

Anna Vinichenko

# Nordmenns reaksjoner på avvik i innlærernorsk

En eye-tracking undersøkelse av effekter av finitt- og tempusavvik på språkprosessering

Masteroppgave i nordisk språkvitenskap

Veileder: Kristin Melum Eide

Medveileder: Mila Vulchanova

Mai 2021



Anna Vinichenko

# **Nordmenns reaksjoner på avvik i innlærernorsk**

En eye-tracking undersøkelse av effekter av finitt- og tempusavvik på språkprosessering

Masteroppgave i nordisk språkvitenskap  
Veileder: Kristin Melum Eide  
Medveileder: Mila Vulchanova  
Mai 2021

Norges teknisk-naturvitenskapelige universitet  
Det humanistiske fakultet  
Institutt for språk og litteratur



Kunnskap for en bedre verden



# Sammendrag

Språklige avvik – grammatiske feil, staving- og bøyingsfeil – fører til spesifikke hjerne- og atferdsresponser i mennesker. Denne studien har undersøkt hvordan morsmålstalere av norsk reagerer på forskjellige typer av avvik i verbmorfologi ved å gjennomføre et eye-tracking-eksperiment. Informantene leste en rekke setninger som inneholdt avvik i to morfosyntaktiske kategorier: finittet og tempus. Stimulisetningene inneholdt infinitivsformer av verb som fremstod som avvikende i konteksten, med enten finitt- eller tempusavvik, eller både finitt- og tempusavvik samtidig. Informantene leste også kontrollsetningene der infinitivsformen var grammatikalsk riktig, til sammenligning.

Basert på Eides (2009) teori om at finittet og tempus er to adskilte trekk, forventet jeg å finne forskjellige reaksjoner på finitt- og tempusavvikene. I tillegg, basert på bevismateriale som antyder at finittet er et mindre fremtredende trekk enn tempus (mest sannsynlig fordi finittet mangler semantisk innhold), forventet jeg at tempus ville oppfattes som et «viktigere» trekk, og at nordmenn ville prosessere finittavvik fortere enn tempusavvik. Jeg forventet også at prosessering av setninger med begge typer av avvik ville kreve en større kognitiv innsats enn prosessering av setninger med bare én type av avvik.

Resultatene av eye-tracking-eksperimentet støttet delvis disse prediksjonene: setninger med finitt- og tempusavvik produserte forskjellige reaksjoner, og tempusavviket førte til lengre lesetid, noe som indikerer at dette avviket var vanskeligere å prosessere enn finittavvik. Imidlertid produserte ikke setninger med dobbelt-avviket en større effekt enn setninger som inneholdt bare én type av avvik.

Selv om videre forskning klart er nødvendig for å klarlegge kilden til effektene oppdaget i denne studien, påviser resultatene ulike lesemønstre i setninger med finitt- og tempusavvik, noe som synes til å være empirisk evidens som korroborerer Eides (2009) teori om at finittet og tempus er to distinkte trekk.



# Abstract

Linguistic errors – grammar mistakes, incorrect conjugation and spelling – all produce specific responses in human behaviour and brain activity. This study investigated the reactions of native Norwegian speakers to errors in verb morphology by studying their eye movements and reading patterns. The participants in this eye-tracking experiment read a number of sentences, which contained violations of two functional features: finiteness and tense. The stimuli contained infinitive verb forms used in three different types of context, which made the verb appear erroneous, either violating the finiteness or the tense distinction, or both. In the control condition the infinitive verb form was used in a context which made it appear grammatical.

Based on Eide's (2009) theory stating that finiteness and tense are two separate and distinct features, I predicted that the participants would demonstrate different reactions to violations of finiteness and tense. In addition, based on a body of evidence suggesting that finiteness is a less salient feature than tense (likely due to its semantic vacuity), I predicted that Norwegian speakers would perceive the tense feature as more "important", and would therefore recover quicker from the violations of finiteness rather than tense. The combined condition, containing both violations, was hypothesized to be the most cognitively taxing for the participants.

The results partially supported these predictions: the finiteness and tense conditions did in fact produce distinct reading patterns, with the tense violation resulting in significantly longer reading times, indicating more cognitive effort required to repair the anomaly. However, the combined condition did not differ significantly from either of the pure violation conditions.

Although further research is necessary in order to confirm the underlying source of the effects observed in these results, the implication of this study is that reactions to finiteness and tense violations differ, which suggests that these two functional categories may in fact be distinct, corroborating Eide's theory with empirical evidence.





# Forord

To all the zealous defenders of the Norwegian language in academia: I apologise for writing the acknowledgments in English, but all of these thanks come from my heart, and you know as well as I do that we do not get to choose our *hjertespråk*.

First of all, I would like to thank the Language and Literature department at NTNU for giving me the opportunity to work on this project, when other universities believed a foreigner does not stand – or deserve – a chance of completing a master's degree in Norwegian linguistics.

I want to thank Kristin Melum Eide and Evelyn Arko Milburn for all the invaluable help and guidance they offered, as well as for genuinely caring, not just about this project but also about me.

I would like to thank each and every one of the participants that took part in my experiment despite the global pandemic, the bus strikes, and having to go all the way to Dragvoll. Without you this study would not have been possible.

I want to thank Charlotte Jensen for doing everything to make me feel at home here in Norway, for her selflessness, and for just being Charlotte. You truly are the nest.

Most of all, I would like to thank my incredible parents, Tatiana and Dmitry. Мама и папа, спасибо за вашу неиссякаемую поддержку и заботу, за наши долгие разговоры и за то, какие вы классные. Без вас у меня ничего бы не получилось.

Last but not least, I would like to thank myself – for being strong, and brilliant, and for not giving up.

Trondheim, May 2021

Anya Vinichenko



# Innholdsfortegnelse

Sammendrag.....	v
Abstract .....	vii
Forord .....	ix
Innholdsfortegnelse.....	xi
Figurer og tabeller.....	xiii
<b>1 Innledning.....</b>	<b>1</b>
1.1 Denne studien.....	2
1.2 Oppbygging av oppgaven.....	3
1.3 Terminologi, forkortelser og symboler .....	4
<b>2 Finitthet .....</b>	<b>7</b>
2.1 Finitthet, hva er det egentlig? .....	7
2.2 Eides teori om finitthet.....	9
2.3 Finit- vs. tempusdistinksjonen .....	11
2.3.1 Finit- og tempusavvik i norsk som andrespråk.....	12
2.3.2 Tap av finittdistinksjonen i engelsk .....	14
2.4 Chapter summary: Finiteness.....	16
<b>3 Eye-tracking i språkforskning.....</b>	<b>19</b>
3.1 Metoder i språkforskning .....	19
3.2 Eye-tracking .....	24
3.2.1 Teknologiske aspekter .....	25
3.2.2 Visuell språkprosessering.....	28
3.2.3 Hva kan eye-tracking avsløre om språk? .....	31
3.3 Chapter summary: Eye-tracking .....	38

<b>4 Metode</b> .....	<b>41</b>
4.1 Hypoteser og prediksjoner .....	41
4.2 Materialet .....	43
4.2.1 Verbutvalget i eksperimentet .....	43
4.2.2 Design av stimuli .....	45
4.2.3 Design av eksperiment .....	47
4.3 Informanter .....	48
4.4 Utstyr .....	49
4.5 Gjennomføring .....	50
4.6 Chapter summary: Method .....	51
<b>5 Dataanalyse og resultater</b> .....	<b>53</b>
5.1 Kvalitetssikring av data .....	53
5.2 Avhengige variabler .....	56
5.3 Rådata .....	57
5.4 Statistisk analyse .....	59
5.4.1 Analysemodell .....	59
5.4.2 Transformasjon av data .....	61
5.4.3 Effektstørrelser .....	63
5.4.4 Resultater for det verbale interessefeltet .....	64
5.4.5 Resultater for det post-verbale interessefeltet .....	71
5.5 Chapter summary: Results .....	76
<b>6 Diskusjon</b> .....	<b>79</b>
6.1 Finitt- vs. tempusavvik .....	80
6.2 Finitt- pluss tempusavvik .....	82
6.3 Videre forskning .....	84
6.4 Konklusjoner .....	87
6.5 Chapter summary: Discussion .....	88
Referanser .....	91
Vedlegg A: Stimuliliste .....	101
Vedlegg B: Parvise sammenligninger .....	107

# Figurer og tabeller

Figur 3.1: Hornhinnerefleksjonen og den lyse pupillen som de vises av et infrarødt kamera_____	26
Figur 3.2: Dataenheter i eye-tracking_____	28
Figur 3.3: Øyebevegelser under lesing_____	30
Figur 3.4: Eksempel på et lesemønster_____	34
Tabell 1.1: Symboler og forkortelser_____	6
Tabell 2.1: Eides skjema av finitt- og tempusdistinksjoner_____	10
Tabell 2.2: Tempusparadigme for det norske verbet <i>like</i> _____	11
Tabell 2.3: Tempusparadigme for svake verb i moderne engelsk_____	14
Tabell 2.4: Tempusparadigme for sterke verb i moderne engelsk_____	14
Tabell 2.5: Tempusparadigme i gammel engelsk_____	15
Table 2.6: Eide's (2009) finiteness and tense distinctions_____	16
Tabell 5.1: Blikkdata for det verbale interessefeltet_____	58
Tabell 5.2: Blikkdata for det post-verbale interessefeltet_____	58



# Kapittel 1

## Innledning

«Språkfeil» er et fenomen som alle støter på nesten hver dag: i tekstmeldinger fra venner og i sosiale medier, i avisartikler og til og med i bøker. Mange blir lett irritert over feilstaving og bøyingsfeil, mens andre legger knapt merke til det; noen gruer seg for at det norske språket er i ferd med å bli ødelagt, mens andre synes at det ikke er så farlig å blande *og* og *å* en gang iblant. Selv om meninger om viktigheten av grammatikk varierer mye fra person til person, finnes det ubevisste biologiske reaksjoner på språkfeil som er ganske like i alle mennesker. Når man hører eller ser en grammatisk feil, reagerer hjernen umiddelbart ved å produsere spesifikke elektriske signaler. De prosessene som foregår i hjernen når den forsøker å prosessere en skriftlig språkfeil påvirker også menneskets kropp og atferd, først og fremst øyebevegelsene: mennesker ser lengre på de ordene som inneholder grammatikkavvik og trenger ofte å lese setningen flere ganger for å forstå den. Fordi mennesker ikke kan kontrollere slike reaksjoner som hjerneaktivitet og øyebevegelser, kan disse tilby en unik og mer objektiv innsikt i hvordan mennesker faktisk oppfatter språkfeil. I tillegg avslører disse hvordan hjernen prøver å kompensere språklige avvik ved å finne ut hva som egentlig er meningen med setningen, for å «reparere skaden» feilen gjør.

Det finnes mange grunner til at forskere er interessert i prosessering av språklige avvik, alt fra en rent vitenskapelig interesse for indre prosesser som foregår i menneskets sinn til en mer praktisk bruk i for eksempel språkundervisning. Selv om morsmålstalere også produserer avvikende språk, er «språkfeil» virkelig et kjennetegn på innlærerspråk. Voksne mennesker klarer sjelden (og ifølge noen språkforskere – aldri) å nå morsmålslik kompetanse i et fremmedspråk, og i tidlige stadier av andrespråkstilegnelse er språkfeil helt klart uunngåelig. Disse «feilene» (ofte kalt «avvik», som er en mindre belastende term) har derimot en direkte påvirkning på livet til innvandrere. Som nevnt blir mange morsmålstalere irritert over språklige feil og kan

ofte bedømme språkinnlærere negativt basert på språkferdighetene deres (Kløve & Husby, 2008). Derfor fremstår det som interessant å undersøke hvordan morsmålstalere reagerer på ulike språklige avvik som kjennetegner S2-produksjon, og hvorvidt noen typer språkfeil er «verre» enn andre. Språk er den vanligste kommunikasjonsmåte for mennesker, og språklige avvik kan derfor være avgjørende for vellykket kommunikasjon. Gjennom analyse av menneskets øyebevegelser er det mulig å trekke konklusjoner om hvordan språkfeil påvirker kognisjon. I tillegg kan analyse av reaksjoner på språklige feil brukes til å lære mer om strukturen til et gitt språk: fordi språket har sin eksistens i menneskets sinn/hjerne, er det viktig å undersøke hvorvidt formelle teorier om språk og grammatikk kan underbygges og korroboreres med observerbar kognitiv evidens.

### 1.1 Denne studien

Målet med denne studien var å undersøke hvordan nordmenn prosesserer avvik i verbmorfologi mens de leser, samt sammenligne reaksjoner på forskjellige typer av avvik. Avvik i verbmorfologi er fokuset i denne studien fordi tilegnelse av morfosyntaks er en av de største utfordringene for S2-innlærere: det finnes mye interindividuell variasjon i grammatisk bruk av bøyingsaffikser blant innlærere som ellers har samme nivå av språkkompetanse, samt intraindividuell variasjon, som kan være avhengig av slike ting som tid på dagen og hvor mye kaffe personen drakk den morgenen (Slabakova, 2016, s. 189). Bruk av et avviksparadigme tillot meg å teste ut Eides teori (2002 og senere) om finittet som et distinkt, primitivt og usammensatt trekk, uavhengig av tempus- og kongruensmarkeringer på verbet. Hvis det er vitenskapelig hold i denne teorien, forventer jeg at nordmenns reaksjoner på avvik i finitt- og tempusmorfologi skal være forskjellige. Problemstillingen er derfor som følger:

- Oppfatter nordmenn finittet og tempus som to distinkte trekk?

For å undersøke denne problemstillingen, gjennomførte jeg et eye-tracking-eksperiment, der et infrarødt kamera fanget opp øyebevegelsene til deltakere som leste setninger med forskjellige typer av avvik i verbmorfologien, for eksempel:

- a. *Ola blir hjemme og passe ungene sine selv* (finittavvik)
- b. *Ola har i tre år passe ungene sine selv* (tempusavvik)



En mer detaljert beskrivelse av stimulisetninger brukt i dette eksperimentet er presentert i seksjon 4.2.2, og den komplette listen over alle stimuliene er å finne i vedlegg A. Avvik i finitt- og tempusmorfologi forekommer regelmessig i innlærernorsk (Wik, 2014), og basert på teoriene og studiene diskutert i kapittel 2, finner jeg grunn til å tro at finittet oppfattes som et distinkt morfosyntaktisk trekk i norsk grammatikk. Dermed forventer jeg å se en forskjell mellom øyebevegelsene til nordmenn som leser setninger med avvik i finittmorfologi (a.) og de som leser setningene med avvik i tempusmorfologi (b.). I tillegg forventer jeg at finittavvik er lettere å prosessere enn tempusavvik, grunnet semantiske forskjeller mellom disse to kategoriene. Jeg legger frem en detaljert beskrivelse av hypoteser og prediksjoner i seksjon 4.1.

### Relevans

- (i) Dersom jeg kan påvise ved hjelp av eksperimentelle metoder at finittet og tempus oppfattes ulikt, sier dette noe om disse kategoriens ontologiske status og øker våre kunnskaper om norsk grammatikk;
- (ii) Dersom jeg kan føre evidens for at ulike typer av avvik oppfattes som mer og mindre plagsomme for norske morsmålsbrukere, bør dette legge føringer for fremtidig undervisning i norsk som andrespråk;
- (iii) Dersom jeg kan korroborere en språkvitenskapelig teori utelukkende basert på en teoretisk formalistisk modell ved hjelp av empiriske resultater fra eksperimentelle forsøk, er dette et eksempel på hvordan forskningsfronten bringes fremover via en kombinasjon av ulike tilnærminger.

## 1.2 Oppbygging av oppgaven

Denne oppgaven består av 6 kapitler. Kapittel 2 gir først en oversikt over tradisjonelle syn på begrepet «finittet» (seksjon 2.1), og deretter presenteres Eides finittteori (seksjon 2.2). Jeg vil så fremlegge evidens som synes å støtte Eides teori og samtidig utgjør bakgrunnen for de prediksjonene som testes i denne studien (seksjon 2.3).

I kapittel 3 vil jeg kort presentere forskjellige metoder som brukes i språkvitenskapelig forskning og grunngi mitt valg av eye-tracking som metode for dette prosjektet. Seksjon 3.2 gir en oversikt over metoden eye-tracking og hvorfor den er velegnet for språkvitenskapelige studier. Jeg vil også forklare de viktigste tekniske aspektene ved eye-tracking, diskutere hva slags data man kan få ved hjelp av eye-

tracking, samt fordeler og ulemper ved denne metoden, slik at resten av oppgaven skal være forståelig og tilgjengelig også for forskere uten erfaring med eye-tracking.

I kapittel 4 presenterer jeg først studien min i mer detalj og formulerer en klar, testbar hypotese. Seksjon 4.2 beskriver eksperiment- og stimulidesignet, mens delene 4.3 og 4.4 beskriver henholdsvis informantene som deltok i eksperimentet og hvilke utstyr som ble brukt. Seksjon 4.5 beskriver utføringen av eksperimentet.

I kapittel 5 presenterer jeg dataene som er eksperimentets resultater, hvorpå jeg presenterer statistiske analyser av disse dataene. Seksjon 5.1 beskriver hvordan dataene ble forberedt for analysen. Seksjon 5.2 gir oversikt over de avhengige variablene, altså de målingene som ble analysert i denne studien. Seksjon 5.3 gir en begrunnelse for mitt valg av analysemodell, samt videre databearbeiding som var nødvendig for bruk av den valgte modellen. Til slutt blir resultatene av analysen presentert i detalj i seksjon 5.4.

Kapittel 6 er en drøfting om funnene fra studien og hvordan disse kan fortolkes, og inneholder noen anbefalinger for videre forskning.

Hvert kapittel avsluttes med et kort sammendrag på engelsk, for å gjøre oppgaven tilgjengelig også for forskere som ikke behersker norsk. Detaljnivået i begrepsforklaringen kan for noen kanskje fremstå som overdrevent og overflødig, men fordi denne oppgaven beskriver en tverrfaglig studie i grensesnittet mellom teoretisk språkvitenskap og eksperimentell psykologi, syntes det nødvendig og definitivt mest forsvarlig å forklare begrepsapparatet såpass grundig at både psykologer og lingvister kunne forstå alle aspekter ved denne studien uten å måtte lese mye ekstramateriale.

### 1.3 Terminologi, forkortelser og symboler

Fordi oppgaven er skrevet på norsk, mens kildematerialet nesten utelukkende er skrevet på engelsk, og fagtermer i overveiende grad benyttes i sin engelske form, er det en utfordring å finne et nivå for et godt og flytende språk som samtidig beholder meningsinnholdet og presisjonsnivået i formuleringene. Å renske ut alle engelske låneord har da heller ikke vært et mål, og jeg har prioritert å presentere materialet på en forståelig måte i norsk språkdrakt. Jeg valgte å beholde engelske varianter av de veletablerte termene hvor en direkte oversettelse ville ha vært mer forvirrende enn hjelpsom, for eksempel navn på ulike blikksporingsmålinger (f.eks. *regressions out*) eller statistiske begrep (f.eks. *bootstrapping*). Ved siden av de termene som jeg likevel valgte å oversette skal det gis den originale engelske varianten i parentes, for eksempel:

kamuflasjesetninger (*fillers*), oppfattelsesrekkevidde (*perceptual span*), osv. «Eye-tracking» og «blikksporing» skal brukes synonymt.

I tillegg valgte jeg å unngå bruk av termen «mellomspråk» i denne oppgaven, ettersom den har fått en nedsettende konnotasjon i det at «mellomspråk» ikke oppfattes som et ekte, fullverdig språk. Innlærerspråk er imidlertid ekte og fullverdig, og jeg vil derfor bruke termer «innlærernorsk» og «S2-norsk» fremfor «norsk mellomspråk». Se Domínguez et al. (2019) for en diskusjon om mulige negative konnotasjoner i terminologien som brukes i andrespråksforskning.

Fordi kapittel 5 beskriver resultater av statistiske analyser, bruker jeg en del greske og latinske symboler som er svært vanlige og standardiserte innen statistikkfeltet. For at oppgaven kan være godt forståelig selv for dem som ikke bruker statistikk til vanlig, gir jeg her en oversikt over de symbolene og forkortelsene som skal brukes i denne oppgaven og forklarer hva de betyr.

<b><i>n</i></b>	Antall enheter i utvalget, fra engelsk <i>number</i> .
<b><i>M</i></b>	Gjennomsnittet, fra engelsk <i>mean</i> .
<b><i>SD</i></b>	Standardavvik, fra engelsk <i>standard deviation</i> ; angir hvor mye skårene i et utvalg varierer fra gjennomsnittet.
<b><math>H_0/H_1</math></b>	Nullhypotese/alternativ hypotese.
<b><i>p</i></b>	<i>p</i> -verdi, fra engelsk <i>probability</i> ; gjenspeiler sannsynligheten for en observasjon gitt at nullhypotesen er sann.
<b><i>df</i></b>	Frihetsgrader, fra engelsk <i>degrees of freedom</i> ; gjenspeiler i hvilken grad en modell kan variere. Frihetsgradene i en modell for forventede verdier av stokastiske variable er lik differensen mellom antall variable og antall parametre i den statistiske modellen.
<b><i>df<sub>M</sub></i></b>	Modellfrihetsgrader; tilsvarende $(k - 1)$ , der $k$ er antall grupper/betingelser.
<b><i>df<sub>W</sub></i></b>	Innengruppefrihetsgrader; tilsvarende $(n \times df_M)$ .
<b><i>df<sub>R</sub></i></b>	Residualfrihetsgrader; tilsvarende $(df_W - df_M)$ .
<b><i>t</i></b>	<i>t</i> -verdi; beregnes ved hjelp av en <i>t</i> -test og uttrykker forholdet mellom systematisk og tilfeldig variasjon: jo større verdi, desto sikrere er det at den observerte forskjellen er reell og ikke tilfeldig. Brukes for å sammenligne to grupper. Tallet som skrives i parentes etter <i>t</i> er frihetsgrader.

$F$	$F$ -verdi; uttrykker forholdet mellom systematisk og tilfeldig variasjon: jo større verdi, desto sikrere er det at effekten skyldes manipulasjonen. Brukes for å sammenligne flere grupper. Det første tallet som skrives i parentes etter $F$ er $df_M$ , mens det andre tallet er $df_R$ .
$d_{rm}$	Cohens $d$ for repeterte målinger; en effektstørrelse.
$\eta_p^2$	Partiell eta kvadrert; en effektstørrelse.

Tabell 1.1: Symboler og forkortelser

# Kapittel 2

## Finittthet

I dette kapitlet skal det presenteres den teoretiske bakgrunnen for denne oppgaven. Første underkapittel beskriver eksisterende tilnærminger til begrepet «finittthet» og tradisjonelle teorier om hva finittthet omfatter. Underkapittel 2.2. beskriver Eides (2009) teori om finittthet, som er grunnlaget for dette prosjektet. Det siste underkapitlet drøfter evidens som støtter Eides påstanden om at finittthet og tempus er to adskilte trekk, samt antyder at finittthet er et mindre semantisk fremtredende trekk enn tempus.

### 2.1 Finittthet, hva er det egentlig?

«Finittthet er ett av de begrepene som alle bruker og ingen forstår» (Klein, 2009, s. 336). Selv om denne påstanden kan virke som en overdrivelse, er det ingen tvil om at veldig få av de språkforskerne som bruker begrepet «finittthet» er enige om begrepets definisjon (Cowper, 2002). I dagens språklitteratur innebærer den vanligste måten å identifisere finittthet tilstedeværelse av morfologisk uttrykt tempus- og samsvarstrekk på verbet. Klein (2009) påstår imidlertid at finittthet ikke kan anses som bare en bøyingskategori av verb, både fordi forskere snakker om finittthet også i språk der finittthet ikke er markert på verb (f.eks. mandarinkinesisk, se Grano (2017)), og fordi noen ubøyde verb også regnes som finitte (f.eks. *ought* og *can* i engelsk, se Klein (2009)). Verbmorfologi er altså bare én av måtene å uttrykke finittthet på.

Forker (2011) kategoriserer tre forskjellige tilnærminger til begrepet «finittthet» i lingvistisk litteratur:

- (i) Finittthet som verbets morfologiske (dvs. hovedsakelig bøyings-) trekk: finitte verb er verb som har morfologisk uttrykt tempus- og samsvarsbøying. Denne tilnærmingen anses imidlertid som svært

problematisk, fordi det er umulig å identifisere universelle morfologiske kategorier som uttrykker finittthet.

- (ii) Finittthet som verbets distribusjonstrekk: denne tilnærmingen knytter finittthet til verbets posisjon i setningen, ettersom det er kun finitte verb som kan stå alene som verbal i en setning, og det kan være kun ett finitt verb i predikatet. De ikke-finitte formene krever at et annet, finitt verb inngår som del av predikatet, og predikatet kan inneholde flere ikke-finitte verb. Denne tilnærmingen har også noen svakheter: Nikolaeva (2007, s. 3) henviser til en rekke språk der antageligvis ikke-finitte verbformer fungerer som det eneste verbalet i en setning. Dette fenomenet forekommer også i norsk, i setninger som for eksempel *Jeg ikke drikke øl på en fredag? Særlig.* (Eide, 2009)
- (iii) Finittthet som en egenskap av setninger i diskurs: Givón (1990, s. 853) påstår at «finittthet er det systematiske grammatiske uttrykket som viser hvor sterkt innføydd en setning er i sin umiddelbare setningskontekst». Dette betyr at jo mindre finitt en setning er, jo mer trenger den syntaktisk og semantisk/tematisk integrasjon i konteksten.

Givón (1990) anser altså finittthet som en skala i stedet for en binær kategori (finitt vs. ikke-finitt). Dette synspunktet er det dominerende i funksjonell grammatikk, der graden av finittthet bestemmes av en kombinasjon av flere trekk (tempus, modalitet, aksjonsart, negasjon, osv.). En annen debatt gjelder hvorvidt finittthet er en rent syntaktisk eller også en semantisk kategori: Barron (2000) påstår at finittthet ikke er et morfosyntaktisk, men heller et semantisk trekk som står i forbindelse med tempus. I generativ lingvistikk regnes finittthet som et syntaktisk trekk med semantiske effekter (Platzack, 1996), mens enkelte forskere, særlig innenfor funksjonell grammatikk, analyserer finittthet fra strukturell, semantisk og pragmatisk synspunkt (Gretsch & Perdue, 2007).

For å oppsummere finnes det mange forskjellige tilnærminger til å forske på finittthet og lite enighet om hvordan finittthet kan defineres. Likevel finnes det noen påstander som mesteparten av forskere er enige om: (i) finittthet er som oftest ledsaget av morfologiske markeringer, (ii) finittthet har syntaktiske egenskaper, og (iii) finittthet påvirker tolkningen av en ytring. I denne oppgaven skal jeg spesielt ta for meg Eide (2009) sitt perspektiv på finittthet, som jeg beskriver nærmere i neste seksjon.

