste seg å

kunne brukes til å predikere leseflyt (Roembke, Hazeltine, Reed & McMurray, 2018). Hvorvidt dette kan være et valid mål på automatisert ordgjenkjenning kan imidlertid diskuteres, da lesing som aktivitet innebærer at en alltid forholder seg til flere ord som presenteres samtidig. Altani et al. (2020) synes ved sin forskning å bidra med argumenter for at utviklingen av leseflyt avhenger av mer enn kun hurtig og korrekt individuell ordgjenkjenning alene, og at automatisering i stor grad kan knyttes til det å kunne prosessere ord sekvensielt. Gjeldende oppgave tar dermed utgangspunkt i en hypotese om at individer som ikke oppnår leseflyt, i realiteten kan ha vansker med såkalt sekvensiell prosessering. Dette er noe som ikke ser ut til å være undersøkt innenfor gjeldende kontekst, men baseres på forskning som har vist at nære bokstaver later til å forstyrre prosesseringen av en målbokstav (Bouma, 1970; Eriksen & Eriksen, 1974).

Målet med denne studien er derfor å ta i bruk et nytt mål på automatisert ordgjenkjenning, hvor ord i tillegg til å bli presentert individuelt og isolert, også blir presentert med tilstøtende ord, noe som vil være nærmere en ordinær lesesituasjon. En sammenligning vil dermed gjøres med utgangspunkt i hastighet, for å undersøke hvorvidt de to ulike måtene å presentere ord på vil differensiere i å predikere leseflyt.

## 1.1 Problemstilling

Oppgavens forskningsspørsmål er som følger:

*I hvilken grad kan leseflyt predikeres av ordgjenkjenningshastighet på isolerte ord sammenlignet med ord med ord-flanker?*

Det ønskes i tillegg å sjekke om ords frekvens kan ha innvirkning på resultatet, og underspørsmålet er dermed følgende:

*Er resultatet det samme når en differensierer mellom høyfrekvente og lavfrekvente ord?*

## 1.2 Oppgavens oppbygging

*Gjeldende kapittel* introduserer valg av temaet leseflyt, med et særlig fokus på automatisert ordgjenkjenning, samt belyser hvordan temaet har relevans for det spesialpedagogiske feltet. Avslutningsvis presenteres oppgavens forskningsspørsmål.

*Kapittel 2* redegjør for oppgavens teoretiske grunnlag. Her blir sentrale begreper, definisjoner og ulike teorier knyttet opp til tema presentert, og hele kapittelet leder på denne måten frem til utformingen av forskningsspørsmålet.

*Kapittel 3* beskriver valg knyttet til metode, utvalg og datainnsamling. Spesielt vil kapittelet omhandle den eksperimentelle oppgaven som er benyttet for å måle automatisert ordgjenkjenning. Siste del av kapittelet vil beskrive kriterier for validitet og reliabilitet på generelt grunnlag, samt potensielle trusler mot disse i kvantitativ forskning. Dette som utgangspunkt for diskusjonen i kapittel 5. Helt til slutt beskrives forskningsetiske hensyn.

*Kapittel 4* presenterer først resultatene fra deskriptiv analyse, hvorpå dette påfølges av en vurdering av variablenes normalfordeling. Visuelle fremstillinger av variablenes normalfordelingskurve er også inkludert. Videre vil resultater fra bivariat korrelasjonsanalyse og hierarkisk regresjonsanalyse bli presentert. Til slutt blir analysenes resultater kort oppsummert.

*Kapittel 5* diskuterer først resultatene i lys av oppgavens teorigrunnlag. Deretter vil resultatene vurderes i lys av studiens validitet og reliabilitet, og synliggjør her i hvilken grad funnene i gjeldende studie kan tenkes å være valide.

*Kapittel 6* er en kort konklusjon basert på foregående diskusjon.

## 2. Teoretisk bakgrunn

### 2.1 Lesing = avkoding x forståelse?

Det foreligger utallige teorier, definisjoner og modeller som ønsker å forklare leseprosessen, noe som understreker at dette er et komplekst fenomen som ikke nødvendigvis lar seg enkelt forklare. Likevel er det nødvendig med en presentasjon av én av de mest sentrale forklaringsmodellene innenfor feltet, både som nødvendig rammeverk for oppgavens tematikk og til videre diskusjon.

En av de mer kjente modellene som forsøker å forklare hva lesing er, er «The Simple View of Reading», presentert av Gough og Tunmer (1986). I deres modell blir lesing forklart som et produkt av to komponenter: avkoding og språkforståelse. Det argumenteres for at begge komponentene er nødvendige for å kunne lese, og at de sammen bidrar til dette. Med avkoding menes det sammenhengen mellom språklyder og bokstaver, også kjent som fonem-grafem-korrespondanse (Gough & Tunmer, 1986). Avkoding kan ifølge deres forklaring forstås som: “The ability to rapidly derive a representation from printed input that allows access to the appropriate entry in the mental lexicon, and thus, the retrieval of semantic information at the word level” (Hoover & Gough, 1990, s. 130). Avkoding regnes slik sett som å være avgjørende for å kunne utvikle adekvate leseferdigheter, men er ifølge Gough og Tunmer (1986) likevel ikke tilstrekkelig for å kunne anses å være en funksjonell leser. Dette illustreres ved å argumentere for at det vil være mulig å avkode ord på andre språk, uten at den som leser nødvendigvis kan si noe om hva ordene betyr. For å kunne forstå det som leses må ordene også være meningsbærende for leseren, dette i form av at leseren kan knytte dem til semantisk innhold. Det er denne kompetansen som regnes som språkforståelse i Gough og Tunmers (1986) forklaring.

Kort oppsummert argumenterer «The Simple View of Reading» for at avkoding og språkforståelse kan foregå uavhengig av hverandre, men at det først kan kalles funksjonell lesing når begge komponentene er til stede. Det kommer imidlertid ikke frem av forklaringsmodellen til Gough og Tunmer (1986) *hvordan* lesingen i begynnelsen går fra å

være en langsom og arbeidsom prosess, til å kunne gjennomføres både hurtig og uanstrengt, også kjent som *leseflyt*.

## 2.2 Leseflyt

Underveis i skoleløpet er det forventet at barn skal gå fra å kunne lære seg å lese rent teknisk, til å kunne være i stand til å lese for å lære (Gabrielsen og Lundetræ, 2013). Dette motivet finner en igjen i kompetansemål beskrevet i læreplanen (Utdanningsdirektoratet, 2020). Her blir det nevnt at elever i løpet av andre klasse skal kunne lese tekster med «sammenheng og forståelse på papir og digitalt». Flytbegrepet dukker opp i kompetansemål knyttet til fjerde klasse, hvor det påpekes at elever skal «lese tekster med flyt og forståelse og bruke lesestrategier målrettet for å lære» (Utdanningsdirektoratet, 2020). Beskrivelsene hentet fra kompetansemålene forutsetter at det har skjedd en progresjon i elevens leseferdigheter, hvorpå det å lese med flyt synes å være en eksplisitt indikator på dette.

Leseflyt kan forklares som å lese tekst hurtig og nøyaktig, med hensiktsmessig prosodi (NICHD, 2000). Denne definisjonen ble fremlagt i forbindelse med publiseringen av rapporten fra National Reading Panel, et resultat av et større arbeid med å kartlegge og vurdere ulike tilnærminger til leseopplæring. Leseflyt ble valgt ut som ett av tre hovedelementer for undersøkelse, på bakgrunn av økende bekymring knyttet til skolebarn som ikke utviklet leseflyt (NICHD, 2000). Rapporten fra National Reading Panel fastslo at leseflyt hadde nær sammenheng med leseforståelse, og at skolebarn som hadde lav grad av leseflyt også hadde vansker med å forstå tekst, og understreket at leseflyt fremstår som en avgjørende ingrediens i utviklingen av gode leseferdigheter (NICHD, 2000).

I rapporten fra National Reading Panel (NICHD, 2000) foreligger det en tydelig melding om at leseflyt ikke kan forklares av effektive og nøyaktige avkodingsferdigheter alene. Denne oppfatningen har også stått helt sentralt i forskning knyttet til leseflyt, hvor det lenge har vært et ønske om å finne svar på hvordan lesing kan gå fra å være langsom og arbeidsom, til hurtig, nøyaktig og uanstrengt, herunder forstått som *leseflyt* (Samuels, 2006). Dette har ført til flere forklaringsmodeller som ønsker å illustrere nettopp dette. For denne oppgavens

fokus vil det muligens være hensiktsmessig å begynne med Ehri (1995) sin forklaring på hvordan avkodning etter hvert utvikles til ordgjenkjenning.

## 2.3 Utviklingen av leseflyt - fra avkodning til ordgjenkjenning

Ehri (1995) sin modell forsøker å forklare hvordan begynnende lesere utvikler gode, funksjonelle leseferdigheter. Hun forklarer dette gjennom ulike utviklingsfaser, som illustrerer progresjonen som skjer hos leseren i denne prosessen. Ehri (1995) sin modell har totalt fire faser: før-alfabetisk fase, delvis alfabetisk fase, fullverdig alfabetisk fase og konsolidert alfabetisk fase. Kjernen i modellen til Ehri (1995) er at leseren danner koblinger mellom det talte og det skrevne språket, og at disse koblingene blir gradvis sterkere. Hver fase er slik sett tenkt å gi en beskrivelse over et typisk utviklingsforløp av leseferdigheter hos barn. Det gjøres likevel oppmerksom på at fasene ikke nødvendigvis opptrer fullstendig sekvensielt, og at både interne og eksterne faktorer ved barnet vil være med på å påvirke denne prosessen i et komplekst samspill (Ehri, 1995).

Fasene til Ehri (1995) beskriver et forløp hvor barn gjerne starter med å uttale ord på bakgrunn av at ordene er knyttet til bestemte visuelle markører, som eksempelvis at det kjenner igjen og kan uttale skilt og logoer. Etter hvert vil det opparbeide forbindelser mellom noen bokstaver og lyder i språket, og kan for eksempel lese «sol» fordi det har kjennskap til lyden knyttet til bokstavene «s» og «l». (Ehri, 1995). Videre vil fullstendige koblinger mellom skrevne bokstaver og lyder i språket opparbeides, og barnet kan bruke denne kunnskapen til å avkode nye ord. Etter hvert vil ord gjenkjennes umiddelbart, noe Ehri (1995) kaller for «sight word reading» (SWR). Ord som barnet møter ofte vil da lagres i barnets «sight word vocabulary» (SWV), og vil bidra til at barnet kan gjenkjenne ord raskt, uten å benytte seg av fonologisk avkodingsstrategi. Til slutt vil barnet opparbeide seg erfaring med mange ulike ord gjennom lesing, og stadig flere ord vil slik lagres i barnets SWV. Dette innebærer at barnet også vil opparbeide en forståelse for at ord er bygd opp av mindre ordenheter, gjennom erfaring med gjentakende stavelses- og lydmonstre. Disse mønstrene vil også internaliseres og lagres, som igjen vil forenkle lagring av ord i SWV (Ehri, 1995).

Mens Ehri (2005) hevder at *alle* ord kan etableres som «sight words» så lenge leseren har møtt disse ofte nok, viser andre til at ords frekvens ser ut til å spille en rolle når det kommer til hvor effektivt de gjenkjennes. Brysbaert, Mander og Keuleers (2018) viser til forskning som argumenterer for at høyfrekvente ord prosesseres raskere enn lavfrekvente ord. Dette hevder Samuels (2006) er fordi høyfrekvente ord i større grad gjenkjennes som holistiske enheter. Ordfrekvens har av denne grunn blitt tillagt å ha en sentral rolle i forskning som vedrører ordgjenkjenning (Brysbaert et al., 2018).

Ehri (1995) sin modell er spesielt nyttig fordi den differensierer tydelig mellom ordavkodning og ordgjenkjenning. Mens *ordavkodning* beskrives som strategier knyttet til den begynnende leser som ikke har opparbeidet tydelige forbindelser mellom fonemer og grafemer, beskrives *ordgjenkjenning* som når forbindelser er så etablerte at leseren er i stand til å gjenkjenne ord umiddelbart. Høien og Lundberg (2012) påpeker at ordgjenkjenning er en mer hensiktsmessig måte å lese ord på, fordi den gir leseren omgående tilgang til ords mening og uttale. Både Ehri (1995) og Høien og Lundberg (2012) trekker videre frem en viktig karakteristikk som skiller ordgjenkjenning fra ordavkodning. Dette er at ordlesingen later til å være *automatisert*.

## 2.4 LaBerge og Samuels automatiseringsteori – et viktig bidrag

LaBerge og Samuels (1974) var tidlig ute med å poengtere at *automatisering* synes å være helt essensielt for å kunne gjennomføre noe så komplekst som å lese tekst med flyt, forklart med at gode lesere er i stand til å gjennomføre denne aktiviteten både *hurtig, nøyaktig og uanstrengt*. De argumenterte for at lesing med flyt involverer så mange simultane prosesser at det vil være hensiktsmessig å plukke denne ferdigheten fra hverandre for å kunne forstå hvilke delprosesser og delferdigheter som er med på å utgjøre den. LaBerge og Samuels (1974) presenterte derfor deres «Theory of Automatic Information Processing in Reading» som et forsøk på å forklare hvordan komponenter involvert i lesing blir *automatiske*.

LaBerge og Samuels (1974) forklarer at lesing innebærer å prosessere informasjon i tre ulike stadier – først gjennom *visuelt minne*, deretter *fonologisk minne*, og til slutt *semantisk minne*. Informasjon som prosesseres på hvert stadium vil danne seg i det som LaBerge og Samuels

(1974) kaller for «koder». Dette forklares som indre representasjoner, som lagres i hukommelsen. At disse kodene læres og lagres, anses i deres forklaring som avgjørende for å kunne gjenkjenne ord hurtig og nøyaktig.

På det første stadiet prosesseres tekst gjennom visuell informasjon som innhentes. Dette kan være linjer, kurver og andre visuelle kjennetegn, for eksempel at leseren legger merke til hvilke visuelle differensieringer det er mellom grafemene «p» og «b». Disse visuelle kjennetegnene blir lagret som koder i leserens visuelle minne. Etter hvert som leseren får mer erfaring med ulike bokstaver, stavelsesmønstre og ord, antas det at flere koder dannes (LaBerge & Samuels, 1974). På det andre stadiet prosesseres tekst gjennom fonologisk minne. Dette stadiet opererer på bakgrunn av informasjon i foregående stadium. Ifølge LaBerge og Samuels (1974) vil leseren bruke visuell informasjon for å danne tilhørende auditive representasjoner. Det antas derfor at fonologiske koder først dannes på fonemnivå, deretter stavelsesnivå, før det til slutt vil kunne dannes fonologiske koder for hele ord. Det siste stadiet i modellen til LaBerge og Samuels (1974) er prosessering av tekst gjennom semantisk minne. Også dette stadiet er helt avhengig av de foregående, da ords betydning først vil kunne hentes ut på bakgrunn av korrekt kombinasjon av visuell- og fonologisk ordkode. Fordi ords betydning i stor grad først læres gjennom muntlig kommunikasjon, antas det at «kontakten» mellom det fonologiske minnet og det semantiske minnet allerede er sterk. Semantiske koder vil på samme måte som i de foregående stadier dannes i det semantiske minnet etter hvert som leseren får mer erfaring. Alle koder som leseren har etablert i visuelt-, fonologisk- og semantisk minne antas å kunne aktiveres og «hentes ut» i møte med visuelt stimuli, og slik sett bistå i det som kalles hurtig ordgjenkjenning.

Helt sentralt i teorien til LaBerge og Samuels (1974) er at lesing i stor grad må ses i lys av *oppmerksomhet*, samt en oppfatning av at menneskets evne til oppmerksomhet er begrenset. Dette innebærer en naturlig begrensning for mengde informasjon som kan prosesseres samtidig. Deres automatiseringsteori argumenterer for at de overnevnte prosessene involvert i informasjonsprosessering utvikles i to ulike stadier: nøyaktighetsstadiet, og automatiseringsstadiet. Ved det første stadiet er det behov for fokusert oppmerksomhet ved handlingen som utføres, men ved det andre stadiet skjer prosessen automatisk. Ifølge LaBerge og Samuels (1974) er gode lesere et resultat av at de har automatisert en del av de



innledende prosessene ved lesing av ord, som igjen åpner opp for at kognitive ressurser kan brukes for å forstå tekst. Begynnende lesere, samt lesere som ikke oppnår leseflyt antas å bruke for mye ressurser på å avkode ord, noe som innebærer at oppmerksomheten overbelastes i den grad at det ikke vil være tilstrekkelig ressurser igjen å bruke for å forstå tekst. Lesevansker knyttes slik sett opp mot manglende automatisering i tilsvarende prosesser, altså at evne til ordgjenkjenning ikke er automatisert.

Teorien til LaBerge og Samuels (1974) har blitt ansett for å være et viktig bidrag for å kunne forstå leseflyt, og det virker som at det er konsensus om at *automatisert ordgjenkjenning* er sentralt i utviklingen av denne ferdigheten (Kuhn, Schwanenflugel & Meisinger, 2010; Samuels, 2006; Rasinski, Reutzel, Chard & Linan-Thompson, 2011). Det kommer imidlertid ikke tydelig nok frem av teorien akkurat når og hva det er som avgjør om en ferdighet regnes som å være automatisert eller ikke.

## 2.5 Logans kriterier for automatiserte ferdigheter

Logan (1997) har også hatt lesing som utgangspunkt i sitt arbeid med å utarbeide kriterier for når en ferdighet kan regnes som automatisert. Logan sine kriterier baserer seg på at erfaringer stadig danner seg som separate koder i minnet. I nye erfaringer vil aktivering av koder skje gjennom assosiasjoner. Erfaringer/koder som er relevante for hverandre vil på denne måten hentes ut fra minnet, og jo mer kunnskap i form av koder som er tilgjengelig – jo mer er i stand til å aktiveres. Lesing forklares slik å kunne foregå gjennom minne-basert prosessering, altså en automatisert respons i form av minneaktivering. Hensikten bak kriteriene var å kunne bruke dem i arbeidet med å skille automatiserte fra ikke-automatiserte ferdigheter. Logan (1997) utarbeidet fire kriterier for automatisk prosessering. Det første kriteriet er *hastighet*. Automatisk prosessering skjer hurtig, og det synes som økt erfaring fører til økt hastighet. Det andre kriteriet er at automatisk prosessering later til å være *uanstrengt*. Dette forklares ved at vi eksempelvis kan gjøre to ting samtidig – som å holde en samtale gående mens vi sykler, uten at dette er verken anstrengende eller vanskelig å gjøre. Det tredje kriteriet er *autonomi*. Dette innebærer at prosesseringen skjer uten at dette gjøres med hensikt. Denne autonomien har blant annet blitt påvist gjennom forskning på det som

kalles Stroop-effekten, der deltakere har fått instruksjoner om å navngi fargen på ord de får presentert (f.eks. **grønn rød blå**). Resultater har vist at deltakere ikke er i stand til å la være å lese ordene, til tross for at dette ikke kreves av deltakerne. (Logan, 1997). Det siste kriteriet for automatisk prosessering er at ferdigheten kan utføres *ubevisst*. Dette illustreres ved at erfarne bilkjørere eksempelvis kan gire opp eller ned under kjøring, uten å være klar over det (Logan, 1997).

## 2.6 Leseflyt = automatisering?

Logan (1997) sine kriterier samsvarer godt med det National Reading Panel kaller for «true fluency», og det er at ekte leseflyt karakteriseres som å være både hurtig, uanstrengt og autonomt (NICHD, 2000). Leseflyt har på bakgrunn av en slik forståelse ofte blitt målt som en kombinasjon av korrekthet og hurtighet (antall riktige leste ord per minutt ved høytlesing), og dette har generelt også vært brukt som en indikator på automatisering (Hudson, Pullen, Lane & Torgesen, 2009). Dette problematiseres imidlertid av Rasinski og Young (2015), som påpeker at dette har ført til et u hensiktsmessig fokus på å øke lese hastighet i intervensjonssammenheng. De gjør oppmerksom på at det økt lese hastighet ikke fremmer automatisering, men at det heller er automatisering som fremmer økt lese hastighet (Rasinski & Young, 2015). Cummings (2016) ser ut til å støtte en slik forståelse, og argumenterer for å sette likhetstegn mellom automatisering og leseflyt. Dette begrunnes med at leseflyt virker å være et resultat av automatisering i en spesifikk ferdighet. Videre argumenteres det for at fokuset derfor bør ligge på *automatisering* i stedet for *leseflyt*, og foreslår at for å forstå automatisering er en også nødt til å studere det utenfor kontekst, i et forsøk på å isolere det en ønsker å undersøke. I lys av flyt må dette kunne anses å være automatisert ordgjenkjenning.