## 2.2 Eides teori om finitthet

I dette underkapitlet refererer jeg til Eides finitthet-teori slik den fremstilles i en artikkel fra 2009 (med noen henvisninger også til Eide (2016)). En viktig påstand fra denne teorien er at Eide anser finitthet som et grunnleggende trekk, adskilt fra tempus. Selv om mange tradisjonelle teorier beskriver finitthet-trekket som sammensatt av andre, mer «primære» trekk (som for eksempel tempus og kongruens), er Eide ikke alene om å skille finitthet fra tempus: Holmberg og Platzack (1995, s. 23) påstår også at selv om finitthet er en nødvendig forutsetning for tempus, er ikke finitthet og tempus identiske, fordi det finnes finite verb uten tempus (imperativ) og ikke-finite verb som likevel har tempus (preteritum partisipp).

En annen viktig og original påstand fremsatt av Eide (2009) er at alle verbformer (i hvert fall i germanske språk) er tempusformer, inkludert infinitiv og perfektum partisipp. Dette står i strid med det tradisjonelle synet i nordisk lingvistik: i *Norsk referansegrammatikk* hevder Faarlund et al. (1997, s. 468) at det kun er de finite verbformene, presens og preteritum, som har tempusbøying. Benevnelsene «finit» og «ikke-finit» antas altså til å bety «bestemt» eller «ikke bestemt» med hensyn til en bøyingskategori som for eksempel tempus (i skandinaviske språk) eller tall og person (i språk som f.eks. fransk og spansk). Eides originale påstand er basert på Comrie (1985) sitt perspektiv på tempus som et grammatikalisert uttrykk for temporale forhold. Ifølge Comrie er tempus en deiktisk kategori, som brukes til å forankre en situasjon eller et saksforhold: saksforholdet får sin tidsmessige bestemmelse ved å forankres relativt til en annen situasjon eller et gitt tidspunkt, vanligvis talesituasjonen eller taletidspunktet. På dette grunnlaget hevder han at perfektum partisipp i likhet med presens og preteritum har tempus, ettersom partisippet uttrykker fortid relativt til tidspunktet uttrykt av hjelpeverbet. Comrie regner imidlertid ikke infinitiv for å inneha tempus, og dette er Eide uenig i. Hun hevder at infinitiv også uttrykker tempus, nærmere bestemt «ikke-fortid». En innvending mot Eides teori er at infinitiv kan synes å uttrykke fortid i setninger som for eksempel *Han hadde sluttet å røyke* (Næs, 1972). Eide (2012) påstår imidlertid at også i denne setningen uttrykker infinitiven ikke-fortid, ettersom den relaterer seg temporalt til verbet som styrer den (akkurat som perfektum partisipp). *Røyke-* og *slutte-*handlinger opphører samtidig: røykingen er ikke slutt før sluttingen er slutt, og infinitiven kan derfor sies til å uttrykke relativ samtid.

Eide kombinerer påstanden om at alle verbformer er tempusformer med Comrie (1985) sin påstand om at de fleste europeiske språk bare har et tempusparadigme der fortid er skilt fra ikke-fortid (det finnes altså ingen spesiell form for futurum). Dette gir grunnlag for følgende paradigme for de vanlige germanske språkene:

	+ <i>finittthet</i>	- <i>finittthet</i>
+ <i>fortid</i>	preteritum	partisipp
- <i>fortid</i>	presens	infinitiv

Tabell 2.1: Eides skjema av finitt- og tempusdistinksjoner

En oppmerksom leser vil legge merke til at verken imperativ eller presens partisipp er inkludert i dette skjemaet. Grunnen til dette er at presens partisipp blir regnet som adjektiv i norsk (Faarlund et al., 1997), mens imperativ ofte regnes som en finitt verbform uten tempus, og den passer derfor ikke i skjemaet. Selv om imperativ ikke har tempus, uttrykker den likevel modalitet, og Eide (2016) foreslår at en leksikalsk rot må uttrykke enten tempus eller modalitet for å regnes som et verb. Imperativformen skal ikke diskuteres videre i denne oppgaven, ettersom denne studien er begrenset til de fire verbformene som er presentert i tabellen ovenfor.

Eide videreutvikler teorien sin ved å bygge på Comrie (1985) sin iakttagelse av at det finnes absolutte og relative tempusformer i naturlige språk: det deiktiske senteret av absolutte tempusformer er taletidspunktet, mens relative tempusformer relaterer seg temporalt til et tidspunkt i konteksten (som oftest til tempuset uttrykt av verbet som styrer dem). Eide foreslår at skillet mellom absolutte og relative tempusformer tilsvarende skillet mellom finitte og ikke-finitte former i germanske språk. Denne påstanden danner grunnlaget for sammenligning av finittthet med pronominalitet i det nominale domenet. Denne idéen er utbygd i Eide (2016), hvor hun foreslår at finitte verb i moderne skandinavisk oppfører seg som temporale pronomener, mens ikke-finitte former oppfører seg som temporale anaforer. Eide bruker som støttende evidens for dette forslaget det faktum at et ikke-finitt verb må hente sin temporale referanse fra et lokalt, strukturelt høyere verb, akkurat som en anafor i det nominale domenet må hente referanse fra en lokal, c-kommanderende antesedent. Ytterligere evidens som underbygger denne sammenligningen er at samme verbform kan brukes i både finitte og ikke-finitte kontekster i språk som mangler finittdistinksjonen (f.eks. kreolspråk),



slik som samme klitiske pronomen kan brukes som både pronomen og anafor i visse romanske språk (f.eks. europeisk portugisisk).

For å oppsummere antar jeg i denne oppgaven Eides (2009) perspektiv på finitthet og tempus, som innebærer at infinitiv og perfektum partisipp er ekte og fullverdige elementer av det grunnleggende tempussystemet. Dette innebærer at tempus og finitthet oppfattes som to distinkte trekk også fra et kognitivt perspektiv, noe som jeg forsøker å finne bevis på i denne oppgaven. I neste underkapittel skal jeg redegjøre for mulige forskjeller mellom finitt- og tempusdistinksjoner.

### 2.3 Finitt- vs. tempusdistinksjon

Alle norske verb uttrykker to syntaktiske trekk ( $[\pm\text{FINITTHET}]$  og  $[\pm\text{FORTID}]$ ) morfologisk. Dette betyr at norske verb får distinkte morfologiske markeringer som gjør det åpenbart og utvetydig hvorvidt verbformen er finitt eller ikke-finitt, og om den uttrykker fortid eller ikke-fortid, for eksempel:

	+ <i>finitthet</i>	- <i>finitthet</i>
+ <i>fortid</i>	likte	likt
- <i>fortid</i>	liker	like

Tabell 2.2: Tempusparadigme for det norske verbet *like*

Hvis en ikke-finitt verbform brukes i en kontekst der en finitt form er forventet, oppfattes dette som en feil. Slike feil er nokså vanlige i innlærernorsk: ifølge Næss (2006) og Wik (2014) viser innlærere brudd på både finittdistinksjonen (ved å blande infinitiv med presens eller perfektum partisipp med preteritum) og på tempusdistinksjonen (ved å blande infinitiv med perfektum partisipp eller presens med preteritum). I noen tilfeller bryter språkinnlærere på begge distinksjoner samtidig, for eksempel ved å blande presens med perfektum partisipp eller infinitiv med preteritum. Forskningsspørsmålet som undersøkes i denne oppgaven er hvordan morsmålstalere reagerer på slike avvik i verbmorfologi. Hvis kategoriene tempus og finitthet faktisk er adskilte, slik som Eide beskriver det, er det rimelig å forvente at det finnes en forskjell i hvordan nordmenn prosesserer avvik i disse to kategoriene.

I dette underkapitlet skal jeg fremlegge evidens som støtter både Eides teori om finitthet og tempus, og prediksjonen om at finittavvik skal oppfattes mindre støtende og derfor være lettere å prosessere enn tempusavvik. Denne prediksjonen (først nevnt

i innledningskapitlet og utvidet i seksjon 4.1) er basert på idéen at tempus semantisk er et klart mer fremtredende trekk enn finittthet. Neste to seksjoner presenterer evidens fra innlærernorsk og engelsk som synes å støtte denne idéen, og utgjør dermed den teoretiske bakgrunnen til denne oppgaven.

### 2.3.1 Finitt- og tempusavvik i norsk som andrespråk

Tilegnelse av morfologiske ferdigheter er en utfordring i både første- og andrespråkstilegnelse (Kjærbæk, 2018; Libben & Archibald, 2019). Selv om norsk ikke er like rikt på verbmorfologi som mange andre europeiske språk, produserer norskinnlærere (selv avanserte innlærere med lang oppholdstid i Norge) verbformer med morfologiske avvik, som viser brudd på finitt- eller tempusdistinksjonen, eller begge to samtidig. Her er noen autentiske eksempler<sup>1</sup> på disse avvikene i innlærernorsk:

- *Selvølgelig skal jeg **beholder**...* (finittavvik: presens i stedet for infinitiv)
- *Jeg **kjøpt** min første Norske mobiltelefon...* (finittavvik: perfektum partisipp i stedet for preteritum)
- *Når jeg skulle skrive om positive sider ved privatbilisme **kommer** jeg på at det er en veldig fleksibel måte å komme seg frem...* (tempusavvik: presens i stedet for preteritum)
- *...kan det **ble** veldig dyrt* (finitt- og tempusavvik: preteritum i stedet for infinitiv)
- *I den tiden **tenke** folk sikkert ikke...* (finitt- og tempusavvik: infinitiv i stedet for preteritum)

En grundig og detaljert studie av tempus- og finittavvik i innlærernorsk ble utført av Wik (2014): hun analyserte 80 tekster skrevet av avanserte innlærere med engelsk og tysk som morsmål; tekstene var tilfeldig plukket ut fra det norske andrespråkskorpuset *ASK*. Ut fra denne analysen konkluderte Wik med at norskinnlærere som oftest blander infinitiv med presens eller perfektum partisipp med preteritum, og at det er mer uvanlig med andre typer feil. Dette betyr at norskinnlærere oftere viser brudd på finittdistinksjonen enn på tempusdistinksjonen. Wik identifiserer tre mulige grunner for denne asymmetrien.

Den første mulige grunnen Wik nevner er tolkbarheten av syntaktiske trekk: ifølge Chomsky (1995) kan syntaktiske trekk klassifiseres som tolkbare (trekk som

---

<sup>1</sup> Eksempler kommer fra det norske andrespråkskorpuset *ASK*, sitert i Wik (2014)

bidrar til semantisk tolkning av ordet) og utolkbare (trekk som ikke bidrar til semantisk tolkning). På denne måten kan finitthet regnes som et utolkbart trekk, mens tempus regnes som et tolkbart trekk. For å illustrere denne forskjellen kan vi se på frasene *jeg spiser* vs. *jeg spiste*. På grunn av forskjellige bøyingsendelser tolker vi verbformene *spiser* og *spiste* forskjellig: i den første frasen betegner tempusaffikset samtidig, noe som medfører at spisesituasjonen kan tolkes som samtidig med talesituasjonen, mens preteritumsaffikset i den andre frasen tvinger oss til å tolke utsagnet slik at spisesituasjonen er fortidig i forhold til talesituasjonen. På den andre siden kan vi se på frasene *jeg spiste* vs. *jeg har spist* – her tolkes begge verbformene *spiste* og *spist* som fortidige. Morfologisk uttrykkelse av finitthet indikerer hvorvidt verbets tempusform er absolutt eller relativ, men bidrar ikke til semantisk tolkning av selve verbet. Wik (2014) foreslår at finitthet er vanskeligere å tilegne seg for norskinnlærere nettopp fordi finitthet er et utolkbart trekk. Denne påstanden er basert på flere andrespråksteorier som foreslår at utolkbare trekk i et fremmedspråk er vanskeligere eller rett og slett umulig å tilegne seg for språkinnlærere (se f.eks. «tolkbarhetshypotesen» av Tsimpli og Dimitrakopoulou (2007)). Jeg vil anta her at siden finitthet ikke bidrar til semantisk tolkning av verbet, kan dette føre til at finittdistinksjonen er mindre robust ikke bare hos norskinnlærere, men også hos morsmålstalere.

Den andre potensielle forklaringen Wik (2014) fremlegger for sine funn er finitt-trekkets manglende semantiske basis. Denne begrunnelsen kan anses som en mer intuitiv versjon av tolkbarhetshypotesen. Både Wik og min egen studie bygger på Eides teori om finitthet, der finitthet fremstilles som et distinkt trekk, men et rent syntaktisk trekk som mangler eget semantisk innhold. Glahn et al. (2001) foreslår at syntaktiske trekk uten semantisk innhold er vanskeligere å konseptualisere enn trekk som faktisk har semantisk innhold. Glahn og kollegaer utførte en studie der de sammenlignet trekkene *numerus* og *genus*, og fant evidens for at *numerus* er konseptuelt lettere å gripe og derfor lettere å tilegne seg for andrespråksinnlæreren: forskjell mellom «én» og «flere» er lettere å konseptualisere enn grammatisk kjønn, ettersom det ikke finnes noen logisk grunn til at det skulle hete «*et hus*» men «*en stol*». På samme måte er det sannsynligvis lettere å konseptualisere forskjellen mellom «fortid» og «ikke-fortid» enn mellom «finit» og «ikke-finit». Påstanden jeg fremsetter i denne oppgaven er at tempus er lettere å konseptualisere også for morsmålstalere, noe som fører til at dette trekket oppfattes som «viktigere» eller mer fremtredende enn finitthet.

Transfer er den tredje og siste forklaringen Wik griper til: ifølge resultatene fra Wik sin studie, gjør de engelske innlærerne flere finittfeil (52%) enn de tyske

informantene (34%). Wik påstår at grunnen til dette er at engelsk mangler finittdistinksjonen, i motsetning til tysk. Denne forklaringen har imidlertid noen svakheter: hvis bare innlærere uten finittthet i morsmålet hadde hatt større vansker med finittthet enn tempus, burde de tyske informantene ha produsert like mange finittfeil som tempusfeil, ettersom begge trekkene finnes i tysk. Derimot var resultatene annerledes: også tyske informanter produserte vesentlig flere finittfeil enn tempusfeil. Dette indikerer at asymmetrien mellom finitt- og tempusavvik i ulike innlæreres norsk ikke bare kan gå tilbake på transfer, men heller egenskaper ved selve de morfosyntaktiske trekkene, som for eksempel semantisk innhold. Hvis denne antagelsen stemmer, er det rimelig å forvente at disse ulikhetene ikke bare påvirker norskinnlærere, men også morsmålstalere og deres oppfattelse av trekkene finittthet og tempus. Derfor forventer jeg å se forskjellige reaksjoner på finitt- og tempusavvik hos nordmenn, nærmere bestemt fortere prosessering av finittavvik.

### 2.3.2 Tap av finittdistinksjonen i engelsk

I motsetning til andre germanske språk, mangler moderne engelsk finittmarkeringer på nesten alle verb.<sup>2</sup> Svake verb har bare to former<sup>3</sup> (se tabell 2.3); hvorvidt disse verbformene har en absolutt eller en relativ tolkning må bestemmes ut fra konteksten. Sterke verb følger et paradigme med tre celler, der preteritum- og partisippverbformer er morfologisk forskjellige (se tabell 2.4). Modale verb har bare finite former, altså presens og preteritum (f.eks. *can* og *could*), mens ikke-finite former *\*to can* og *\*have could* er uakseptable.

+ fortid	preteritum/partisipp: <i>loved, talked</i>
– fortid	presens/infinitiv: <i>love, talk</i>

Tabell 2.3: Tempusparadigme for svake verb i moderne engelsk

	+ finittthet	– finittthet
+ fortid	preteritum: <i>went, saw</i>	partisipp: <i>gone, seen</i>
– fortid	presens/infinitiv: <i>go, see</i>	

Tabell 2.4: Tempusparadigme for sterke verb i moderne engelsk

<sup>2</sup> Med unntak av hjelpeverbet *be*, som skal ikke diskuteres i denne oppgaven; for en detaljert beskrivelse se Eide (2009).

<sup>3</sup> S-endelsen i andre person entall antas til å være en samsvarsmarkering som spiller ingen rolle i tempussystemet.

Dette har imidlertid ikke alltid vært slik – finittdistinksjonen eksisterte i gammel- og middelengelsk (Eide, 2009), der tempusparadigmet for alle verb (tabell 2.5) hadde mer til felles med moderne skandinaviske språk. Dette fireveis-paradigmet falt sammen til et toveis-paradigme for svake verb (tabell 2.3) og et slags treveis-paradigme for sterke verb (tabell 2.4) i moderne engelsk.

	+ finitthet	– finitthet
+ fortid	preteritum: <i>sang, lufode</i>	partisipp: <i>gesungen, lufod</i>
– fortid	presens: <i>singe, lufie</i>	infinitiv: <i>singan, lufian</i>

Tabell 2.5: Tempusparadigme i gammel engelsk

Eide (2009) argumenterer for at selv om det finnes noen verbformer i moderne engelsk der finitthet er morfologisk uttrykt (fortidsformer av sterke verb), er disse formene for få og tilbyr utilstrekkelig evidens for finittdistinksjonen i moderne engelsk. Eide påstår at fortidsformer av sterke verb i moderne engelsk er kun levninger av finittdistinksjonen, i likhet med rester av genitiv i norsk (f.eks. «til fjell<sub>g</sub>» og «til fot<sub>g</sub>»). En viktig observasjon fremsatt av Eide er at forenkling av tempusparadigmet i engelsk fremdeles pågår: etterlevninger av finittdistinksjonen forsvinner også fra sterke verb i mange engelske dialekter (f.eks. Sydney bydialekt (Eisikovits, 1987), flere amerikanske dialekter (Pinker, 2000, s. 86) og en rekke britiske dialekter (Sampson, 2002)), som fører til at dagens normerte treveis-paradigme for sterke verb faller sammen til et toveis-paradigme, i tråd med det generelle systemet i moderne engelsk (Eide, 2009, s.367). I disse dialektene brukes det bare en form (enten den finitte eller den ikke-finitte verbformen) som verbets generelle fortidsform, i både absolutt og relativ tolkning, for eksempel *Here's what I would've **wrote*** eller *They **taken** my Sundays away*. Noen ganger kan en helt ny verbform erstatte både den finitte og den ikke-finitte verbformen, for eksempel *seed* som en fortidsform av verbet *to see* (Sampson, 2002).

Det er ikke utenkelig at finittdistinksjonen forsvinner fra alle hovedverb i engelsk innen relativt kort tid: eksisterende sterke verb er i ferd med å miste finittmarkeringer (som vi ser i dialektale endringer), mens nye verb som kommer inn i engelsk alltid bøyes i samsvar med det «nye» toveis-tempusparadigmet (finittdistinksjonen er altså ikke produktiv). Uansett hva som skjer med engelsk verbmorfologi i fremtiden, er det ukontroversielt at finittdistinksjonen eksisterte i

gammel- og middelengelsk og er nesten ikke-eksisterende i moderne engelsk. Dette støtter Eides teori om at finittthet er adskilt fra tempus, og kan i tillegg være et tegn på at finittdistinksjonen er mindre robust enn tempusdistinksjonen: finittdistinksjonen i engelsk begynte å tøres bort for noen hundre år siden, og de etterlevningene av denne distinksjonen som fortsatt eksisterer i dag utsettes ofte for sammenfall hos engelske språkbrukere (se over), selv om de aldri viser brudd på tempusdistinksjonen.

## 2.4 Chapter summary: Finiteness

First section of this chapter covers the existing approaches to the concept of finiteness in modern linguistic literature, of which Forker (2011) identifies three main types: finiteness as a morphological property of the verb, finiteness as a distributional property of the verb, and finiteness as a property of the clause in discourse. Other topics for disagreement include whether finiteness is a binary category or a scale, and whether it is a purely syntactic category or not. There are, however, some statements that most linguists agree on: that finiteness is usually accompanied by specific morphological markings, that it has syntactic characteristics, and that it gives rise to specific interpretation effects (Eide, 2009).

Section 2.2 presents Eide's (2002 and later) perspective on finiteness, which I will adopt in this thesis. Eide claims that finiteness is a primitive distinction that cuts across tense, mood, and agreement markings. She compares the finiteness distinction to Comrie's (1985) distinction between absolute and relative tenses, where finite verb forms correspond to absolute tenses, the deictic centre of which is the moment of speech, while non-finite verbs correspond to relative tenses, which take some other contextually-given time point as their point of reference. This comparison is the basis for Eide's four-way tense paradigm for a garden-variety Germanic language, in which each cell contains a distinct verb form, and in which the infinitive and the past participle are classified as tensed verb forms:

	+ finiteness	– finiteness
+ past	preterit <i>likte, sang</i>	participle <i>likt, sunget</i>
– past	present <i>liker, synger</i>	infinitive <i>like, synge</i>

Table 2.6: Eide's (2009) finiteness and tense distinctions

Section 2.3 presents evidence supporting the two claims investigated in this study: the idea that finiteness and tense are in fact two distinct features, and that finiteness is a less salient feature than tense. The first reason to suspect an asymmetry in the salience of the two features (section 2.3.1) is that learners of Norwegian violate the finiteness distinction considerably more often than that of tense (Wik, 2014), likely because finiteness lacks semantic content, which makes it harder to conceptualise. It is reasonable to assume that this semantic vacuity would make the feature less salient also for native speakers. The second reason (section 2.3.2) to assume that finiteness is distinct from and less salient than tense is the fact that the finiteness distinction existed in Old and Middle English but has almost entirely disappeared in Present Day English. All that remains of it are the preterit and past participle forms of irregular verbs, which, however, are not productive verb forms and are regularly confused and/or used interchangeably in many English dialects. This evidence, combined with Eide's theory, forms the theoretical background for this study.





# Kapittel 3

## Eye-tracking i språkforskning

Målet med dette kapitlet er å forsvare valg av eye-tracking som metode for dette prosjektet, samt beskrive de viktigste aspektene ved eye-tracking for lesere som ikke er godt kjent med denne metoden.

### 3.1 Metoder i språkforskning

Språk er et mangfoldig fenomen som følgelig kan undersøkes fra mange perspektiver: språkhistorie, dialektal variasjon, første- og andrespråkstilegnelse, språkets kognitive og nevrobiologiske grunnlag, språkpatologier, språkteknologi, osv. Alle disse feltene benytter seg av ulike metoder. Valg av metode kommer an på forskningsspørsmålet, forskerens ferdigheter og ikke minst finansielle ressurser. Dette betyr imidlertid ikke at det finnes kun én perfekt metode for hvert forskningsspørsmål: moderne vitenskap beveger seg bort fra monometodologi, og flere og flere forskere hevder at pluralisme (bruk av flere forskjellige metoder) er å foretrekke, ettersom alle metoder har sine fordeler og begrensninger (van Peer et al., 2012). Slik blir det også i det språkvitenskapelige feltet: språkforskere anerkjenner i stadig økende grad at lingvistikk overlapper med samfunnsfag, biologi, humaniora og informatikk, og bør derfor studeres på en tverrfaglig måte. Forskere fra forskjellige domener kan lære fra hverandre og dra nytte av nye metoder og tilnærminger til datainnsamling.

Metoder som brukes i språkforskning kan klassifiseres på flere forskjellige måter med hensyn til (i) hva slags data som blir samlet inn eller hvordan dataene analyseres (kvalitativ vs. kvantitativ dataanalyse), (ii) hvordan dataene samles inn (observasjon vs. elisitering), og (iii) hva som er fokuset i forskningen (faktiske menneskespråk vs. menneskets generelle språkevne). I dette underkapitlet skal jeg gi en oversikt over noen av de mest brukte metodene i språkvitenskap, og grunngi mitt valg av eye-tracking fremfor disse.

### Korpora

Korporaanalyse har blitt brukt mer og mer i de siste par tiårene, særlig på grunn av teknologiske fremskritt. De største fordelene ved denne metoden er at korpora ikke bare inneholder skriftlig tekst, men også tale og tegnspråk. I tillegg finnes det korpora for mange forskjellige språk, samt innlærerspråk, barnetale, oversettelseskorpora, osv. Andre fordeler er at korpora er søkbare på mange plan, at dataene allerede er tilgjengelige, og at man slik kan spare tid og penger ved å unngå datainnsamling. Dessuten fremstår dataene ikke isolert, men i kontekst. Dette mangfoldet av informasjon gjør korpora til en nyttig ressurs for mange domener: teoretisk lingvistikk, sosio- og datalingvistikk, andre- og førstespråkstilegnelse, samt historisk språkvitenskap, hvor korpora kan være den eneste tilgjengelige datakilden. Det finnes imidlertid også ulemper: man kjenner ikke alltid den språklige bakgrunnen til forfatterne av korpustekstene, det kan være umulig å skille variasjon fra skrivefeil, og mange korpora er annotert automatisk, noe som gjerne fører til feil i annotasjon. I tillegg mangler korpora «negative bevis»: det er umulig å si hvilke strukturer som *ikke* er akseptable i et gitt språk gjennom korpusanalyse.

Fordi korpora inneholder kun språklig produksjon og ikke sier noen ting om nordmenns holdninger om språklige avvik, kunne ikke korpora brukes som hovedmetode i denne oppgaven. Korpora var derimot en veldig nyttig ressurs for å identifisere hvilke språklige avvik som finnes i innlærernorsk, samt hvor frekvente forskjellige typer av avvik er.

### Introspeksjon

Introspeksjon handler om å bruke intuisjon til å bedømme hva som er grammatikalsk og ikke grammatikalsk i eget språk. Forskeren kan bruke sin egen intuisjon (noe som ofte kalles *armchair linguistics*) eller benytte seg av andres intuisjon (akseptabilitetsvurderinger). Disse har lenge vært i bruk i teoretisk (formell) lingvistikk, fordi introspeksjon er velegnet til studie av det som Chomsky (1965) kaller *competence*, altså formelle aspekter ved forskjellige menneskespråk og den mekanismen som ligger bak menneskets språkevne (i motsetning til *performance*, som er det språket mennesker faktisk produserer på hverdagsbasis), og det er nettopp dette som er fokuset i teoretisk lingvistikk. Til tross for dette er bruk av egen intuisjon mye kritisert i dag. Johannessen (2003) påstår at fra et vitenskapelig synspunkt feiler introspeksjon på nærmest alle punkter:

a) Det er umulig for andre forskere å etterprøve dataene, [...] d) forskeren går lett i den fella at han eller hun farger evalueringen av dataene ut fra hypotesen eller sitt teoretiske ståsted, e) mange, også forskere, har en feilaktig oppfatning av hvordan ens egen språkbruk egentlig er, f) mangelen på input fra andre gjør at man overser andre relevante data. (s.139)

I tillegg er utvalgsstørrelsen i introspeksjonsstudier oftest svært liten: bare én person. Andre samfunns- og atferdsvitenskaper ville aldri ha akseptert dataene fra et så lite utvalg, ettersom slike studier ikke kan produsere robuste og generaliserbare resultater.

Bruken av andres introspeksjon kan løse noen av disse problemene: ved hjelp av akseptabilitetsvurderinger kan man studere andre språk enn sitt eget morsmål; man får innsikt i andre dialekter og idiolekter enn sin egen; man får en større utvalgsstørrelse og dermed mer generaliserbare resultater (Johannessen, 2003).

Imidlertid syntes både introspeksjon og akseptabilitetsvurderinger lite egnet for denne studien: i tillegg til de ovennevnte ulempene kan forskere kun bruke introspeksjon for å undersøke morsmålet sitt, og dette var ikke tema for den foreliggende studien. Når det gjelder akseptabilitetsvurderinger, ligger problemet i det faktum at informanter kan ha normative oppfatninger som det viser seg i praksis at de ikke følger. Flere studier har funnet at mennesker faktisk har en ganske dårlig intuisjon om sitt eget språk (Gibbs, 2006; Gonzalez-Marquez, 2007), noe som gjør dataene produsert i akseptabilitetsvurderingsstudier lite objektive og upålitelige (men se Sprouse og Almeida (2013, 2017) for et motstående synspunkt). I tillegg var målet ved denne studien å få innsikt i ubevisste prosesser som underligger språkprosessering, og derfor var introspeksjon og akseptabilitetsvurderinger dårlig egnet som metode, ettersom begge to produserer data basert på lingvistisk metakunnskap.

## **Eksp<sup>er</sup>iment**

«Eksperimentelle metoder» er en veldig bred kategori som omfatter også eye-tracking. Selv om eksperimentdesign og metode for datainnsamling kan variere svært mye, innebærer alle eksperimenter at en forsker manipulerer betingelser (eller «uavhengige variabler») for å undersøke virkningene av dette, altså effekten av manipulasjonen på avhengige variabler. Eksperimenter gjennomføres i et nøye kontrollert miljø, slik at det er mulig å konkludere at den oppdagede effekten faktisk skyldes den eksperimentelle manipulasjonen og ikke noe annet (en forstyrrende variabel). Nærvær – eller fravær – av effekten tolkes opp mot hypotesen og andre eksisterende teorier om språket.

Forskere prøver å utføre eksperimenter på et utvalg av mennesker som er representativt for en større populasjon, slik at resultatene skal kunne generaliseres til å være gyldige ikke bare i laben men også i verden.

Eksperimentelle metoder brukes i språkvitenskap for å samle inn informasjon om språk uten å stille direkte spørsmål til informantene, som for eksempel «Er dette grammatisk?», «Hvordan ville du sagt...?», osv. Ved hjelp av eksperimenter er det mulig å samle inn data om ubevisste og automatiske prosesser knyttet til språket. Eksperimenter er også den beste måten å påvise kausalitet: observasjonsmetoder kan *beskrive* språklige fenomener og deres korrelasjon, men de kan ikke *forklare* disse, altså identifisere hva som forårsaker et gitt fenomen. Faarlund (1987) påstår at «å forklare noe» betyr enten a) å føre det inn under en generalisering, eller b) å relatere det til noe annet og utenforliggende. Han hevder også at den første varianten ikke er særlig interessant som forklaring, ettersom det egentlig er bare en beskrivelse av en regularitet. Forklaring må relatere noe til noe annet, det vil si at *explanandum* og *explanans* må tilhøre ulike domener. Som følge av dette mener jeg at for å virkelig forklare språkets formelle egenskaper, må man søke *explanans* i et annet domene, for eksempel i biologi eller psykologi.

Det er nettopp psyko- og nevrolingvistikk som er kjent for bruk av eksperimentelle metoder. De mest brukte psykolingvistiske metodene inkluderer *lexical decision task* (der deltakere får se ett ord om gangen på en dataskjerm og må bestemme seg om ordet er ekte eller oppdiktet), *self-paced reading* (der deltakere leser en tekst ett ord om gangen og må trykke på en tast for å få se neste ord) og eye-tracking (der deltakere leser hele setninger eller tekstavsnitt mens et kamera sporer øyebevegelsene deres). I nevrolingvistikk brukes det flere måter å måle hjerneaktivitet på, for eksempel elektroencefalografi (EEG), magnetencefalografi (MEG), funksjonell magnetresonanstomografi (fMRI) og funksjonell nær-infrarød spektroskopi (fNIRS). Disse kan vise hvordan hjernen reagerer på forskjellige språklige stimuli, og hvilke deler av hjernen som aktiveres under forskjellige språklige hendelser, ved å måle hjernens elektriske aktivitet (EEG), magnetiske felter produsert av den nevronale aktiviteten (MEG) eller nivå av oksygen i blodet (fMRI og fNIRS). Disse metodene er derimot ganske kostbare, krever mye teknisk kunnskap for å forberede og gjennomføre eksperimentet, samt teoretisk kunnskap for å analysere dataene og tolke resultatene.

### Valg av metode i denne studien

Jeg valgte å bruke den hypotetisk-deduktive modellen i min studie, som innebærer følgende steg (Godfrey-Smith, 2021, s. 236):

- (i) Innsamling av data via observasjon
- (ii) Formulering av en hypotese basert på observasjoner
- (iii) Formulering av konkrete (testbare) prediksjoner baserte på hypotesen
- (iv) Testing av prediksjoner
- (v) Forkasting eller understøttelse av hypotesen

Det første steget var ikke utført av meg, men av Eide (2009), som observerte og analyserte formelle egenskaper av norsk og engelsk (og andre germanske språk), samt historiske endringer av engelsk, og Wik (2014), som analyserte korpusdataene av avvik i norsk innlærerspråk. Eides hypotese om finitthet og tempus som to distinkte trekk utgjør en del av det andre steget; hypotesen i denne studien ble utvidet for å inkludere sammenligning av finitt- og tempusdistinksjoner. Målet med denne studien var å teste Eides hypotese fra et kognitivt perspektiv, derfor handlet prediksjonene (steg tre) om språkprosessering, nærmere bestemt om at finitthet og tempus oppfattes som to distinkte trekk, og at finittavvik er lettere å prosessere enn tempusavvik (prediksjoner er beskrevet i detalj i seksjon 4.1). Steg fire og fem var målet ved dette eksperimentet og blir presentert i følgende kapitler.

Eye-tracking var ikke den eneste eksperimentelle metoden som passet denne studien: disse prediksjonene kunne også ha blitt testet ved hjelp av for eksempel *self-paced reading* (SPR) eller EEG. Imidlertid tilbyr SPR ikke en like naturlig innsikt i språkprosessering-under-lesing som eye-tracking, ettersom deltakerne får se bare ett ord om gangen, og de må trykke på en tast etter hvert ord. Fordi deltakerne ikke kan vende tilbake i setningen, er det ikke mulig å få innsikt i senere språkprosessering, altså de kompenseringprosessene som foregår når personen prøver å forstå hva meningen bak det språklige avviket egentlig var. I tillegg var jeg interessert i å se hvordan språkprosessering utfolder seg i sanntid, mens SPR-dataene gjenspeiler kun resultatet av tidlig syntaktisk analyse (lesetiden for hvert segment). EEG er en fantastisk metode for å få nevrologisk data, altså informasjon om hva som foregår i selve hjernen og ikke bare om effekter av disse på atferden. EEG er derimot mindre likt en naturlig situasjon enn eye-tracking, fordi deltakerne får lese bare ett ord om gangen, de kan ikke kontrollere hvor lenge de ser på hvert ord, og de får instruksjoner om å ikke blunke eller bevege øynene mens de leser. I tillegg innebærer EEG-eksperimenter en mer

komplisert prosedyre enn eye-tracking og er vesentlig mer tidkrevende, noe som gjør det vanskeligere å rekruttere deltakere og fører til et mindre utvalg og mindre generaliserbare resultater.

Med tanke på alle de ovennevnte aspektene, bestemte jeg meg for å bruke eye-tracking i denne studien. Andre fordeler med denne metoden og teknologiske aspekter av eye-tracking er beskrevet i mer detalj i neste underkapittel.

### 3.2 Eye-tracking

Eye-tracking, eller «blikksporing», er en teknologi som måler øyebevegelser og brukes i mange forskjellige domener. Eye-tracking er ikke en ny oppfinnelse – forskere begynte å analysere øyets bevegelser under lesing allerede på slutten av 1800-tallet. En fransk optiker Émile Javal (1839-1907) studerte øyebevegelser ved direkte observasjon av øyet og var den første til å fremme en teori om at øyet ikke beveger seg i en kontinuerlig og jevn bevegelse, men helst i en serie av raske bevegelser og pauser. Teorien ble underbygget av den amerikanske psykologen Edmund B. Huey (1870-1913), ved hjelp av det vi kan kalle den første eyetrackeren: en slags kontaktlinse i plast som ble festet på leserens øye. En tynn metallstang var koblet til linsen og «tegnet» øyebevegelsene med sot på en papirrull. Huey var interessert i den «øye-sinn»-forbindelsen som ligger bak lesing og trodde at lesing var en komplisert kognitiv prosess: et gjennomspill mellom øyets bevegelser, hjernens fysiologi og menneskets sinn (Reed & Meyer, 2007). Senere kom mindre inngripende metoder for blikksporing, som gjerne involverte lysrefleksjonen på hornhinnen som ble markert på en fotografisk plate eller film (se f.eks. Dodge og Cline (1901) og Judd et al. (1905)).

Rayner (1998) karakteriserer tre bølger av eye-tracking-forskning: tidlig forskning som fokuserte på grunnleggende fakta om øyebevegelser (1879-1920); behaviorisme-epoke (ca. 1930-1950) der fokuset lå på anvendt forskning; og moderne eye-tracking-forskning (1970-i dag). De store fremskrittene i eye-tracking-feltet som er gjort siden 1970-tallet skyldes både fremskrittet i psykologi, noe som forårsaket en økt interesse for bedre metodologi i kognitiv forskning, og det teknologiske fremskrittet som ga oss moderne datamaskiner (Płużyczka, 2018). Fokuset i eye-tracking-feltet skiftet over til dataprosessering, nøyaktighet og presisjon av eyetrackerne, og metoder for å redusere belastningen på deltakernes øyne (Jacob & Karn, 2003). Bedre og billigere utstyr betyr at eye-tracking kan brukes i mange universiteter og laboratorier verden over, og at eksperimenter kan gjennomføres i mer naturlige omgivelser, der

deltakerne kan oppføre seg mer som vanlig (f.eks. bevege hodet sitt og av og til også kroppen så mye de vil). Dette fører til at resultatene av eye-tracking-studier har høyere økologisk validitet, ettersom de er mindre påvirket av de kunstige lab-omstendighetene. Bruk av datamaskiner gjør at selv de minste og raskeste øyebevegelsene blir tatt opp med høy nøyaktighet og synkronisert med bilde fra PC-skjermen. Dette gjør det mulig å bruke eye-tracking i dynamiske scenarioer, der deltakere for eksempel ser på en video eller navigerer en nettside. I tillegg til psykologi er det mange andre disipliner som benytter seg av eye-tracking, for eksempel medisin, nevrovitenskap, matematikk og IKT, brukervennlighet, undervisning, idrett, sosiologi, kunst og ikke minst språkvitenskap (Carter & Luke, 2020).

### 3.2.1 Teknologiske aspekter

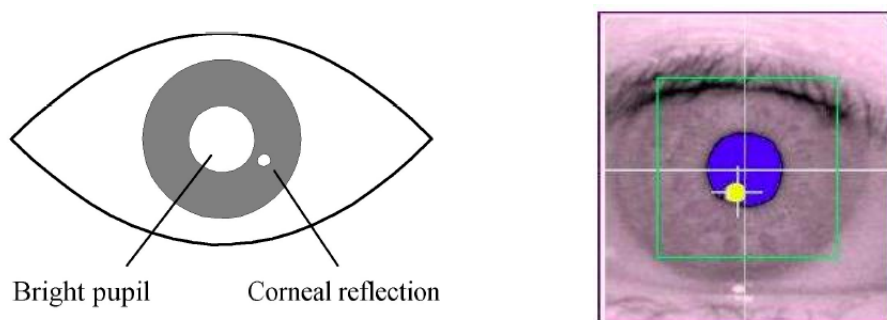
Det finnes mange forskjellige typer eyetrackere, og de kan fordeles i tre kategorier: eyetrackere med og uten hodestabilisering som plasseres ved eller er integrert i dataskjermen, og eyetrackere som festes på hodet.

Den første typen genererer de mest presise og nøyaktige dataene, både på grunn av hodestabilisering og fordi disse eyetrackerne pleier å ha den høyeste samplingsfrekvensen, samt størst nøyaktighet (*accuracy*) og presisjon (*precision*). Samplingsfrekvensen er antallet målinger tatt per sekund og måles i hertz (Hz). Dette betyr at en eyetracker med samplingsfrekvens på 300Hz tar opp øyebevegelser 300 ganger per sekund. En høyere samplingsfrekvens gir både høyere datakvalitet og en mulighet til å samle inn data fra små interessefelter (*areas of interest*, se definisjon nedenfor) på skjermen. Samplingsfrekvensen i moderne eyetrackere varierer fra 30Hz til 2000Hz. Nøyaktighet og presisjon er også avgjørende for datakvaliteten: «nøyaktighet» henviser til forskjellen mellom den faktiske og den målte blikkposisjonen (måles i grader av synsvinkel), mens «presisjon» handler om målingenes nærhet til hverandre, altså om hvorvidt gjentatte målinger av en og samme ting ligner på hverandre.

Med den andre typen eyetrackere er deltakernes hode ikke stabilisert, men det betyr ikke nødvendigvis at de kan bevege seg så mye de vil uten at eyetrackeren «mister» blikkposisjonen deres. Den største fordelen ved denne typen eyetrackere er at de kan brukes med deltakere som er følsomme for kroppskontakt eller som ikke kan sitte rolig, for eksempel barn eller personer med autisme. Generelt sett tilbyr denne typen eye-tracking også høyere økologisk validitet, fordi eksperimentene skjer i mer naturlige omgivelser.

Den siste typen eyetrackere festes på deltakernes hode og er som regel en slags hatt eller briller. Med disse kan deltakere bevege seg fritt, men dataanalyse blir mye vanskeligere ettersom forskere ikke kan kontrollere hva deltaker ser på. I tillegg har disse eyetrackere lav samplingsfrekvens og lav nøyaktighet, som uvegerlig produserer data av lavere kvalitet (Conklin et al., 2018).

Mesteparten av moderne eyetrackere bruker teknologien som kalles «pupillsenter/hornhinnerefleksjon» (Goldberg & Wichansky, 2003). Denne teknologien fungerer ved at usynlig infrarødt<sup>4</sup> lys rettes mot øyet, hvor lyset reflekteres og fanges opp igjen av et kamera, som avbilder pupillen som en lys disk og hornhinnerefleksjonen som en liten skarp prikk (se figur 3.1). Vektoren mellom pupillsenteret og hornhinnerefleksjonen brukes til å regne ut øyeeplets bevegelser. Ved å bruke to referansepunkter på øyeeplet kan eye-tracking-utstyret adskille øyebevegelser fra hodebevegelser, som gjør dataene enda mer presise og ikke behøver at deltakeren sitter helt rolig.



Figur 3.1: Hornhinnerefleksjonen og den lyse pupillen som de vises av et infrarødt kamera. Venstre: illustrasjon fra Poole og Ball (2004); høyre: bilde fra EyeLink 1000 bruksanvisning.

Før eksperimentet begynner må kameraet fokuseres på deltakerens øye og illuminasjonsnivå må justeres slik at både pupillen og hornhinnerefleksjonen kan identifiseres av eyetrackeren uansett hvilken del av dataskjermen deltakeren ser på. Det er viktig at verken skjermen eller belysningen i laben er for lys, ettersom begge to kan føre til at pupillen blir for liten, noe som gjør det vanskeligere for eyetrackeren å identifisere pupillen. Etterpå må kameraet kalibreres: dette gjøres ved at deltakeren fokuserer på enkelte punkter av skjermen (som oftest 9 eller 13 punkter markerte med svarte prikker), og eyetrackeren identifiserer og «husker» hvordan deltakerens øye ser

---

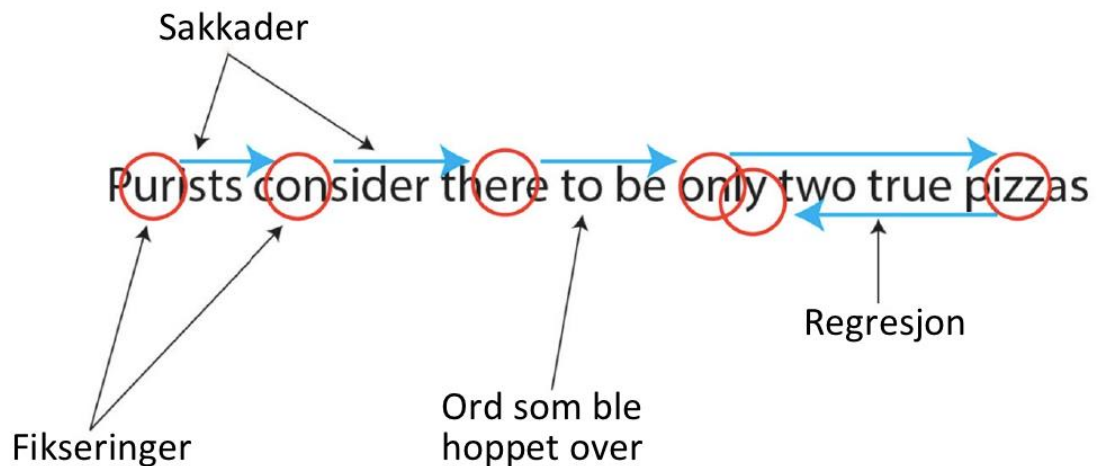
<sup>4</sup> Ved å bruke infrarødt lys i stedet for synlig lys, kan man skape refleksjoner på hornhinne uten å blende deltakere.



ut når blikket er rettet mot disse punktene. God kalibrering er veldig viktig for å få nøyaktige og presise data, derfor er det fornuftig å gjenta kalibrering noen ganger under eksperimentet, særlig om deltakeren beveger seg mye (Holmqvist et al., 2011).

Selv om informanter kan se hvor som helst på skjermen under eksperimentet, er det bare noen få områder på skjermen som interesserer forskere, for eksempel enkelte ord, fraser eller deler av et bilde. Disse kalles «interessefelter», og det er dataene fra disse feltene som blir analysert i psykolingvistiske studier. Opptak av øyebevegelser blir lagret av eye-tracking-programvare som forskjellige typer av dataenheter:

- **Fikseringer** er den mest brukte dataenheten og gjengir stans i øyebevegelser. Disse kan vare fra 10-20 millisekunder og opp til flere sekunder, men en gjennomsnittlig fiksering i lesing er ca. 250ms. Imidlertid er blikket ikke helt stillestående under fikseringer: øyet beveger seg sakte bort fra senteret av fikseringen (dette kalles *ocular drift* eller okulær vandring), og små raske bevegelser (mikrosakkader) retter øyet tilbake til dets utgangsposisjon.
- **Sakkader** er raske bevegelser fra en fiksering til en annen. Sakkader er den raskeste bevegelsen menneskekropp kan gjøre: de varer 30-80ms, og hastigheten av en sakkade kan variere fra 30 til 500 grader av synsvinkel per sekund. Når det gjelder lesing, kalles sakkader som vender tilbake i teksten «regresjoner».
- **Smooth pursuit** er en saktere bevegelse som skjer når personen følger en gjenstand i bevegelse med blikket. *Smooth pursuit* er veldig forskjellig fra sakkader: den behøver en gjenstand som blikket kan følge, og den kontrolleres av en forskjellig del av hjernen.
- **Blunkefrekvens** og **pupillstørrelse** er andre typer av målinger som kan samles inn av eyetrackeren, men disse brukes mest i psykologiske studier. Både størrelsen på pupillen og blunkefrekvens er påvirket av visse kognitive prosesser, som kan derfor undersøkes gjennom analyse av disse dataenhetene.



Figur 3.5: Dataenheter i eye-tracking (adaptert fra Carter og Luke (2020)).

### 3.2.2 Visuell språkprosessering

Synet er en av de viktigste sansene for mennesker: nesten 80% av all sanseinformasjon når hjerne via den visuelle kanalen (Gollücke, 2009). Språkprosessering er noe vi bruker synet til hver eneste dag i lesing, tegnspråk, og andre visuelle modaliteter av språket. Derfor er det ikke overraskende at øyebevegelser kan gi innsikt i hvordan mennesker prosesserer språket. Grunnlaget for bruk av eye-tracking i språk- og særlig leseforskning er antagelsen at det finnes en forbindelse mellom øyebevegelser og mentale prosesser, den såkalte «øye-sinn»-forbindelsen. Denne antagelsen tilsier at øynene hviler på et ord så lenge som det trengs for å prosessere dette ordet (Just & Carpenter, 1980). For å forstå hvordan eye-tracking kan best benyttes i språkforskning, er det nødvendig å forstå hvordan hjernen vår prosesserer språket mens vi leser.

Det finnes en relativt alminnelig enighet om at visuell språkprosessering skjer i flere faser. Under den første fasen blir den visuelle inputen overført til en slags «buffer». Den andre fasen er analyse av inputen i arbeidsminnet,<sup>5</sup> mens den tredje og siste fasen er integrasjonen av den analyserte inputen, sammen med lingvistisk og kognitiv fortolkning av teksten. Som nevnt tidligere består lesing av en rekke korte fikseringer og raske bevegelser, eller «sakkader». Selv om fikseringer varer i gjennomsnitt bare rundt 250ms, er dette faktisk lengre enn nødvendig for å gjenkjenne et ord (ord blir identifisert etter ca. 50ms). Grunnen til at mennesker må fikser på

---

<sup>5</sup> Arbeidsminnet (også kalt korttidsminne) er den umiddelbare hukommelsen, som har som hovedoppgave å bearbeide ny informasjon. Arbeidsminnet er ganske begrenset og kan bare jobbe med noen få ting av gangen (Lezak, 1982).

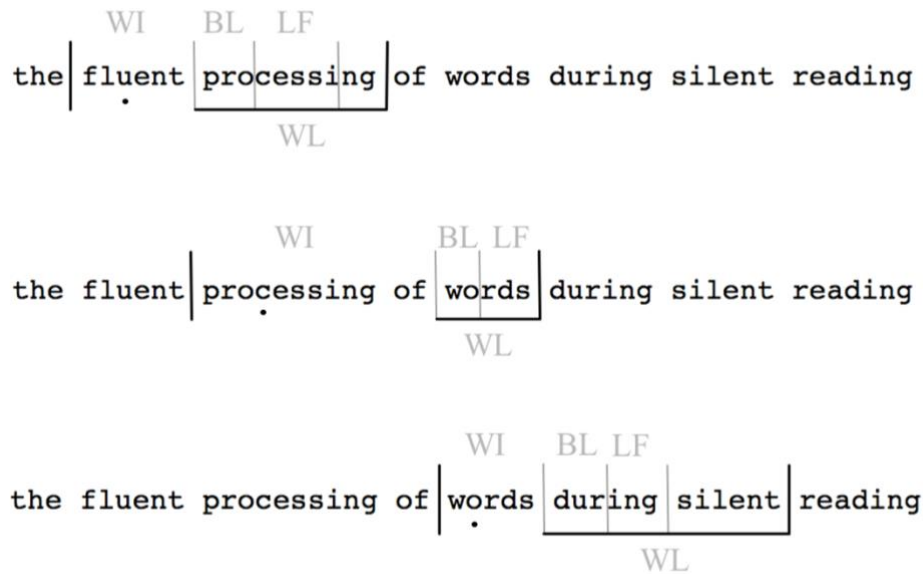
hvert ord i så lang tid er fordi inputen må overføres fra den visuelle bufferen til arbeidsminnet. Hvis ny input blir lagt inn i bufferen før overføringen av tidligere mottatt informasjon til arbeidsminnet er ferdig, kan den nye inputen blandes sammen med den tidligere inputen, eller overskrive den (Warren, 2012, s. 110).

Øynene våre beveger seg ikke mesteparten av tiden mens vi leser: typiske fikseringer varer rundt 250ms, mens sakkader i lesing er veldig raske, rundt 30ms (Carter & Luke, 2020). Hver sakkade beveger blikket vårt ca. 8-9 bokstaver frem,<sup>6</sup> men menneskets oppfattelsesrekkevidde (*perceptual span*) er imidlertid nesten to ganger større enn dette. Oppfattelsesrekkevidde er også asymmetrisk: mennesker oppfatter mer informasjon fra en side av fikseringen enn fra den andre. Dette betyr at i språk som norsk (der man leser fra venstre til høyre) får man se 3-4 bokstaver til venstre fra den aktuelle fikseringen og opp til 15 bokstaver til høyre fra fikseringen, mens i språk som for eksempel hebraisk (der man leser fra høyre til venstre) ser man mer til venstre fra fikseringen enn til høyre (Warren, 2012). I tillegg er forskjellige typer av informasjon tilgjengelige fra ulike deler av oppfattelsesrekkevidden (se figur 3.3): mennesker får informasjon som kan brukes til ordidentifisering (WI, *word identification*) fra området nærmest fikseringen (innen 4-8 bokstaver); lengre frem er det fortsatt mulig å få informasjon om første bokstaver (BL, *beginning letters*) av neste ord; etter det er det kun informasjon om bokstavstrekk (LF, *letter features*) som er tilgjengelig.<sup>7</sup> Informasjon som er tilgjengelig over den største rekkevidden er informasjonen om ordlengde (WL, *word length*). Dette betyr at selv om man kan se ca. 15 bokstaver frem fra fikseringen, ser man alt ikke like klart. Derfor beveger blikket seg bare 8-9 bokstaver av gangen, som regel til det neste uidentifiserte ordet (Rayner & Balota, 1989).

---

<sup>6</sup> Dette gjelder de språkene som bruker alfabetiske skriftsystem.

<sup>7</sup> *Letter features* er bokstavstrekk som kan hjelpe å identifisere en bokstav. For eksempel hvis man oppfatter at en bokstav inneholder en loddrett linje, blir alle bokstavene som inneholder loddrett linjer aktivert i minnet (f.eks. «t», «h», «b»), mens bokstavene som ikke inneholder et slikt trekk blir deaktivert (f.eks. «o», «x», «g»).



Figur 3.6: Øyebevegelser under lesing (adaptert fra Rayner & Balota, 1989, s.268)

Tegnforklaring: • = fiksering; WI = ordidentifisering; BL = førstebokstaver;

LF = bokstavstrekk; WL = ordlengde

Det faktum at mennesker kan se videre frem fra fikseringen (fenomenet som på engelsk kalles *preview effect*) og få informasjon fra dette området er veldig viktig i lesing. For eksempel gir informasjonen om ordlengde mulighet til å skille funksjonsord fra innholdsord (de førstnevnte er som regel ganske korte) og regne ut hvilket ord neste fiksering skal ligge på. Derfor blir mange funksjonsord hoppet over i lesing – de blir prosessert som en del av en tidligere fiksering. Dette er også grunnen til at forskere må passe på at de kritiske ordene i eye-tracking-eksperimenter er lange nok (5+ bokstaver), for å være sikker at disse blir fiksert av blikket. Identifisering av det følgende ordets første bokstaver, samt noen bokstavstrekk av de andre bokstavene hjelper til å utløse ordsøket i det mentale leksikonet før selve ordet blir fiksert (Warren, 2012). Det finnes mange faktorer som kan påvirke hvor langt frem blikket skal bevege seg: tekstens kompleksitet, forutsigbarhet av det neste ordet (med hensyn til kontekst), det neste ordets frekvens og potensielle tvetydighet, og hvorvidt leseren har fått «priming»<sup>8</sup> for det følgende ordet fra tidligere deler av teksten.