## 2.7 Automatisering – nødvendig med andre perspektiver?

Både LaBerge og Samuels (1974) og Logan (1997) ble tidlig ansett som viktige bidragsytere i arbeidet med å forstå hvordan leseferdigheter oppleves som automatiserte, og deres bidrag blir fortsatt betraktet som relevante, også i lys av nyere forskning på leseflyt (Hudson et al., 2009; Kuhn et al., 2010; Rasinski et.al, 2011). Likevel ser teoriene ut til å komme til kort i intervensjonssammenheng, da ingen av dem synes å tilby en virkningsfull behandling i situasjoner der individer ikke oppnår automatisert ordgjenkjenning. Andre fagfelt som kognitiv psykologi har imidlertid valgt å undersøke automatisering mer i retning av det Cummings (2016) foreslår, nemlig å undersøke hva automatisering innebærer som isolert konsept. Det kan først nevnes at en del av diskusjonen rundt automatisering har her handlet om hvorvidt automatisering skal forstås som et enten-eller-konsept, eller som et kontinuum, med automatisert i den ene enden, og ikke-automatisert i den andre (Moors & De Houwer, 2006). Besner, Stolz og Risko (2013) problematiserer det å forstå automatisering gjennom et sett med faste kriterier, da dette har en tendens til å implisere at noe er automatisert gjennom å samsvare med en gitt standard– eller ikke. Det argumenteres videre for at dette vil føre til at ingen prosesser kan kvalifiseres som automatiske, fordi det er ingen prosesser som møter alle kriterier, delvis fordi det ikke foreligger en enighet om hva disse kriteriene er. Forskning gjort av Bargh (2007) underbygger en slik forståelse, som argumenterer for at komplekse psykologiske fenomener er satt sammen av både automatiske og ikke-automatiske prosesser. Besner et al. (2013) foreslår derfor en kontinuum-forståelse av automatisering, hvor det kan tenkes at en ferdighet i praksis kan (ved gjentatt øving) bevege seg fra å være ikke-automatisert til automatisert. Det er viktig å understreke at en slik forståelse ikke nødvendigvis avfeier å konseptualisere automatisering inn i kjennetegn, men at disse ikke burde være for rigide. Gjennom en kontinuum-forståelse åpner en slik sett opp for at én ferdighet kan forstås som mer automatisert enn en annen, og at automatiske ferdigheter kan variere i beskrivelser (Besner et al., 2013). Moors & De Houwer (2006) har også denne innstillingen til automatisering, og vektlegger at en såkalt «deconstructional» tilnærming til automatisering kan være hensiktsmessig. Her argumenteres det for at automatisering fint kan forstås som et kontinuum, samtidig som en «bryter ned» det som undersøkes ved å identifisere hvilke karakteristikk som kan tillegges automatisering. Deres arbeid

konkluderer med at det er fullt mulig å separere automatisering inn i kjennetegn, og at en kan undersøke disse isolert, men også i kombinasjoner (Moors & De Houwer, 2006).

## 2.8 «Interference» som rettesnor?

Innen leseflyt har en del av uoverensstemmelsen knyttet til konseptet automatisering omhandlet «interference», altså hvorvidt automatiserte prosesser skal kunne bli forstyrret av, eller forbli upåvirket av andre prosesser (Moors & De Houwer, 2006). Argumentasjonen som ligger til grunn for at automatiske prosesser ikke kan forstyrres av, og kan pågå parallelt med andre prosesser er ideen om at automatiske prosesser bruker minimalt med ressurser, helt i tråd med teorien til Laberge og Samuels (1974). Logan (1997) står sådan på motsatt side i diskusjonen om «interference», da han ved sine studier mener å ha bidratt til evidens for at automatiserte ferdigheter kan forstyrres – noe som her argumenterer for at automatiserte prosesser ikke nødvendigvis bør knyttes til ressursbruk, men heller til hvordan disse opererer uavhengig av kontroll. Dette synes også å være perspektivet som har ligget til grunn for en del av den psykologiske forskningen som har blitt gjort på automatisering.

Innen kognitiv psykologi har det eksempelvis typisk vært tradisjon for å bruke manipulasjon for å teste hvorvidt prosesser kan kategoriseres som automatiske (Besner et al., 2013). Dette har ofte blitt gjort med utgangspunkt i hypoteser om at automatiske prosesser opererer uavhengig fra x. Særlig tre komponenter som antas å kunne forklare automatisering har vært under testing. Den første hypotesen er at automatisering ikke ser ut til å kreve *oppmerksomhet* (Besner et al., 2013). Dette har gjerne blitt testet gjennom en manipulasjon som tilbakeholder oppmerksomhet fra å påvirke det en undersøker. Dette innebærer en idé om at hvis noe er automatisk, så burde prosessen man undersøker bli upåvirket av denne manipulasjonen. Forskning har eksempelvis vist at testsubjekter er i stand til å prosessere informasjon til tross for at oppmerksomheten er viet til en annen sentral operasjon (Lien et al., 2006). Den andre hypotesen er at automatiserte prosesser ikke kan *kontrolleres* (Besner et al., 2013). Dette betyr at en automatisk prosess ikke nødvendigvis skjer med hensikt, men som en automatisk respons. Blant annet blir betegnelsen ballistisk brukt for å eksemplifisere ideen om at en automatisk prosess må anses som en fri kraft, som ikke kan stoppes når den

først er mobilisert (Besner et al., 2013). Stroop-effekten (Logan, 1997) må nevnes også her, da det er denne som ofte blir brukt som et eksempel på automatiseringens autonomi. Den tredje og siste hypotesen er at automatiske prosesser ikke krever *bevissthet* (Besner et al., 2013). Dette har spesielt blitt testet gjennom såkalte «masked priming paradigms». Dette involverer å presentere stimuli for testobjekter, som deretter blir maskert. I lys av hypotesen bør maskeringen ikke hindre prosesseringen av stimuli, da det tenkes at bevissthet ikke er nødvendig i informasjonsprosessering. «Priming», som betyr forhåndspåvirkning, innebærer i denne sammenhengen å undersøke om prosessering kan fremmes gjennom at en såkalt maskert «prime» presenteres forut for målstimuli (Besner et al., 2013). For eksempel har det blitt konkludert med at semantisk prosessering av ord (f.eks. lærer) kan fremmes gjennom at det først presenteres et relatert «prime»-ord (f.eks. elev) som maskeres, sammenlignet med et «prime»-ord som ikke kan relateres (f.eks. hammer), noe som har blitt brukt som evidens for at prosessering opererer uavhengig av bevissthet – altså at dette skjer automatisk.

## 2.9 Bruk av «masking» i selvstendig mål på automatisert ordgjenkjenning?

Foregående avsnitt avdekker at kognitiv psykologi som fagfelt har jobbet iherdig med å finne ut hvordan en best mulig skal kunne måle automatisering som et isolert konsept, hvorpå ulike former for manipulasjon har stått sentralt i denne tilnærmingen. Det kan se ut som om nyere publisert forskning på leseflyt har latt seg inspirere av denne tilnærmingens måten i arbeidet med å bringe klarhet i automatisert ordgjenkjenning (Roembke, Hazeltine, Reed and McMurray, 2018). Roembke med kolleger (2018) utviklet en ny måte å måle automatisert ordgjenkjenning på. Et ønske om å isolere automatisk ordgjenkjenning fra generell prosesseringshastighet og kjennskap til grafem-fonem-korrespondanse banet vei for å ta i bruk «backward masking», i deres studie, som på norsk kan oversettes til bakvendt maskering. Roembke med kolleger (2018) sin studie baserte seg på ulike oppgavetyper som alle hadde kortvarig presentasjon av et ord før ordet ble dekket for med en maske. Deltakere skulle deretter velge respons til ordet avhengig av oppgavetype, som differensierte mellom «klikk på korrekt bilde», «klikk på ordet som rimer» og «verifiser ordet». Ved å ta utgangspunkt i korrekthet som mål på automatisert ordgjenkjenning fant studien sterke

korrelasjoner mellom grad av automatisert ordgjenkjenning og leseflyt, noe som indikerer at det er mulig å gå fra subjektive indikatorer til et mer allment, uavhengig mål på når ordgjenkjenning kan sies å være automatisert.

## 2.10 Automatisering som «multi-item processing»

Studien til Roembke et al. (2018) må likevel forstås som et startskudd i denne sammenheng, da det er flere av de overnevnte aspekter ved automatisering som ikke ser ut til å være undersøkt. En mulig begrensning ved studien til Roembke et al. (2018) er at den har kun undersøkt automatisert ordgjenkjenning i lys av *individuelle* ord som ble presentert *isolert*. Altani et al. (2020) stiller spørsmål om hvorvidt en hurtig og korrekt gjenkjenning på *individuelle* ord i det hele tatt er en god måte å predikere leseflyt på. Dette knyttes delvis til at hurtighet og korrekthet tyder på å være to egenskaper ved ordgjenkjenning som ikke utvikles parallelt, og at hurtighet først kan oppnås etter at grad av korrekthet har nådd et visst nivå. Altani et al. (2020) refererer videre til forskning som indikerer at sammenhengen mellom individuell ordgjenkjenning og leseflyt synes å være mest til stede i de tidlige fasene av leseutviklingen, men at sammenhengen stadig reduseres etter hvert som barn utvikler mer avanserte leseferdigheter (Protopapas, Altani & Georgiou, 2013). Til gjengjeld har hurtig benevning av tallserier vist seg å gi en bedre prediksjon på leseflyt hos eldre barn, noe som impliserer at utviklingen av leseflyt ser ut til å være avhengig av noe *mer* enn å kunne gjenkjenne individuelle ord hurtig og korrekt (Altani, Georgiou & Protopapas, 2017; Protopapas et al., 2013). Hurtighet og nøyaktighet blir heller foreslått til å være to helt essensielle byggesteiner for å kunne gjennomføre det Altani et al. (2020) betegner som «multi-item processing», altså at flere elementer kan prosesseres simultant. Dette forklares gjennom at leseflyt gjerne måles som hurtig og korrekt lesing av *ordlister*, altså flere ord presentert *samtidig*, hvor ordene må behandles i sekvenser. Slik sett kan hurtig benevning av tallserier muligens kreve en tilsvarende form for prosessering. Altani et al. (2020) utfordrer slik den rådende oppfatning om at leseflyt kommer til uttrykk gjennom hurtig og korrekt ordgjenkjenning alene.

En slik forståelse legger dermed opp til at automatisering ikke nødvendigvis bare handler om hurtig og korrekt gjenkjenning av individuelle ord, men at automatisering også handler om hurtig og korrekt gjenkjenning av flere ord samtidig– gjennom en form for sekvensiell prosessering. Førstnevnte forstås likeledes som en forutsetning for sistnevnte. Med tanke på dem som ikke oppnår leseflyt, kan dette ifølge Altani et al. (2020) slik sett dreie seg om vansker med «multi-item processing». En mulig forklaring på dette kan kanskje finnes i «interference», som foreløpig må sies å være et ubesvart spørsmål knyttet til automatiske prosesser. Antagelsen om at automatiske prosesser opererer uavhengig av kontroll, og dermed kan forstyrres, blir særlig relevant i sammenheng med «multi-item processing», da det kan tenkes at forstyrrelser kan skapes av nærliggende elementer. Forskning har belyst at nære bokstaver kan forstyrre prosesseringen av en målbokstav (Bouma, 1970; Eriksen & Eriksen, 1974). Til tross for at denne forskningen ikke angår automatisk ordgjenkjenning direkte, kan det fortsatt tenkes å være interessant i denne konteksten. Sett fra et leseflyt-perspektiv kan det derfor tenkes at problemer med å lese flere ord som blir presentert samtidig, i realiteten er et bilde på forstyrrelser som skapes av nærliggende ord, altså en mangelfull sekvensiell prosessering. Dette er noe som ikke ser ut til å være undersøkt i relevant faglitteratur.

## 2.11 Å bruke «Visual World Paradigm» i avdekking av automatisert ordgjenkjenning

En tilnærming som er mye brukt i forskning som vedrører språkprosessering er «the visual world paradigm». Dette refererer til en eksperimentell metode som involverer oppgaver der bilder presenteres med lingvistisk input, med et premiss om at begge elementer kan bli linket til den virkelige verden (Conklin, Pellier-Sánchez, & Carrol, 2018). Metoden ble opprinnelig introdusert i 1974 av Cooper og Tanenhaus med kolleger, men har senere blitt brukt i utallige studier som undersøker prosessering av språk (Huettig, Rommers & Meyer, 2011). Metoden baserer seg ofte på å la deltakere se ulike bilder på en dataskjerm, samtidig som de får presentert auditivt stimuli i form av setninger. Vanligvis vil bildene illustrere objekter som nevnes i setningen som blir presentert, samt irrelevante bilder som skal fungere som «distraktorer». Øyebevegelser blir dermed kartlagt for å undersøke hvorvidt auditivt stimuli

fører til at deltakere fokuserer på relevante bilder (Huettig et al., 2011). Målet for undersøkelse er derfor det som skjer umiddelbart etter at stimuli er presentert, og hensikten er å forstå samspillet mellom språklig og visuell informasjonsbehandling. Det er imidlertid ikke bare bilder som har blitt brukt i «the visual world paradigm», og det finnes utgaver som involverer skrevne ord. Salverda og Tanenhaus (2010) argumenterer for at en slik variant kan være hensiktsmessig i arbeidet med å forstå ortografisk prosessering. Paradigmat anses derfor som svært fleksibelt, og kan slik sett være egnet å bruke til å undersøke overnevnte hypotese.

## 2.12 Teoretisk oppsummering og implikasjoner for oppgavens forskningsspørsmål

Gjeldende teorikapittel belyser at leseflyt anses som en helt sentral ingrediens i utviklingen av effektive, funksjonelle leseferdigheter. Videre tyder utviklingen av leseflyt å være avhengig av det som kan betegnes som *automatisert ordgjenkjenning*. Flere teorier har prøvd å forklare hvordan automatisert ordgjenkjenning bidrar til utviklingen av leseflyt.

Automatisert ordgjenkjenning har med utgangspunkt i disse ofte blitt kartlagt gjennom antall korrekt leste ord per minutt. Kritikere har imidlertid påpekt at dette har ført til at manglende automatisert ordgjenkjenning gjerne har blitt behandlet med intervensjoner som tar sikte på å øke lesehastighet, og at fokuset heller burde dreies mot *automatisering* – for å kunne avdekke hva automatisert ordgjenkjenning egentlig innebærer. Automatiserte prosesser forklares ofte gjennom ulike karakteristikk, men det er noe manglende konsensus både når det gjelder en tydelig definisjon og identifikasjon på når noe kan sies å være automatisert. I hvilken grad automatiserte prosesser kan forstyrres eller ikke er også noe det foreløpig ikke foreligger enighet om. Dette har ført til at fagfelt som eksempelvis kognitiv psykologi har forsøkt å undersøke automatisering på et mer isolert grunnlag, gjerne ved bruk av hypotesetesting og manipulasjon. Dette ser ut til å ha satt sitt preg på nyere forskning knyttet til automatisert ordgjenkjenning og leseflyt, som har gjort forsøk med å bruke bakvendt maskering som et selvstendig mål på automatisert ordgjenkjenning. Denne har konkludert med at dette tilsynelatende kan brukes til å predikere leseflyt. Det har imidlertid blitt trukket frem at sammenhengen mellom individuell ordgjenkjenning og leseflyt synes å reduseres



etter hvert som barn blir eldre, og at hurtig benevning av tallserier heller gir en bedre prediksjon på leseflyt. Det argumenteres dermed for at leseflyt ikke bare kan baseres på en hurtig og korrekt gjenkjenning av individuelle ord, men at det kreves en form hurtig og korrekt sekvensiell prosessering i tillegg. For de individer som ikke oppnår leseflyt synes det nødvendig å undersøke om dette tilsynelatende skapes av «interference», altså om det kan være slik at prosessen forstyrres av tilstøtende elementer, forstått som mangelfull sekvensiell prosessering. Med tanke på å oppnå leseflyt fremstår dette som særlig viktig, da tekstlesing som aktivitet *alltid* innebærer å lese ord som er omringet av andre ord.

I lys av gjeldende oppgave er det imidlertid manglende evidens for overnevnte hypotese, noe som genererer et imperativ for å bidra til å etablere mer sikker kunnskap i denne sammenheng. Oppgaven baserer seg dermed på et eksperiment utviklet ut ifra prinsipper fra «the visual world paradigm» og bruk av bakvendt maskering med hensikt om å sammenligne ordgjenkjenningshastighet på *individuelle ord* med ordgjenkjenningshastighet *på ord med såkalte «ord-flanker»*. Leseflyt vil her være sammenligningsgrunnlaget, gjennom å undersøke i hvilken grad disse to ulike presentasjonsformene ser ut til å differensiere i å predikere leseflyt. I lys av overnevnte teori antas det en hypotetisk relasjon mellom automatisering og «interference», hvilket impliserer at ordgjenkjenningshastighet på ord med ord-flanker sådan vil predikere leseflyt bedre enn ordgjenkjenningshastighet på individuelle ord. Funnene i denne oppgavens undersøkelse vil bli sett i lys av denne diskusjonen, samt drøftes opp mot teorien som har blitt presentert.

Oppgavens forskningsspørsmål er som følger:

*I hvilken grad kan leseflyt predikeres av ordgjenkjenningshastighet på isolerte ord sammenlignet med ord med ord-flanker?*

Det ønskes i tillegg å finne ut om ords frekvens kan ha innvirkning på resultatet, og underspørsmålet er dermed følgende:

*Er resultatet det samme når en differensierer mellom høyfrekvente og lavfrekvente ord?*

## 3. Metode

### 3.1 Utvalg

Utvalget besto av 62 voksne (40 kvinner) mellom 19-30 år. Gjennomsnittsalder på utvalget var 24,8 år med et standardavvik på 4,9. Hovedsakelig besto utvalget av studenter fra universitetet i Oslo, samt unge voksne i arbeid. Utvalgsmetoden var en kombinasjon av et såkalt bekvemmelighetsutvalg og selvseleksjon (Cresswell, 2014), hvilket innebærer at hovedvekten av deltakere var medstudenter, venner og bekjente som deltok frivillig.

Kriterier for utvalget var at alle deltakere hadde norsk som morsmål og normalt syn, noe som ble indikert gjennom at deltakere ikke brukte briller eller linser. Dette ble kartlagt gjennom selvrapportering. Før deltakelse signerte samtlige deltakere samtykkeskjema (se appendiks 1) utarbeidet i samsvar med Norsk senter for forskningsdata (NSD) sine retningslinjer.

### 3.2 Innsamling av data

#### *Kartleggingsverktøy*

For å kartlegge leseflyt ble den norske versjonen av Test of Word Reading Efficiency (TOWRE) brukt (se appendiks 2). Kartleggingstesten er todelt, og består av to lister med ord som skal leses høyt på 45 sekunder, så hurtig og korrekt som mulig. Administrering av hele testen tar derfor 2x45 sekunder. Den ene listen består av virkelige ord, mens den andre består av nonsensord. Begge lister øker i vanskelighetsgrad. Ord som leses riktig skåres med ett poeng, og ord som leses feil eller ikke leses, skåres med null poeng. Kombinert skal resultatene fra disse to deltestene gi en indikasjon på leseflyt i form av effektiv ordlesing, da den måler både lesehastighet og lesenøyaktighet (Furnes & Samuelsson, 2010).

## *Utstyr*

En øyesporingsenhet av typen SR Research Eyelink 1000 Plus ble brukt. Utstyret består av et kamera med høy oppløsning, samt en avlesningshastighet på opp til 2000 Hz og en nøyaktighet ned til  $0,15^\circ$  (Conklin et al., 2018). Dette innebærer at kameraet er i stand til å spore hvor deltakeren har blikket sitt på en skjerm, gjennom å konvertere bildet av deltakerens pupille til blikklokasjon. Kameraet var utstyrt med et 35mm objektiv og ble plassert i et feste beregnet for skrivebord. Deltakere fikk presentert stimuli på en 24" LCD-skjerm (1920 x 1080) med en avstand på 88-90 cm mellom skjerm og deltaker. Avstand fra øyesporingsenhet til deltaker ble målt opp til å være på 50-55 cm. Alle kalibreringer og øyesporinger ble gjort med deltakere sittende i en hodestabilisator, såkalt støttet hodemodus. Øyesporing var monokulær, men deltakerne fikk se med begge øyne mens de gjennomførte eksperimentet. I tillegg til å måle øyebevegelser med øyesporingsenheten ble det også brukt en datamus til å måle deltakernes responstid, dette gjennom at deltakere selv skulle klikke på valgt stimuli.

### 3.3 Eksperimentell oppgave

Hele eksperimentet var konstruert som 120 bildevalgs-oppgaver, der valg av bilde skulle fungere som en respons på et «target word» som ble presentert visuelt, men på tre ulike måter. Da det ikke virker som om det foreligger et tilsvarende etablert ord på norsk, blir derfor den direkte oversettelsen «målord» brukt videre. For hver oppgave fikk deltakeren først presentert fire ulike bilder i hvert skjermhjørne i 1500 ms, før det dukket opp en rød sirkel på skjermens sentrum. Den røde sirkelen forandret deretter farge til blå etter 520 ms, dette indikerte at deltakere kunne klikke med musepekeren (se appendiks 6, figur 1). Sirkelen hadde så til funksjon å sikre at deltakere hadde et fokusert blikk på skjermområdet hvor målord skulle presenteres, ved å holde igjen dette helt til deltakerens fokuserte blikk ble oppdaget. For å sikre fokusert blikk var deltakere nødt til å fiksere på den blå sirkelen i minimum 100 ms.