---

<sup>8</sup> Semantisk priming eller semantisk forhåndspåvirkning er i språkvitenskap det fenomenet at tidligere leste eller hørte ord «aktiverer» visse deler av det mentale leksikonet, slik at semantisk relaterte ord eller uttrykk er lettere tilgjengelig. For eksempel vil ordet «hund» bli fortere identifisert hvis man først leser/hører ordet «katt» enn hvis man først leser/hører ordet «skip» (Heyman et al.,

Grunnen til at mennesker kan lese bare det de fokuserer på ligger i øyets biologi. Øyet ligner på mange måter på et kamera: det har en åpning (pupillen), en linse og et lysfølsomt område som kalles «fovea». Selv om synsfeltet for hvert øye strekker seg til nesten 90 grader på utsiden av fikseringspunktet, er synsstyrken størst i fovea, som av og til kalles også skarpsynsområdet. Synsstyrken faller raskt fra fikseringspunktet utover mot de ytre delene av synsfeltet (Smythies, 1996). Selve fovea er bare 1,5mm stor og dets mest følsomme område er enda mindre, rundt 250–300  $\mu\text{m}$  i diameter. Dette tilsvarer bare  $1^{\circ}20'$  av synsvinkel, og er det eneste området der mennesker ser skarpt nok for å kunne lese. Rundt fovea ligger området som kalles parafovea, hvor synet er svakere, men noen typer informasjon kan fortsatt bli innhentet (rundt  $5^{\circ}$  av synsvinkel). Utenfor parafovea ligger den perifere delen av netthinnen, som bare kan oppfatte visuell informasjon av lav frekvens (Carter & Luke, 2020).

### 3.2.3 Hva kan eye-tracking avsløre om språk?

Denne seksjonen gjør rede for hvordan eye-tracking kan brukes i språkforskning, og diskuterer hva slags språkdata det er mulig å få ved hjelp av blikksporingseksperimenter, samt hva som er viktigst å ta i betraktning under eksperimentplanlegging for å få valide data.

Hittil har jeg for det meste diskutert bruk av eye-tracking i lesestudier, men eye-tracking kan også brukes for å undersøke taleprosessering. I dette tilfellet bruker forskere både visuelle- og lydstimuli: deltakere får høre tale (enkelte ord, lengre fraser eller hele historier), mens de ser på bilder på en dataskjerm. Denne typen eye-tracking kalles *Visual World paradigm*, som er basert på to viktige funn av Cooper (1974): (i) visuelle stimuli aktiverer tilsvarende konsepter i menneskets hjerne, og (ii) blikket beveger seg til det visuelle objektet som tilsvarende det personen hører (språklig input), samt til andre objekter som er fonologisk eller semantisk relatert til inputen. *VW*-paradigmet gir mulighet for å undersøke en rekke fenomener, som for eksempel ordidentifisering, prediksjon og referanseprosessering. *VW*-paradigmet sammen med leseeksperimenter utgjør mesteparten av psykolingvistiske eye-tracking-studier. Likevel vil jeg ikke diskutere *VW*-paradigmet videre, fordi fokuset i denne studien ligger på prosessering av skriftlig språk.

---

2014). Priming kan også være fonologisk eller ortografisk, slik at ordene som høres eller staves på en lignende måte også blir aktivert (Frisson et al., 2014).

Prosessering av skriftlig språk undersøkes gjennom leseeksperimenter, der man bruker enkelte ord, setninger eller lengre tekstavsnitt som stimuli, og deltakere får som oppgave å lese teksten på en vanlig måte. Det er også mulig å inkludere andre typer oppgaver, som for eksempel grammatikalitetsvurderinger eller svar på spørsmål om setningenes innhold, men det er som regel ønskelig å få de mest naturlige dataene fra eksperimentet, og det får man best ved å be deltakere om å lese teksten på en helt naturlig måte, uten å fokusere på grammatikken eller andre egenskaper av teksten (Godfroid, 2020). Mange forskjellige aspekter av språk kan undersøkes ved hjelp av leseeksperimenter, for eksempel:

- Prosessering av grammatiske avvik: stimuli i slike studier er setninger som inneholder et syntaktisk, semantisk, eller en annen type grammatisk avvik (se f.eks. Sagarra og Ellis (2013), Clahsen et al. (2013))
- Disambiguering (*ambiguity resolution*): stimuli i slike studier er setninger som inneholder en syntaktisk eller en semantisk tvetydighet, og målet er å undersøke når og hvordan leseren bestemmer seg for hva som er den riktige tolkningen av setningen (se f.eks. Chamorro et al. (2016); Dussias og Sagarra (2007)).
- Bruk av undertekster i filmer: denne typen studier undersøker multimodal språkprosessering, altså hvordan mennesker integrerer samtidig input fra flere kilder (video, skriftlig språk og tale) i en sammenhengende og multimodal representasjon av verden. Mulige forskningstemaer i slike studier er hvordan multimodal input påvirker lesing, kognitiv belastning og læring av fremmedspråk (Godfroid, 2020).

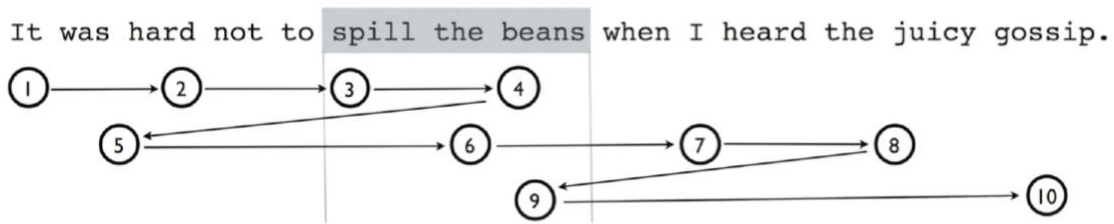
### **Hvilke målinger bør inkluderes i analysen?**

Valg av riktige målinger er avgjørende for en vellykket eye-tracking-studie, men det er også vanskeligere enn det fremstår, ettersom det finnes hundrevis av forskjellige målinger som kan analyseres. Det er ikke nødvendig å bestemme på forhånd hvilke målinger som skal brukes i analysen, fordi alle dataene blir samlet inn av eyetrackeren, og hvilke som helst målinger kan eksporteres for analyse etter datainnsamlingen. Det er imidlertid nødvendig å velge passende målinger for hvert gitt forskningsspørsmål: man risikerer å gå glipp av effekten hvis man ser på de gale målingene, men samtidig er det ikke fornuftig å analysere for mange målinger, fordi da vil man mest sannsynlig finne en eller annen tilfeldig effekt i en av disse målingene, og man risikerer å feilaktig tilskrive effekten til den eksperimentelle manipulasjonen (Conklin et al., 2018).

Fordi det viktigste antagelse i eye-tracking er den «øye-sinn»-hypotesen (Just & Carpenter, 1980), som antyder at varigheten på fikseringen gjenspeiler innsatsen som kreves for å prosessere stimulusen, er det vanlig for forskere å inkludere minst én varighetsmåling i analysen (Godfroid, 2020, s. 88). I lesestudier fokuserer forskere på fikseringer også fordi disse er mer følsomme overfor lingvistiske faktorer enn sakkader (Staub & Rayner, 2007). Én type sakkader spiller derimot en viktig rolle i lesestudier: regresjoner, altså sakkader i motsatt retning av teksten, gjenspeiler vansker i integrasjon av ordet i konteksten.

Fikseringsmålinger er som regel fordelt i «tidlige» og «sene» målinger, som gjenspeiler henholdsvis tidlige og sene faser av visuell språkprosessering. Tidlige målinger kan sies å gjenspeile automatiske prosesser for ordidentifisering og leksikalsk tilgang (*word recognition and lexical access processes*), mens sene målinger gjenspeiler som regel mer bevisste, kontrollerte og strategiske prosesser (Inhoff & Rayner, 1986). Her er en liste over noen av de mest brukte tidlige målingene i lesestudier (tall i klammeparenteser indikerer hvilke fikseringer hver måling består av, illustrert i figur 3.4):

- *First fixation duration* [3] er varigheten på første fiksering gjort på et ord eller et interessefelt. Rapporteres i millisekunder.
- *Gaze duration* eller *first pass reading time* [3, 4] er varigheten på alle fikseringer som er gjort på et interessefelt under førstegangslesing (før blikket forlater interessefeltet til høyre). Denne målingen kan brukes som en tidlig måling i stedet for *first fixation duration* for lengre ord eller ordsekvenser, som skal mest sannsynlig fikseres flere ganger. Rapporteres i millisekunder.
- *Skipping rate* brukes for å beskrive andel ord som blir hoppet over, altså ikke fiksert under førstegangslesing. Ord kan bli hoppet over på grunn av både visuelle og lingvistiske faktorer, og de blir ofte prosessert i parafovea mens blikket er fiksert på et tidligere ord. *Skipping rate*-målingen rapporteres som sannsynlighet eller prosent.



Figur 3.4: Eksempel på et lesemønster (Conklin & Pellicer-Sánchez, 2016).

Tegnforklaring: Interessefeltet er markert i grått. Et illustrativt lesemønster er fremstilt under setningen. Fikseringer er illustrert med sirkler der tall indikerer fikseringenes rekkefølge.

Mest brukte sene målinger er de følgende:

- *Total reading time* eller *dwell time* [3, 4, 6, 9] er varigheten på alle fikseringer gjort på et interessefelt. Denne målingen påvirkes av både tidlig ordidentifiseringsprosesser og senere ordintegrasjon. Rapporteres i millisekunder.
- *Fixation count* [3, 4, 6, 9] er totalt antall fikseringer gjort på et interessefelt. Denne målingen er en alternativ måte å vurdere hvor mye oppmerksomhet et gitt interessefelt har fått gjennom hele forsøket. Rapporteres som tall.

Regresjonsmålinger er vanskelig å karakterisere som tidlige eller sene, fordi de kan gjenspeile integreringsvansker som oppstår når man først fikserer på et ord, samt den tiden som trenges for å kompensere for disse vanskene (Clifton Jr et al., 2007). De to vanligste regresjonsmålingene er:

- *Regression path duration* eller *go past time* [3, 4, 5, 6] er en oppsummert varighet på alle fikseringer gjort etter at blikket landet i interessefeltet og før blikket forlot interessefeltet til høyre (altså fikseringer gjort på selve interessefeltet og under regresjonen fra det). Rapporteres i millisekunder.
- *Regressions out* beskriver sannsynlighet for at blikket vender tilbake i teksten fra det aktuelle interessefeltet. Rapporteres som prosentandel, som indikerer i hvor mange forsøk blikket forlot dette interessefeltet og søkte til venstre.

De fleste eye-tracking-studiene bruker en kombinasjon av tidlige og sene målinger, og ofte målinger fra flere interessefelter.



### Viktige hensyn ved planlegging av et eye-tracking eksperiment

Minutiøs planlegging av eksperimentet er avgjørende for en vellykket studie. Som nevnt tidligere i seksjon 3.1 er det viktig å kontrollere alle mulige forstyrrende variabler som kan påvirke resultatene: i verste fall risikerer forskeren å trekke feilaktige konklusjoner ved å påstå at den oppdagede effekten skyldes den eksperimentelle manipulasjonen, mens i realiteten var det en ekstern faktor som forårsaket effekten. Selv hvis effekten faktisk skyldes den eksperimentelle manipulasjonen, kan det være vanskelig å oppdage den når det er mye tilfeldig variasjon i dataene, som ofte er forårsaket av forstyrrende variabler. De vanligste forstyrrende variablene i eye-tracking-eksperimenter – og hvordan disse kan kontrolleres – er beskrevet i denne seksjonen.

Først og fremst er det viktig at stimulisetningene ligner på hverandre så mye som mulig: hvis den eneste forskjellen mellom setningene er det kritiske ordet, og blikkmønster er ulike i disse setningene, blir det lettere å konkludere med at forskjellen i øyebevegelser er forårsaket av den eksperimentelle manipulasjonen. Ideelt sett må setningene ligne på hverandre med hensyn til syntaktisk struktur, forutsigbarhet og lengde. Interessefelter (som inneholder de kritiske ordene) må forekomme i samme posisjon i alle setninger, fordi lesehastigheten blir redusert gjennom teksten (Conklin & Pellicer-Sánchez, 2016). Det er også vanlig i psykolingvistiske studier å unngå plassering av kritiske ord på slutten av setningen på grunn av det som kalles avslutningseffekter (*wrap-up effects*). Grunnen til dette er antagelsen om at syntaktisk og semantisk integrasjon skjer på slutten av setningen, og at dette uansett kan føre til lengre lesetid for det siste ordet i setningen. Imidlertid finnes det ikke noe avgjørende bevis på hva slags integreringsprosesser som finner sted på slutten av setningen, hvordan disse eventuelt påvirker øyebevegelsene, og hvorvidt det i det hele tatt skjer (Stowe et al., 2018). Noen forskere påstår at det faktisk kan være best å fokusere på at interessefeltene ikke ligger nær skjermkanten, ettersom øyebevegelsesregistrering pleier å være mindre nøyaktig når blikket er rettet mot skjermkanten (Godfroid, 2020, s. 172).

Interessefeltene må ligne på hverandre på flere nivå: først og fremst må de være like lange i alle testsetninger, ettersom lengre ord forårsaker lengre fikseringer (Baddeley et al., 1975). I tillegg må interessefeltene være lange nok (5+ bokstaver), slik at de ikke blir hoppet over under førstegangslesing. De kritiske ordene må være av samme ordklasse og ha samme bruksfrekvens: mer frekvente ord identifiseres forttere og er ofte kortere enn mindre frekvente ord (Warren, 2012). For de ordene som har flere stavingsmåter, er det best å bruke den mest frekvente stavingsmåten, noe som kan bestemmes ved hjelp av korporaanalyse.

Utvalg av informanter er også viktig i eye-tracking. Utvalget må først og fremst være representativt for den populasjonen som undersøkes i en gitt studie, altså hvis målet er å undersøke språkprosessering hos den «normale» populasjonen, kan for eksempel personer med lese- og skrivevansker ikke rekrutteres. Alle informantene må ha normal synsskarphet enten med eller uten briller/linser. Imidlertid kan visse briller og linser gjøre øyeidentifisering vanskeligere, noe som kan gjøre at en del dataene fra deltakere med slike briller/linser må forkastes etter innsamling. Personer med skjeling kan som regel ikke delta i eye-tracking-eksperimenter. Det kan også være vanskelig å spore blikket til personer med tunge øyelokk, samt personer med veldig mørke regnbuehinner eller øyevipper. Det er dessuten viktig å samle inn informantenes bakgrunnsinformasjon, ettersom deres språkferdigheter, utdanning, alder, jobb, og mange andre faktorer kan være potensielle forstyrrende variabler.

### **Fordeler og ulemper ved eye-tracking**

Noen av de største fordelene ved bruk av eye-tracking som metode i språkvitenskapelige studier er de følgende:

- Allsidighet: eye-tracking kan brukes i forskjellige paradigmer for å undersøke en stor mengde språklige fenomener, som for eksempel lesing, taleprosessering, øyebevegelser i dialog, skriving, etc. (Godfroid, 2020).
- Høy økologisk validitet: lesing i eye-tracking eksperimenter ligner mye på vanlig lesing, i det at deltakere ser hele setninger og kan bruke så lang tid som nødvendig for å lese disse. Det er mulig å trekke konklusjoner bare fra øyebevegelser, uten behov for spesielle oppgaver eller vurderinger fra deltakere. Dataene gjenspeiler altså naturlige prosesser og kan generaliseres til andre omstendigheter (utenfor laben).
- Tilgang til ubevisste prosesser: fordi øyebevegelsene skjer ubevisst, er det mulig å trekke konklusjoner om ubevisst språkprosessering (i motsetning til metakunnskap om språk) (Carter & Luke, 2020).
- Høy datakvalitet: eye-tracking produserer data av høy kvalitet og oppløsning, både i tid og rom: dette gjør eye-tracking mer sensitivt for visse fenomener enn for eksempel *self-paced reading* (selv-styrt leseeksperiment).
- «Online»-språkprosessering: eye-tracking-data gjenspeiler språkprosessering mens den skjer, altså «i sanntid». Tilgang til et så detaljert opptak av

lesemønster betyr at eye-tracking-studier kan undersøke veldig subtile forskningsspørsmål.

- Barnevennlighet: eye-tracking kan brukes med små barn, ettersom blikksporingsdata er veldig sensitive og implisitte, og kan derfor brukes med enkle naturlige oppgaver, som for eksempel å plukke opp eller flytte et objekt. Moderne eyetrackere er også mindre påtrengende enn for eksempel EEG, ettersom de ikke må festes på barnets hode.

Som alle andre metoder har eye-tracking visse ulemper og begrensinger:

- Korrekt bruk av eye-tracking-utstyr krever opplæring og erfaring, noe som kan være problematisk for unge forskere og for andre uerfarne forskere som har begrenset tid til opplæring.
- Eye-tracking-eksperimenter kan være tidkrevende, ettersom bare én person kan delta om gangen, gjentakelse av utstyrkalibrering er ofte nødvendig, osv.
- Det er lett å overvurdere betydning av eye-tracking-data: disse dataene gir ingen direkte tilgang til kognisjon, mentale representasjoner eller hjerneaktivitet; eye-tracking er en atferdsmetode som kun måler øyebevegelser. Dataene om disse bevegelsene sammen med eksisterende teorier (f.eks. «øye-sinn»-forbindelsen) må tolkes av forskere for å trekke konklusjoner om kognisjon og språkprosessering. I tillegg har «øye-sinn»-hypotesen sine begrensninger (Anderson et al., 2004): blikket gjenspeiler ikke alle kognitive prosesserer, og det skjer også at mennesker «soner ut» og stirrer på et punkt uten å prosessere den visuelle inputen.
- Utvalget i eye-tracking-studier er ofte ikke representativt nok, ettersom mange personer ikke kan delta i eksperimenter (på grunn av øyefarge, andre aspekter av øyets biologi eller linser/briller de bruker). I tillegg gjennomføres eye-tracking-eksperimenter som oftest i en lab, noe som betyr at det er kun personer som bor i nærheten som kan delta. Dette kan derimot rettes ved hjelp av nyere teknologier, som for eksempel webkamerabasert eye-tracking og bærbare eyetrackere.

### 3.3 Chapter summary: Eye-tracking

This chapter describes the use of eye-tracking in language research and justifies the choice of eye-tracking as the method for the current study. Because language is a complex, multifaceted phenomenon, an interdisciplinary approach is often necessary: whereas theoretical linguistics focuses on descriptive methods such as introspection, interview, and acceptability judgements, the field of psycholinguistics is usually associated with experimental methodology. Such techniques as self-paced reading, eye-tracking, and EEG provide high-quality data about the subconscious processes underlying language processing, which are impossible to obtain via the traditional approaches relying on the speakers' metalinguistic knowledge and intuition about language. These methods also provide a large quantity of data, which makes inferential statistical analysis possible.

Section 3.2.1 of this chapter explains the technical aspects of eye-tracking. Most modern eye-trackers rely on calculating the participants' gaze position by measuring the distance between the pupil centre and the corneal reflection created by infrared light. Eye-trackers can be placed by the computer screen (in which case the participant's head usually needs to be stabilized) or mounted on the participant's head (as a headband or glasses). The sampling frequency, precision and accuracy of the eye-tracker, as well as proper calibration are all crucial for obtaining accurate, valid data. The two main types of eye-tracking data are fixations (moments when the eyes are still) and saccades (quick movements in between fixations).

Section 3.2.2 covers visual processing of language, which happens in three stages: the visual input is first stored in a kind of a "buffer", then transferred to and analysed in the working memory, before finally being integrated with the linguistic and cognitive interpretation of the text (Warren, 2012). An important feature of reading is the so-called "preview effect": humans can see about 15 characters to the right of the fixation, albeit not all of them equally clearly. Still, perceiving the beginning letters of the following word and estimating that word's length determines the following fixation and explains why short function words are often skipped in reading: they are processed in the parafovea during the previous fixation.

The last section of this chapter describes the use of eye-tracking in linguistics research, which is based on the "eye-mind" hypothesis. This hypothesis implies that what is being fixated by the eye is what is being processed in the given moment, and that the length of the fixation corresponds to the cognitive effort necessary to process

the stimulus. Advantages of the use of eye-tracking for language research include the diverse applications of the method, its high ecological validity, as well as the high quality and resolution of the data. In addition, eye-tracking can provide insight into online, subconscious language processing, which was one of the main reasons to choose eye-tracking as method for this study. The limitations of this method are the fact that researchers require training and experience to successfully operate eye-tracking equipment, and that the experiments can be quite time-consuming, as well as the fact that the participant samples are oftentimes insufficiently representative of the general population. Moreover, eye-tracking does not provide any direct information about the cognitive processes involved in processing various linguistic phenomena, which means that eye-tracking data have to be carefully interpreted in regard to their meaning.



# Kapittel 4

## Metode

Første delen av dette kapitlet fremstiller i detalj hypoteser og prediksjoner som ble testet i denne studien. Underkapittel 4.2 beskriver materialet brukt i eksperimentet, samt begrunner designet av både stimuliene og selve eksperimentet. Underkapitler 4.3 og 4.4 fremstiller henholdsvis infromantutvalget og utstyret brukt i eksperimentet. Den siste delen av kapitlet beskriver gjennomføring av eksperimentet.

### 4.1 Hypoteser og prediksjoner

Målet med denne studien var å undersøke om det var mulig å finne evidens for at finitthet-trekket oppfattes som distinkt fra andre verbspesifikke trekk, noe som kan tolkes som at finitthet har en ontologisk status og en form for kognitiv representasjon. I denne studien sammenligner jeg hvorvidt, og i hvilken grad, norske språkbrukere oppfatter finitthet og tempus som distinkte trekk. Norsk er godt egnet til å undersøke dette spørsmålet, fordi norske verb mangler mange andre typer morfologiske trekk, som for eksempel tall og person (trekk som tradisjonelt har blitt sett som det som sammen med tempus-trekket utgjør finitthet).

Problemstillingen i denne studien er som følger:

- Oppfatter nordmenn finitthet og tempus som to distinkte trekk?

Et eye-tracking eksperiment ble utført for å undersøke dette spørsmålet, der nordmenn leste setninger med forskjellige typer avvik i verbmorfologi, mens et infrarødt kamera fanget opp øyebevegelsene deres. Følgende er et eksempel på de fire typene av stimulusetningene som informanter leste under eksperimentet:

- (i) Camilla får ikke lov til å gjøre lekser foran TV-en.
- (ii) Hver kveld sitter Camilla og gjøre lekser foran TV-en.
- (iii) Camilla har alltid gjøre lekser foran TV-en.

(iv) Forrige helg satt Camilla og gjøre lekser foran TV-en.

De eksperimentelle betingelsene (*conditions*) inneholdt tre forskjellige typer avvik: avvik som innebar feil finittmarkering (ii), avvik som innebar feil tempusmarkering (iii), og avvik som innebar feil på begge typer markering, både finittet og tempus, på det samme verbet (iv). Grammatikalske setninger (i) ble brukt som kontrollsetninger, for å etablere et sammenligningsgrunnlag for hvordan nordmenn prosesserer verb uten morfologiske avvik.

Basert på Eides (2009) teori om finittet og tempus som to adskilte kategorier, forventet jeg først og fremst at deltakere ville ha forskjellige reaksjoner (altså en forskjell i øyebevegelser) på verbene med finitt- og tempusavvik. En forskjell i blikkmønster kan anses som et bevis på at informantene oppfatter disse to trekktypene forskjellig, noe som kan tolkes som en kognitiv ontologisk statusforskjell mellom disse to trekkene. Fordi øyebevegelser påvirkes av forskjellige typer avvik i grammatikken, forventet jeg å se mer langvarige fikseringer på ugrammatikalske verb, samt flere fikseringer på og flere regresjoner (tilbakevendinger) til disse verbene.

Samtidig kan vi spekulere om hvorvidt finittdistinksjonen ikke er bare adskilt fra tempusdistinksjonen, men om den også kan være mindre fremtredende, fordi finittet er et utolkbart trekk (Cowper, 2016; Wik, 2014), noe som gjør det vanskeligere å konseptualisere enn det tolkbare trekket tempus. Dette burde vi kunne forvente å se også hos norske informanter, selv om finittdistinksjonen forøvrig er robust i norsk. Hvis denne prediksjonen stemmer, kan et avvik som involverer finittdistinksjonen være mindre avstikkende enn et avvik som bare involverer tempusdistinksjonen, noe som gjør finittavviket relativt sett lettere å prosessere. Disse forventningene ble omformulert til hypoteser som ble testet ved hjelp av et eye-tracking-eksperiment:

- $H_0$ : Finittet- og tempus-trekkene er uløselig knyttet sammen; både finitt- og tempusavvikene påvirker derfor øyebevegelser på samme måte.
- $H_1$ : Finittet- og tempus er to adskilte trekk. Avvik i finitt- og tempusmorfologi oppfattes som to distinkte typer av avvik, noe som resulterer i forskjellige lesemønstre.

I tillegg formulerer jeg en underhypotese som hevder at finittet-trekket er mindre robust enn tempus, og finittavvik er derfor lettere å prosessere enn tempusavvik. Øyebevegelser som følger disse avvikene bør i så fall skille seg fra hverandre på følgende måter:



- Finittavvik: lengre og flere fikseringer, samt flere regresjoner enn i kontrollsetninger, men mindre og kortere fikseringer og færre regresjoner enn i setninger med brudd på bare tempusdistinksjonen eller setninger med brudd på både tempus- og finittdistinksjonen.
- Tempusavvik: lengre og flere fikseringer og flere regresjoner enn både i kontrollsetninger og i setninger med finittavvik. Mindre og kortere fikseringer og færre regresjoner enn i setninger med både tempus- og finittavvik.
- Tempus- pluss finittavvik: lengre og flere fikseringer og flere regresjoner enn både i kontrollsetninger og i setninger med bare én type brudd.

## 4.2 Materialet

### 4.2.1 Verbutvalget i eksperimentet

Det første steget i å lage stimulisetninger var å gjøre et utvalg av verb til bruk i eksperimentet, altså de verbene som skulle inneholde avvikene i tempus- og finittmorfologi. Som beskrevet i seksjon 3.2, er ordlengde svært viktig i eye-tracking, ettersom lengre ord i seg selv utløser lengre og flere fikseringer. For å minimere tilfeldig variasjon i dataene, måtte alle målverbene inneholde like mange bokstaver. I tillegg måtte målverbene være minst 5 bokstaver lange: kortere ord blir ofte «hoppet over» under lesing, fordi de kan opptre i synsfeltet og bli prosessert mens personen faktisk fokuserer på det forrige ordet.

En annen utfordring var å finne balansen mellom ytre og indre validitet i eksperimentet: en studie har ytre validitet ettersom resultatene kan generaliseres, og dermed gjelde en større mengde data enn det studien undersøkte. Indre validitet forutsetter at man har god kontroll over mulige feilkilder i studien, og at funnene kan forklares gjennom den antatte hypotesen (Nunan, 1992, s. 15). For å beholde høy indre validitet, burde jeg bruke ett og samme verb i alle stimulisetningene, slik at all variasjon i funnene skulle kunne skyldes den eksperimentelle manipulasjonen. På den andre siden, for å ha høy ytre validitet, burde jeg bruke så mange verb som mulig i eksperimentet, slik at funnene kunne generaliseres til hele det norske språket.