Den blå sirkelen ble da erstattet med stimuli i form av et målord, presentert i små bokstaver med skriftstørrelse 20 og skrifttype Consolas i normal stil. Målordet ble enten presentert isolert (se appendiks 6, figur 2) uten såkalte «flanker», med «visuelle flanker» i form av prosenttegn (se appendiks 6, figur 3), eller med flanker i form av andre ord (se appendiks 6, figur 4). Disse tre ulike presentasjonsformene vil videre i oppgaven bli referert til som tilstander. Stimuli ble etter 75 ms erstattet av en maske bestående av 10 emneknaggsymboler (se appendiks 6, figur 5). Masken ble etter 100 ms borte, og deretter skulle deltakere klikke på ett av fire bilder som avbildet målordet som ble presentert. En ny oppgave startet umiddelbart etter at deltakeren hadde klikket på et bilde.

Før deltakere gjennomførte oppgavene tilknyttet selve eksperimentet fikk de seks tilsvarende bildevalgoppgaver som fungerte som introduksjon og øving. Deretter gjennomførte deltakerne de 120 oppgavene, fordelt på tre blokker med 40 oppgaver i hver blokk. Innledningsvis ble binokulære kalibreringer gjort, med 13-punkts-kalibrering for nøyaktighet. Videre ble øyesporingsenheten satt til å spore det øyet som hadde mest nøyaktige data, ut ifra første kalibrering. For hver oppgaveblokk ble øyet kalibrert på nytt. All kalibrering ble gjort manuelt, og resultatene ble deretter bekreftet med validering, hvor en feiltoleranse ble satt til maks  $0.5^\circ$  på gjennomsnitt, samt  $1.0^\circ$  på maksimal feil. For eventuelle problemer knyttet til valideringsresultater ble det forsøkt justeringer, før ny kalibrering og validering ble gjort til ønsket resultat ble oppnådd.

### *Målord*

Alle de 120 ordene som skulle fungere som målord til hver oppgave ble valgt til å være substantiver med to stavelser, samt at de skulle ha et bokstavantall fra minimum tre og maksimum åtte bokstaver. Årsaken til at substantiver ble valgt som målord, var at ordene skulle kunne avbildes i eksperimentet, og være lett gjenkjennelige. I tillegg ble det bestemt at et likt antall høyfrekvente og lavfrekvente substantiv skulle inngå i ordutvalget, hvilket innebar en fordeling i to grupper; én med 60 høyfrekvente substantiver og én med 60 lavfrekvente substantiver. Resterende bestemmelser om bokstavlengde, antall stavelser og frekvensfordeling hadde til hensikt å gjenspeile naturlig variasjon i språket og slik sett sikre variert datagrunnlag. For å kunne gjennomføre en systematisk inndeling av ordene inn i ulike

frekvensgrupper ble Norwegian Orthographic Analyzer (<http://npa.staging2.scify.org/>) brukt. Dette er elektronisk programvare som kan brukes for å få en tallverdi på et ords frekvens, såkalt Zipf -frekvens. Et ords frekvens blir nevnt som en viktig variabel i forskning som omhandler ordgjenkjenning, dette fordi man raskere prosesserer de ordene man har møtt ofte, sammenlignet med ord som man sjeldent møter (Van Heuven, Mandera, Keuleers & Brysbaert, 2014). Zipf-skalaen ble utarbeidet med hensikt om å lage en mer brukervennlig skala i denne sammenheng, og er en skala som går fra verdiene 1 (lavfrekvente ord) til 7 (høyfrekvente ord). Dette innebærer at midtverdiene på skalaen er det som skiller det lavfrekvente fra de høyfrekvente ordene (Van Heuven, Mandera, Keuleers & Brysbaert, 2014). Basert på Zipf-skalen ble høyfrekvente ord derfor satt til å ha en Zipf-frekvens mellom 4-6, mens alle lavfrekvente ord skulle ha en Zipf-frekvens på 1-3,5 (se appendiks 3 og 4). Under er en deskriptiv tabell som viser gjennomsnittsverdi, samt minimum- og maksimumsverdi på Zipf-frekvens for hver gruppe. Frekvensgruppe (FreqGroup) 1 tilsvarer lavfrekvente ord, og frekvensgruppe 2 tilsvarer høyfrekvente ord. Tabellen gir også informasjon om antall bokstaver (Nlet). OLD20 står for «Orthographic Levenshtein Distance 20, og er et mål på ortografisk likhet (Yarkoni, Balota & Yap, 2008) Verdien representerer antall ortografiske naboer et ord har, altså antall nye ord som kan dannes med utgangspunkt i å bytte ut en bokstav med en annen (Adelman, 2012). Bigram referer til antall ord som inneholder lignende bokstavpar (Adelman, 2012).

Tabell 1: Deskriptiv statistikk for eksperimentell oppgave

	FreqGroup	Nlet	bigram	OLD20	fZipf
N	1	60	60	60	60
	2	60	60	60	60
Missing	1	0	0	0	0
	2	0	0	0	0
Mean	1	5.58	3.31	1.61	2.81
	2	5.42	3.40	1.45	4.36
Median	1	6.00	3.34	1.63	2.93
	2	5.00	3.43	1.45	4.25
Minimum	1	4	2.36	1.00	1.65
	2	3	2.89	1.00	3.98
Maximum	1	9	3.64	2.75	3.35
	2	8	3.83	2.75	5.62

### Flanker

I tillegg skulle hvert målord kunne bli presentert med én flanke på hver side, noe som innebar et behov for 240 flanke-ord. Flanke-ord skulle også være substantiv med et bokstavantall på minimum tre og maksimum åtte bokstaver. Kriterier for disse ordene var imidlertid at de ikke kunne dele samme stavelse, forbokstav eller siste bokstav som målordet de sto sammen med. For å innfri disse kriteriene ble derfor ordene som skulle fungere som flanker tildelt til hvert enkelt målord.

### *Fordeling av ord mellom tilstander*

I tildelingen av målord til tilstander ble utjevning en nødvendighet. Dette innebar å lage et system som gjorde at et målord som ble vist uten flanker for én deltaker, ville vises sammen med visuelle- eller ord-flanker for andre deltakere. Hensikten med dette var å sikre at hvert enkelt målord ville bli representert i alle tre tilstandene. For å kunne oppnå utjevning måtte de høy- og lavfrekvente ordene først stokkes hver for seg. Det ble deretter laget tre identiske kopier av denne fordelingen for hver av gruppene, hvorpå ulike målord ble tildelt tilstander. I den første kopien var de 20 første målordene tildelt isolert tilstand. I den andre kopien var de 20 første målordene tildelt symboler som «visuelle flanker», mens i den siste kopien var de 20 første målordene tildelt ord-flanker. De tre ulike kopiene kunne deretter stokkes innad, slik at målord ville bli presentert ulikt for ulike deltakere, men med en likevektig fordeling av tilstander under hver gjennomføring.

### *Bilder*

Totalt 480 ulike bilder ble innhentet, da det var behov for fire ulike bilder knyttet til hver av de totalt 120 oppgavene i eksperimentet. For hvert målord var det behov for et tilhørende bilde som korresponderte med ordet. I tillegg skulle hver oppgave ha minimum ett bilde som skulle fungere som en «competitor». Dette bildet skulle avbilde et ord som besto av samme forbokstav eller stavelse som målordet (f.eks. hvis målordet var flygel, kunne «competitor»-bildet illustrere ordet flue). Dette bildet hadde til hensikt å demonstrere om deltakere gjengjente målordet fremfor kun å gjenkjenne de første bokstavene i det. De to gjenstående bildene på hver oppgave skulle ha ingen likheter med verken målord, andre bilder eller ord-flanker, betegnet som «distractor» (f.eks. hvis målord var flygel kunne de to irrelevante bildene avbilde ordene dusj og diamant). Noen av oppgavene hadde to «competitor»-bilder og ett irrelevant «distractor»-bilde. Alle bilder var ulike illustrasjoner med farge, i ensartet tegneseriestil. De fleste bildene kom fra en bildedatabase opprettet av Bob McMurray på Universitetet i Ohio, som har omfattende erfaring med lignende oppgaver. Gjenstående bilder ble innhentet ved å bruke avansert bildesøk på Google, og på denne måten innhente bilder med fri bruksrettighet. Dette ble gjort av undertegnede og medstudent Dzan Zelihic.

En fullstendig liste over alle elementer som inngikk i den eksperimentelle oppgaven er vedlagt i appendiks 5.

### 3.4 Gjennomføring

Innledningsvis ble alle deltakere informert om at formålet med eksperimentet var å undersøke en spesifikk forestilling knyttet til automatisert ordgjenkjenning, og signerte deretter samtykkeskjema. Alle deltakere gjennomførte først de to deltestene av TOWRE. Her ble deltakere instruert til å lese de to listene høyt, så hurtig og korrekt som mulig. Alle leste listen med virkelige ord først, før listen med non-ord. Videre ble deltakere instruert til å gjennomføre eksperimentet som involverte øyesporing. Før gjennomføring ble nødvendige justeringer gjort på bordhøyde og hodestabilisator, slik at alle deltakere satt komfortabelt. Alle fikk deretter informasjon om at eksperimentet ville starte med oppgaver som skulle fungere som øving, dette for å gjøre seg kjent med eksperimentets oppbygging. Etter øving gjennomførte deltakerne alle oppgaver med tilhørende pauser. Både TOWRE-testene og den eksperimentelle oppgaven ble utført i én og samme økt, med en total varighet på ca. 30 minutter per deltaker.

### 3.5 Validitet og reliabilitet

Til tross for at kvantitativ forskning gjerne har hatt for vane å bli fremstilt med stor autoritet, baseres den likevel på beslutninger gjort av mennesker – hvilket impliserer at feil mest sannsynlig kan forekomme (Tuftes, 2018). All forskning bør derfor vurderes i lys av validitet og reliabilitet, to helt sentrale komponenter for å kunne vurdere dens grad av troverdighet og pålitelighet (Cumming & Colin-Jageman, 2017). Reliabilitet refererer til data, og i hvilken grad disse er frie for såkalte «tilfeldige målefeil». Tilfeldige målefeil er feil ved målingen som ikke har en systematisk forklaring, men som likevel kan gjøre data mindre troverdig (Tuftes, 2018). Med andre ord handler en vurdering av reliabilitet å diskutere hvor sikkert en kan oppnå tilsvarende resultater ved nye målinger. Validitet knyttes til hvorvidt en kan trekke gyldige slutninger fra forskning (Cresswell, 2014). Lund (2002) foreslår Cook og Campbell sitt validitetssystem i arbeidet med å vurdere validitet, og deres system inkluderer



fire typer validitet; indre, ytre, statistisk og begrepsvaliditet. Disse blir nærmere beskrevet under.

### 3.5.1 Indre validitet

Indre validitet tar utgangspunktet i sammenhengen mellom uavhengig og avhengig variabel, og i hvilken grad studiens funn kan forklares gjennom antatt relasjon mellom variablene (Lund, 2002). Truslene mot god indre validitet knyttes derfor opp mot studiens design (Cresswell, 2014). Ekte eksperiment blir gjerne løftet frem som et ønskelig design i lys av god indre validitet, og randomiserte kontrollerte tester (RCT) blir gjerne sett på som «gullstandarden» i denne sammenheng (Tuft, 2018). Dette innebærer manipulasjon, sterk grad av kontroll og minst to forsøksbetingelser (Lund, 2002). Mens manipulasjon dreier seg om at forskeren intervensjoner i forsøkssituasjonen, handler kontroll om flere ting, som kontrollgruppe, tilfeldig individfordeling over forsøksbetingelser og standardiserte prosedyrer og målinger. Flere forsøksbetingelser handler om å kunne kontrollere uavhengig variabel i større grad, ved å kontrollere for eventuelle irrelevante komponenter. Jo mer designet avviker fra «gullstandarden», jo dårligere er den indre validiteten (Tuft, 2018). Gjeldende studie bruker et eksperimentelt design, men avviker likevel noe fra å være en såkalt «RCT». Hvilke trusler mot indre validitet som kan være gjeldende for denne studien blir derfor vurdert og diskutert i diskusjonskapittelet.

### 3.5.2 Ytre validitet

God ytre validitet oppstår når det er mulig å foreta generaliseringer til eller fra relevante individer, situasjoner og tider (Lund, 2002). Trusler mot ytre validitet er dermed forhold som vanskeliggjør muligheten til dette, som et ikke-representativt utvalg og individhomogenitet (Lund, 2002). Ytre validitet må derfor ses i lys av studiens utvalg, og hvorvidt dette kan være representativt utover seg selv (Cresswell, 2014). God ytre validitet fremskaffes gjerne ved å gjennomføre tilleggsforskning med et annet utvalg, i en annen setting, samt med et opphold i tid mellom studiene. En vurdering av den ytre validiteten må derimot alltid ses i lys av formålet med forskningen og teoretisk bakgrunn, da god ytre validitet anses viktigere i

anvendt forskning enn i grunnforskning (Lund, 2002). I hvilken grad denne studien synes å ha god ytre validitet blir nærmere vurdert i kapittel 5.

### 3.5.3 Statistisk validitet

Gjeldende studie bygger på statistiske analysemetoder, og spørsmålet om statistisk validitet er dermed av høy relevans. God statistisk validitet er knyttet til sikkerhet i våre statistiske slutninger, altså om tendensen kan regnes som statistisk signifikant, eller «rimelig sterk» (Lund, 2002). Dette handler ikke om i hvilken grad det er en kausal sammenheng mellom variabler, men at vanlige krav til slutningsstatistikk er innfridd. Lav statistisk validitet oppstår gjennom brudd på disse kravene, og ved lav statistisk styrke (Lund, 2002). Ved god statistisk validitet vil en redusere sannsynligheten for at såkalte typefeil oppstår. En type I-feil oppstår når en nullhypotese forkastes selv om den er sann. En type II-feil oppstår når en aksepterer en nullhypotese som er feilaktig.

Statistisk validitet kan styrkes gjennom flere faktorer. At de statiske analysene er utført på en korrekt og nøyaktige måte er sentralt, og stiller høye krav til forskeren som gjennomfører dem. Videre er det nødvendig at de statistiske forutsetningene for valgte analyser også er innfridd. God statistisk validitet må også kunne sies å være helt avhengig av nøyaktige data, hvilket indikerer at et pålitelig datagrunnlag uten målefeil er ideelt. Å drøfte grad av reliabilitet inngår derfor som en viktig del av vurderingen av studiens statistiske validitet.

### 3.5.4 Begrepsvaliditet

God begrepsvaliditet er når fenomenet som undersøkes virkelig blir målt gjennom verktøyet som brukes (Field, 2009). Da foreliggende studie baserer seg på måling av fenomener som ikke er direkte målbare blir en vurdering av begrepsvaliditeten viktig i denne sammenheng. Å bruke allerede standardiserte verktøy kan styrke begrepsvaliditeten, samt å bruke ulike kartleggingsverktøy for å måle ett og samme fenomen (Lund, 2002). Dette krever imidlertid en tydelig definisjon på fenomenet. Begrepsoperasjonalisering står svært sentralt her, hvilket handler om å finne en god nok indikator på det som skal måles (Lund, 2002). I foreliggende studie er variablene knyttet til automatisert ordgjenkjenning særlig utfordrende med tanke på

begrepsvaliditet, da dette konseptet ikke gir inntrykk av å være entydig definert. Dette vil derfor diskuteres mer inngående i diskusjonskapittelet.

### 3.6 Forskningsetiske hensyn

Gjeldende studie er godkjent av Norsk senter for forskningsdata (NSD), som vurderer forskningsprosjekter i lys av personlige opplysninger, datainnsamling – og lagring, samt etiske hensyn. Søknaden ble godkjent 08.10.2019. Alle deltakere i studien var voksne individer over 18 år, og informert samtykke ble innhentet i tråd med NSD sine krav og retningslinjer. Ingen personlige data ble lagret i forbindelse med gjennomføringen, og publisert materiale er fullstendig anonymisert. Alle deltakere hadde muligheten til å trekke seg fra studien til enhver tid, uten negative følger. Dette er også i tråd med retningslinjene for forskningsetiske hensyn for samfunnsvitenskap, humaniora, juss og teknologi utarbeidet av Den nasjonale forskningsetiske komité for samfunnsvitenskap og humaniora (NESH), og må knyttes til *hensynet til mennesket* (De nasjonale forskningsetiske komiteene, 2016) Videre vektlegger de samme retningslinjene viktigheten av at *hensynet til forskersamfunnet* også ivaretas. Dette handler om å bedrive god vitenskapelig praksis gjennom tydelig henvisningsteknikk, samt etterstrebe vitenskapelig redelighet gjennom transparent og nøyaktig forskning som kan etterprøves. I rollen som forsker anser jeg det som viktig å forplikte meg til disse prinsippene, og disse er forsøkt innfridd etter beste evne.

## 4. Resultater

Dette kapitlet redegjør for resultatene av de ulike statistiske analysene som ble gjort for å kunne svare på oppgavens forskningsspørsmål. Innledningsvis vil deskriptive analyser bli presentert, hvor en vurdering av gjennomsnitt, standardavvik, skjevhetsverdi og kurtosisverdi for hver av variablene blir gjort. Visuelle fremstillinger av variablenes normalfordelingskurve er også inkludert her. Videre vil resultatene fra bivariat korrelasjonsanalyse bli presentert, før resultatene fra hierarkisk multipel regresjonsanalyse fremlegges. Avslutningsvis blir resultatene fra disse analysene kort oppsummert.

### 4.1 Deskriptive analyser

Deskriptiv statistikk har til hensikt å gi en oversikt over de viktigste aspektene ved gjeldende forskningsdata (Cumming & Colin-Jageman, 2017). Dette inkluderer en beskrivelse av sentrale tendenser og spredning i data, og eventuelle avvik fra normalfordelingen. Avvik i fordelingen kan forekomme på to ulike måter, enten ved *skjevhet* eller *kurtosis*. Skjevhet refererer til en fordeling der flere verdier ligger i én av halene sammenlignet med den andre, og fordelingen sies derfor å kunne være enten venstre- eller høyreskjev (Field, 2009). Skjevhet må derfor forstås som mangel på symmetri i fordelingen. En høyreskjev fordeling indikerer en lengre hale på høyre side. Dette kalles også *positiv skjevhet*. En *negativ skjevhet* er dermed skjev mot venstre, og indikerer en lengre hale på venstresiden (Field, 2009). Kurtosis refererer også til halene i fordelingen, gjennom hvor fyldige eller tynne disse er (Field, 2009). En fordeling kan ha positiv kurtosisverdi, noe som tilsvarer en spissere fordeling med korte fyldige haler, eller en negativ kurtosisverdi, som tilsvarer en flatere fordeling med lange tynne haler (Field, 2009). Akseptable verdier for kurtosis og skjevhet anses å være  $\pm 1$  (Field, 2009). Jo større avstand disse verdiene er fra 0, jo mer avviker fordelingen fra å være normal. Skjevhets- og kurtosisverdier for gjeldende variabler blir supplert med histogrammer, hvor fordelingen er visualisert. Q-Q-diagrammer kan også visualisere grad av normalitet, og er dermed også lagt ved. Q-Q-diagrammer ser observerte verdier i et datasett i lys av normale (forventede) verdier ved hjelp av en rett linje og punkter. Linjen representerer den forventede plasseringen, og punktene representerer de faktiske

verdiene i datasettet. I en normalfordeling vil punktene falle direkte på linjen, mens det ved avvik vil vises gjennom punkter utenfor linjen (Field, 2009). Jo mer punktene avviker fra linjen, jo mer avviker fordelingen fra å være normal.