Som følge av dette iboende dilemmaet prøvde jeg å velge de verbene som er mest representative for norsk. Ifølge Simonsen og Bjerkan (1998) må både *type*-frekvens

(antall verb i hver verbklasse) og *token*-frekvens (bruksfrekvens av verbene i hver verbklasse) tas i betraktning. Jeg brukte samme klassifisering av verb som i studien av Simonsen og Bjerkan: verb som får en endelse i preteritum regnes som svake (f.eks. kaste – kastet), mens verb som ikke får en endelse i preteritum regnes som sterke (f.eks. fryse – frøs<sub>\_</sub>). Svake verb inndeles videre i to klasser: verb som får en endelse som begynner med en vokal (*-et* eller *-a*) tilhører den store svake klassen, mens verb som får en endelse som begynner med en dental konsonant (*-te* eller *-de*) tilhører den lille svake klassen. Andelen av verb fra hver verbklasse i eksperimentet ble bestemt i samsvar med størrelsen på verbklassene: 56% av alle norske verb tilhører den store svake klassen, 37% tilhører den lille svake klassen og bare ca. 4% er sterke verb ifølge *Bokmålsordboka*.<sup>9</sup> Imidlertid er mange sterke og modale verb svært frekvente i norsk, noe som altså gjør disse verbene relevante og viktige for denne studien. Likevel kunne bare sterke verb, ikke modalverb, inkluderes i eksperimentet, fordi modale verb (*måtte*, *burde*, osv.) har samme form i infinitiv og preteritum. Hvis en infinitivform av et modalt verb brukes i stedet for en preteritumsform, er avviket på finitthet og tempus dermed ikke observerbart i verbformen. Til slutt var det også viktig å velge frekvente verb, ettersom sjeldent brukte verb forårsaker lengre lesetid og derved øker innslaget av usystematisk variasjon i dataene.

Alt i alt valgte jeg fem verb for dette eksperimentet: to verb fra den store svake klassen (*å passe*, *å vaske*), to verb fra den lille svake klassen (*å bruke*, *å gjøre*) og ett sterkt verb (*å synge*). Bruksfrekvensen av selve verbene (*token*-frekvens) ble undersøkt for å velge riktige verb fra hver verbklasse. *Passe*, *vaske*, *bruke*, og *gjøre* er blant 10 mest frekvente verb i henholdsvis den store og den lille svake klassen, ifølge *Frekvensordlister for norsk talespråk og skriftspråk*, som er basert på Oslo-korpuset av taggedede norske tekster. *Synge* tilhører «a-preteritum»-gruppe<sup>10</sup> av sterke verb (som inneholder flest sterke verb blant alle andre grupper), og *synge* er i tillegg en av 10 mest frekvente sterke verb i norsk.

---

<sup>9</sup> Disse dataene er baserte på 1986-utgave av *Bokmålsordboka*; prosentandeler beregnet av Simonsen og Bjerkan (1998).

<sup>10</sup> Navnet på denne gruppen henviser til vokalveksling mellom infinitiv- og preteritumsverbformer: verb som tilhører denne gruppen får rotvokalen /a/ i preteritum.

#### 4.2.2 Design av stimuli

Stimulisetningene er skrevet på moderat bokmål, ettersom bokmål er den dominerende målformen for tekst i Trøndelag (Utdanningsdirektoratet, 2014), og det var svært sannsynlig at mesteparten av deltakerne i eksperimentet behersket bokmål bedre enn nynorsk. Til sammen skapte jeg 40 stimulissett<sup>11</sup>, med 8 sett per utvalgt verb. Hvert stimulissett består av fire versjoner av samme setning:

- (i) Kontrollsetning (heretter kalt «betingelse K»), der infinitivsformen av målverbene er brukt på en grammatikalsk måte, f.eks. *Camilla får ikke lov til å gjøre lekser foran TV-en*. Denne versjonen er tenkt å skape et sammenligningsgrunnlag som ugrammatikalske setninger kan sammenlignes med.
- (ii) Avvik i finittmarkering (heretter kalt «betingelse F»), der infinitivsformen er brukt i stedet for presensformen, som er et brudd på finittdistinksjonen, f.eks. *Hver dag sitter Camilla og gjøre lekser foran TV-en*.
- (iii) Avvik i tempusmarkering (heretter kalt «betingelse T»), der infinitivsformen er brukt i stedet for partisippformen, som er et brudd på tempusdistinksjonen, f.eks. *Camilla har alltid gjøre lekser foran TV-en*.
- (iv) Avvik i både finitt- og tempusmarkering (heretter kalt «betingelse F+T»), der infinitivsformen er brukt i stedet for preteritumsformen, som er et brudd på både tempus- og finittdistinksjoner, f.eks. *Forrige helg satt Camilla og gjøre lekser foran TV-en*.

I eye-tracking, som i mange andre typer eksperimentelle metoder, er det avgjørende at alle varianter (betingelser) av samme stimulus ligner på hverandre så mye som mulig, for å minimere tilfeldig variasjon i dataene. Foruten testverbene måtte resten av setningen også kontrolleres. Den delen av setningen som kommer før testverbet er litt ulik i forskjellige versjoner av samme setning, hovedsakelig fordi det var nødvendig å plassere en tempusmarkør (f.eks. et tidsadverbial)<sup>12</sup> før testverbet, for å skape en forventning om en spesifikk verbform i deltakerne. Det var imidlertid veldig viktig å beholde den post-verbale delen av setningen lik i alle fire betingelsene, for å kunne

<sup>11</sup> Fullstendig stimuliliste kan man finne i vedlegg A.

<sup>12</sup> Betingelser K og T inneholder ikke alltid et tidsadverbial som tempusmarkøren, ettersom hjelpeverb og infinitivmerke i disse setningene fungerer som en tempusmarkør (dvs. skaper en forventning om en spesifikk verbform).

sammenligne disse nøyaktig og finne potensielle *spillover*-effekter (i.e. prosesseringsvansker forårsaket av avvik i verbmorfologien som blir synbare i følgende setningsdel). Derimot utløser topikalisering av det temporale adverbialet flytting av det finitte verbet (altså test verbet i betingelsene F og F+T) frem i setningen, noe som gjør at den post-verbale delen av setningen da blir ulik i forskjellige betingelser. For å unngå dette, inkluderte jeg pseudokoordineringsverb<sup>13</sup> i tillegg til testverbene i slike setninger, for eksempel *Hver dag **sitter** Camilla og gjøre lekser foran TV-en*. Denne strategien tillot meg å beholde den post-verbale delen av setningen identisk i alle betingelser, men bidro til ulikheter i den pre-verbale setningsdelen. Det hadde riktignok vært teoretisk mulig å omarbeide disse delene slik at de lignet mer på hverandre (hadde samme antall ord og syntaktisk struktur), men grunnet svært begrenset tid til utprøving, landet jeg på at stimuliene likevel var tilstrekkelig like til å kunne gi valide resultater. Selv om forskjeller i den pre-verbale delen av stimuli ikke egentlig er avgjørende for denne studien, er dette noe som må tas i betraktning i dataanalysen og i fremtidige studier.

Et annet aspekt ved stimuliene som må kontrolleres, er hvor naturlige de fremstår. Rare eller unaturlige formuleringer eller merkelig innhold forårsaker også lengre fikseringer og bidrar dermed til støy i dataene, noe som gjør det vanskeligere å finne effekten. Et potensielt problem med stimuliene i denne studien er pseudokoordineringsverbene, ettersom disse vanligvis blir brukt i muntlig norsk og derfor kan fremstå som overflødige eller altfor uformelle i skriftlig språk. I tillegg inneholder mer enn halvparten (57.5%) av stimuliene topikalisering, noe som er en vesentlig større andel topikalisererte setninger enn det som er vanlig i spontant produsert skriftlig og muntlig norsk (ca. 30-35% ifølge Eide og Hjelde (2018, s. 55)). Dette kan også virke overflødig, rart eller unaturlig for informantene, og slik bidra til uønsket tilfeldig variasjon. Generelt sett bør stimuli i eye-tracking være ganske enkle og logiske, slik at bare den eksperimentelle manipulasjonen påvirker prosesseringen. Alle ordene, inkludert navn som ble brukt i stimuliene i dette eksperimentet er frekvente og vanlige for nordmenn. Testverbene er alltid brukt i hovedsetninger og alltid uten verbpartikler, for å beholde setningenes syntaktiske struktur så enkel som mulig.

---

<sup>13</sup> I «vanlig» verbkoordinering regnes begge verb som like viktige (f.eks. *Vi danser og synger*), men i pseudokoordinering er det en iboende asymmetri. Det første verbet har en viss karakter av hjelpeverb og det andre verbet fremstår som det innholdsmessig viktigste (f.eks. *Hun sitter og leser*). Det er altså når for eksempel positur- og bevegelsesverb brukes i slike konstruksjoner, at de kalles for pseudokoordineringsverb (Lie, 2017).

Alle stimuliene ( $n=160$ , 40 sett  $\times$  4 betingelser) ble testet av 120 morsmålstalere av norsk og vurdert etter hvor naturlige og velformede de var. Ingen av disse 120 personene deltok i hovedeksperimentet. Disse pilotinformantene ble fordelt i fire grupper, der hver gruppe leste 40 setninger (én setning fra hver stimulisset). Alle setningene ble presentert i sin grammatikalsk korrekte form, altså uten avvik i verbmorfologi. Undersøkelse ble gjennomført på nett ved hjelp av Google Forms, og deltakere bedømte setningenes naturlighet ved hjelp av 1–4 Likert-skala<sup>14</sup> (der 1 = «helt unaturlig, jeg ville aldri ha sagt noe slikt», 2 = «ganske unaturlig», 3 = «ganske naturlig», 4 = «helt naturlig, noe jeg eller en venn kunne ha sagt»). Under hver setning var det et tomt felt der deltakerne kunne forklare hva som var rart eller unaturlig med setningen og hvordan de ville ha endret den. Setninger med en gjennomsnittlig skår på mindre enn 3 ble revidert og korrigert i samsvar med deltakernes forslag.

#### 4.2.3 Design av eksperiment

Et av de første valgene i planlegging av en psykologisk studie er valget mellom innengruppedesign og mellomgruppedesign. I mellomgruppedesignet blir deltakerne fordelt i to eller flere grupper, og hver gruppe får ulike testforhold. Resultatene fra alle grupper sammenlignes, og en av gruppene fungerer som regel som en «kontrollgruppe» som danner et utgangspunkt for sammenligning. I innengruppedesignet eksponeres alle deltakerne for alle testforhold på forskjellige tidspunkter. Det er resultater for hver type av testforhold som sammenlignes i dette designet, og sammenligningsgrunnlaget dannes av dataene som ble samlet inn før første test. Fordelen med innengruppedesignet er en høyere statistisk styrke, det vil si sannsynligheten for å oppdage en effekt og forkaste nullhypotesen: fordi man bruker samme deltakere i alle testforhold, unngår man at variasjonen mellom deltakere bidrar til uønsket tilfeldig variasjon. Når dette aspektet er ivarettatt, er det lettere å oppdage den systematiske variasjonen som utgjør den eksperimentelle effekten (Field, 2013, s. 388). I tillegg trenger man færre deltakere for et innengruppeeksperiment, ettersom alle deltakere bidrar med data for alle testforhold.

---

<sup>14</sup> Selv om den vanligste varianten av Likert-skala har 5 eller 7 svarmuligheter, valgte jeg å bruke formatet «tvunget valg», der deltakeren ikke kan gi et «nøytralt» svar. Ifølge Allen og Seaman (2007) pleier deltakere å velge det nøytrale alternativet når de ikke har lyst til å tenke så mye på spørsmålet eller når de mister interesse for undersøkelsen. I så tilfellet betyr svaret «jeg vet/bryr meg ikke», noe som ikke er særlig informativt for denne studien.

Som følge av dette valgte jeg å bruke innengruppedesignet. Imidlertid var det ikke mulig for alle informantene mine å lese alle stimulisetningene, både fordi setningene var så mange ( $n=160$ ), og fordi deltakerne mest sannsynlig ville ha oppdaget målet med eksperimentet, hvis de leste flere versjoner av samme setningen der bare testverbets morfologi varierte. Motbalansering (*counterbalancing*) av type «latinsk kvadrat» var en god løsning på dette problemet: stimulisetningene ble fordelt på 4 stimulilister, der hver liste inneholdt bare én setning fra hvert stimuliset, men et likt antall setninger for hver betingelse.<sup>15</sup> Informantene ble også tilfeldig fordelt i 4 grupper, og hver gruppe fikk sin egen stimuliliste. Som følge av dette ble alle informantene eksponert for alle fire betingelser uten å skape feilkilder i eksperimentet. Hver stimuliliste inneholdt 40 testsetninger og 40 kamuflasjesetninger (*fillers*).

Kamuflasjesetninger brukes regulært for å unngå at deltakerne oppdager målet med eksperimentet. De inkluderes ikke i analysen, ettersom det eneste målet med disse er å distrahere deltakerne fra teststimuliene. Kamuflasjesetninger i dette eksperimentet lignet på testsetningene med tanke på syntaktisk struktur, men hadde ingen morfologiske avvik i verbene. Halvparten av kamuflasjesetningene var grammatikalske, mens den andre halvdel inneholdt to typer av avvik: genus-feil i (u)bestemte artikler og V2-brudd. Disse avvikene ble inkludert for å tentativt lede deltakernes oppmerksomhet bort fra verbavvikene i testsetningene.

### 4.3 Informanter

Deltakere i dette eksperimentet var 64 voksne nordmenn (31 kvinner og 33 menn) i alderen fra 18 år til 69<sup>16</sup> år ( $M = 28$ ,  $SD = 12$ ). Alle informantene har vokst opp i Norge

---

<sup>15</sup> Eksempel på en stimuliliste: én setning fra sett 1 i betingelse K, én setning fra sett 2 i betingelse F, 3-T, 4-FT, 5-K, 6-F, 7-T...40-FT.

<sup>16</sup> Aldring innebærer visse fysiske (Salvi et al., 2006) og kognitive (Crawford et al., 2017) endringer som kan påvirke menneskets øyebevegelser. På grunn av dette er det vanlig å begrense informantutvalget for eye-tracking-eksperimenter til mennesker mellom 18 og 40 år. Jeg valgte likevel å samle inn data også fra eldre informanter, hovedsakelig fordi aldring påvirker først og fremst de eye-tracking-målingene som ikke ble analysert i denne studien (f.eks. sakkader og *smooth-pursuit*, se Dowiasch et al. (2015)). I tillegg ble det under dataanalysen klart at alle de 9 informantene som var eldre enn 40 år produserte helt vanlige data som ikke skilte seg fra dataene produsert av yngre deltakere.

og hadde norsk som morsmål.<sup>17</sup> Alle utenom én deltaker brukte hovedsakelig bokmål som skriftspråk. Alle informantene hadde utdanning på videregående nivå eller høyere, og gjennomsnittlig antall år med høyere utdanning hos informantene var 2.6 år (fra 0 til 8 år). Alle deltakerne hadde normalt eller korrigert-til-normalt syn. En av deltakerne informerte uoppfordret om sin dysleksidiagnose etter deltakelsen og måtte dermed få sine data forkastet fra analysen, ettersom personer med lesevansker har særegne øyebevegelser; de produserer derfor data som ikke kan generaliseres til resten av befolkningen (Rayner, 1998, s. 22). Rekruttering skjedde ved hjelp av plakater på Dragvoll universitetscampus, samt via sosiale medier og metoden «venn-av-en-venn». Deltakelse var frivillig og gratis, men etter loddtrekning fikk tre informanter et gavekort som takk for deltakelsen. Eksperimentet ble utført i høstsemester 2020.

#### 4.4 Utstyr

Tre typer av eyetrackere var tilgjengelige på Språkprosessering- og språktilegnelseslaben på NTNU på tidspunktet eksperimentet ble satt opp: Tobii T120, Tobii T60XL og EyeLink 1000. Fordelen med Tobii eyetrackere er at de kan samle inn blikkdata uten hodestøtte, men samtidig har disse eyetrackere en mye lavere samplingsfrekvens enn EyeLink 1000, noe som gir lavere datakvalitet (Conklin et al., 2018). Høy samplingsfrekvens er viktig når man undersøker blikkmønster i små områder, som for eksempel enkelte ord – derfor valgte jeg å bruke EyeLink 1000-eyetrackeren.

EyeLink 1000 har en samplingsfrekvens på 1000Hz og romoppløsning på under 30' bueminutt, som betyr at den kan fange opp veldig korte og raske øyebevegelser med høy nøyaktighet. Programmering av eksperimentet ble gjort i «Experiment Builder»-programvare (SR Research Ltd., Toronto, Canada). Eyetrackeren fulgte bevegelsene til deltakernes høyre øye<sup>18</sup> mens de leste setninger binokulært på en 47.5 × 30cm PC-skjerm med oppløsning på 1680 × 1050 piksler, plassert ca. 67cm fra dem. Hakestøtte ble brukt for å redusere hodebevegelser.

---

<sup>17</sup> Imidlertid hadde seks informanter en utenlandsk forelder og var dermed tospråklige: tre deltakere var norsk-franske, én var norsk-amerikansk, én norsk-nederlandsk og én norsk-russisk.

<sup>18</sup> Binokulær eye-tracking gjennomføres som regel med lavere samplingsfrekvens. Monokulær eye-tracking passer for de fleste studier, fordi et menneskes øyne beveger seg synkront i tid og rom (Conklin et al., 2018).

Stimulisetninger ble presentert i svart farge på en grå<sup>19</sup> bakgrunn, med tanke på å redusere belastning på øynene, som kan føre til blinking. En ikke-proporsjonal font<sup>20</sup> ble brukt, der bredden av ett skrifttegn tilsvarte 0.34 grader av synsvinkel. Alle stimuliene ble presentert på samme sted på skjermen (i cirka øyehøyden), og alle setninger var en linje lange, for å unngå vertikale øyebevegelser.<sup>21</sup> En blank skjerm med en fikseringsmarkør (en liten svart prikk) ble fremvist før hver eneste setning for å gjennomføre en nøyaktighetssjekk. Deltakerne måtte fokusere på fikseringsmarkøren, og de fikk se neste setning kun hvis blikket deres lå innen 1° fra fikseringsmarkøren. Hvis målingen ikke var nøyaktig nok, måtte eyetrackeren recalibreres.

### 4.5 Gjennomføring

Eksperimentet tok rundt 20-25 minutt per gjennomføring. Deltakerne fikk instruksjoner om å lese på en vanlig måte og ble fortalt at de kunne bruke så mye tid de trengte på hver setning. De fikk vite at det var viktig å forstå setningene, fordi de skulle få enkle ja-nei spørsmål etter noen av setningene; disse spørsmålene forekom etter 50% av kamouflasjesetningene (halvparten av spørsmålene skulle besvares med «ja», den andre halvparten med «nei»). Informantene svarte på spørsmål ved å trykke på venstre/høyre pilknapp på tastaturet; under hvert spørsmål var det en påminnelse om hvilke knapper som skulle brukes. Målet med disse spørsmålene var å sikre at deltakere ikke skumleste, men virkelig forstod setningene. Ved få riktige svar ville deltakernes data forkastes fra den videre analysen.

Etter at deltakere leste informasjonsbrevet og signerte samtykkeskjemaet, samt fikk instruksjoner om hvordan eksperimentet skulle foregå, ble eyetrackeren kalibrert (med 9-punkts kalibrering) og senere recalibrert hvis nødvendig, for eksempel hvis blikket avvek mer enn 1 grad av synsvinkel fra fikseringsmarkøren. Kalibreringen ble akseptert hvis gjennomsnittlig blikkavvik var mindre enn 0.7°. Eksperimentlederen kontrollerte når stimuliene kom på og forsvant fra skjermen, for å unngå problemer med koordinasjon, noe som kunne ha oppstått hvis deltakerne måtte fokusere på fikseringsmarkøren og trykke på en knapp samtidig. Deltakerne viste at de var klare for den neste setningen ved å fokusere på fikseringsmarkøren på den blanke skjermen; de

---

<sup>19</sup> RGB [153, 153, 153]

<sup>20</sup> En font der hvert enkelt tegn tar like stor plass. I eye-tracking er det viktig at ord som er like lange tar opp like stor plass, uansett hvilke bokstaver de består av.

<sup>21</sup> Eyetrackere er mindre nøyaktige med vertikale enn horisontale øyebevegelser (Godfroid, 2020).



viste at de var ferdige med å lese setningen ved å fokusere på en annen markør, en svart prikk nederst til høyre på skjermen.

Stimuliene ble presentert i tre blokker: første blokk inneholdt tre øvingssetninger og ett innholdsspørsmål, som viste deltakerne hvordan eksperimentet skulle foregå og ga mulighet for å stille spørsmål. To andre blokker inneholdt 40 setninger hver: 20 testsetninger og 20 kamuflasjesetninger. Etter hver blokk vistes en skjerm med frasen «ta en pause»; informantene fikk vite at de i tillegg kunne be om flere pauser når som helst i løpet av eksperimentet. Stimulirekkefølgen ble randomisert innen hver blokk for hver deltaker for å unngå rekkefølgeeffekter, altså for å sikre at stimulirekkefølgen ikke influerte på sluttresultatene.

Etter at deltakerne fullførte eksperimentet, fikk de vite hva studien handlet om. Deretter svarte de på et spørreskjema om språkbakgrunnen sin (sin egen dialekt, foreldrenes dialekter, foretrukket skriftspråk, ferdigheter i fremmedspråk), samt generell bakgrunnsinformasjon (alder, utdanning, osv.).

#### **4.6 Chapter summary: Method**

The alternative hypothesis in this study states that native Norwegian speakers perceive finiteness and tense as two distinct, separate features. Additionally, the finiteness distinction is presumably less salient than that of tense, due to its lack of semantic content, which makes finiteness more difficult to conceptualise than tense. Therefore, I predicted that violations of finiteness and tense would result in different eye-movements, and that violations of finiteness would be perceived as less offensive by the native speakers and be processed faster than those of tense. A combination of the two violations was predicted to be the most cognitively taxing to process.

An eye-tracking experiment was conducted in order to investigate these predictions. 64 adult native Norwegians were recruited to participate (aged 18 to 69,  $M=28$ ,  $SD=12$ ), although one participant later had to be excluded due to a dyslexia diagnosis. EyeLink 1000 eye-tracker with a sampling frequency of 1000Hz and a spatial resolution of less than 30 minutes of arc was used in this experiment.

The stimuli consisted of 40 quadruplets, each containing a sentence in 4 different conditions: control condition (a grammatical sentence), finiteness violation condition (infinitive verb form used in place of the present tense form), tense violation condition (infinitive in place of past participle), and a combined violation condition,

containing anomalies in both finiteness and tense morphology (infinitive in place of preterit), for example:

Control condition: *Camilla får ikke lov til å gjøre lekser foran TV-en.*

Camilla is not allowed to do homework in front of the TV.

Finiteness violation: *Hver kveld sitter Camilla og gjøre lekser foran TV-en.*

Every evening Camilla do homework in front of the TV.

Tense violation: *Camilla har alltid gjøre lekser foran TV-en.*

Camilla has always do homework in front of the TV.

Combined violation: *Forrige helg satt Camilla og gjøre lekser foran TV-en.*

Last weekend Camilla do homework in front of the TV.

The sentences were matched in terms of length and content, with the post-verb area of interest being identical across the four conditions in each quadruplet. Five target verbs were chosen as most representative of Norwegian verbs in terms of both type and token frequency: *passe* (take care of), *vaske* (wash), *bruke* (use), *gjøre* (do/make), *synge* (sing). The stimuli were counterbalanced across 4 presentation lists using a Latin square design. Each list also contained 40 filler sentences, in order to prevent the participants from discovering the aim of the experiment. Participants were randomly assigned to the presentation lists. This way, each participant was exposed to all four conditions and could serve as their own control. During the course of the experiment, the participants had to answer 20 simple yes/no questions pertaining to the content of the previous sentence, in order to ensure that they were focused on the task.

Each experiment session lasted approximately 20-25 minutes. The eye-tracker was calibrated using 9-point calibration before the start of the experiment, and later recalibrated as necessary. The stimuli were presented in three blocks: the first block contained three practice sentences, and each of the other two blocks contained 20 experimental items and 20 filler sentences. The order of presentation was randomised within each block. A break was scheduled after every block; the participants were informed that they could request additional breaks at any time. After the completion of the experiment the participants were debriefed, and then filled out a background questionnaire (age, education, native dialect, parents' dialects, preferred written form, competence in foreign languages, etc.).

# Kapittel 5

## Dataanalyse og resultater

Dette kapitlet presenterer dataene som ble samlet inn under eksperimentet, hvordan disse ble analysert, samt resultatene som fremstod i analysen. Første underkapittel beskriver datasettet og hvilke dataenheter som måtte forkastes før analysen var gjennomført. Underkapittel 5.2 gir oversikt over avhengige variabler som ble analysert i denne studien og hva de betyr med hensyn til språkprosessering. Gjennomsnittene og standardavvikene for disse variablene er presentert i seksjon 5.3. Det siste underkapitlet beskriver den statistiske analysen i detalj: valg av analysemodellen, transformasjon av data, og de endelige resultatene for hvert interessefelt.

### 5.1 Kvalitetssikring av data

Før dataene kan analyseres, må de kvalitetssikres: dette innebærer å forkaste og korrigere visse prøvesegmenter<sup>22</sup> (*trials*) og fikseringer. Denne kvalitetssikringen ble gjennomført ved hjelp av «Data Viewer»-programvare (SR Research Ltd., Toronto, Canada).

Fordi én av deltakerne måtte forkastes på grunn av en dysleksidiagnose, ble bare dataene fra de resterende 63 deltakerne inkludert i analysen. Først og fremst forkastet jeg dataene fra kamouflasje- og øvingssetninger, ettersom disse var uvesentlige for denne studien. Dataene fra testsetningene utgjorde 2520 prøvesegmenter. Ett segment måtte forkastes på grunn av en feil i eksperimentet, der opptaket ble avsluttet før deltakeren hadde lest hele setningen.

---

<sup>22</sup> Opptaket av øyebevegelser ble automatisk fordelt i segmenter av eyetracker-programvare. Hvert opptakssegment tilsvarer én testsetning: segmentet begynner når setningen blir synlig på skjermen og avsluttes når setningen forsvinner fra skjermen.

Selv om blikksporingsdata ble tatt opp for hele setninger, er det bare to «interessefelter» i hver setning som ble analysert i denne studien: selve test verbet (heretter kalt «verbalt interessefelt») og delen av setningen rett etter test verbet (heretter kalt «post-verbalt interessefelt»), som inneholdt enten ett eller to<sup>23</sup> ord som fulgte test verbet. Det var viktig å inkludere det post-verbale interessefeltet i analysen, ettersom effekter av morfosyntaktiske avvik kommer til uttrykk også etter at personen har lest det ordet der avviket oppstod (såkalte *spillover*-effekter), ifølge Conklin og Pellicer-Sánchez (2016). Det var tilstrekkelig med kun to interessefelter, fordi effekter av morfosyntaktiske avvik (i motsetning til effekter av semantiske avvik) er umiddelbare, men kortvarige (Braze et al., 2002). Dette betyr at morfosyntaktiske avvik forårsaker regresjoner fra og lengre lesetid på selve avviket og et par ord etter det. Snart etterpå blir øyebevegelsene «vanlige» igjen, altså slik som de er i kontrollsetninger. Interessefeltene i dette eksperimentet ble definert på forhånd, under programmering av eksperimentet.

Alle prøvesegmentene som inneholdt «sporingstap» (*track loss*), altså tilfeller der kameraet ikke kunne identifisere deltakerens pupill<sup>24</sup> i minst ett av disse to interessefeltene,<sup>25</sup> måtte forkastes. Disse tilsvarte 77 prøvesegmenter eller 3.06% av det originale datasettet. I de fleste eye-tracking-studier velger forskere å forkaste opptak med sporingstap, fordi fikseringer som kommer før og etter et blunk ikke er pålitelige, ettersom de kan være kunstige eller påvirket av blunket: «blinks are always preceded and followed by partial occlusion of the pupil, causing artificial changes in pupil position» (EyeLink bruksanvisning, 2009, s.100). Til slutt måtte noen fikseringer i 62 av prøvesegmentene justeres langs vertikalaksen, fordi det virket som om deltakerne leste «over» eller «under» teksten. Dette kan skje i enkelte tilfeller på grunn av biologiske særegenheter eller kalibreringsproblemer. Justering ble aldri utført på enkeltfikseringer, men helst på hele fikseringsrekker, i de tilfellene der det var klart at deltakeren leste setningen på en normal måte.

---

<sup>23</sup> Fordi blikket ofte hopper over korte ord, måtte det post-verbale interessefeltet inneholde to ord, dersom det første ordet etter test verbet var kortere enn 5 bokstaver.

<sup>24</sup> Sporingstap skjer som oftest på grunn av blinking, men kan også forårsakes av sminke, mysing, noen typer kontaktlinser, osv.

<sup>25</sup> Hvis sporingstapet begynte og/eller sluttet i ett av interessefeltene, ble dette regnet som grunn for å forkaste prøvesegmentet.

Få riktige svar på ja-nei spørsmålene om setningenes innhold var ett av kriteriene som ble brukt for å forkaste deltakere, fordi dette var en indikasjon på at deltakeren ikke fokuserte godt nok på oppgaven. Nøyaktighetsterskelen for forkasting av deltakere ble satt på 80%. Alle deltakerne i eksperimentet fikk 85% riktige svar eller mer ( $M = 95\%$ ), derfor ble ingen faktisk forkastet på dette grunnlaget.

Etter korrigerings og vraking av problematiske prøvesegmenter, ble trinn 1 og 4 av «Data Viewer firetrinns kvalitetssikringsprosedyre» utført. Under trinn 1 av denne prosedyren ble alle fikseringer kortere enn 80ms slått sammen med lengre fikseringer som lå innen 0.5 grader av synsvinkelen (tilsvarende ca. 1.5 bokstav) fra dem. Under trinn 4 ble alle gjenstående fikseringer kortere enn 80ms og alle fikseringer lengre enn 800ms slettet. Totalt var det 1561 fikseringer (4.18% av det originale datasettet) som ble slettet eller slått sammen med andre fikseringer. Terskelen for hvilke fikseringer som må slettes kan variere ganske mye avhengig av hva som undersøkes i en gitt studie. Lignende terskler (80ms og 800ms) har blitt brukt i mange andre eye-tracking-studier med lignende design, som for eksempel Foppolo og Staub (2020), Knickerbocker et al. (2019), og Kuperman og Van Dyke (2011). En viktig årsak til å utelate veldig korte/lange fikseringer fra analysen er at altfor lange fikseringer antas å være et tegn på redusert oppmerksomhet, mens altfor korte fikseringer ikke representerer kognitiv prosessering, men helst andre hendelser, som for eksempel mikrosakkader og falske fikseringer (Godfroid, 2020). Dessuten er mennesker rett og slett blinde i løpet av sakkaden og denne blinde fasen sprer seg delvis til fikseringen. Post-sakkadisk blikkhemming varer i rundt 100-120ms. Dermed kan inntak og prosessering av fikseringen begynne tidligst 70-80ms etter at blikket ble festet på et punkt (Volkman, 1986).

Etter at datasettet var kvalitetssikret, ble dataene for de to interessefeltene eksportert for analyse. Fikseringer utenfor disse feltene ble ikke analysert. I noen av prøvesegmentene ble ett av feltene (test verbet eller ordet etter det) hoppet over, og disse områdene ble derfor ikke inkludert i analysen. Til sammen var det 184 verbale og 75 post-verbale interessefelter som ble hoppet over. Antall post-verbale interessefelter som ble hoppet over var ganske likt i alle betingelser (21, 17, 17, 20), mens antall verbale interessefelter som ble hoppet over var en del høyere for kontrollsetningene (60) enn for testbetingelsene (41, 36, 47). Dette er ikke overraskende, ettersom ugrammatikalske verb tiltrækker seg mer oppmerksomhet fra leseren, og det er mindre sannsynlig at disse blir hoppet over. Det var også mindre sannsynlig at post-verbale

interessefelter skulle blitt hoppet over i dette eksperimentet, fordi de som regel var lengre enn de verbale interessefeltene.

## 5.2 Avhengige variabler

«Betingelse» var den eneste uavhengige variabelen i denne studien, og denne inneholdt fire nivå: betingelse K (grammatisk kontrollsetning), F (finittavvik), T (tempusavvik) og F+T (finitt- pluss tempusavvik). De avhengige variablene i denne studien var forskjellige typer blikkdata samlet inn under eksperimentet. Følgende er en oversikt over de seks blikksporingsmålingene som ble analysert for begge interessefeltene.

To av disse målingene representerer tidlig kognitiv prosessering og kalles derfor «tidlige målinger» (*early measures*):

- *First fixation duration* er varigheten på den første fikseringen i interessefeltet. Denne målingen er det tidligste tidspunktet man kan se effekter av ordets egenskaper (Liversedge et al., 1998) og *spillover*-effekter (Clifton Jr et al., 2007).
- *Gaze duration* er den oppsummerte varigheten på alle fikseringer gjort i et interessefelt under førstegangslesing før blikket forlot interessefeltet til høyre eller til venstre. Denne målingen er sensitiv for både semantiske og syntaktiske avvik (Rayner et al., 2004).

De gjenstående fire målingene er «sene målinger» (*late measures*) som gjenspeiler sen<sup>26</sup> og veldig sen kognitiv prosessering:

- *Regression path duration* er den oppsummerte varigheten på alle fikseringer i selve interessefeltet og fikseringer gjort i løpet av regresjonen tilbake i setningen, før blikket forlot interessefeltet til høyre. Denne målingen gjenspeiler de leksikalske og integreringsmessige vanskene som var grunnen til regresjonen (Rayner et al., 2004).
- *Dwell time* er den oppsummerte varigheten på alle fikseringer i interessefeltet.

---

<sup>26</sup> Det er viktig å huske at kumulative målinger, som f.eks. *regression path duration* og *dwell time* er en sum av tidlige og sene fikseringer og kan derfor bli påvirket av både tidlig og sen prosessering. En signifikant effekt i disse kumulative målingene uten effekt i tidlige målinger er en indikasjon på en sen effekt i prosessering. (Conklin & Pellicer-Sánchez, 2016)

- *Fixation count* er antallet fikseringer i interessefeltet. Både denne målingen og *dwelling time* er en måte å vurdere oppmerksomheten tiltrukket av et interessefelt gjennom hele forsøket (Conklin et al., 2018).
- *Regressions out* er sannsynligheten for at blikket vender tilbake i teksten fra et gitt interessefelt under førstegangslesing. Denne målingen kan være et tegn på vanskeligheter på høyere tekstlingvistisk nivå, for eksempel vanskeligheter med å forstå hvordan flere ord står i forhold til hverandre (Rayner & Pollatsek, 1989).

Ifølge Conklin og Pellicer-Sánchez (2016) er effekter av syntaktisk analyse (*parsing*) som oftest observerbare i sene målinger, men kan også oppstå i tidlige målinger (se f.eks. Staub (2007)). Derfor anbefaler forfatterne å inkludere en rekke forskjellige målinger i analysen, samt undersøke dataene fra flere interessefelter (for å sjekke for *spillover*-effekter). Jeg valgte å analysere de seks målingene som er nevnt ovenfor, fordi effekter av morfosyntaktiske avvik har blitt observert i én eller flere av disse målingene i lignende studier (se f.eks. Ni et al. (1998) og Braze et al. (2002)).

### 5.3 Rådata

Denne seksjonen skal handle om deskriptiv statistikk, altså beskrivelse av de rådataene jeg har fått fra eksperimentet. I statistisk analyse må man ofte transformere og bearbeide dataene på forskjellige måter, slik at den valgte statistiske modellen er bedre tilpasset dataene. Dette var også tilfellet i denne studien, men det kan likevel være nyttig å granske rådataene først for å få innsikt i hva slags data som ble produsert av eksperimentet. I tabeller 5.1 og 5.2 kan man finne gjennomsnittene og standardavvikene for de seks variablene beskrevet i forrige seksjon. Tabellene gjelder henholdsvis det verbale og det post-verbale interessefeltet. Radene representerer den uavhengige variabelen «betingelse»: K – kontrollbetingelse, F – finittavvik, T – tempusavvik, FT – finitt- pluss tempusavvik.

	First fixation duration (ms)		Gaze duration (ms)		Regression path duration (ms)		Dwell time (ms)		Fixation count		Regressions out
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	%
<b>K</b>	214	62	242	97	286	206	296	181	1.40	0.76	8.4
<b>F</b>	237	84	265	125	302	192	376	244	1.62	0.94	7.1
<b>T</b>	239	96	274	134	345	260	437	277	1.89	1.10	10.7
<b>FT</b>	241	93	276	137	324	234	407	252	1.74	1.00	8.4

Tabell 5.1: Blikkdata for det verbale interessefeltet

	First fixation duration (ms)		Gaze duration (ms)		Regression path duration (ms)		Dwell time (ms)		Fixation count		Regressions out
	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	<i>M</i>	<i>SD</i>	%
<b>K</b>	228	67	351	208	415	300	429	273	1.91	1.14	9.3
<b>F</b>	235	82	336	206	470	358	450	311	1.97	1.24	23.5
<b>T</b>	233	81	331	187	560	471	480	310	2.14	1.26	31.8
<b>FT</b>	234	75	338	200	495	376	470	307	2.09	1.27	27.5

Tabell 5.2: Blikkdata for det post-verbale interessefeltet

Allerede her kan vi se at i det verbale interessefeltet har nesten alle variabler høyere verdier i testbetingelser enn i kontrollbetingelsen. Dette bekrefter at informantene la merke til avvikene og at de gjorde det veldig fort, siden vi kan se en betydelig forskjell mellom grammatikalske og ugrammatikalske setninger allerede i varigheten på første fiksering i interessefeltet. Interessant nok er det bare tempus-betingelse som har en



høyere andel regresjoner fra det verbale interessefeltet enn kontrollbetingelsen, mens finitt-betingelse har faktisk færre regresjoner enn kontrollbetingelsen. Lesetidsforskjell i sene målinger og forskjell i antall fikseringer mellom ulike testbetingelser indikerer at visse typer avvik behøver mer tid enn andre for å bli kompensert.

I dataene for det post-verbale interessefeltet finner man et nokså forskjellig mønster: varigheten på første fiksering er nesten identisk i alle betingelser, mens *gaze duration* viser lengre lesetid for kontrollbetingelsen enn for noen av testbetingelsene. En mulig forklaring på et slikt «omvendt» resultat kan være at jo «verre» avviket er, jo fortere forlater personens blikk det post-verbale interessefeltet, for å gå tilbake i setningen og prøve å «reparere bruddet», noe som fører til kortere førstegangslesetid. Når det gjelder sene målinger, ser vi samme mønster som i det verbale interessefeltet: ugrammatikalske setninger får større verdier enn kontrollsetninger, og samtidig har tempus-betingelse de høyeste verdiene av alle testbetingelsene. I motsetning til det verbale interessefeltet, vedvarer dette mønsteret også i *regressions out*-variabelen.

En annen viktig observasjon er at betingelse F+T (tempus- pluss finittavvik) ikke har de høyeste verdiene av de tre testbetingelsene i noe interessefelt, slik det var forventet ifølge den alternative hypotesen. Det er altså klart – selv uten statistisk analyse – at prediksjonen om at betingelse F+T skal være den vanskeligste å prosessere og at den skal forårsake de lengste lesetidene ikke er i overenstemmelse med de faktiske resultatene fra studien.

Store standardavvik i alle målinger kan skyldes mye variasjon både mellom deltakere og stimulisetninger. Dette var en av grunnene til å transformere dataene på en måte som jeg skal beskrive nærmere i neste seksjon. Standardavvikene gir også en indikasjon av hvor mye «støy» det er i dataene, og det er derfor man trenger å bruke statistiske analysemodeller fremfor å bare sammenligne gjennomsnittene, ettersom statistisk analyse tar i betraktning forholdet mellom signal og støy for å trekke generaliserbare slutninger om effektens signifikans (Svartdal, 2015).

## 5.4 Statistisk analyse

### 5.4.1 Analysemodell

Valg av analysemodell avhenger av mange kriterier, som for eksempel forskningsspørsmål, antall avhengige og uavhengige variabler, variabelenes karaktertrekk og målenivå, studiedesign, forskerens kunnskap om statistikk, osv. I denne studien brukte jeg en enveis variansanalyse for repeterte målinger (heretter kalt

rm-ANOVA, fra engelsk *repeated-measures ANalysis Of VAriance*). Variansanalyse benyttes i nullhypotesetesting for å sammenligne flere grupper, og i enveistilfellet tester man kun én egenskap som varierer mellom gruppene. «Repeterte målinger» referer til innengruppe-studiedesignet som jeg beskrev i seksjon 4.2.3. Dette betyr at jeg ikke sammenlignet dataene fra flere forskjellige utvalg, men helst effekter av fire forskjellige betingelser i dataene fra ett og samme utvalg. Som navnet tilsier, sammenligner variansanalysen systematisk og tilfeldig varians i dataene for så å beregne sannsynligheten for at resultatene skyldes den eksperimentelle manipulasjonen.

Den uavhengige variabelen i denne studien er kategorisk og blir regnet i analysen som en innengruppefaktor med fire nivå (betingelser K, F, T og F+T). Avhengige variabler er de seks målingene som er beskrevet i seksjon 5.2; fem av disse variablene er kontinuerlige (målt i millisekunder eller antall), mens *regressions out* er en binær variabel (enten oppstår det en regresjon fra et gitt interessefelt eller ikke). Når det gjelder binære utfallsvariabler, er logistisk regresjonsanalyse å foretrekke fremfor ANOVA. Imidlertid, grunnet begrenset tid og erfaring, valgte jeg å utføre en rm-ANOVA også for denne variabelen. ANOVA har blitt brukt for binære variabler i lignende studier (f.eks. Warren et al. (2008)), og Mai (2018) gir en detaljert forklaring på hvordan man kan utføre en variansanalyse med binære- og tellelige data.

I de tilfellene der rm-ANOVA viste en signifikant effekt ( $\alpha = .05$ ),<sup>27</sup> altså at det var en signifikant forskjell mellom noen av betingelsene, utførte jeg også parvise sammenligninger for å beregne interaksjonen mellom alle betingelsene. Parvise sammenligninger er i bunn og grunn en rekke tohalet *t*-tester<sup>28</sup> som sammenligner dataene fra hvert par betingelser. Fordi multiple sammenligninger øker risiko for type I feil<sup>29</sup> (Abdi, 2007), som i slike tilfeller kalles *familywise error rate*, brukte jeg Šidák-korreksjon for å justere *p*-verdien.

---

<sup>27</sup> Signifikansnivå  $\alpha$  er en grenseverdi for å forkaste nullhypotesen. *p*-verdien må være lik eller lavere enn alfa-nivået for å forkaste nullhypotesen.

<sup>28</sup> *T*-test er en parametrisk test utviklet for å identifisere forskjeller mellom to gjennomsnittskårer. Tohalet *t*-test brukes for å teste nullhypotesen mot en ikke-retningsbestemt hypotese (i motsetning til en enhalet *t*-test, som brukes for å teste  $H_0$  mot en retningsbestemt  $H_1$ ). En tohalet *t*-test regnes som signifikant hvis *t*-verdien faller under en av de to endene av *t*-fordelingen (herav navnet «tohalet»).

<sup>29</sup> En type I feil eller en forkastningsfeil er en statistisk feil som består i en feilaktig avvisning av nullhypotesen (Wooldridge, 2012).

Statistikkpakke IBM SPSS Statistics, versjon 27 (IBM Corp., Armonk, N.Y., USA) ble brukt for å utføre analysen.

#### 5.4.2 Transformasjon av data

Alle statistiske modeller har visse forutsetninger angående dataene som skal analyseres. Hvis dataene unnlater å tilfredsstillere noen av forutsetningene, kan dette føre til feilkilder i modellen og feilaktige slutninger. Fordi ANOVA bygger på den generelle lineære modellen, har ANOVA samme forutsetninger, inkludert linearitet, additivitet, normalfordeling og homoskedastisitet av residualer (Field, 2013). Additivitet og linearitet betyr at det er et lineært forhold mellom variabler, og at en modell basert på flere variabler best representeres ved at effektene av disse variablene legges sammen. Normalfordeling og homoskedastisitet av residualer<sup>30</sup> betyr at avhengige variabler må være omtrent normalfordelt innenfor de ulike kategoriene angitt av den uavhengige variabelen, og at variansen til residualer innenfor de ulike kategoriene bør være tilnærmet lik. Det er vanlig for eye-tracking-data å tilfredsstillere de to første forutsetningene, men disse dataene pleier å ha en betydelig positivt skjev fordeling og heterogen varians, som skyldes interindividuelle variasjoner i lesetempo og lesemønstre (Holmqvist et al., 2011).

Slik skjevhet og heteroskedastisitet av residualer var også til stede i mitt datasett: dette var klart både fra visuell datautforskning (histogrammer og plotter) og fra resultater av signifikanstester (Kolmogorov-Smirnov og Levenes tester). I slike tilfeller kan man transformere dataene slik at de tilfredsstillere forutsetningene, bruke ikke-parametriske tester (f.eks. Friedmans ANOVA), eller bruke «bootstrapping».<sup>31</sup> Jeg valgte å transformere datasettet mitt ved hjelp av den naturlige logaritmefunksjonen, fordi slik transformasjon kan korrigere heteroskedastisitet, samt både positiv skjevhet og positiv kurtose (Field, 2013, s. 203). Ved å ta den naturlige logaritmen til hver variabel, klarte jeg å endre skalaen og fikk variablene til å nærme seg normalfordeling og homoskedastisitet. Dette gjør at modellen er bedre tilpasset dataene. Etter transformasjonen bør dataene sjekkes igjen for normalitet og homoskedastisitet. De fleste av logaritmetransformerte målingene fikk fortsatt signifikante resultater i Kolmogorov-Smirnov og Levenes tester, men det er viktig å huske at disse testene er

---

<sup>30</sup> Residualer er forskjellen mellom observerte og forventede verdier (altså middelverdien) for den avhengige variabelen (Field, 2013).

<sup>31</sup> «Bootstrapping» er imidlertid ikke tilgjengelig for rm-ANOVA i SPSS.

forutinntatte med hensyn til utvalgsstørrelse – de viser signifikans selv for de minste effektene, når utvalget er stort nok (Field, 2013, s. 184). Derfor valgte jeg å sjekke histogrammer og Q-Q plotter (som viste en fordeling tilnærmet normal) og tallene for skjevhet og kurtose.<sup>32</sup> Skjevheten for alle variabler var mellom 0 og 1, og kurtoseverdier var mellom -0.5 og 3. Til sammenligning var skjevhetsverdiene for utransformerte variabler mellom 1 og 5, og kurtoseverdier var mellom 3 og 30. Når det gjelder homoskedastisitet, utforsket jeg plotter av residualer mot predikerte verdier (*zpred vs. zred* plotter), som viste at residualer av transformerte variabler hadde en ganske homogen varians. Jeg regnet derfor med at etter transformasjonen var disse variablene tilstrekkelig normalfordelte og homoskedastiske for å tilfredsstillere forutsetningene av ANOVA-modellen.

En annet potensiell kilde av bias i dataene kan være ekstremverdier. Noen av disse kan være forårsaket av naturlig variasjon i utvalget, mens andre indikerer en feil i målinger. Slike feilmålinger i dataene kan forkastes individuelt, mens skårene som skiller seg veldig mye fra resten av dataene kan forkastes ved hjelp av trimming, der man kan enten ta bort de øverste og nederste 5%, 10%, osv. fra datasettet, eller de skårene som ligger mer enn  $2.5SD$  eller  $3SD$  fra middelveidien (Field, 2013). I mitt datasett var det ingen «feilmålinger», og ingen av ekstremverdiene var veldig utpreget – alle forskjeller i lesetidsskårer var akseptable som interindividuell variasjon i lese mønstre. Jeg valgte derfor å ikke trimme dataene mine, også fordi ANOVA er robust mot milde avvik fra dets forutsetninger. I tillegg kan jeg utnytte sentralgrenseteoremet (Field, 2013, s. 169), som tilsier at fordi datautvalget mitt er så stort, skal det ha en normal samplingsfordeling, uansett hvilken fordeling selve utvalget faktisk har. Dette er viktig å ta i betraktning, fordi det er nettopp denne samplingsfordelingen som må være normalfordelt for å beregne middelveidier, signifikansnivåer og konfidensintervaller nøyaktig.

Den siste forutsetningen for rm-ANOVA er sfæriskitet, som betyr at variansen til differansen mellom alle mulige par av betingelser bør være tilnærmet lik for alle parkombinasjoner (Field, 2013). Denne forutsetningen sjekkes ved hjelp av Mauchly test, og hvis testen er signifikant, må  $F$ -forholdet justeres. Jeg brukte Greenhouse-Geisser-justeringen i de tilfellene der Mauchly test viste signifikans.

---

<sup>32</sup> En normalfordelt variabel får skjevhets- og kurtoseverdi mellom -1 og 1 i SPSS (Ghasemi & Zahediasl, 2012).

Til slutt vil jeg nevne at identiske tester gjennomført med uttransformerte variabler viste samme resultatmønster som testene med transformerte variabler. Dette bekrefter at effektene funnet i denne analysen faktisk er til stede i dataene, og at de ikke var forårsaket av datatransformasjonen.

### Sammenslåing av data

Grunnet temmelig komplisert motbalansering som ble brukt i dette studiedesignet, inneholder datasettet mitt flere datapunkter for hver stimulus og hver deltaker. I variansanalysen er det ikke mulig å bruke alle disse datapunktene; derfor måtte datasettet slås sammen på tvers av enten stimuliene eller informantene, selv om dette gjør dataene mindre detaljerte. Gonzalez-Marquez (2007) fremhever at i de tilfellene hvor data sammenslås på tvers av informanter, kan resultatene generaliseres til andre personer som ikke har deltatt i studien, mens i de tilfellene hvor data blir sammenslått på tvers av stimuli, kan resultater generaliseres til andre stimuli, som ikke ble brukt i studien. Jeg valgte å slå sammen dataene mine på tvers av deltakere: variabler som ble analysert i rm-ANOVA var altså middelerdiene for hver stimulisett ( $n = 40$ ) i fire betingelser.

#### 5.4.3 Effektstørrelser

I håp om å unngå den alt-eller-ingen-tenkningen som er forfremmet av nullhypotese signifikanstesting der fokuset ligger kun på  $p$ -verdien (Sullivan & Feinn, 2012), vil jeg også rapportere standardiserte effektstørrelser for alle mine funn. Effektstørrelser kan bli brukt for å sammenligne vekten til de effektene som jeg fant i studien min, selv om de er målt på forskjellige måter (f.eks. i millisekunder og i antall), samt for å sammenligne mine funn med effekter funnet i andre lignende studier. Jeg vil rapportere partiell eta kvadrert for hovedanalyseeffekter (rm-ANOVA). Denne effektstørrelsen gjenspeiler hvor mye av den totale variasjonen er forårsaket av den uavhengige variabelen. Imidlertid, fordi hypotesen min gjelder samspillet mellom de fire forskjellige betingelsene, altså hvordan de er forskjellige fra hverandre, ble de viktigste effektstørrelsene beregnet for parvise sammenligninger. Jeg valgte å beregne Cohens  $d$ , og jeg tok i bruk en av formlene anbefalt av Lakens (2013, s. 4) som er tilpasset design med repeterte målinger:

$$d_{rm} = \frac{|M_1 - M_2|}{\sqrt{SD_1^2 + SD_2^2 - 2 \times r \times SD_1 \times SD_2}} \times \sqrt{2(1-r)}$$

I denne formelen er  $r$  korrelasjonen mellom to betingelser, som betyr at denne effektstørrelsen tar i betraktning korrelasjonen mellom ulike målinger. Cohen (1988) klassifiserer relativ størrelse på partiell eta kvadrert som 0.01 = liten, 0.06 = moderat, 0.14 = stor effekt, og størrelsen på Cohens  $d$  som 0.2 = liten, 0.5 = moderat, 0.8 = stor effekt. Imidlertid er det best å fokusere på sammenligninger av effektstørrelser i forskjellige funn, siden «store» og «små» effekter kan ha ganske forskjellige verdier innen forskjellige fag, samt i forskjellige studiedesign innen samme fag.

En annen måte å anslå effektstørrelsen på i en gitt studie, er å se på konfidensintervallene til middelerverdiene (Durlak, 2009). Konfidensintervallet er det sannsynlighetsområdet som et nytt middeltall bør befinne seg i, hvis eksperimentet ble gjentatt. Usikkerhetsstolper i alle stolpediagrammer i seksjoner 5.4.4 og 5.4.5 representerer 95% konfidensintervallene, som betyr at 95% av de utvalgene som i prinsippet kan trekkes fra populasjonen skal ha en middelerverdi innen dette intervallet. Som følge av dette indikerer kortere feilfeltene større effekter.