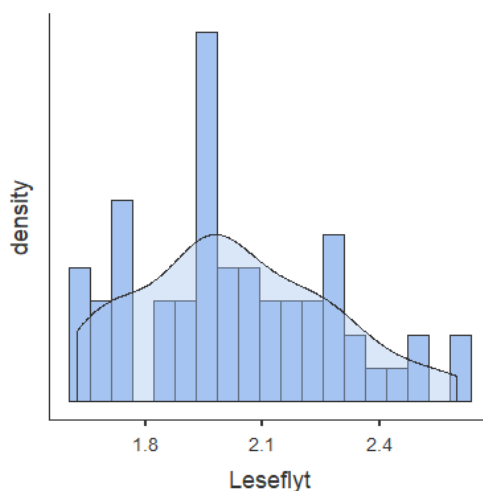
Tabell 2 viser en oppsummering av informasjon som inngår i den deskriptive analysen. *N* referer her til antall deltakere. *Mean* refererer til gjennomsnitt, mens *Range* henviser til rekkevidden mellom laveste og høyeste verdi ved gjeldende variabel. *Standard deviation* referer til standardavvik, som er et mål på hvor langt en verdi omtrentlig befinner seg fra gjennomsnittet for hver enkelt variabel (Cumming & Colin-Jageman, 2016). Slik sett vil standardavviket påvirkes i stor grad av ekstreme verdier, og et lavt standardavvik innebærer derfor at en stor andel av dataen ligger nærme gjennomsnittet. På grunn av utvalgets størrelse er Shapiro-Wilk-testen også inkludert for å gi informasjon om variablene avviker fra normalfordelingen. Denne testen sammenligner skårene i utvalget med et normalt fordelt utvalg med tilsvarende gjennomsnitt og standardavvik, og gir så et mål på om det er signifikante avvik (Field, 2009). Etter påfølgende tabell vil hver enkelt variabel beskrives mer inngående, også med visuelle representasjoner.

Tabell 2: Deskriptiv statistikk

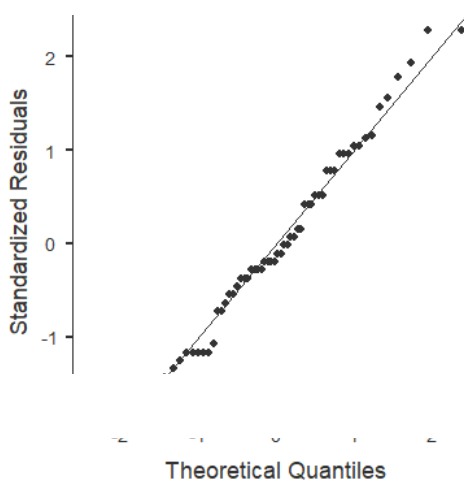
	Responstid ms				
	Høyfrekvent			Lavfrekvent	
	Leseflyt	Isolert	Flanker	Isolert	Flanker
N	60	60	60	60	60
Missing	0	0	0	0	0
Mean	2.03	1106	1375	1148	1370
Standard deviation	0.251	226	356	209	297
Range	0.978	970	1577	835	1310
Minimum	1.62	745	867	765	810
Maximum	2.60	1714	2444	1601	2121
Skewness	0.347	0.899	0.961	0.356	0.495
Std. error skewness	0.309	0.309	0.309	0.309	0.309
Kurtosis	0.457	0.489	0.765	-0.546	-0.494
Std. error kurtosis	0.608	0.608	0.608	0.608	0.608
Shapiro-Wilk p	0.116	0.003	0.002	0.175	0.060

### 4.1.1 Vurdering av variabelen leseflyt

Variabelen leseflyt kartlegges gjennom antall korrekt leste ord per sekund på deltest *Sight Word Efficiency* (TOWRE), i norsk versjon. Deltesten består av 104 virkelige ord som skal leses høyt på maks 45 sekunder. Gjennomsnittet på antall korrekt leste ord per sekund er 2.03, med et standardavvik på 0.251. Shapiro-Wilk-testen viser en p-verdi på 0.116 og indikerer dermed at variabelen avviker fra normalfordelingen med et .05 signifikansnivå. Fordelingen er videre noe høyreskjev, med en positiv skjevhetsverdi på 0.347, og en negativ kurtosisverdi på  $-0.457$ , hvilket indikerer en flatere fordeling. Både skjevhetsverdi og kurtosisverdi er akseptable, men samlet sett avviker variabelen noe fra normalfordelingen.



Figur 1: Histogram av variabelen leseflyt



Figur 2: Q-Q-diagram av variabelen leseflyt

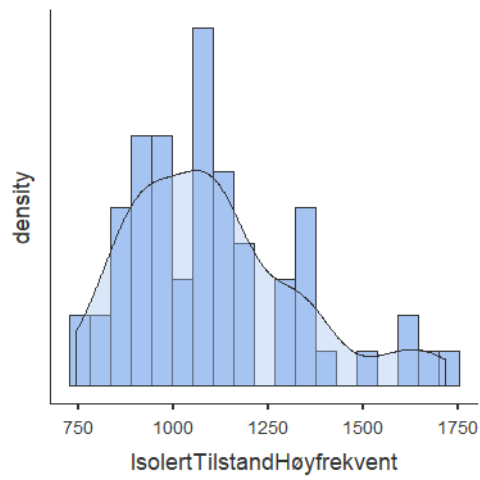
#### 4.1.2 Vurdering av variablene som representerer automatisk ordgjenkjenning

Responstid ble brukt som mål på ordgjenkjenningshastighet. Denne ble målt fra tidspunktet deltakere fikk presentert et målord og til de hadde klikket på det bildet de mente korresponderte til ordet de fikk presentert. For oppgavens fokus er det tatt utgangspunkt i responstid på målord presentert i isolert tilstand, samt målord presentert med ord-flanker. Gjeldende data tar utgangspunkt i deltakeres gjennomsnittlige respons per tilstand. Da oppgaven også har som formål å undersøke om responstid synes å påvirkes av ords frekvens er de to tilstandene videre inndelt i én høyfrekvent og én lavfrekvent variabel, noe som produserer totalt fire variabler knyttet til ordgjenkjenningshastighet. Tabell 1 viser dermed en total oversikt over reaksjonstid i ms for de ulike tilstandene, klassifisert etter frekvens. Under vil hver av variablene som representerer ordgjenkjenningshastighet bli vurdert nærmere.

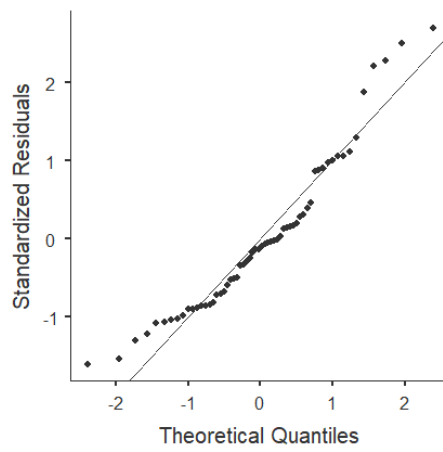
#### 4.1.3 Vurdering av variabelen ordgjenkjenningshastighet på isolert høyfrekvente ord

Gjennomsnittet for ordgjenkjenningshastighet på isolerte, høyfrekvente ord er 1106 ms, med et standardavvik på 226. Shapiro-Wilk-testen viser en p-verdi på 0.003, og indikerer at resultatene er normalt fordelt, med et signifikansnivå på .05. Fordelingen har en positiv skjevhetsverdi på 0.899, og er noe høyreskjev. Også kurtosis er positiv, med en verdi på 0.489, og indikerer en noe spissere fordeling. Verdiene på skjevhet og kurtosis er begge akseptable i henhold til Field (2009), og variabelen ser ut til å være relativt normalfordelt.





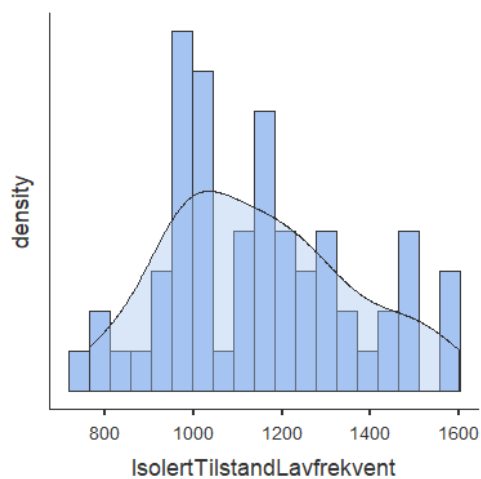
*Figur 3: Histogram av variabelen ordgjenkjenningshastighet på isolerte høyfrekvente ord*



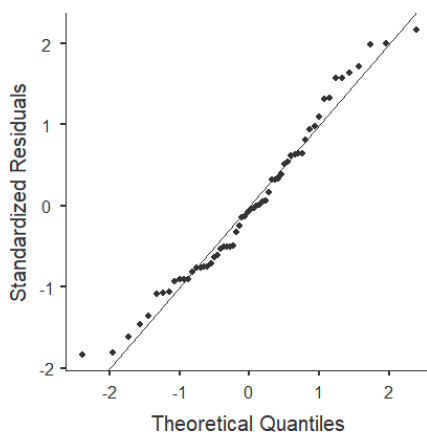
*Figur 4: Q-Q-diagram av variabelen ordgjenkjenningshastighet på isolerte høyfrekvente ord*

#### 4.1.4 Vurdering av variabelen ordgjenkjenningshastighet på isolerte lavfrekvente ord

Gjennomsnittlig ordgjenkjenningshastighet på isolerte lavfrekvente ord er 1148 ms med et standardavvik på 209. Shapiro-Wilk-testen indikerer at variabelens fordeling avviker noe fra normalfordelingen, med en p-verdi på 0.175. Variabelen er noe høyreskjev, med en skjevhetsverdi på 0.356. Variabelens kurtosisverdi er  $-0.546$ , og indikerer en noe flat fordeling. Skjevhets- og kurtosisverdier er begge akseptable.



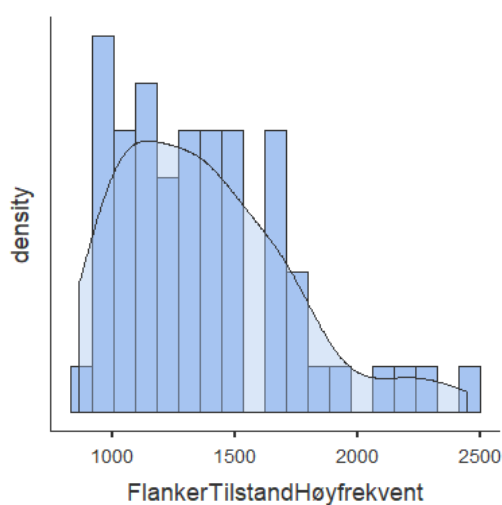
Figur 5: Histogram av variabelen ordgjenkjenningshastighet på isolerte lavfrekvente ord



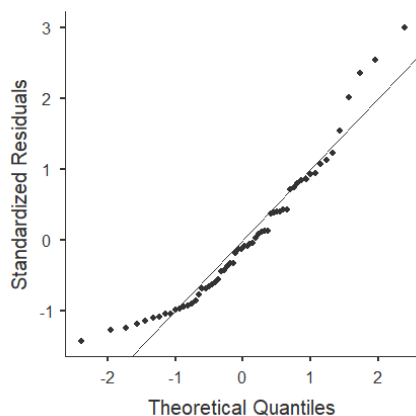
Figur 6: Q-Q-diagram av variabelen ordgjenkjenningshastighet på isolerte lavfrekvente ord

#### 4.1.5 Vurdering av variabelen ordgjenkjenningshastighet på høyfrekvente ord med ord-flanker

Gjennomsnittet på ordgjenkjenningshastighet på høyfrekvente ord med ord-flanker er 1375 ms med et standardavvik på 356. Fordelingen har en skjevhetsverdi på 0.961, og er derfor skjev mot høyre. Kurtosisverdien antyder at variabelen har en spissere fordeling, med en verdi på 0.765. Shapiro-Wilk-testen viser en p-verdi på 0.002, hvilket indikerer at variabelen ikke avviker signifikant fra normalfordelingen. Både skjevhets- og kurtosisverdi er akseptable.



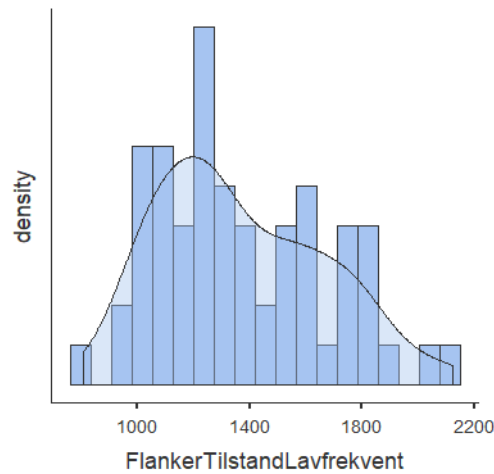
Figur 7: Histogram av variabelen ordgjenkjenningshastighet på høyfrekvent ord med ord-flanker



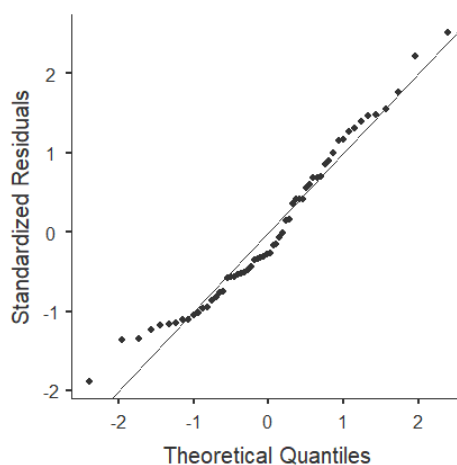
Figur 8: Q-Q-diagram av variabelen ordgjenkjenningshastighet på høyfrekvent ord med ord-flanker

## Vurdering av variabelen ordgjenkjenningshastighet på lavfrekvente ord med ord-flanker

Gjennomsnittlig ordgjenkjenningshastighet på lavfrekvente ord med ord-flanker er 1370 ms med et standardavvik på 297. Shapiro-Wilk-testen viser en p-verdi på 0.060, og signaliserer avvik fra normalfordelingen. Skjevhetsverdien er 0.459, og fordelingen er en anelse høyreskjev. Kurtosisverdien indikerer en flatere fordeling, da denne er på  $-0.494$ .



Figur 9: Histogram av variabelen ordgjenkjenningshastighet på lavfrekvente ord med ord-flanker



Figur 10: Q-Q-diagram av variabelen ordgjenkjenningshastighet på lavfrekvente ord med ord-flanker

## 4.2 Samlet vurdering av variablenes fordeling

Samlet sett vurderes alle variablene til å avvike fra normalfordelingen i varierende grad. Dette kan også ses på Q-Q-diagrammer for hver enkelt variabel. Ved å se på resultatene fra Shapiro-Wilk-testen i tabell 1 er det signifikante avvik på variablene leseflyt, isolerte lavfrekvente ord og lavfrekvente ord med ord-flanker med et signifikansnivå på .05. Samlet sett er derfor ingen av variablene normalt fordelt, og denne informasjonen må tas hensyn til i videre analyser.

## 4.3 Bivariat korrelasjonsanalyse

For å kunne undersøke sammenhengen mellom to variabler er en bivariat korrelasjonsanalyse et viktig statistisk redskap (Navarro & Foxcroft, 2019). En bivariat korrelasjonsanalyse kan vise i hvilken grad verdien på den ene variabelen synes å endre seg i lys av endringer i verdi hos den andre, altså hvordan to variabler samvarierer (Field, 2009). Grad av sammenheng beskrives vanligvis gjennom korrelasjonskoeffisienten Pearsons  $r$ , hvorpå  $r$  kan ha en verdi mellom  $-1$  og  $+1$  (Navarro & Foxcroft, 2019). Hvis  $r$  er  $\pm 1$ , regnes dette som perfekt korrelasjon. Hvis  $r$  betegnes som  $0$  anses dette som ingen korrelasjon. Korrelasjonskoeffisienten er i den forstand forstått som et effektmål, og jo nærmere verdien er  $\pm 1$ , jo sterkere sammenheng er det mellom de undersøkte variablene (Cumming & Calin-Jageman, 2017). Korrelasjonen mellom variablene kan være både positiv og negativ. En negativ korrelasjon vil ha en verdi mellom  $0$  og  $-1$ , og innebærer at verdiøkning i én variabel assosieres med en verdireduksjon i den andre. I en positiv korrelasjon vil derimot en verdiøkning i én variabel assosieres med verdiøkning i den andre, og vil ha en korrelasjonskoeffisient mellom  $0$  og  $+1$  (Field, 2009). I innledende analyse vil både korrelasjonskoeffisientene Pearsons  $r$  og Spearmans  $\rho$  bli inkludert. Spearmans  $\rho$  er et ikke-parametrisk mål som kan benyttes på data som avviker fra normalfordelingen, betegnet med den greske bokstaven  $\rho$ , uttalt rho (Field, 2009). Av den grunn at variablene hadde varierende avvik fra normalfordelingen virker det hensiktsmessig å inkludere Spearmans  $\rho$  for å minimere denne effekten. Det er nødvendig å understreke at til tross for at en bivariat korrelasjonsanalyse kan være et nyttig redskap for å innhente kunnskap om sammenheng

mellom variabler, så sier korrelasjonskoeffisienten derimot ingenting om kausalitet (Cumming & Calin-Jageman, 2017; Navarro & Foxcroft, 2019). Dette betyr at resultatene av analysen må tolkes med forsiktighet. Størrelsen på korrelasjonen kan imidlertid antyde hvor betydningsfulle resultatene virker å være, og Field (2009) foreslår å bruke Cohens (1992) inndeling som mulig rettesnor i denne sammenheng. Ifølge Cohen (1992) antyder  $r = .10$  en liten effekt,  $r = .30$  en moderat effekt, mens  $r = .50$  antyder en sterk korrelasjon.

#### 4.3.1 Resultater bivariat korrelasjonsanalyse

Påfølgende tabell viser resultatene fra bivariat korrelasjonsanalyse. Signifikante resultater er markert med \*, hvorpå flere stjerner uttrykker lavere signifikansnivå. Resultater som er signifikante på .05-nivå er markert med \*, .01-nivå med \*\*, mens et signifikansnivå på .001 er markert med \*\*\*. For hver korrelasjon mellom variablene blir Pearsons  $r$  med tilhørende  $p$ -verdi presentert øverst, mens Spearmans  $\rho$  med tilhørende  $p$ -verdi blir presentert nederst. Tendensene ser ut til å være at de to korrelasjonskoeffisientene er relativt like, men av hensyn til at variablene i ulik grad avviker fra normalfordelingen vil Spearmans  $\rho$  bli benyttet i videre analyser, beskrevet som  $\rho$ .

Tabell 3: Korrelasjonsmatriks

		Høyfrekvent			Lavfrekvent		
		Lese-flyt	Isolert	Flanker	Isolert	Flanker	
Høyfrekvent	Lese-flyt	Pearsons` s r	-				
		p-value	-				
		95% CI Upper	-				
		95% CI Lower	-				
		Spearman`s rho	-				
		p-value	-				
	Isolert	Pearsons`s r	-0.378 **	-			
		p-value	0.003	-			
		95% CI Upper	-0.138	-			
		95% CI Lower	-0.577	-			
		Spearman`s rho	-0.390 **	-			
		p-value	0.002	-			
	Flanker	Pearsons`s r	-0.230	0.751 ***	-		
		p-value	0.077	<.001	-		
		95% CI Upper	0.025	0.844	-		
		95% CI Lower	-0.458	0.614	-		
		Spearman`s rho	-0.234	0.745 ***	-		
		p-value	0.072	<.001	-		
Lavfrekvent	Isolert	Pearsons`s r	-0.327 *	0.913 ***	0.733 **	-	
		p-value	0.011	<.001	<.001	-	
		95% CI Upper	-0.080	0.948	0.832	-	
		95% CI Lower	-0.537	0.859	0.589	-	
		Spearman`s rho	-0.309 *	0.883 ***	0.741 **	-	
		p-value	0.016	<.001	<.001	-	
	Flanker	Pearsons`s r	-0.286 *	0.769 ***	0.782 ***	0.786 ***	-
		p-value	0.027	<.001	<.001	<.001	-
		95% CI Upper	-0.035	0.856	0.864	0.867	-
		95% CI Lower	-0.503	0.640	0.659	0.665	-
		Spearman`s rho	-0.318 *	0.744**	0.797 ***	0.760 ***	-
		p-value	0.013	<.001	<.001	<.001	-

Note: \* p <.05, \*\* p <.01, \*\*\* <.001

De bivariate korrelasjonsanalysene viser at leseflyt korrelerer høyest med ordgjenkjenningshastighet på isolerte høyfrekvente ord ( $\rho = -0.390$ ), med et signifikansnivå på .01. Dette betyr en delt variasjon på 15,2 % mellom de to variablene. Leseflyt korrelerte også med ordgjenkjenningshastighet på isolerte lavfrekvente ord med ( $\rho = -0.309$ ), samt ordgjenkjenningshastighet på lavfrekvente ord med ord-flanker ( $\rho = -0.318$ ), begge på et .05-nivå. Leseflyt har derfor 9,5 % delt varians med ordgjenkjenningshastighet på isolerte lavfrekvente ord, og 10,1 % delt varians med ordgjenkjenningshastighet på lavfrekvente ord med ord-flanker. Leseflyt hadde ikke en signifikant korrelasjon med ordgjenkjenningshastighet på høyfrekvente ord med ord-flanker ( $\rho = -0.234$ ). Korrelasjonene er noe lavere enn forventet, i lys av relevant teori. Alle korrelasjonskoeffisientene av interesse for analyse er negative, hvilket indikerer at en verdiøkning i den ene variabelen vil føre til en verdireduksjon i den andre (Field, 2009). I denne sammenheng betyr det at høyere grad av leseflyt assosieres med kortere responstid, altså at deltakere trykker raskt på bildet som korresponderer til målordet. Disse resultatene stemmer imidlertid godt overens med innledende antagelser. Til tross for at korrelasjonsanalyser er nyttige i den grad de kan identifisere denne form for sammenheng mellom variabler, vil de ikke være i stand til å si noe om hvilke variabler som har størst innvirkning på det konseptet vi ønsker å forstå, nemlig leseflyt. Konfidensintervallene for variablene viser at de overlapper i stor grad, og de er dermed for nærme til å kunne trekke konklusjoner om hvilken av variablene som ser ut til å assosieres mest med leseflyt. For å kunne få et innblikk i hvilken grad de ulike tilstandene kan forklare variasjon i leseflyt ble det videre gjennomført en regresjonsanalyse. En slik analyse er nyttig, fordi den gjør oss i stand til å kunne forstå våre data i større grad.



## 4.4 Regresjonsanalyse

En regresjonsanalyse er en statistisk metode som kan brukes for å undersøke hvorvidt en variabel er i stand til å forutsi en annen (Field, 2009). Regresjon baserer seg på at variabler har ulike roller, der den ene kategoriseres som en forklaringsvariabel (betegnet X), mens den andre kategoriseres som en responsvariabel (betegnet Y). Regresjon handler slik sett om hva X kan fortelle om Y, og dette forholdet modelleres gjennom en rett linje som uttrykker hvordan Y endrer seg når X skifter verdier, forstått som simpel lineær regresjon (Cumming & Colin-Jageman, 2017). Om resultatene fra regresjonsanalysen er statistisk signifikante kan regresjonslinjen brukes til å predikere verdien av Y for en gitt verdi av X, hvilket blir ansett for å være et svært viktig statistisk verktøy (Navarro & Foxcroft, 2019). Hvis analysen inkluderer to eller flere forklaringsvariabler betegnes dette som en multippel regresjonsanalyse (Field, 2009).

Hierarkisk multippel regresjonsanalyse refererer til å gjennomføre analysen i en selvvalgt prioritert rekkefølge (Field, 2009). En slik tilnærming anses som hensiktsmessig hvis en ønsker å finne ut hvilke variabler eller kombinasjoner av variabler som i størst grad synes å predikere avhengig variabel (Tabachnick & Fidell, 2014). Uavhengige variabler som antas å bidra til mest variasjon i avhengig variabel bør inkluderes først i analysen, før man inkluderer supplerende variabler (Field, 2009). En hierarkisk multippel regresjonsanalyse gjennomføres ved å legge inn uavhengige variabler i ulike sekvenser, og kan på denne måten brukes til å vurdere hver enkelt variabels forklaringsverdi på avhengig variabel. En vil da være i stand til å kunne si noe om hvorvidt en uavhengig variabel har direkte påvirkning på avhengig variabel.

Da oppgaven har som formål å undersøke hvorvidt leseflyt predikeres gjennom isolert ordgjenkjenning sammenlignet med ordgjenkjenning på ord med flanker-ord, samt om resultater gir inntrykk av å være tilsvarende ved ulik ordfrekvens, er det gjennomført en hierarkisk multippel regresjonsanalyse. Oppgavens spørsmål impliserer derfor fire forklaringsvariabler; to for ordgjenkjenning i hver av de to tilstandene (isolert og med ord-flanker), og to for ordfrekvens (høy og lav), og regresjonsanalysen er organisert i tre sekvenser. Første analysesekvens sammenligner ordgjenkjenning på høyfrekvente ord i

begge tilstander, før andre analysesekvens sammenligner tilsvarende på lavfrekvente ord. Til slutt blir alle fire forklaringsvariablene sett under ett i lys av avhengig variabel.

#### 4.4.1 Transformering av data

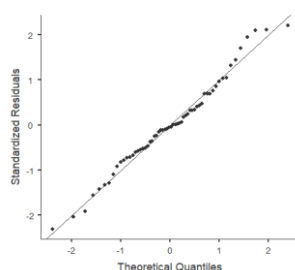
Før regresjonsanalyser ble gjennomført ble uavhengige variabler transformert gjennom funksjonen  $\log_{10}$ . Å transformere betyr i statistisk sammenheng å bruke en funksjon for å omskalere data, og kan være særlig hensiktsmessig for data som er skjev mot høyre (Diez, Barr & Çetinkaya-Rundel, 2019). Log10-transformasjonen komprimerer den øvre halen og strekker ut den nedre halen, slik at data synes mer symmetriske. Å transformere data slik at de oppleves som mer normalfordelte er nyttig, for da vil det bli lettere å kunne ta i bruk statistiske modeller og metoder (Diez, Barr & Çetinkaya-Rundel, 2019). En kvadratrotransformasjon ble først brukt, men dette var ikke nok til å hente frem en normal fordeling. Avhengig variabel er ikke transformert, da verken kvadratrotransformasjon eller  $\log_{10}$  utgjorde vesentlig forskjell på fordelingen. Tabell 4 viser deskriptiv statistikk over transformerte uavhengige variabler.

Tabell 4: Deskriptiv statistikk over transformerte variabler med LOG10

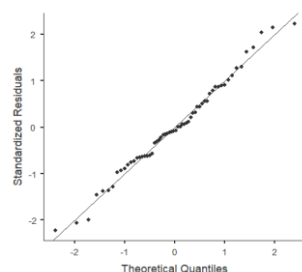
	LOG10				
	Høyfrekvent			Lavfrekvent	
	Leseflyt	Isolert	Flanker	Isolert	Flanker
N	60	60	60	60	60
Missing	0	0	0	0	0
Mean	2.03	3.04	3.13	3.05	3.13
Standard deviation	0.251	0.0847	0.107	0.0788	0.0932
Range	0.978	0.362	0.450	0.320	0.418
Minimum	1.62	2.87	2.94	2.88	2.91
Maximum	2.60	3.23	3.39	3.20	3.33
Skewness	0.347	0.440	0.399	-0.0137	0.0990
Std. error skewness	0.309	0.309	0.309	0.309	0.309
Kurtosis	0.457	-0.199	-0.401	-0.548	-0.694
Std. error kurtosis	0.608	0.608	0.608	0.608	0.608
Shapiro-Wilk p	0.116	0.176	0.200	0.524	0.414

#### 4.4.2 Statistiske forutsetninger for regresjonsanalyse

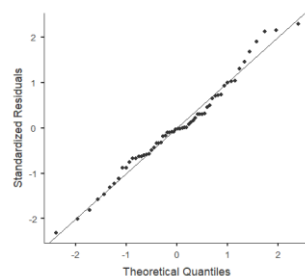
Å kunne trekke valide slutninger fra en regresjonsanalyse avhenger av at flere statistiske forutsetninger har blitt innfridd (Navarro & Foxcroft, 2019). Gjeldende avsnitt tar utgangspunkt i forutsetningene for regresjonsanalyse påpekt i Field (2009) og Navarro og Foxcroft (2019), med mindre annet er spesifisert. Forutsetningen om *intervallskalerte variabler* er innfridd. Det samme gjelder for *ikke-null-varians*, som er en forutsetning om at de uavhengige variablene har verdier som varierer. En forutsetning om normalfordelte residualer synes også å være innfridd, da q-q-diagrammer viser at punktene faller relativt fint langs linjen for alle tre sekvenser av regresjonsanalysen (figur 11, 12 og 13)



Figur 11: Q-Q-diagram, første sekvens



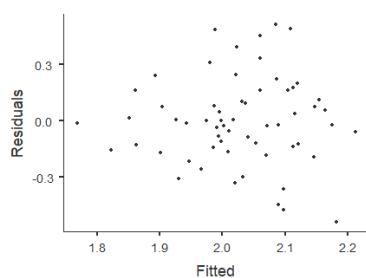
Figur 12: Q-Q-diagram, andre sekvens



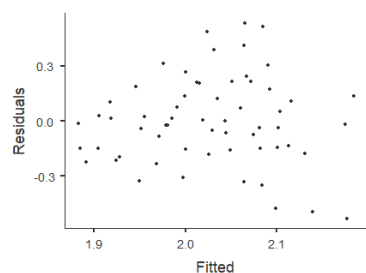
Figur 13: Q-Q-diagram, tredje sekvens

Videre kan forutsetningen om linearitet undersøkes gjennom punktdiagrammer (figur 14, 15 og 16) som viser fordelingen av residualer på forventede verdier. Her viser diagrammene at punktene ser ut til å være jevnt og tilfeldig fordelt for hver sekvens. Disse viser også at forutsetningen om homoskedastisitet er oppfylt, noe som illustreres gjennom nokså lik variasjon i hvordan punktene er fordelt utover, uten et spesielt synlig mønster.

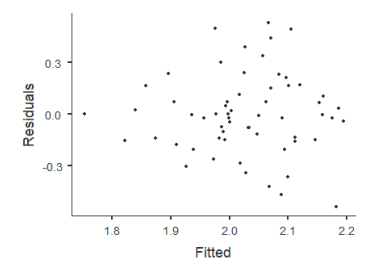
**Punktdiagrammer over residualer:**



*Figur 14: Punktdiagram over residualer, første sekvens*



*Figur 15: Punktdiagram over residualer, andre sekvens*



*Figur 16: Punktdiagram over residualer, tredje sekvens*

Forutsetningen om *ikke-perfekt multikollinearitet* mellom uavhengige variabler identifiseres gjennom variance inflation factors (VIF). VIF-verdien forteller om variablene i modellen er for høyt korrelert med hverandre. En for høy VIF-verdi vanskeliggjør å vurdere den individuelle påvirkningen en forklaringsvariabel har på avhengig variabel (Field, 2009). Det foreligger flere synspunkt omkring hva som er akseptabel VIF-verdi. Field (2009) argumenterer for at en VIF over 10, samt en toleranseverdi på mindre enn 0,1 anses som problematisk. Hair et al. (2010) foreslår imidlertid en øvre grense på 5, og mener denne kan være like problematisk, særlig ved utvalg av mindre størrelse. Verdiene på kollinearitetsstatistikken (appendiks 7) i Jamovi viser at VIF er under 10 for alle variablene, hver av dem med en toleranse over 0,1. Ved siste sekvens i regresjonsanalysen har imidlertid to av variablene en VIF-verdi på over 6, hvilket kan indikere en utfordring med multikollinearitet i lys av grensen foreslått av Hair et al. (2010). Hvordan VIF-verdier synes å påvirke gjeldende analyse vil bli nærmere diskutert i oppsummeringen av resultatene.

### 4.4.3 Resultater av regresjonsanalyser

#### Høyfrekvent ordgjenkjenningshastighet

Den første regresjonsanalysen ble gjennomført for å undersøke i hvilken grad høyfrekvent ordgjenkjenningshastighet på isolerte ord kan predikere leseflyt, sammenlignet med høyfrekvent ordgjenkjenningshastighet på ord med ord-flanker. Leseflyt ble dermed lagt inn som avhengig variabel, mens høyfrekvent isolert ordgjenkjenningshastighet ble lagt til som uavhengig variabel i første blokk. I andre blokk ble høyfrekvent ordgjenkjenningshastighet på ord med ord-flanker lagt inn. Koeffisienter og resultater fra gjennomføringen gjengis i tabell 5 og 6.

Tabell 5: Koeffisienter fra regresjonsanalyse, høyfrekvent ordgjenkjenningshastighet

	Predictor	Estimate	SE	t	p
		5.243	1.111	4.718	<.001
Høyfrekvent	Isolert	-1.400	0.564	-2.483	0.016
	Flanker	0.330	0.448	0.738	0.463

Tabell 6: Oppsummerte resultater regresjonsanalyse, høyfrekvent ordgjenkjenningshastighet

Overall Model Test						
Model	R	R <sup>2</sup>	F	df1	df2	p
1	0.364	0.133	8.88	1	58	0.004
2	0.375	0.141	4.68	2	57	0.013

Tabell 6 viser at ordgjenkjenningshastighet på høyfrekvente ord bidrar signifikant til forklaring av variasjonen i leseflyt hos utvalget, med  $p = 0.004$  og  $0.013$  for henholdsvis modell 1 og 2. Modell 1 viser at 13,3 % av variasjonen i leseflyt forklares av isolert ordgjenkjenningshastighet på høyfrekvente ord. Det betyr at de resterende 86,7 % av variansen skyldes andre faktorer. Når ordgjenkjenningshastighet på høyfrekvente ord med ord-flanker inkluderes (modell 2) er det kun 0.8 % ytterlig variasjon som kan forklares. Ordgjenkjenningshastighet i isolert tilstand ser dermed ut til å være den mest betydningsfulle variabelen i gjeldende analyse.

For å kunne undersøke om tilstandene differensierer i hvorvidt de kan predikere leseflyt er ytterligere analyse gjort ved å se på modellsammenligningen som Jamovi produserer (tabell 7). Sammenligningen kan gjøres ved å bruke modell 1 som null-hypotese, mens modell 2 fungerer som alternativ hypotese. I lys av oppgavens forskningsspørsmål er null-hypotesen at det ikke er signifikant forskjell mellom høyfrekvent ordgjenkjenningshastighet på isolerte ord, sammenlignet med høyfrekvent ordgjenkjenningshastighet på ord med ord-flanker, mens alternativ hypotese er at det er signifikant forskjell mellom tilstandene. Tabell 7 viser at  $p = 0.463$ . Med et signifikansnivå på .05 betyr dette en ikke kan forkaste null-hypotesen, som igjen innebærer at effekten ikke er sterk nok til å kunne si at det er signifikant forskjell i hvilken grad tilstandene predikerer leseflyt.

Tabell 7: Modellsammenligning, høyfrekvent ordgjenkjenningshastighet

<b>Comparison</b>							
<b>Model</b>		<b>Model</b>	<b><math>\Delta R^2</math></b>	<b>F</b>	<b>df1</b>	<b>df2</b>	<b>p</b>
1	-	2	0.00822	0.545	1	57	0.463



## Lavfrekvent ordgjenkjenningshastighet

Den andre regresjonsanalysen ble gjennomført for å undersøke i hvilken grad lavfrekvent ordgjenkjenningshastighet på isolerte ord kan predikere leseflyt, sammenlignet med lavfrekvent ordgjenkjenningshastighet på ord med ord-flanker. Analysen ble gjort på tilsvarende måte som den første, med leseflyt som avhengig variabel. Lavfrekvent isolert ordgjenkjenningshastighet ble lagt til som uavhengig variabel i første blokk, og i andre blokk ble lavfrekvent ordgjenkjenningshastighet på ord med ord-flanker lagt inn. Koeffisienter og resultater fra gjennomføringen gjengis i tabell 8 og 9.

Tabell 8: Koeffisienter fra regresjonsanalyse, lavfrekvent ordgjenkjenningshastighet

	<b>Predictor</b>	<b>Estimate</b>	<b>SE</b>	<b>t</b>	<b>p</b>
	Intercept	5.001	1.244	4.021	<.001
lavfrekvent	Isolert	-0.841	0.631	-1.333	0.188
	Flanker	-0.131	0.533	-0.245	0.808

Tabell 9: Oppsummerte resultater fra regresjonsanalyse, lavfrekvent ordgjenkjenningshastighet

<b>Overall Model Test</b>						
<b>Model</b>	<b>R</b>	<b>R<sup>2</sup></b>	<b>F</b>	<b>df1</b>	<b>df2</b>	<b>p</b>
1	0.301	0.0905	5.77	1	58	0.020
2	0.302	0.0914	2.87	2	57	0.065

Tabell 9, modell 1 viser at 9 % av variasjonen i leseflyt hos utvalget kan forklares av ordgjenkjenningshastighet på isolerte lavfrekvente ord, og bidraget regnes som signifikant med  $p = 0.020$ . Dette betyr at 91 % av variansen i leseflyt hos utvalget kan forklares gjennom andre faktorer. Når ordgjenkjenningshastighet på lavfrekvente ord med ord-flanker inkluderes (modell 2) er det kun 0.09 % ytterligere variasjon som kan forklares. Dette resultatet er ikke signifikant ( $p = 0.065$ ). Også i denne regresjonsanalysen later det til at ordgjenkjenningshastighet i isolert tilstand er den mest betydningsfulle variabelen.

Modellsammenligningen ble også inkludert i denne analysen, for å få innsikt i hvilken grad variablene differensierte i å predikere leseflyt (tabell 10). Denne ble brukt på tilsvarende måte som i foregående analyse, hvorpå null-hypotesen er at det ikke er signifikant forskjell mellom lavfrekvent ordgjenkjenningshastighet på isolerte ord, sammenlignet med lavfrekvent ordgjenkjenningshastighet på ord med ord-flanker. Alternativ hypotese er at det er signifikant forskjell mellom tilstandene. Tabell 10 viser at  $p = 0.808$ . Dette betyr at en heller ikke her kan forkaste null-hypotesen, som betyr at effekten ikke er sterk til å kunne si at det er signifikant forskjell i hvilken grad tilstandene predikerer leseflyt.

Tabell 10: Modellsammenligning, lavfrekvent ordgjenkjenningshastighet

Comparison							
Model		Model	$\Delta R^2$	F	df1	df2	p
1	-	2	9.55e-4	0.0599	1	57	0.808

## Endelig regresjonsanalyse – alle variabler

Den tredje og siste regresjonsanalysen ble gjennomført med hensikt om å se alle forklaringsvariablene i lys av avhengig variabel, leseflyt. Tabell 12 gjengir resultatene for modellen der alle fire forklaringsvariabler er inkludert. Koeffisienter fra endelig analyse vises i tabell 11.

Tabell 11: Koeffisienter fra regresjonsanalyse, alle forklaringsvariabler

	Predictor	Estimate	SE	t	p
	Intercept	5.0681	1.232	4.115	<.001
Høyfrekvent	Isolert	-1.7054	0.932	-1.829	0.073
	Flanker	0.3242	0.529	0.613	0.543
Lavfrekvent	Isolert	0.4425	0.982	0.451	0.654
	Flanker	-0.0732	0.624	-0.117	0.907

Tabell 12: Oppsummerte resultater fra regresjonsanalyse, alle forklaringsvariabler

Overall Model Test						
Model	R	R <sup>2</sup>	F	df1	df2	p
1	0.380	0.144	2.32	4	55	0.069

Tabell 12 viser at 14,4 % av variasjonen i leseflyt hos utvalget kan forklares ved å inkludere alle variablene i analysen. Dette betyr at 85,5 % av variasjonen kan forklares av andre faktorer. Resultatet er imidlertid ikke signifikant, med  $p = 0.069$ . Vedlagte regresjonskoeffisienter (tabell 11) viser både effektstørrelse og retning hver enkelt variabel har på leseflyt, mens de andre variablene holdes konstant. Ingen av forklaringsvariablene synes i denne sammenhengen å utgjøre en signifikant forskjell på avhengig variabel.

## 4.5 Oppsummering av resultater

Resultatene fra den bivarierte korrelasjonsanalysen viser at leseflyt korrelerer på .01 eller .05 med tre av fire forklaringsvariabler. Basert på gjennomførte analyser antyder dette en tydelig sammenheng mellom ordgjenkjenningshastighet og leseflyt. Forklaringsvariablene har en ulik felles varians med leseflyt. Felles varians er 15,2 % mellom leseflyt og ordgjenkjenningshastighet på isolerte høyfrekvente ord, 9,5 % mellom leseflyt og ordgjenkjenningshastighet på isolerte lavfrekvente ord, og 10,1 % mellom leseflyt og ordgjenkjenningshastighet på lavfrekvente ord med ord-flanker. Innledende regresjonsanalyser viser først at ordgjenkjenningshastighet på isolerte høyfrekvente ord forklarer ca. 13,3 % av variasjonen i leseflyt, deretter at ordgjenkjenningshastighet på isolerte lavfrekvente ord forklarer ca. 9 % av variasjonen i leseflyt. Endelig regresjonsmodell viser at variablene ikke bidrar signifikant til variasjon, når hver av dem inkluderes i modellen. Da bivariat korrelasjonsanalyse har vist signifikante korrelasjoner mellom avhengig variabel og tre av uavhengige variabler, kan det tenkes at multikollinearitet skaper problemer når alle variablene inkluderes i regresjonsmodellen. Som nevnt innledningsvis til regresjonsanalysen synes VIF-verdien å være for høy for to av forklaringsvariablene, noe som ser ut til å skape støy i den endelige analysemodellen. I lys av Hair et al. (2010) er det vanligvis utvalg av mindre størrelse som korresponderer med høye VIF-verdier. En kan anta at både liten størrelse på utvalget, men også antall forklaringsvariabler påvirker resultatene som fremstilles ved endelig regresjonsanalyse. Ifølge Field (2009) burde utvalget være større jo flere uavhengige variabler som inkluderes i analysen, og refererer til Greens (1991) tommelfingerregel om at utvalgets størrelse burde være minimum  $50 + 8k$ , hvor  $k$  er antall forklaringsvariabler. Da gjeldende regresjonsanalyse baserer seg på fire variabler, skulle utvalget størrelse i lys av overnevnte regel ideelt sett vært 82 eller større. Det må også tillegges at jo flere variabler, jo mer kompleksitet tilføres regresjonsmodellen. På bakgrunn av dette argumenterer Navarro og Foxcroft (2019) for å være forsiktig med å inkludere for mange variabler hvis en har et ønske om en generaliserbar modell. Slik sett kan det antas at alle variablene inkludert i samme modell går på bekostning av dens egnethet for prediksjon.