#### 5.4.4 Resultater for det verbale interessefeltet<sup>33</sup>

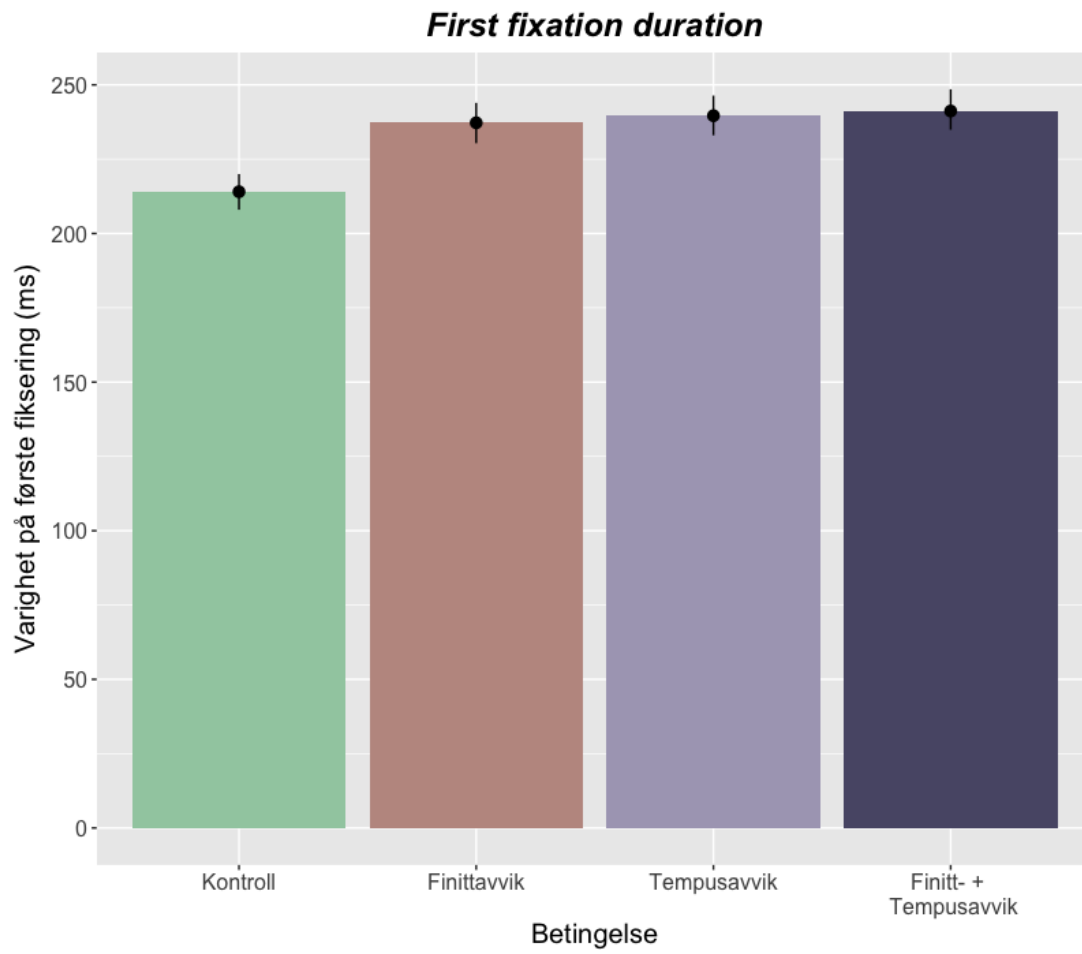
##### *i. First fixation duration*

Resultatene av rm-ANOVA viste at det var en statistisk signifikant forskjell mellom minst to av betingelsene,  $F(3,117) = 13.03$ ,  $p < .001$ ,  $\eta_p^2 = .250$ . Parvise sammenligninger<sup>34</sup> demonstrerte at det kun var en signifikant forskjell mellom betingelsene K og F ( $t(39) = -5.12$ ,  $p < .001$ ,  $d_{rm} = 1.09$ ), betingelsene K og T ( $t(39) = -4.65$ ,  $p < .001$ ,  $d_{rm} = 1.00$ ), og betingelsene K og F+T ( $t(39) = -5.79$ ,  $p < .001$ ,  $d_{rm} = 1.19$ ). Dette betyr at første fiksering på kontrollsetninger var vesentlig kortere enn første fiksering på ugrammatikalske setninger. Det var imidlertid ingen forskjell i varigheten på første fiksering mellom de tre testbetingelsene (alle  $p$ -verdier  $\geq .992$ ).

---

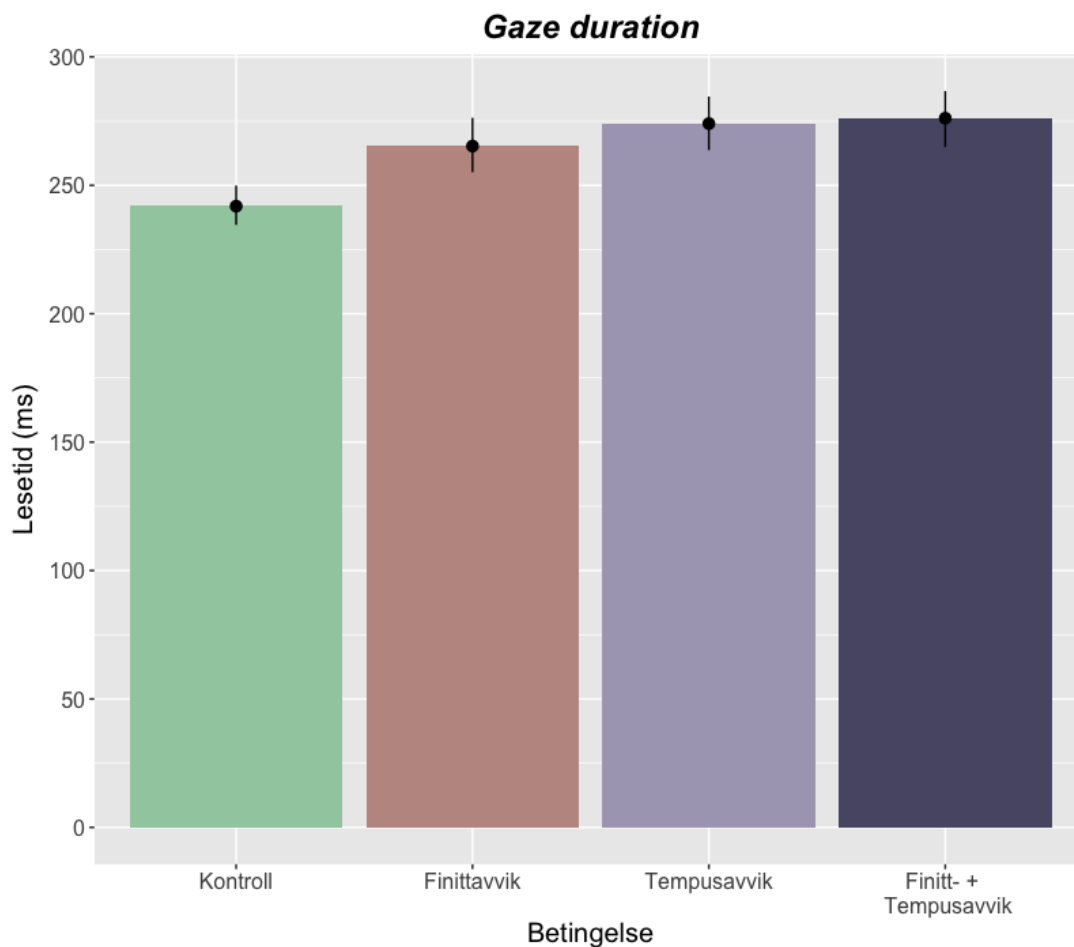
<sup>33</sup> Stolpediagrammer i denne og neste seksjonen fremstiller gjennomsnittene for uttransformerte variabler, ettersom transformerte variabler ikke gir en veldig god illustrasjon av resultatene.

<sup>34</sup> Komplette tabeller for parvise sammenligninger er å finne i vedlegg B.



### ii. Gaze duration

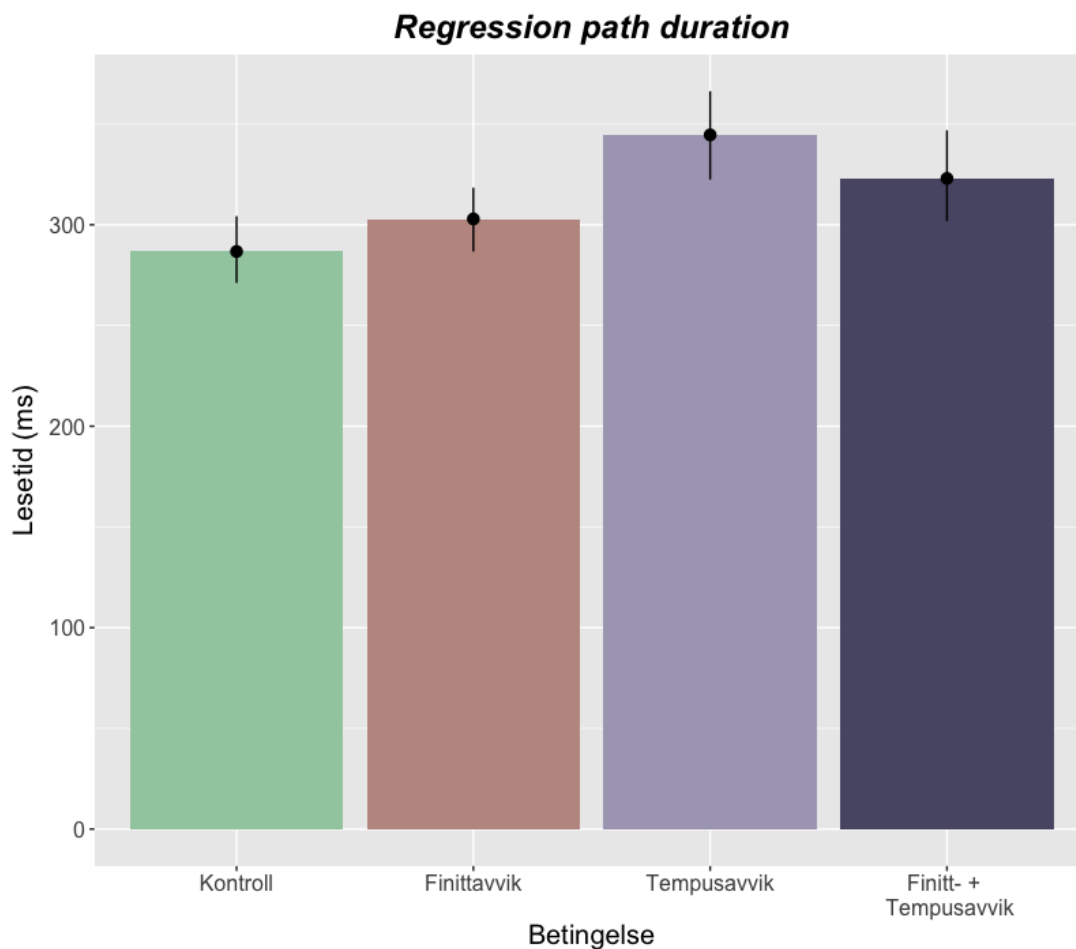
Resultatene av rm-ANOVA viste at det var en statistisk signifikant forskjell mellom minst to av betingelsene,  $F(3,117) = 8.21, p < .001, \eta_p^2 = .174$ . Parvise sammenligninger tilsa at det var en statistisk signifikant forskjell mellom betingelsene K og F ( $t(39) = -3.36, p = .011, d_{rm} = 0.75$ ), betingelsene K og T ( $t(39) = -4.54, p < .001, d_{rm} = 0.93$ ), og mellom betingelsene K og F+T ( $t(39) = -4.50, p < .001, d_{rm} = 1.01$ ). Disse resultatene viste samme mønster som resultatene for *first fixation duration* – den eneste signifikante interaksjonen var mellom kontrollbetingelsen og hver av testbetingelsene.





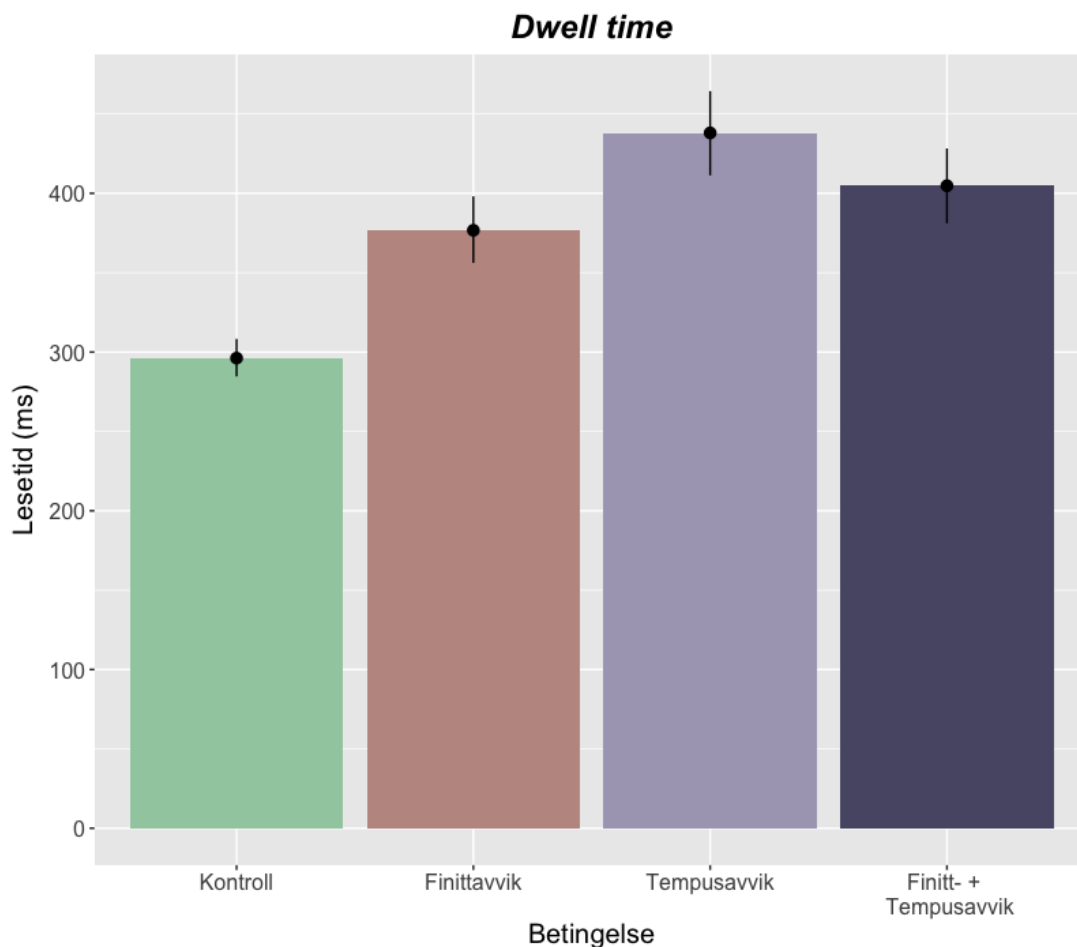
*iii. Regression path duration*

Resultatene av rm-ANOVA viste at det var en statistisk signifikant forskjell mellom minst to av betingelsene,  $F(3,117) = 7.52, p < .001, \eta_p^2 = .162$ . Parvise sammenligninger tilsa at det var en signifikant forskjell mellom betingelsene K og T ( $t(39) = -4.66, p < .001, d_{rm} = 1.00$ ), betingelsene K og F+T ( $t(39) = -2.96, p = .031, d_{rm} = 0.66$ ), samt mellom betingelsene F og T ( $t(39) = -3.28, p = .013, d_{rm} = 0.61$ ). Forskjellen mellom betingelsene K og F var ikke signifikant, men hadde uansett en moderat effektstørrelse ( $t(39) = -2.08, p = .04, d_{rm} = 0.44$ ). Disse resultatene viser forskjeller i lesetiden mellom de tre avvikene som utgjorde de ulike betingelsene: betingelse F forårsaket den korteste lesetiden av alle tre og var ikke vesentlig forskjellig fra kontrollbetingelsen. Betingelse T viste den lengste lesetiden og var den eneste betingelsen som hadde en «stor» effekt. Lesetiden for betingelse T var vesentlig lengre enn lesetiden i både betingelse K og F. Lesetiden for betingelse F+T lå mellom lesetidene for betingelser F og T, men var bare vesentlig forskjellig fra kontrollbetingelsen.



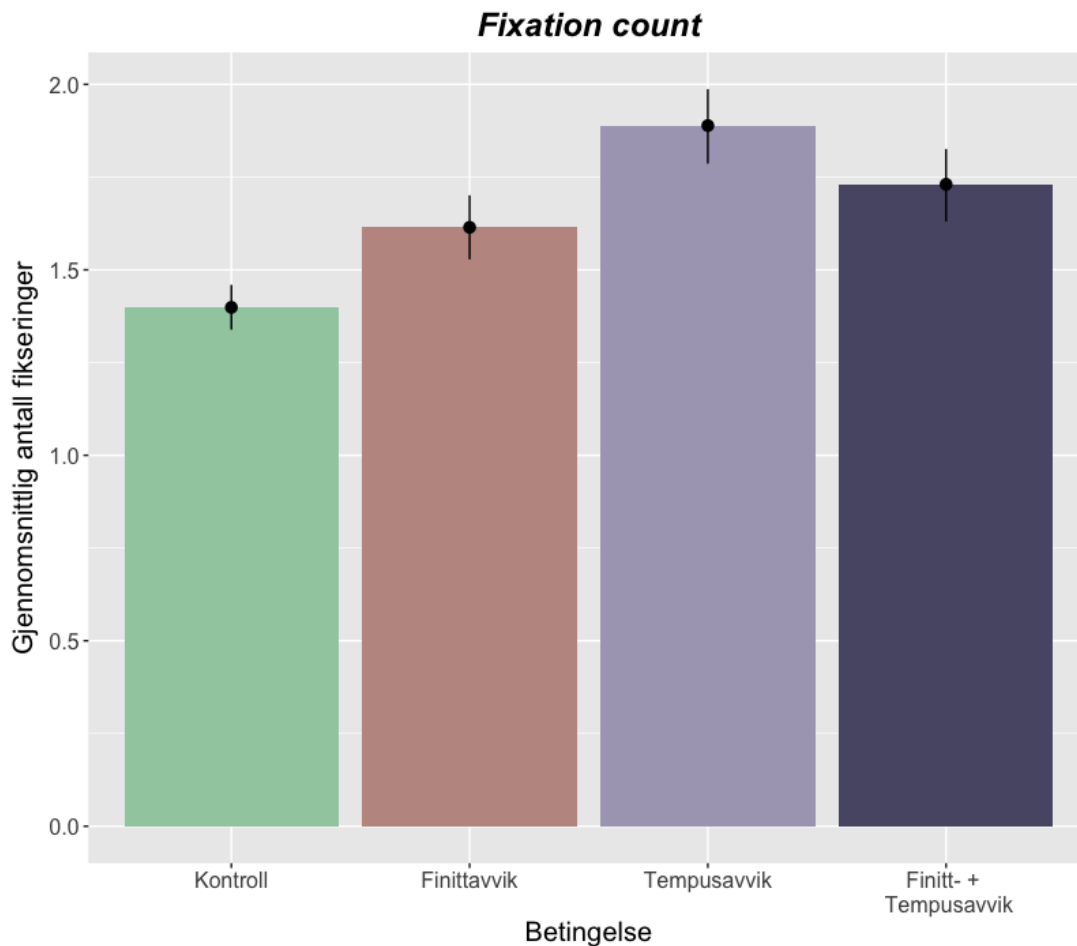
*iv. Dwell time*

For denne målingen anga Mauchlys test at sfærishet ikke var tilfredsstillende,  $\chi^2(5) = 17.08$ ,  $p = .004$ . Derfor måtte frihetsgrader justeres med Greenhouse-Geisser estimat av sfærishet,  $\epsilon = .79$ . Justert  $F$ -verdien viste en signifikant forskjell mellom minst to av betingelsene,  $F(2.37, 92.24) = 36.07$ ,  $p < .001$ ,  $\eta_p^2 = .480$ . Parvise sammenligninger påviste en vesentlig forskjell mellom betingelsene K og F ( $t(39) = -7.23$ ,  $p < .001$ ,  $d_{rm} = 1.56$ ), K og T ( $t(39) = -12.09$ ,  $p < .001$ ,  $d_{rm} = 2.25$ ), K og F+T ( $t(39) = -8.60$ ,  $p < .001$ ,  $d_{rm} = 1.80$ ), samt mellom betingelsene F og T ( $t(39) = -3.98$ ,  $p = .002$ ,  $d_{rm} = 0.80$ ). Her også ser vi samme mønster som i resultatene ovenfor, der alle ugrammatikalske betingelser har en vesentlig lengre total lesetid enn kontrollbetingelsen, med lengste lesetiden i betingelse T. Betingelse F har en vesentlig kortere total lesetid enn betingelse T, men hvis vi sammenligner effektstørrelser, så ser vi en mye større effekt for interaksjonen mellom grammatikalske og ugrammatikalske setninger, enn for interaksjonen mellom finitt- og tempusavvikene.



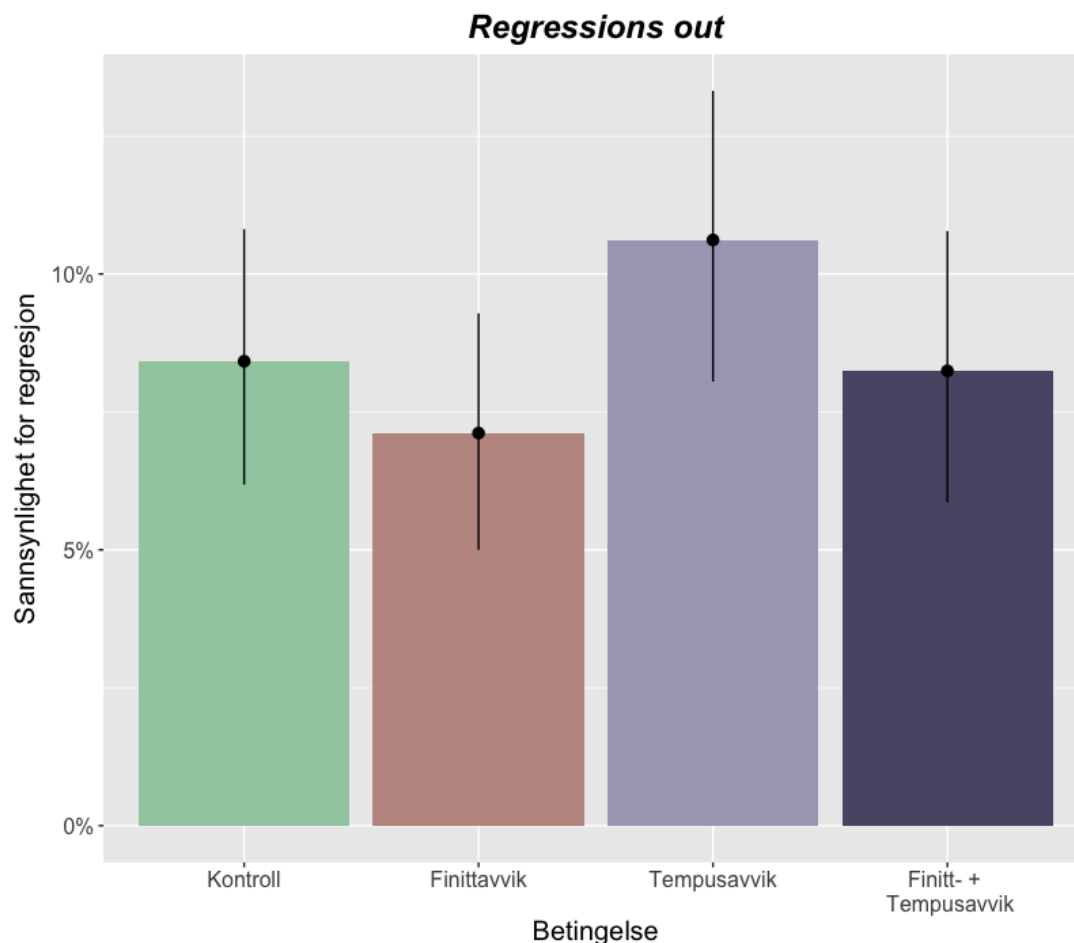
v. *Fixation count*

Resultatene av rm-ANOVA viste at det var en statistisk signifikant forskjell mellom minst to av betingelsene,  $F(3,117) = 23.81$ ,  $p < .001$ ,  $\eta_p^2 = .379$ . Parvise sammenligninger ga evidens for en vesentlig forskjell mellom betingelsene K og F ( $t(39) = -4.10$ ,  $p = .001$ ,  $d_{rm} = 0.92$ ), K og T ( $t(39) = -9.27$ ,  $p < .001$ ,  $d_{rm} = 1.98$ ), K og F+T ( $t(39) = -6.32$ ,  $p < .001$ ,  $d_{rm} = 1.30$ ), samt mellom betingelsene F og T ( $t(39) = -4.63$ ,  $p < .001$ ,  $d_{rm} = 0.96$ ). Dette er samme resultatmønsteret som for *dwelt time*-variabelen. Jeg vil spesielt påpeke at i denne målingen er effektstørrelsen for interaksjoner mellom betingelsene K og F, og F og T nesten identisk: dette tilsier at tempusavviket var nesten like forskjellig fra finittavviket som setninger med finittavvik var forskjellig fra grammatikalske setninger, med tanke på prosesseringsvansker.



### *vi. Regressions out*

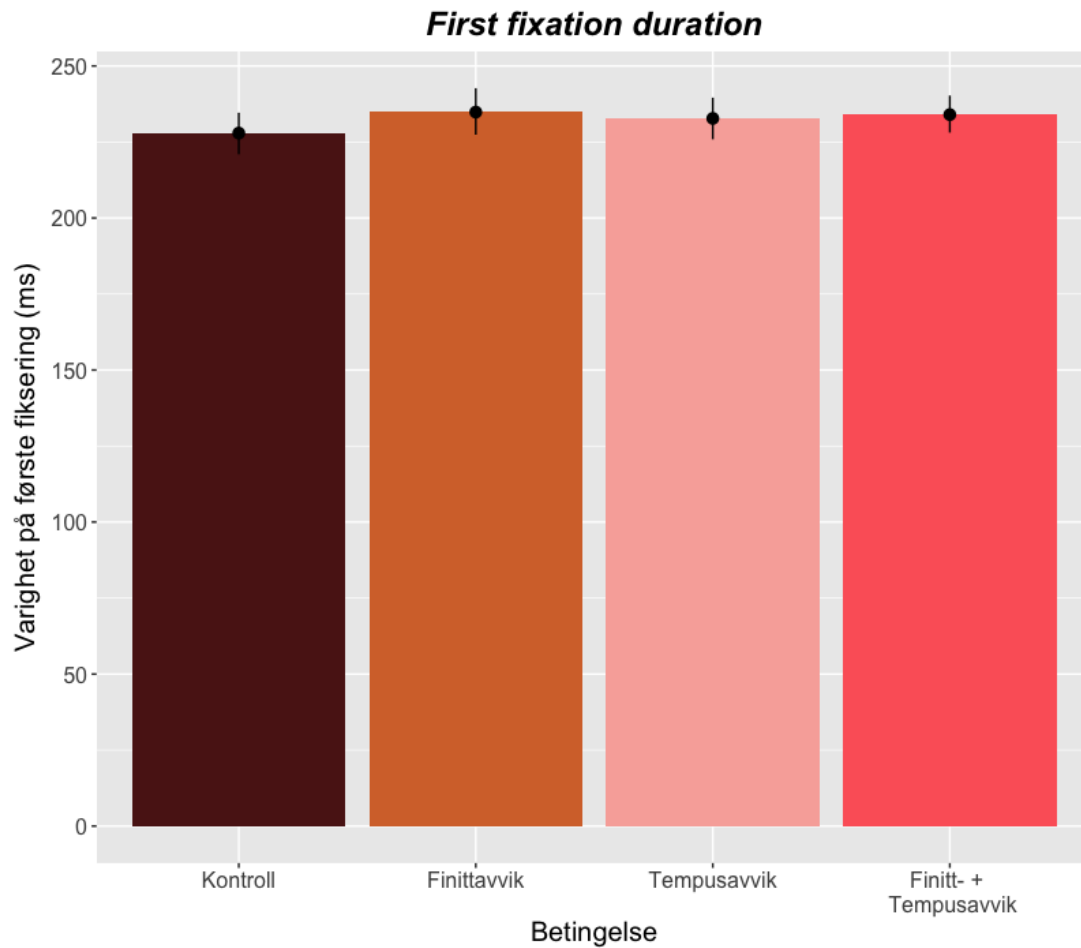
Resultatene av rm-ANOVA viste at det ikke var noen signifikant forskjell mellom noen av betingelsene,  $F(3,117) = 1.30, p = .277, \eta_p^2 = .032$ . Nå i ettertid virker det som at denne målingen er bedre tilpasset det post-verbale interessefeltet, på grunn av ulikheter i de pre-verbale delene av stimulisetningene: sannsynligheten for at en regresjon oppstår fra det verbale interessefeltet var kanskje mer avhengig av hvor mye informasjon som var tilgjengelig før dette feltet enn av testbetingelsen. I de setningene der veldig lite informasjon ble presentert før testverbet, var det mer sannsynlig at informantene ville fortsette å lese etter at de hadde sett den avvikende verbformen for å lete etter klargjørende informasjon videre i setningen. Det var når dette forsøket mislyktes at deltakerne vendte tilbake i setningen, ofte fra det post-verbale interessefeltet.