## 5. Diskusjon

I vitenskapelig forskning blir kvalitet gjerne løftet frem gjennom ulike krav til validitet og reliabilitet. Dette innebærer en diskusjon om hvorvidt det en kan trekke gyldige slutninger fra resultatene som foreligger. I praksis innebærer dette å erkjenne forskningens begrensning ved å synliggjøre og diskutere mulige feilkilder (Lund, 2002). I første del av diskusjonskapittelet vil det foretas en drøfting av resultatene i lys av presentert teori, før det i andre del av kapittelet gjennomføres en drøfting av i hvilken grad gjeldende forskning anses å være valid, hvorpå Cook og Campbells validitetssystem blir brukt som rammeverk. Reliabilitet vil også diskuteres her.

### 5.1 Drøfting av resultatene i lys av teori

Formålet med gjeldende studie var å undersøke ideen om at ordgjenkjenning kan regnes som automatisk i den grad at det ikke synes å forstyrres av tilstøtende ord. Spesielt om dette ser ut til å komme til uttrykk gjennom forholdet mellom ordgjenkjenningshastighet og leseflytskårer. Det hersker liten tvil om at automatisert ordgjenkjenning anses å være en helt essensiell komponent i utviklingen av leseflyt, gjerne forstått som en *umiddelbar* identifisering av ord (Rasinski et al., 2011; Hudson et al., 2009; Kuhn et al., 2010). Den dominerende forestillingen er dermed at det er *hurtig* og *korrekt* gjenkjenning av ord som fremmer flytoppnåelse. I en nyere studie ble automatisert ordgjenkjenning kartlagt gjennom bakvendt maskering av individuelle ord som ble presentert isolert (Roembke et al., 2018), og denne kunne vise til en tydelig sammenheng mellom automatisert ordgjenkjenning og grad av leseflyt. Gjeldende studie kan fastslå at sammenhengen mellom ordgjenkjenningshastighet og leseflyt synes å være til stede, hvilket bekreftet innledende antagelser. De påviste negative sammenhengene stemmer godt overens med presentert teori, da dette impliserer at økning i den ene variabelen vil føre til en reduksjon i den andre. Med tanke på flyt forstås det derfor som at jo høyere grad av flyt, jo kortere tid brukes (på å gjenkjenne ord), eller motsatt – jo lengre tid en bruker (på å gjenkjenne ord), jo lavere grad av flyt. Slik sett er disse funnene ikke overaskende, da dette allerede anses å være etablert kunnskap.

Derimot ønsket studien å bidra til ny kunnskap med bakgrunn i teorien om hvorvidt automatiserte prosesser kan forstyrres. Dette må ses i lys av forskningsfunn som fremlegger at individers evne til å gjenkjenne ord hurtig og korrekt ikke later til å gi en god nok prediksjon på leseflyt etter hvert som elever blir eldre, men heller knyttes opp mot evne til å prosessere flere enheter sekvensielt (Altani et al., 2020). Tankekorset som har blitt presentert i denne sammenheng er at leseflyt ikke kartlegges gjennom ord som presenteres isolert og individuelt, men heller gjennom lesing av ord som presenteres i liste- eller tekstform – noe som i større grad samsvarer med ordinær lesing. En leser som oppnår flyt tenkes derfor å være i stand til å ha automatiserte ordgjenkjenningsferdigheter i den grad at ord leses sekvensielt – og tilstøtende ord vil dermed *ikke* forstyrre denne prosessen. Dette genererte et ønske om å sammenligne hvorvidt ordgjenkjenningshastigheten på ord som presenteres isolert predikerer leseflyt, sammenlignet med ordgjenkjenningshastigheten på ord som presenteres med tilstøtende ord. Ut ifra presentert teori ble det antatt at de to ulike tilstandene (ordgjenkjenningshastighet på isolerte ord og ord med tilstøtende ord) i studien ville være signifikant forskjellig med tanke på å predikere flyt, og at ordgjenkjenningshastighet på ord med ord-flanker ville gi et bedre prediksjonsgrunnlag. Funntilstandene i gjeldende studie synes derimot ikke å kunne vise til slike resultater. Flere momenter kan drøftes i denne sammenheng.

For det første ser deskriptiv statistikk ut til å indikere at deltakere i gjennomsnitt bruker lengre tid på å gjenkjenne ord med ord-flanker, enn individuelle ord - uavhengig av frekvens. Dette stemmer overens med teori som argumenterer for at ord som er presentert med tilstøtende ord antas være mer krevende å gjenkjenne enn individuelle ord som presenteres isolert, og kan antyde at en faktisk sekvensiell prosessering finner sted (Altani et al., 2020). Dette er imidlertid ikke testet statistisk, og en kan heller ikke derfor fastslå at denne forskjellen i tidsbruk er signifikant.

For det andre må funnene ses opp mot studiens utvalg. Noe av det som skiller denne studien med mye av forskningen som er nevnt tidligere, er at gjeldende forskning har undersøkt en hypotese på et utvalg hvor majoriteten besto av voksne universitetsstudenter, fremfor yngre skolebarn i utvikling. Det kan derfor antas at overvekten av utvalget består av mennesker med godt utviklede leseferdigheter. I lys av Altani et al. (2020) foreslås *hurtighet* ved automatisk prosessering å være en byggestein som først kan oppnås etter et visst

korrekthetsnivå er på plass. Studiens deltakere kan tenkes å være godt i stand til å både gjenkjenne ord isolert – og sekvensielt, fordi deres korrektthetsnivå allerede er etablert. Gjeldende utvalg er slik sett kanskje ikke representativt nok for å undersøke hypotesen om forstyrrelser. Funnene kan dog implisere at såkalt sekvensiell prosessering for *gode* lesere ikke synes å kreve betydelig økt ressursbruk, hvilket indikerer at hypotesen bør undersøkes på andre måter. Det kan derfor tenkes at det vil være mer fruktbart å undersøke gjeldende hypotese i lys av lesere i utvikling, eller lesere som ser ut til å ha vansker med å oppnå flyt. Dette vil muligens generere andre resultater som i større grad kan synliggjøre om forstyrrelser synes å påvirke ordgjenkjenningshastigheten, og hvordan dette eventuelt kommer til uttrykk på ulike trinn i leseutviklingen – særlig med fokus på leseflyt.

## 5.2 Drøfting av resultatene i lys av studiens validitet og reliabilitet

Validitetssystemet utarbeidet av Cook og Campbell (1979) kan være nyttig i når en skal gjøre en selvstendig vurdering av egen kvantitativ forskning. Deres system trekker frem fire typer av validitet, eller kvalitetskrav. Disse er statistisk validitet, indre validitet, begrepsvaliditet og ytre validitet, hvor hver av dem synliggjør mulige trusler som potensielt kan gjøre det vanskelig å trekke valide slutninger fra forskningsresultatene. De kommende avsnittene vil bruke rammeverket for å diskutere hvorvidt gjeldende studie later til å ha god validitet.

### 5.2.1 Statistisk validitet

God statistisk validitet omhandler i hvilken grad vi kan anse våre statistiske slutninger som rimelig sikre, og kan slik sett anses som en forutsetning for andre validitetskrav (Lund, 2002). Reliabilitet spiller en sentral rolle her, da god reliabilitet handler om å kunne være sikker på at en har nøyaktige data. En mulig trussel mot reliabiliteten er derfor tilfeldige målefeil (Tuftes, 2018). Dette er feil i registrering, unøyaktighet eller misforståelser knyttet til administrering og gjennomføring. Et av tiltakene for å minimere disse var at undertegnede fikk opplæring og trening i å gjennomføre eksperimentet fra personer med høy kompetanse innenfor tilsvarende metode og design. Manuelle data ble sjekket av flere personer, først av meg selv, og deretter av medstudent Dzan Zelihic.

I gjeldende studie ble det benyttet et «innenindivid»-design, hvilket innebærer at alle deltakere blir eksponert for den samme manipulasjonen. Dette styrker reliabiliteten fordi det tenkes at dette reduserer feil knyttet til individuelle forskjeller i utvalget. Det samme gjelder antall observasjoner, hvor det tenkes at flere observasjoner vil føre til at tilfeldige målefeil jevnes ut (Tuft, 2018). Da deltakere gjennomførte flere blokker som hver av dem inkluderte flere oppgaver kan dette tenkes å styrke studiens reliabilitet.

Noe som kan føre til statistisk invaliditet er å gjøre en såkalt type I-feil, der en konkluderer med at det er sammenheng mellom variablene, til tross for at dette ikke er tilfellet. Resultatene fra de statistiske analysene viser til signifikante korrelasjoner mellom flere av variablene. Vanligvis blir et signifikansnivå på .05 benyttet (Field, 2009). I denne undersøkelsen var noen korrelasjoner signifikante på .05-nivå, men også på .01-nivå. Dette innebærer at sannsynligheten for å ha gjort en type I-feil synes relativt liten.

Styrken på de ulike sammenhengene bør også vurderes (Lund, 2002). Ifølge Lund (2002) må denne vurderingen gjøres i lys av gjeldende forskningsfelt, da «sterke» tendenser later til å være av ulik betydning i ulike felt. Gall, Gall & Borg (2007) påpeker at utdanningsvitenskapelig forskning vanligvis impliserer forskning på svært komplekse og sammensatte fenomener, noe som betyr at en ikke nødvendigvis forventer høyere korrelasjoner enn .40. Funnene i gjeldende analyser viser til korrelasjoner av ulik størrelse. Leseflyt korrelerte høyest med ordgjenkjenningshastighet på isolerte høyfrekvente ord ( $\rho = .390$ ), deretter ordgjenkjenningshastighet på isolerte lavfrekvente ord ( $\rho = .309$ ) og til ordgjenkjenningshastighet på lavfrekvente ord med ord-flanker ( $\rho = .318$ ). Korrelasjonen mellom leseflyt og ordgjenkjenningshastighet på høyfrekvente ord med ord-flanker var imidlertid noe lavere ( $\rho = .234$ ). Basert på korrelasjonene synes disse å vise til en tydelig sammenheng mellom avhengig og uavhengige variabler.

En annen trussel mot statistisk validitet er en såkalt type II-feil, som innebærer å konkludere med at det ikke finnes en sammenheng, selv om denne faktisk eksisterer. I denne oppgaven kunne uavhengige variabler ikke bidra til å forklare variasjon i avhengig variabel da alle ble inkludert i regresjonsmodellen. Basert på inngående analyser synes det som at sammenheng mellom variabler kan være til stede i dette tilfellet, men at den endelige



modellen ikke klarte å fange opp denne på grunn av andre betingelser som tilsynelatende kompliserte den. Eventuelle slutninger må derfor trekkes med forsiktighet.

### 5.2.2 Ytre validitet

Ytre validitet handler om hvorvidt studiens funn synes å kunne generaliseres, spesielt i hvilken grad disse gir inntrykk av å være representative for populasjonen (Lund, 2002). Utvalget i gjeldende studie er voksne studenter og unge i arbeid mellom 19-30 år, alle med norsk som morsmål og uten kjente lesevansker. I lys av dette ser trusselen om *individhomogenitet* å være til stede, hvilket innebærer at deltakere kan sies å være en relativt ensartet gruppe. Dette setter en begrensning på muligheten til å kunne generalisere til andre grupper, og svekker derfor studiens ytre validitet. Likeledes er utvalget hentet gjennom et bekvemmelighetsutvalg, og dermed vil trusselen om et *ikke-representativt individutvalg* også være til stede, hvilket impliserer en risiko for ugyldig generalisering av funn. Dette svekker også studiens ytre validitet. Som allerede nevnt tidligere i diskusjonskapittelet er deltakergrunnlaget hovedsakelig voksne universitetsstudenter, og en potensiell trussel er derfor at resultatene ville vært annerledes hvis en hadde undersøkt tilsvarende på yngre lesere i utvikling.

Ytre validitet må likevel ses i lys av studiens formål og design. Gjeldende studie har til hensikt å bruke et eksperimentelt design for å generere ny kunnskap knyttet til nåværende teorier om automatisert ordgjenkjenning, og kan slik sett plasseres under det som betegnes som «grunnforskning». Lund (2002) belyser at ytre validitet synes å være av mindre viktighet i denne type forskning, sammenlignet med de andre validitetskravene. Dette handler blant annet om at eksperimenter gjerne opererer med det Lund (2002) kaller for «prinsippgeneralisering», noe som innebærer at det som undersøkes ikke alltid er rettet mot en spesifikk populasjon.

Da formålet med gjeldende studie er å undersøke en spesifikk hypotese knyttet til automatisert ordgjenkjenning som det per i dag ikke foreligger et tydelig svar på, argumenterte dette for et utvalg som med nokså stor sikkerhet kunne produsere en mulig indikasjon på denne, altså ved å innhente voksne mennesker med etablerte leseferdigheter og

uten kjente lesevaner. Et ønske om god indre validitet kan derfor sies å ha gått på bekostning av en god ytre validitet. Cresswell (2014) påpeker at den største trusselen på ytre validitet oppstår *idet* forskeren generaliserer utover gruppen i studien. Ved å anerkjenne at en ikke nødvendigvis har sikret representativitet med sine funn, men heller hatt et ønske om å avdekke ny teoretisk kunnskap kan det argumenteres for at en svekket ytre validitet i denne sammenheng er akseptabelt.

### 5.2.3 Indre validitet

Indre validitet handler om hvorvidt det kan trekkes slutninger om relasjonen mellom de inkluderte variablene i studien (Lund, 2002). Å bedømme studiens indre validitet gjøres slik sett gjennom et kritisk blikk på anvendt design og metode. Gjeldende studie er en prediksjonsstudie, hvorpå et eksperimentelt design er benyttet. Dette i seg selv anses ofte som en indikator på god indre validitet (Lund, 2002), men gir likevel ingen garanti. Den tilsynelatende største trusselen mot god indre validitet ser ut til å være *instrumentering*, hvilket innebærer at måleprosedyren som er brukt resulterer i kunstige resultater. Måling av avhengig variabel baserer seg på en standardisert test på engelsk, men i en oversatt norsk versjon. Dette reduserer trusselen for instrumentering *noe*, men det kan fortsatt tenkes at instrumentering kan være en potensiell trussel i denne sammenheng. Når det gjelder måling av uavhengige variabler synes trusselen om instrumentering å være noe større, da det benyttes et ikke-standardisert mål som blant annet baserer seg på eksperimentell manipulasjon. Trusselen kan imidlertid tenkes å reduseres noe, da både teoretiske antakelser og tidligere studier er begrunnelsen for bruken av denne typen måling.

Trusler mot god indre validitet som ikke ser ut til å være spesielt relevante for denne studien er test-retest-effekten, da data er hentet fra kun ett målepunkt. Av samme grunn blir heller ikke frafall eller modning sett på som relevante trusler. Trusselen om modning, blir ytterligere av mindre sannsynlig på bakgrunn av at alle deltakere er voksne mennesker. I tillegg benytter studien et såkalt «innenindivid-design», noe som betyr at det ikke er benyttet en kontrollgruppe. Dette eliminerer dermed trusselen om både seleksjon og atypisk kontrollgruppeatferd. Oppsummert ser det ut som om studien har nokså god indre validitet.

#### 5.2.4 Begrepsvaliditet

Begrepsvaliditet knyttes til selve forskningsspørsmålet, da dette omhandler slutningen fra målevARIABLER til begreper– og slik sett om det synes å være samsvar mellom disse (Lund, 2002) I lys av denne oppgaven vil derfor diskusjonen om begrepsvaliditet knyttes til i hvilken grad de målene som brukes på leseflyt og automatisert ordgjenkjenning virkelig er representative indikatorer på disse begrepene. Da gjeldende forskning undersøker begreper som må anses som såkalte pedagogiske konstruksjoner, skapes en utfordring i lys av begrepsvaliditet allerede her. Denne bygger på at konstruksjonene i seg selv ikke nødvendigvis er målbare. Teorikapittelet har tidligere synliggjort en del av utfordringen ved begrepene som undersøkes, nettopp at de mangler en tydelig konsensus knyttet til definisjon. Dette får naturlige implikasjoner for tilnærming til måling. Fellesnevneren for mye av pedagogisk forskning er at konstruksjoner som undersøkes gjerne kvantifiseres gjennom ulike synlige indikatorer, altså å måle noe ikke-observerbart gjennom ulike observerbare forhold (Kleven, 2002). Dette innebærer et usikkerhetsmoment med tanke på begrepsoperasjonalisering, da en ikke kan si med sikkerhet at en virkelig måler det man er interessert i å undersøke.

I gjeldende forskning brukes en norsk versjon av TOWRE som mål på leseflyt. Denne er laget med utgangspunkt i den opprinnelige engelske versjonen som er standardisert og normert i Amerika, som ifølge Tarar, Meisinger og Dickens (2015) har god begrepsvaliditet. Denne kan imidlertid ikke garanteres for i den norske versjonen, noe som kan være en potensiell trussel.

Når det gjelder automatisert ordgjenkjenning har det allerede i teoridelen blitt løftet frem at dette er et konsept som ikke har en etablert konsensus med tanke på måling. Dette er også en del av fokuset for selve oppgaven, som baserer seg på en relativt ny og lite utforsket måte å måle automatisert ordgjenkjenning på. Da målingen i denne oppgaven også baserer seg på tidsbruk ved museklikk kan en potensiell trussel mot begrepsvaliditeten i dette tilfellet være såkalt «task-impurity», hvilket innebærer at en ikke bare kartlegger konseptet av interesse, men også i realiteten måler andre funksjoner (Snyder, Miyake & Hankin, 2015). I lys av gjeldende studie må det synliggjøres at det motoriske aspektet ved målingen vil tillegge den

noe usikkerhet. Hvorvidt studien har god begrepsvaliditet, må derfor ses i lys overnevnte faktorer.

## **6. Konklusjon**

Ifølge funnene ved gjeldende studie er det ingen signifikant forskjell på hvordan leseflyt kan predikeres av ordgjenkjenningshastighet på isolerte ord, sammenlignet med ord som presenteres med tilstøtende ord. Denne slutningen synes likevel ikke endelig, da oppgavens forskningsspørsmål ikke har blitt forsøkt undersøkt med et utvalg som i større grad preges av være i leseutviklingsfasen, eller et utvalg som inkluderer individer som ikke ser ut til å oppnå tilstrekkelig leseflyt. Dette hadde muligens kunne generert mer fruktbare funn, og bidratt til nyttig kunnskap med tanke på å forstå automatiserte prosesser i lesing bedre.

## Litteraturliste

- Adelman, J. S. (2012). *Visual word recognition: models and methods, orthography and phonology* (1.utg.). London: Psychology press. <https://doi-org.ezproxy.uio.no/10.4324/9780203107010>
- Altani, A., Protopapas, A. & Georgiou, G.K. (2017). The contribution of executive functions to naming digits, objects, and words. *Reading and Writing*, 30, 121–141. <https://doi-org.ezproxy.uio.no/10.1007/s11145-016-9666-4>
- Altani, A., Protopapas, A., Katopodi, K. & Georgiou, K.G. (2020). From Individual Word Recognition to Word List and Text Reading Fluency. *Journal of Educational Psychology*, 112(1), 22-39. <http://dx.doi.org.ezproxy.uio.no/10.1037/edu0000359>
- Bargh, J. A. (2007) *Social Psychology and the Unconscious: The Automaticity of Higher Mental Processes*. New York: Psychology Press.
- Besner, D., Stolz, J. A. & Risko, E. F. (2013). Automaticity. I H. Pashler (Red.), *Encyclopedia of the Mind*, 96-99. <http://dx.doi.org/10.4135/9781452257044.n37>
- Brysbaert, M., Mander, P., & Keuleers, E. (2018). The word frequency effect in word processing: an updated review. *Current directions in psychological science*, 27(1), 45–50.
- Bouma, H. (1970). Interaction effects in parafoveal letter recognition. *Nature*, 226, 177-178.
- Cohen, J. (1992). A Power Primer. *American Psychological Association*, 122(1), 155-159. <https://doi-org.ezproxy.uio.no/10.1037/0033-2909.112.1.155>
- Conklin, K., Pellier-Sánchez, A., & Carrol, G. (2018). *Eye-tracking: A guide for applied linguistics research*. Cambridge: Cambridge University Press.
- Cook, T. D. & Campbell, D. T. (1979). *Quasi-Experimentation, Design & analysis issues for field settings*. Boston: Houghton Mifflin Company.
- Creswell, W. J. (2014). *Research Design: Qualitative, Quantitativ & Mixed Methods Approaches* (4. utg.). California: Sage Publications, Inc.
- Cummings, K. D. & Petscher, Y. (2016) *The Fluency Construct: Curriculum-Based Measurement Concepts and Applications*. <https://doi.org/10.1007/978-1-4939-2803-3>
- Cumming, G. & Calin-Jageman, R. (2017). *Introduction to the New Statistics: Estimation, Open science, & Beyond*. New York: Routledge.
- Ehri, C. L. (1995). Phases of development in learning to read words by sight. *Journal of Reasearch in Reading*, 18(2), 116-125. <https://doi.org/10.1111/j.1467-9817.1995.tb00077.x>

- Ehri, C. L. (2005). Learning to Read Words: Theory, Findings, and Issues. *Scientific Studies of Reading*, 9(2), 167-188. [https://doi-org.ezproxy.uio.no/10.1207/s1532799xssr0902\\_4](https://doi-org.ezproxy.uio.no/10.1207/s1532799xssr0902_4)
- Eriksen, B. A. & Eriksen, C. W. (1974). Effects of noise letters upon the identification of a target letter in a nonsearch task\*. *Perception & psychophysics*, 16(1), 143-149.
- De nasjonale forskningsetiske komiteene. (2016). *Forskningsetiske retningslinjer for samfunnsvitenskap, humaniora, juss og teologi* (4. utg.). Hentet fra <https://www.etikkom.no>
- Diez, D. Barr, C. D., & Çetinkaya-Rundel, M. (2019). *OpenIntro Statistics* (4. utg.). Hentet fra <https://leanpub.com/openintro-statistics>
- Field, A. (2009). *Discovering statistics using SPSS (and sex and drugs and rock 'n' roll)* (3. utg.). Los Angeles: SAGE.
- Furnes, B., & Samuelsson, S. (2010). Predicting Reading and Spelling Difficulties in Transparent and Opaque Orthographies: A Comparison Between Scandinavian and U.S./Australian Children. *Dyslexia*, 16, 119-142.
- Gabrielsen, N. & Lundetræ, K. (2013). Ordavkoding og leseferdighet i PIRLS. I E. Gabrielsen & R. G. Solheim (Red.), *Over kneiken? Leseferdighet på 4. og 5.trinn i et tiårsperspektiv* (s. 117-127). Oslo: Akademika forlag.
- Gall, M. D., Gall, J. P., & Borg, W. R. (2007). *Educational research: an introduction* (8. utg.). Boston: Allyn and Bacon.
- Gough, B. P. & Tunmer, E. W. (1986). Decoding, Reading, and Reading Disability. *Remedial and Special Education*, 7(1), 6-10. <https://doi.org/10.1177/074193258600700104>
- Green, S. B. (1991). How many subjects does it take to do a regression analysis? *Multivariate Behavioral Research*, 26 (3), 499-510.
- Hair, J. F., Black, W. C., Babin, B. J., Anderson, R. E., Tatham, R.K. (2010). *Multivariate Data Analysis* (7. utg.). New York: Pearson.
- Hoover, A. W. & Gough, B. P. (1990). The Simple View of Reading. *Reading and Writing*, 2(2), 127-160. <https://doi.org/10.1007/BF00401799>
- Hudson, F. R., Pullen, C.P., Lane, H. B. & Torgesen, J. K. (2009). The Complex Nature of Reading Fluency: A Multidimensional View. *Reading & Writing Quarterly*, 25, 4–32. <https://doi.org/10.1080/10573560802491208>

- Huettig, F., Rommers, J. & Meyer, A. S. (2011) Using the visual world paradigm to study language processing: A review and critical evaluation. *Acta Psychologica*, 137, 151-171.
- Høien, T. & Lundberg, I. (2012). *Dysleksi: fra teori til praksis*. Oslo: Gyldendal Akademisk
- Kjærnsli, M. & Olsen, R. V. (2013). *Fortsatt en vei å gå: Norske elevers kompetanse i matematikk, naturfag og lesing i PISA 2012*. Oslo: Universitetsforlaget
- Kleven, T. A. (2002) Begrepsoperasjonalisering. I T. Lund (Red.). *Innføring i forskningsmetodologi* (s. 141-183). Bergen: Fagbokforlaget.
- Kuhn, M. R., Schwanenflugel, P. J. & Meisinger, E. B. (2010). Aligning Theory and Assessment of Reading Fluency: Automaticity, Prosody, and Definitions of Fluency. *Reading Research Quarterly*, 45(2), 230-251. <https://doi.org/10.1598/RRQ.45.2.4>
- Kunnskapsdepartementet. (2013). *På rett vei: kvalitet og mangfold i fellesskolen* (Meld. St. 20 (2012-2013)). Hentet fra <https://www.regjeringen.no/>
- Kunnskapsdepartementet. (2020). *Overordnet del – verdier og prinsipper for grunnopplæringen*. Hentet fra <https://www.udir.no/lk20/overordnet-del/prinsipper-for-laring-utvikling-og-danning/grunnleggende-ferdigheter/>
- LaBerge, D., & Samuels, J. S. (1974). Toward a theory of automatic information processing in reading. *Cognitive Psychology*, 6(2), 293-323. [https://doi.org/10.1016/0010-0285\(74\)90015-2](https://doi.org/10.1016/0010-0285(74)90015-2)
- Lien, M-C., Allen, P., Ruthruff, E., Grabbe, J., McCann, R. & Remington, R. (2006). Visual Word Recognition Without Central Attention: Evidence for Greater Automaticity with Advancing Age. *Psychology and Aging*, 21(3), 431-447.
- Lund, T. (2002). Metodologiske prinsipper og referanserammer. I T. Lund (Red.). *Innføring i forskningsmetodologi* (s. 79-124). Bergen: Fagbokforlaget.
- Moors, A. & De Houwer, J. (2006). Automaticity: A Theoretical and Conceptual Analysis. *The American Psychological Association*, 132(2), 297-326. <https://doi.org/10.1037/0033-2909.132.2.297>
- Moors, A & De Houwer J. (2007). What is Automaticity? An Analysis of its Component Features and Their Interrelations. I J. A. Bargh (Red.), *Social Psychology and the Unconscious: The Automaticity of Higher Mental Processes* (s. 11-50). New York: Psychology Press.



- National Institute of Child Health and Human Development. (2000). *Report of the National Reading Panel. Teaching children to read: and evidence-based assessment of the scientific research literature on reading and its implications for reading instruction*. Hentet fra <https://www.nichd.nih.gov/sites/default/files/publications/pubs/nrp/Documents/report.pdf>
- Navarro, D. & Foxcroft, D. (2019). *Learning statistics with jamovi: a tutorial for psychology students and other beginners*. (Version 0.70). <https://doi.org/10.24384/hgc3-7p15>
- Protopapas, A., Altani, A. & Georgiou, K. G. (2013). Development of serial processing in reading and rapid naming. *Journal of Experimental Child Psychology*, 116(4), 914-929. <https://doi.org/10.1016/j.jecp.2013.08.004>
- Rasinski, T. & Young, C. (2015) Fluency Matters. *The Connecticut Reading Association Journal*, 3(1), 21-26.
- Rasinski, T.V., Reutzel, D. R., Chard, D. & Linan-Thompson, S. (2011). Reading Fluency. I M. L. Kamil, P. D. Pearson, E. B. Moje, & P. Afflerbach (Red.), *Handbook of Reading Research* (4.utg., s. 286–319). New York: Routledge
- Salverda, A. P. & Tanenhaus, M. K. (2010). Tracking the time course of orthographic information in spoken-word recognition. *Journal of Experimental Psychology: Learning, Memory and Cognition*, 36, 108-117.
- Samuels, S. J. (2006). Towards a model of reading fluency. I S. J. Samuels & A.E. Farstrup (Red.), *What research has to say about fluency instruction* (s. 24–46). Newark, DE: International Reading Association.
- Snyder, H. R., Miyake, A. & Hankin, H. L. (2015). Advancing understanding of executive function impairments and psychopathology: Bridging the gap between clinical and cognitive approaches. *Frontiers in Psychology*, 6(328), 1-24. <https://doi.org/10.3389/fpsyg.2015.00328>
- Tabachnick, B. G., & Fidell, L. S. (2013). *Using multivariate statistics* (6. utg.). Boston: Pearson.
- Tarar, J. M., Meisinger, E. B. & Dickens, R. H. (2015). Test Review: Test of Word Reading Eddiciency – Second Edition (TOWRE-2). *Canadian Journal of School Psychology*, 30(4), 320-326. <https://doi.org/10.1177/0829573515594334>
- Tufte, P.A. (2018). *Hvordan lese kvantitativ forskning*. Oslo: Cappelen Damm.
- Van Heuven, J.B.W., Mandera, P., Keuleers, E. & Brysbaert, M. (2014). SUBTLEX-UK: A new and improved word frequency database for British English, *The Quarterly Journal of Experimental Psychology*, 67(6), 1176-1190. <https://doi.org/10.1080/17470218.2013.850521>

Wolf, M. & Katzir-Cohen, T. (2001). Reading Fluency and Its Intervention. *Scientific Studies of Reading*, 5(3), 211-239. [https://doi.org/10.1207/S1532799XSSR0503\\_2](https://doi.org/10.1207/S1532799XSSR0503_2)

Yarkoni, T., Balota, D. & Yap, M. (2008). Moving beyond Coltheart's N: A new measure of orthographic similarity. *Psychonomic Bulletin & Review*, 15(5), 971-979. <https://doi.org/10.3758/PBR.15.5.971>

Utdanningsdirektoratet. (2020). *Læreplanen i norsk (NOR01-06)*. Hentet fra <https://www.udir.no/lk20/nor01-06>

# Appendiks 1: Samtykkeskjema

## Vil du delta i forskningsprosjektet

### ”Måling av ordautomatisering ved lesing på norsk”?

Dette er et spørsmål til deg om å delta i et forskningsprosjekt hvor formålet er å undersøke hvorvidt automatisk ordgjenkjenning av et ord blir påvirket av andre nærliggende ord. I dette skrivet gir vi deg informasjon om målene for prosjektet og hva deltakelse vil innebære for deg.

#### Formål

Formålet med forskningsprosjektet er å bidra til å belyse ideen om at nærliggende ord kan forstyrre lesing av enkeltord. Dette er viktig fordi automatisert ordgjenkjenning er nødvendig for å oppnå leseflyt. Tidligere forskning har vist at evne til automatisk ordgjenkjenning vil kunne predikere leseflyt, uavhengig av ord- og bokstavkunnskap. Forskingen som er gjort har tatt utgangspunkt i lesing av isolerte enkeltord, noe som ikke gjenspeiler en reell lesesituasjon der ord vil være omringet av nærliggende ord. Det er imidlertid antatt at nærliggende ord kan skape forstyrrelser som vil påvirke automatisk ordgjenkjenning, men det er ikke gjort noe direkte forskning på akkurat dette.

Forskningsprosjektet er en masteroppgave som baserer seg på en større doktorgradstudie ved Universitetet i Oslo.

#### Hvem er ansvarlig for forskningsprosjektet?

Universitetet i Oslo, Institutt for spesialpedagogikk. Studien er et samarbeidsprosjekt mellom universitetet i Oslo og Department of Psychological and Brain Sciences ved University of Iowa.

#### Hvorfor får du spørsmål om å delta?

Utvalget er trukket ut ifra følgende kriterier:

- Alder (19-30)
- Syn (Deltakere bruker verken briller eller kontaktlinser)
- Språk (Har norsk som hovedspråk)
- Antall deltakere som vil få henvendelse er beregnet til å være 60

Rekruttering av deltakere for prosjektet vil hovedsakelig bestå av både kjente og ukjente studenter, samt unge voksne i arbeid.

## **Hva innebærer det for deg å delta?**

Hvis du velger å delta i prosjektet, innebærer det først at du gjennomfører en enkel lesetest (TOWRE) som måler effektiv ordlesing av både enkeltord og non-ord. Deretter vil du få presentert ord på en skjerm med tilhørende bilder, hvor oppgaven din blir å klikke på det bildet som tilsvarer riktig ord. Lesing på skjerm vil bli kartlagt ved bruk av eye-tracking, dette innebærer at vi bruker sensorteknologi for å kartlegge dine øyebevegelser under lesing. Formålet ved bruk av eye-tracking er å kunne måle hastigheten av informasjonsprosessering, med og uten forstyrrelser. Det vil ta deg ca. 45 minutter å gjennomføre hele prosessen.

### **Det er frivillig å delta**

Det er frivillig å delta i prosjektet. Hvis du velger å delta, kan du når som helst trekke samtykke tilbake uten å oppgi noen grunn. Alle opplysninger om deg er anonymisert. Det vil ikke ha noen negative konsekvenser for deg hvis du ikke vil delta eller senere velger å trekke deg.

### **Ditt personvern – hvordan vi oppbevarer og bruker dine opplysninger**

Vi vil bare bruke opplysningene om deg til formålene vi har fortalt om i dette skrivet. Vi behandler opplysningene konfidensielt og i samsvar med personvernregelverket.

- De som vil ha tilgang til dine data er prosjektansvarlig/veileder og to studenter
- Ingen personlige data vil bli registrert, og alt av data er anonymisert. Vi registrer bare alder og kjønn med brukerkoder for å sikre at ingen uvedkommende vil få tilgang til dine opplysninger. På grunn av brukerkoder vil deltakere heller ikke bli gjenkjent i masteroppgaven.

### **Hva skjer med opplysningene dine når vi avslutter forskningsprosjektet?**

Prosjektet skal etter planen avsluttes 31.01.2020. Videre vil resultatene analyseres og bearbeides i forbindelse med å besvare forskningsspørsmål utarbeidet for masteroppgaven. Ferdigstilling av masteroppgaven forventes i tidsperioden mellom mai – juni 2020. Anonymiserte forskningsdata vil etter dette bli tatt vare på i forbindelse med videre forskning, og muligens delt med andre forskere både i og utenfor EU.

### **Dine rettigheter**

Så lenge du kan identifiseres i datamaterialet, har du rett til:

- innsyn i hvilke personopplysninger som er registrert om deg,
- å få rettet personopplysninger om deg,
- få slettet personopplysninger om deg,
- få utlevert en kopi av dine personopplysninger (dataportabilitet), og
- å sende klage til personvernombudet eller Datatilsynet om behandlingen av dine personopplysninger.

### **Hva gir oss rett til å behandle personopplysninger om deg?**

Vi behandler opplysninger om deg basert på ditt samtykke.

På oppdrag fra Universitetet i Oslo, institutt for spesialpedagogikk har NSD – Norsk senter for forskningsdata AS vurdert at behandlingen av personopplysninger i dette prosjektet er i samsvar med personvernregelverket.

### Hvor kan jeg finne ut mer?

Hvis du har spørsmål til studien, eller ønsker å benytte deg av dine rettigheter, ta kontakt med:

- Institutt for spesialpedagogikk ved Athanasios Protopapas, på e-post: [Athanasios.protopapas@isp.uio.no](mailto:Athanasios.protopapas@isp.uio.no)
- E-post personvernombud: [personvernombud@uio.no](mailto:personvernombud@uio.no)
- NSD – Norsk senter for forskningsdata AS, på epost ([personvertjenester@nsd.no](mailto:personvertjenester@nsd.no)) eller telefon: 55 58 21 17.

Med vennlig hilsen

Prosjektansvarlig  
Athanasios Protopapas

Masterstudenter  
Dzan Zelihic/Caroline Nordli

---

## Samtykkeerklæring

Jeg har mottatt og forstått informasjon om prosjektet [*sett inn tittel*], og har fått anledning til å stille spørsmål. Jeg samtykker til:

- å delta i eye-tracking
- å delta i lesetest (TOWRE)
- at anonymiserte dataopplysninger lagres på ubestemt tid etter prosjektslutt, til eventuell bruk i videre forskning

Jeg samtykker til at mine opplysninger behandles frem til prosjektet er avsluttet, ca. Juni 2020

---

(Signert av prosjektdeltaker, dato)

## Appendiks 2: Norsk versjon av TOWRE

is	verk	kraft	uniform
opp	hopp	bedre	nødvendig
katt	part	insekt	problem
rød	fort	plan	fraværende
min	fin	vakker	reklame
to	melk	berømt	hyggelig
nå	bak	ungdom	eiendom
vi	låst	uten	avstresset
du	finn	endelig	informasjon
tre	papir	sterkt	reprise
og	åpen	budsjett	forståelse
ja	kant	utpresse	følelse
av	eple	kontakt	selvsikker
han	kyss	lakris	intuisjon
som	penger	morgen	skrythals
bok	greit	løse	mulighet
var	far	beskrive	overmodig
hjelp	river	utstyr	fremmedgjort
så	plass	forretning	utryddet
tid	kort	kvalitet	prærie
ved	lært	patent	limousin
la	folk	kollaps	hjertelig
men	nesten	element	detektiv
baby	vinke	pioner	nylig
ny	barn	glemsom	instruksjon
stopp	stort	farefull	gjennomsiktig

## Appendiks 3: Høyfrekvente ord

ORD	N	Bigram		OLD20	ZipfFreq	English	FreqGroup
		letters	w/end				
hale	4	3,628	1,000	3,981	tail	2	
nese	4	3,824	1,000	3,981	nose	2	
robot	5	3,210	1,800	3,981	robot	2	
maske	5	3,625	1,000	3,987	mask	2	
drage	5	3,481	1,350	3,995	dragon	2	
sykkel	6	3,214	1,850	4,011	bicycle	2	
verktøy	7	3,165	2,750	4,025	tools	2	
pølse	5	3,076	1,650	4,029	sausage	2	
pilot	5	3,294	1,650	4,032	pilot	2	
skygge	6	3,183	1,650	4,053	shadow	2	
finger	6	3,705	1,000	4,063	finger	2	
ridder	6	3,541	1,450	4,070	knight	2	
teppe	5	3,498	1,200	4,076	rug	2	
tiger	5	3,687	1,000	4,077	tiger	2	
rotte	5	3,573	1,000	4,084	rat	2	
sukker	6	3,419	1,050	4,094	sugar	2	
jakke	5	3,385	1,000	4,107	jacket	2	
mobil	5	3,077	1,750	4,115	cell phone	2	
demon	5	3,512	1,100	4,117	demon	2	
klokke	6	3,341	1,250	4,132	watch	2	
vampyr	6	2,996	1,850	4,136	vampire	2	
slange	6	3,603	1,400	4,148	snake	2	
billett	7	3,540	1,700	4,150	ticket	2	
vindu	5	3,125	1,750	4,164	window	2	
alarm	5	3,283	1,900	4,190	alarm	2	
søppel	6	3,036	1,950	4,191	garbage	2	
kirke	5	3,347	1,300	4,232	church	2	
olje	4	3,260	1,350	4,236	oil	2	
pakke	5	3,471	1,100	4,245	package	2	
drosje	6	3,362	1,750	4,261	cab	2	
nøkkel	6	3,108	1,750	4,264	key	2	
papir	5	3,198	1,700	4,285	paper	2	
kake	4	3,585	1,000	4,351	cake	2	
stjerne	7	3,543	1,750	4,351	star	2	

kino	5	3,108	1,750	4,354	cinema	2
fotball	7	3,159	1,950	4,358	soccer ball	2
kylling	7	3,313	1,850	4,398	chicken	2
kjole	5	3,439	1,450	4,414	dress	2
engel	5	3,676	1,500	4,418	angel	2
bombe	5	3,248	1,350	4,434	bomb	2
felle	5	3,611	1,000	4,437	mouse trap	2
monster	7	3,701	1,550	4,469	monster	2
rose	4	3,562	1,000	4,511	rose	2
dronning	8	3,473	1,800	4,539	queen	2
skole	5	3,611	1,150	4,578	school	2
soldat	6	3,285	1,750	4,653	soldier	2
pistol	6	3,354	1,700	4,726	gun	2
ansikt	6	3,470	1,850	4,752	face	2
konge	5	3,656	1,450	4,779	king	2
øye	3	3,084	1,000	4,889	eye	2
hjerte	6	3,558	1,600	4,896	heart	2
doktor	6	3,313	1,800	4,965	doctor	2
kaffe	5	3,175	1,400	5,043	coffee	2
nummer	6	3,334	1,200	5,055	number	2
fengsel	7	3,524	1,500	5,129	prison	2
penger	6	3,828	1,000	5,622	money	2
finne	5	3,637	1,000	4,420	fin (fish)	2
kule	4	3,443	1,000	4,503	marble	2
baby	4	2,893	1,350	4,895	baby	2
pose	4	3,467	1,000	4,121	plastic bag	2



## Appendiks 4: Lavfrekvente ord

<b>ORD</b>	<b>N</b>	<b>Bigram token</b>	<b>OLD20</b>	<b>ZipfFreq</b>	<b>English</b>	<b>FreqGroup</b>
	<b>letters</b>	<b>w/end</b>				
snegle	6	3,371	1,650	1,824	snail	1
badstu	6	3,051	2,250	1,824	sauna	1
føner	5	3,382	1,350	1,949	hair dryer	1
lefse	5	3,276	1,450	2,046	(a Norwegian dish)	1
skøyte	6	3,120	1,600	2,301	skate	1
rake	4	3,631	1,000	2,389	rake	1
silo	4	3,212	1,000	2,426	silo	1
flygel	6	3,095	1,900	2,426	grand piano	1
deksel	6	3,453	1,900	2,493	phone cover	1
kongle	6	3,522	1,750	2,523	pinecone	1
giraff	6	2,857	1,900	2,523	giraffe	1
gaupe	5	3,173	1,550	2,578	lynx	1
kjegle	6	3,339	1,750	2,578	cone	1
termos	6	3,491	2,000	2,602	thermos	1
gevir	5	3,333	1,850	2,602	antlers	1
linjal	6	3,162	1,900	2,669	ruler	1
pokal	5	3,212	1,900	2,709	tropy	1
kiwi	4	2,360	1,650	2,838	kiwi	1
børste	6	3,386	1,550	2,852	hair brush	1
skorstein	9	3,630	2,550	2,879	chimney	1
kjele	5	3,567	1,000	2,891	pot	1
fele	4	3,640	1,000	2,891	fiddle	1
tablett	7	3,435	1,950	2,903	pill	1
valnøtt	7	2,974	2,750	2,903	walnut	1
komfyr	6	2,936	2,250	2,915	stove	1
lasso	5	3,247	1,700	2,915	lasso	1
rosin	5	3,519	1,400	2,938	raisin	1
tromme	6	3,451	1,450	2,938	drum	1
moped	5	3,202	1,950	2,949	moped	1
strømpe	7	3,234	1,800	2,949	stocking	1
vaffel	6	3,077	1,900	2,960	waffle	1
binders	7	3,604	1,750	2,960	paper clip	1
krage	5	3,486	1,300	2,960	collar	1
kvis	5	3,422	1,350	2,981	pimple	1
humle	5	3,140	1,200	3,000	bumblebee	1
pære	4	2,926	1,150	3,019	pear	1

bever	5	3,571	1,000	3,046	beaver	1
kaktus	6	3,273	2,000	3,046	cactus	1
mugge	5	3,155	1,050	3,080	pitcher	1
bestikk	7	3,422	1,750	3,146	cutlery	1
bøffel	6	2,795	1,850	3,160	buffalo	1
delfin	6	3,440	1,750	3,167	dolphin	1
trompet	7	3,438	1,750	3,186	trumpet	1
øgle	4	3,018	1,550	3,186	lizard	1
støvel	6	3,196	1,850	3,216	gumboot	1
traktor	7	3,457	1,900	3,234	tractor	1
bluse	5	3,349	1,350	3,261	blouse	1
kjede	5	3,479	1,550	3,266	necklace	1
truse	5	3,480	1,350	3,271	underpants	1
kano	4	3,274	1,000	3,277	canoe	1
potet	5	3,531	1,300	3,297	potato	1
hanske	6	3,628	1,400	3,325	glove	1
kamel	5	3,407	1,300	3,325	camel	1
muskel	6	3,337	1,950	3,334	muscle	1
panne	5	3,643	1,000	3,343	pan	1
krabbe	6	3,232	1,450	3,347	crab	1
tønne	5	3,273	1,250	1,648	barrel	1
sandal	6	3,375	1,600	2,301	sandal	1
planke	6	3,457	1,450	1,648	plank	1
vimpel	6	3,215	1,950	2,389	pennant	1

## Appendiks 5: Liste over alle elementer i eksperimentell oppgave

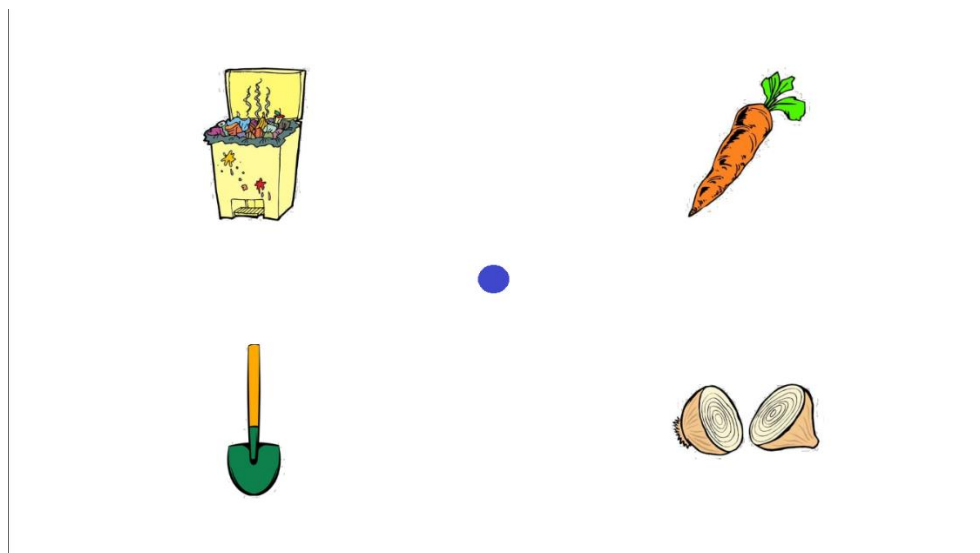
Target	Compet	Distr1	Distr2	Flanker1	Flanker2
snegle	stubbe	pisk	sprøyte	gardin	konfekt
badstu	bunad	knapp	nonne	skulder	drops
føner	fyrstikk	seng	elefant	bensin	eksamen
lefse	leppe	glass	badekar	ansvar	krutt
skøyte	skjelett	vekt	brygge	krutong	anker
rake	ribbe	mopp	batteri	hekk	vogn
silo	suppe	speil	ilder	bygning	frisør
flygel	flue	dusj	diamant	saft	port
deksel	data	gris	paraply	balkong	flåte
kongle	kåpe	bjørn	grevling	flytevest	jord
giraff	skjorte	stol	baguet	furu	vodka
gaupe	gitar	bord	låve	krem	granat
kjegle	kjevle	vott	pizza	brud	stav
termos	tunnel	skjell	hake	kaviar	gløgg
gevir	gebiss	hval	hette	figur	tiara
linjal	lilje	katt	fløyte	busk	grus
pokal	ponny	hund	honning	frimerke	kors
kiwi	kringle	spyd	bacon	rein	tøffel
børste	blyant	hjelm	radio	fjøl	krok
skorstein	skute	skjegg	radar	heis	reim
kjele	kiste	trapp	garasje	undulat	netting
fele	flaske	fjær	toalett	jerv	knagg
tablett	tvilling	hatt	lader	skinke	gjelle
valnøtt	vulkan	hjul	koala	bonde	nakke
komfyr	kompass	buss	kenguru	grotte	melk
lasso	larve	hest	salami	butikk	balsam
rosin	rakett	brød	paprika	tåke	mose
tromme	tavle	kjeks	globus	kartong	sylinder
moped	måne	skjerf	medalje	dirigent	dessert
strømpe	såpe	smokk	flodhest	kalv	korsett
vaffel	vugge	struts	appelsin	katalog	røyskatt
binders	bleie	sopp	ananas	plakat	teater
krage	krykke	grill	kamera	atlas	skap
kvise	krone	kniv	stadion	gang	flokk
humle	hytte	skje	mikrofon	karamell	autograf

pære	pipe	slips	parasoll	kjeller	alfabet
bever	belte	brev	lysekrone	sigarett	religion
kaktus	kanon	kost	knute	manet	fossil
mugge	motor	benk	leopard	akvarium	reptil
bestikk	ballong	telt	viking	gruve	potte
bøffel	bukse	tang	gaffel	oppskrift	medisin
delfin	dukke	ring	esel	sang	dans
trompet	tommel	vest	hammer	meny	atom
øgle	ørret	shorts	lego	melon	skjerm
støvel	stige	flagg	basseng	kommode	loft
traktor	turban	kurv	genser	antenne	juletre
bluse	blomkål	tann	popkorn	lakris	molekyl
kjede	kjøkken	sverd	spiker	parti	brikke
truse	tunge	munk	gjerde	offer	lypsyl
kano	kanin	skjold	hulle	fjell	mulighet
potet	puddel	mark	håndkle	skorpe	sorg
hanske	høvel	maur	pute	labyrint	klut
kamel	kalkun	frosk	jeger	stue	nyre
muskel	musling	visp	smultring	kantine	ramme
panne	palme	svamp	spagetti	kryss	røyk
krabbe	kvadrat	mygg	gele	bevis	angrep
tønne	tomat	ratt	sennep	fortau	rekkverk
sandal	svane	stein	krystall	drøm	teori
planke	pingvin	bart	lampe	verdi	middel
vimpel	vifte	teip	viskelær	fest	resultat
hale	hare	biff	pennal	logo	lunsj
nese	neve	heks	magnet	leilighet	optiker
robot	reke	prest	astronaut	orkan	bølge
maske	mage	klovn	bibel	ørken	sykehus
drage	dyne	lokk	sushi	måned	skjørt
sykkel	sitron	perm	bobil	linje	slave
verktøy	vinyl	skål	brille	mynt	smil
pølse	panda	strikk	høne	klem	spion
pilot	pinne	garn	hane	løype	lomme
skygge	skjære	skall	dynamitt	klasse	sommer
finger	pille	kritt	peanøtt	spøkelse	høst
ridder	rifle	geit	edderkopp	maskin	bokstav
teppe	terning	tårn	spinndelvev	vinter	snor

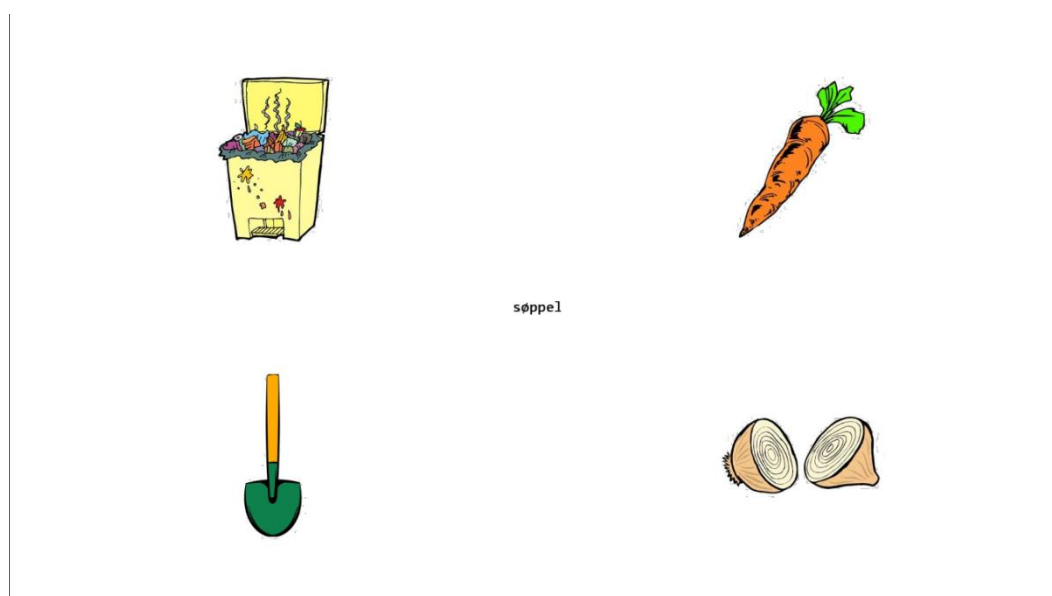
tiger	tyggis	lykt	manual	skum	syre
rotte	rustning	krans	pinsett	tegn	lemur
sukker	salat	ørn	bikini	premie	kobra
jakke	jolle	egg	trillebår	dress	spytt
mobil	madrass	negl	hyssing	brosjyre	toll
demon	diplom	ost	påfugl	krybbe	åker
klokke	kabel	tog	parfyme	fryser	motor
vampyr	vinge	fly	kleshenger	politi	tornado
slange	skuter	lue	kalender	grein	mansjett
billett	bjelle	penn	tastatur	regning	kode
vindu	veske	and	gorilla	gladiator	sikte
alarm	albue	elg	propell	uniform	spill
søppel	spade	løk	gulrot	gribb	menneske
kirke	kjelke	sko	bringeber	sirkus	plen
olje	okse	bål	hamburger	klima	kontroll
pakke	parykk	øse	sløyfe	samfunn	skog
drosje	dråpe	slott	rekkert	president	avløp
nøkkel	navel	tanks	syltetøy	media	blære
papir	padde	veps	løve	nebb	chili
kake	kajakk	spray	hyene	piknik	pels
stjerne	stativ	gips	kirsebær	gress	melding
kino	kikkert	bro	knekkebrød	syndrom	ungdom
fotball	flette	mus	sebra	saus	planet
kylling	kjetting	saks	drone	system	univers
kjole	kjerre	skilt	mandel	valg	ledning
engel	eple	hjort	marsvin	vann	adresse
bombe	bolle	fyr	flaggermus	mønster	video
felle	flamme	nisse	mikroskop	vinkel	krydder
monster	måke	sau	teleskop	gjest	vitne
rose	rede	iglo	trekkspill	venn	kosthold
dronning	drue	caps	tuba	verden	student
skole	sklie	smør	ubåt	voks	dikt
soldat	singlet	rev	satellitt	pedagog	navn
pistol	pirat	kort (spill)	presang	natur	trikk
ansikt	ankel	disc	omelett	himmel	maling
konge	koffert	sekk	åre	middag	visker
øye	øre	dør	puma	vorte	meisel
hjerter	hjerne	peis	skilpadde	ikon	blad

doktor	dommer	ved	pass	bakterie	signal
kaffe	kanne	øks	muffins	dato	stasjon
nummer	nuddel	sag	pensel	kanal	pumpe
fengsel	filter	drill	troll	bryllup	spark
penger	perle	tre	serviett	grense	konsoll
finne	fakkell	bok	krakk	seil	oter
kule	kube	kork	yoghurt	fisk	brus
baby	bamse	nål	grøt	sminke	pledd
pose	pote	kran	ris	innlegg	fjær

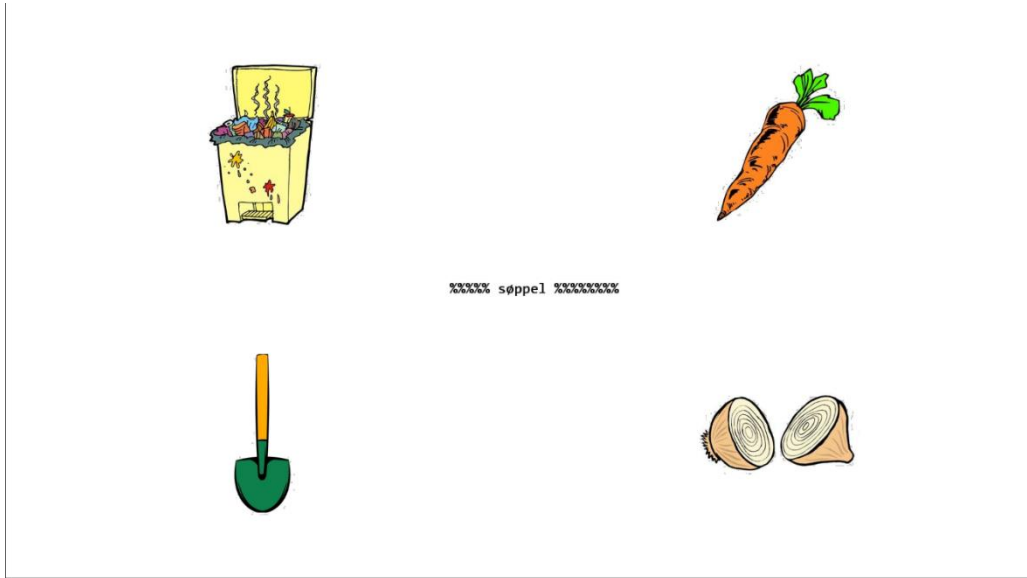
## Appendiks 6: Illustrasjoner fra eksperimentell oppgave



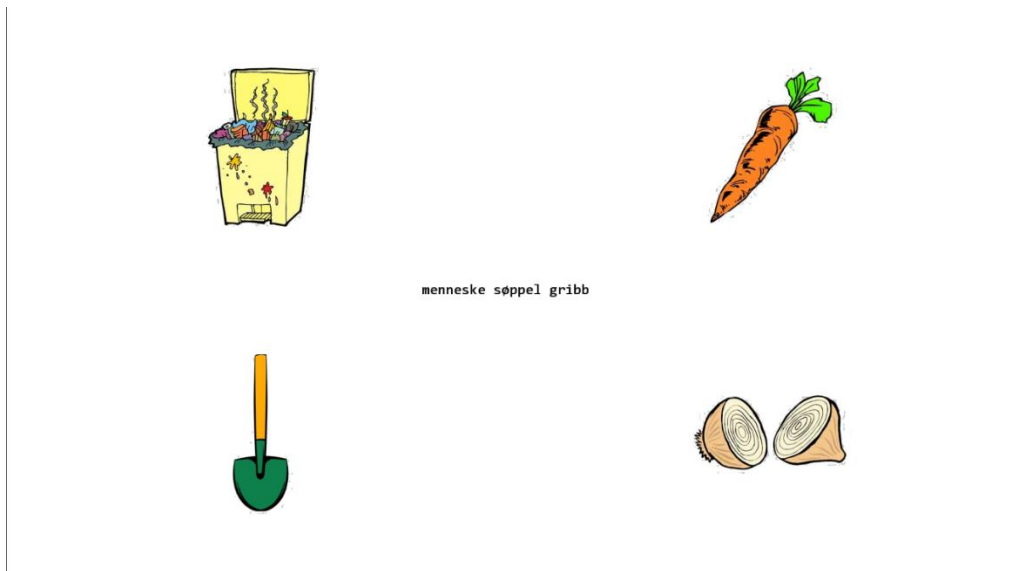
Figur 1: Med blå sirkel



Figur 2: Målord i isolert tilstand

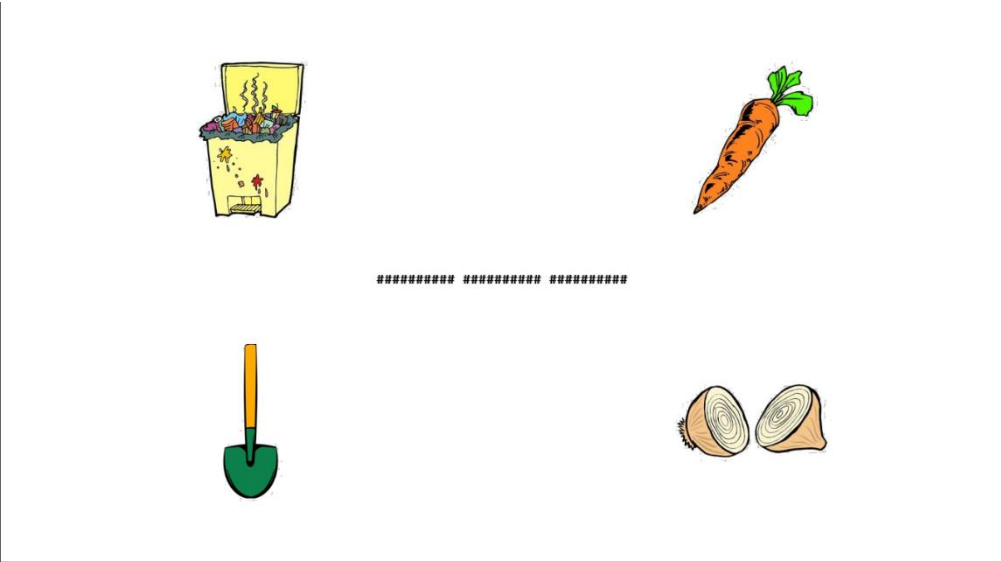


Figur 3: Målord med visuelle flanker



Figur 4: Målord med ord-flanker





Figur 5: Maskering av stimuli

## Appendiks 7: Kolinearitetsstatistikk

---

	VIF	Tolerance
Log10_IsolertTilstandHøyfrekvent	2.39	0.418
Log10_FlankerTilstandHøyfrekvent	2.39	0.418

---

### Kollinearitetsstatistikk for første sekvens av regresjonsanalyse, høyfrekvent ordgjenkjenningshastighet

---

	VIF	Tolerance
Log10_IsolertTilstandLavfrekvent	2.45	0.407
Log10_FlankerTilstandLavfrekvent	2.45	0.407

---

### Kollinearitetsstatistikk for andre sekvens av regresjonsanalyse, lavfrekvent ordgjenkjenningshastighet

---

	VIF	Tolerance
Log10_IsolertTilstandHøyfrekvent	6.33	0.158
Log10_FlankerTilstandHøyfrekvent	3.23	0.309
Log10_IsolertTilstandLavfrekvent	6.09	0.164
Log10_FlankerTilstandLavfrekvent	3.44	0.291

---

### Kollinearitetsstatistikk for siste sekvens av regresjonsanalyse, alle variabler