### 5.4.5 Resultater for det post-verbale interessefeltet

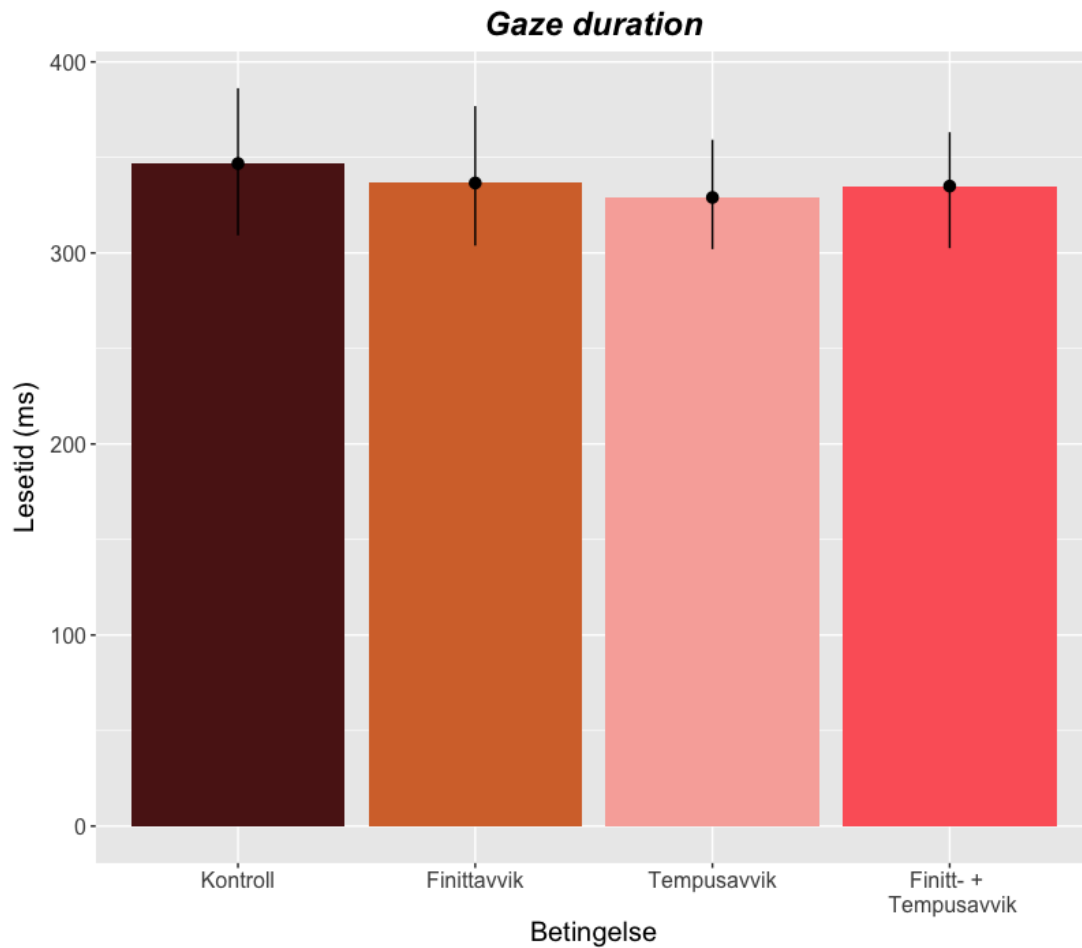
#### *i. First fixation duration*

Resultatene av rm-ANOVA viste ingen signifikant forskjell mellom noen av betingelsene,  $F(3,117) = 0.43$ ,  $p = .73$ ,  $\eta_p^2 = .011$ .



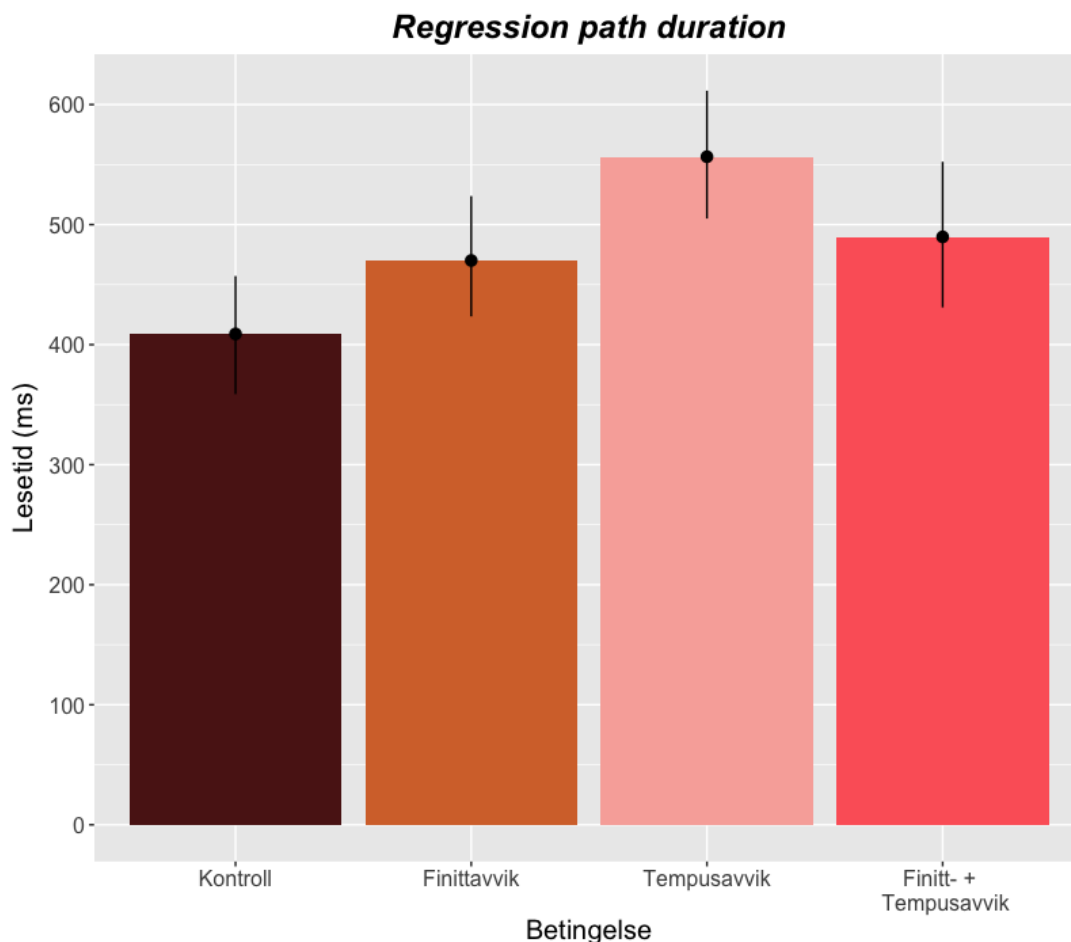
*ii. Gaze duration*

Resultatene av rm-ANOVA viste at det ikke var noen signifikant forskjell mellom noen av betingelsene,  $F(3,117) = 0.80, p = .49, \eta_p^2 = .020$ .



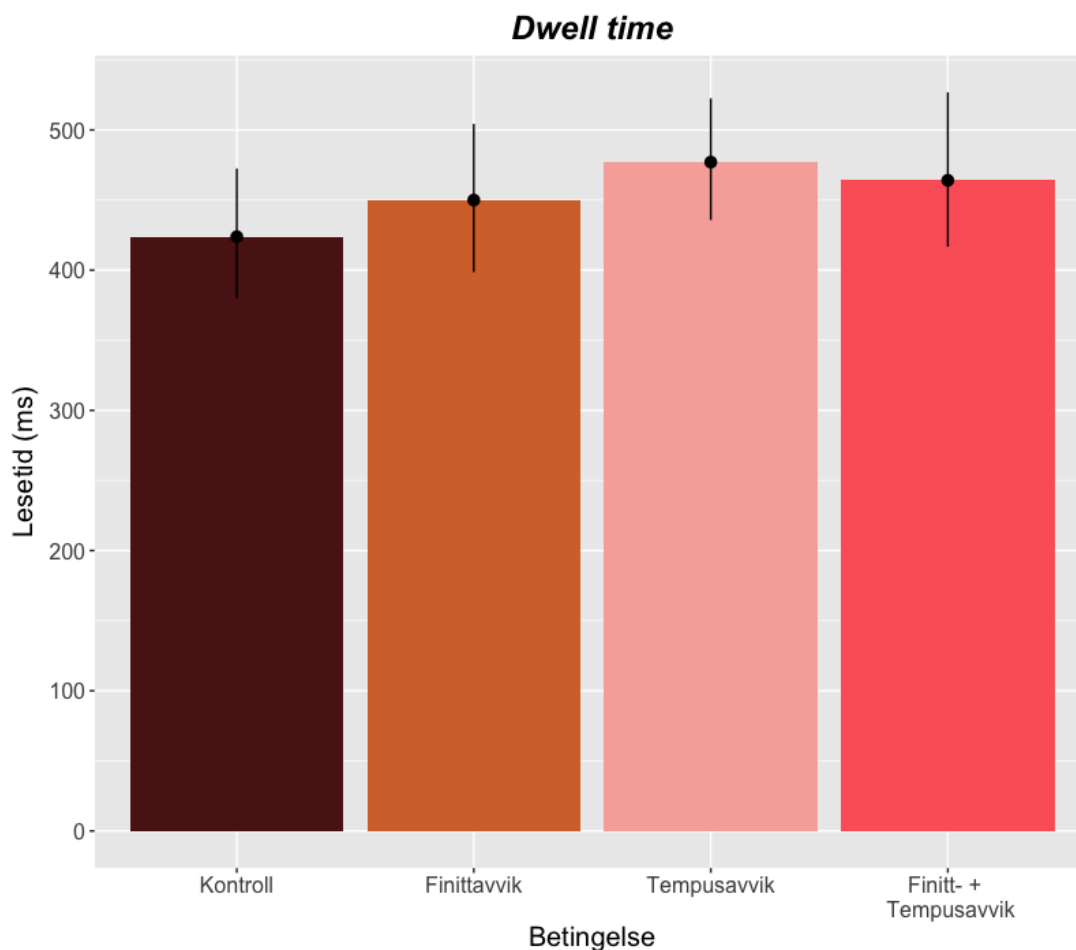
*iii. Regression path duration*

Resultatene av rm-ANOVA viste at det var en statistisk signifikant forskjell mellom minst to av betingelsene,  $F(3,117) = 14.28$ ,  $p < .001$ ,  $\eta_p^2 = .268$ . Parvise sammenligninger påviste en vesentlig forskjell mellom betingelsene K og F ( $t(39) = -3.47$ ,  $p = .008$ ,  $d_{rm} = 0.34$ ), K og T ( $t(39) = -6.33$ ,  $p < .001$ ,  $d_{rm} = 0.71$ ), K og F+T ( $t(39) = -4.55$ ,  $p < .001$ ,  $d_{rm} = 0.45$ ), samt mellom betingelsene F og T ( $t(39) = -3.90$ ,  $p = .002$ ,  $d_{rm} = 0.40$ ). Selv om vi ser samme mønster i disse resultatene som i resultatene for den tilsvarende variabelen i det verbale interessefeltet, er effektstørrelsene vesentlig mindre. Derimot, akkurat som i *fixation count*-variabelen for det verbale interessefeltet, kan vi se like effektstørrelser for interaksjoner mellom betingelsene K og F, og F og T, noe som antyder at forskjell i den kognitive innsatsen som kreves for å prosessere en grammatisk setning vs. en setning med finittavvik er nesten likt forskjellen mellom prosesseringsvansker forårsaket av setninger med finitt- og tempusavvik.



### *iv. Dwell time*

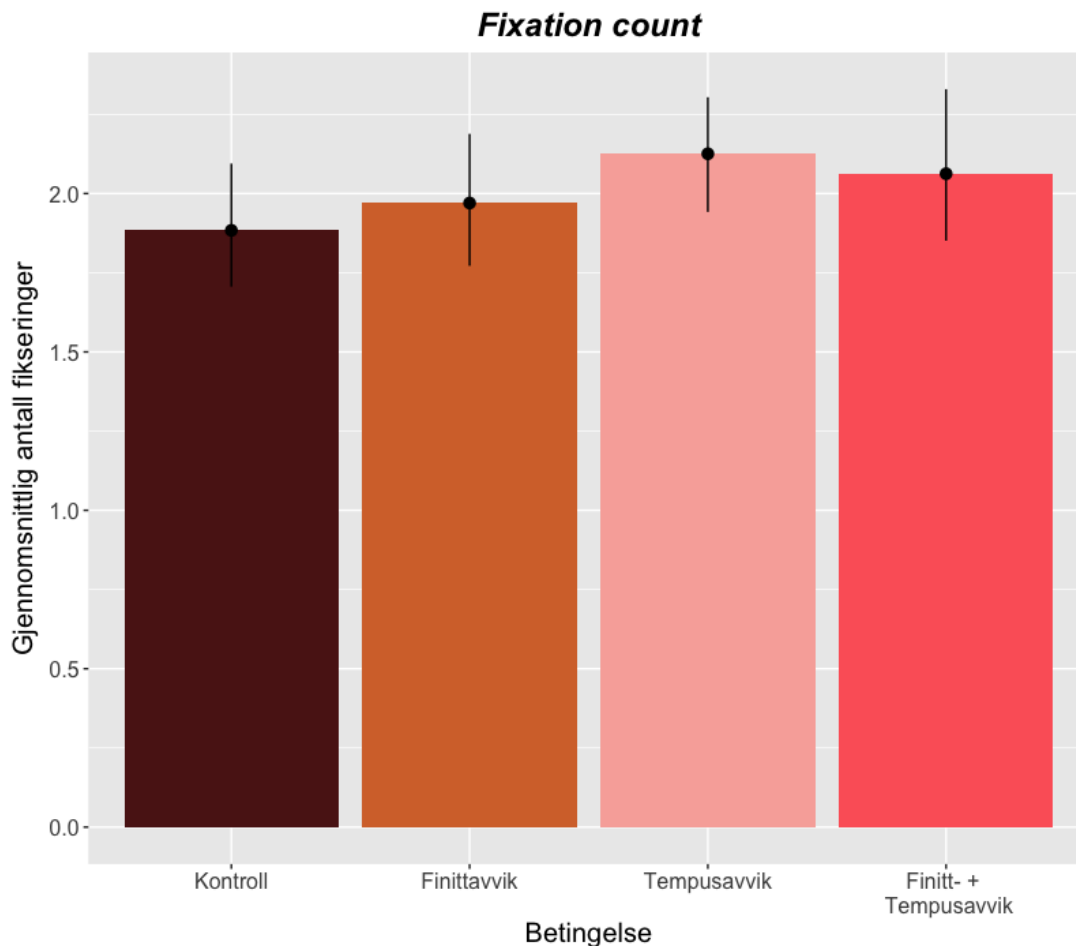
Resultatene av rm-ANOVA viste at det var en statistisk signifikant forskjell mellom minst to av betingelsene,  $F(3,117) = 3.82, p = .012, \eta_p^2 = .089$ . Parvise sammenligninger ga evidens for en vesentlig forskjell mellom betingelsene K og T ( $t(39) = -3.25, p = .014, d_{rm} = 0.33$ ) og mellom betingelsene K og F+T ( $t(39) = -2.83, p = .043, d_{rm} = 0.25$ ). Det var ikke noen signifikant interaksjon mellom betingelsene K og F ( $t(39) = -1.41, p = .666, d_{rm} = 0.14$ ), og heller ikke mellom noen av testbetingelsene (alle  $p$ -verdier  $\geq .418$ ). Den eneste signifikante forskjellen var altså mellom setninger som inneholdt tempusavvik og grammatikalske setninger, men også her er effektstørrelsene ganske små. Det er mulig at ufullkommenheter i eksperimentdesignet har bidratt til små effektstørrelser: det post-verbale interessefeltet varierte ganske mye i lengde og innhold i forskjellige stimuli, noe som kunne ha ført til økt støy i dataene og mindre effektstørrelser.





v. *Fixation count*

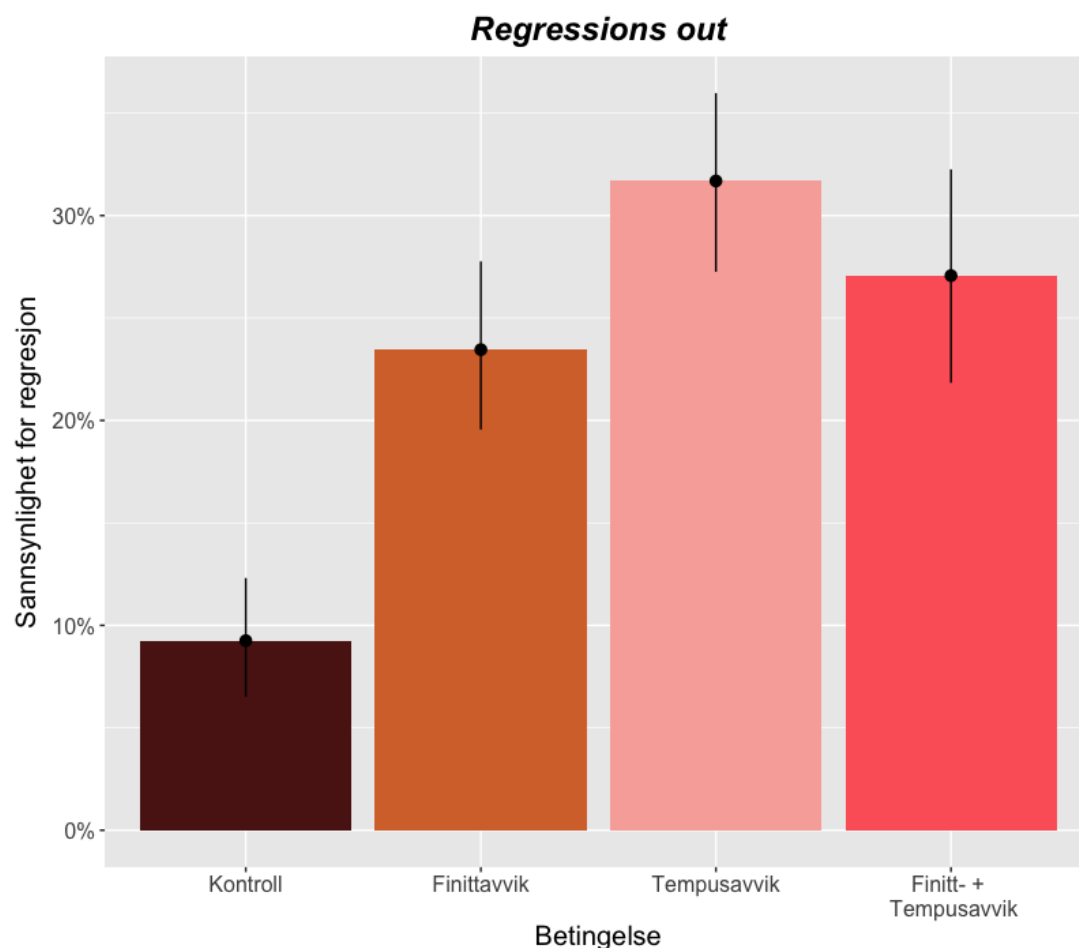
Resultatene av rm-ANOVA viste at det var en statistisk signifikant forskjell mellom minst to av betingelsene,  $F(3,117) = 6.29, p = .001, \eta_p^2 = .139$ . Parvise sammenligninger ga evidens for en vesentlig forskjell mellom betingelsene K og T ( $t(39) = -3.95, p = .002, d_{rm} = 0.38$ ) og mellom betingelsene K og F+T ( $t(39) = -3.28, p = .013, d_{rm} = 0.27$ ). Interaksjonen mellom betingelsene F og T var nær grensen for å være signifikant,  $t(39) = -2.68, p = .063, d_{rm} = 0.26$ . Det var ingen signifikant forskjell mellom betingelser K og F,  $t(39) = -1.60, p = .523, d_{rm} = 0.12$ . Generelt sett ligner disse resultatene på resultatene for variabelen *dwelt time*. Her også er det interessant å se en sterkere interaksjon mellom betingelsene F og T enn mellom betingelsene K og F.



vi. *Regressions out*

Mauchlys test anga at sfærisitetsforutsetningen ikke var tilfredsstillt,  $\chi^2(5) = 21.78, p = .001$ . Frihetsgrader måtte derfor justeres med Greenhouse-Geisser estimat av sfærisitet,  $\epsilon = .76$ . Resultatene av rm-ANOVA viste at det var en statistisk signifikant forskjell mellom minst to av betingelsene,  $F(2.28,88.90) = 22.67, p < .001, \eta_p^2 = .368$ . Parvise sammenligninger tilsa en vesentlig forskjell mellom betingelsene K og F ( $t(39)$ )

=  $-4.99$ ,  $p < .001$ ,  $d_{rm} = 1.08$ ), K og T ( $t(39) = -9.89$ ,  $p < .001$ ,  $d_{rm} = 1.82$ ), og K og F+T ( $t(39) = -7.26$ ,  $p < .001$ ,  $d_{rm} = 1.22$ ), samt mellom betingelsene F og T ( $t(39) = -3.28$ ,  $p = .013$ ,  $d_{rm} = 0.60$ ). Denne målingen viste de største effektene for det post-verbale interessefeltet: veldig store effekter for interaksjoner mellom kontroll- og testbetingelser, og en moderat effekt for interaksjonen mellom betingelser F og T. Som nevnt før er disse resultatene mest sannsynlig forårsaket av deltakernes forsøk å klargjøre avvikene i målverbets morfologi. Forskjellen i effektstørrelser mellom de to regresjonsmålingene for det post-verbale interessefeltet kan indikere at sannsynligheten for regresjon er avhengig først og fremst av typen av avvik, mens varigheten på regresjonen kommer an på både betingelsen og på mengde informasjon tilgjengelig i den tidligere delen av setningen (ulike pre-verbale setningsdeler kunne ha bidratt med tilfeldig varians i dataene, noe som også kunne ha ført til mindre effektstørrelser for *regression path duration*-variabelen).



### 5.5 Chapter summary: Results

The first step in the data analysis was cleaning the data: because one participant had to be excluded from the analysis, the original data pool was comprised of 2520 trials (63

participants  $\times$  40 sentences). One trial had to be discarded due to experimenter error, and 77 trials due to track loss in one of the areas of interest (target verb or post-verb AOI). Fixations shorter than 80ms were merged with longer fixations located within  $0.5^\circ$  from them, other fixations under 80ms, as well as fixations longer than 800ms were removed (4.18% of the original data pool). After the data for the verb and the post-verb AOIs were extracted, the AOIs with missing data (that were skipped during first pass reading) were removed (184 verb and 75 post-verb AOIs).

Six eye-tracking measures were calculated for both AOIs: first fixation duration, gaze duration, regression path duration, dwell time, fixation count, and regressions out. Descriptive statistics for these measures are presented in tables 4.1 (verb AOI) and 4.2 (post-verb AOI) in section 4.3.

One-way repeated-measures ANOVA tests were performed on the data. In cases where ANOVA returned significant results ( $\alpha = .05$ ), pairwise comparisons were used to analyse the relationship between each pair of conditions. Šidák correction was used to control for the familywise error rate. In addition, effect sizes were calculated: partial eta-squared for the main effects and Cohen's  $d$  for repeated-measures for the pairwise comparisons. The data were collapsed across participants before the analysis; the dataset used for the rm-ANOVA consisted of the means for each item ( $n = 40$ ) in four conditions. Due to a heavy positive skew in the data distribution and heterogenous variances of the residuals, the data were natural log transformed, which made them approach a normal distribution. No trimming or winsorizing was performed. Identical tests performed on untransformed data showed the same pattern of results.

Sections 5.4.4 and 5.4.5 present the results of the rm-ANOVA and pairwise comparisons for the verb and post-verb AOIs, respectively. For the verb AOI, early measures only showed a significant interaction between the control conditions and each of the experimental conditions. 'Regressions out' did not show a significant interaction between any of the conditions. Other late measures showed significantly shorter reading times for the control condition than the experimental conditions, as well as significantly longer reading times for the tense condition than finiteness condition. For the post-verb AOI, early measures showed no significant interactions. 'Regressions out' and 'regression path duration' showed the same pattern of results as late measures in the verb AOI. 'Dwell time' and 'fixation count' showed a reliable interaction only between the control and the tense condition, and between the control and the combined condition.



# Kapittel 6

## Diskusjon

I dette kapitlet vil jeg prøve å se resultatene som har kommet frem fra de ulike målingene i sammenheng, samt tolke og diskutere tendenser og mønster i dataene. Jeg vil også drøfte resultatene opp mot forskningsspørsmålene og de teoretiske perspektivene som danner utgangspunkt for undersøkelsen.

Hypoteser og prediksjoner presentert i kapittel 4 gjentas her for klarhets skyld:

- Nullhypotese: Finittet- og tempus-trekkene er uløselig knyttet sammen; både finitt- og tempusavvikene påvirker derfor øyebevegelser på samme måte.
- Alternativ hypotese, del 1: Finittet- og tempus er to adskilte trekk. Avvik i finitt- og tempusmorfologi oppfattes som to distinkte typer av avvik, noe som resulterer i forskjellige lesemønstre.
- Alternativ hypotese, del 2: Finittet-trekket bærer mindre semantisk informasjon enn tempus, og finittavvik er derfor lettere å prosessere enn tempusavvik. Øyebevegelsene som følger disse avvikene bør i så fall skille seg fra hverandre på følgende måter:
  - Finittavvik: lengre og flere fikseringer, samt flere regresjoner enn i kontrollsetninger, men mindre og kortere fikseringer og færre regresjoner enn i setninger med brudd på bare tempusdistinksjonen eller setninger med brudd på både tempus- og finittdistinksjonen.
  - Tempusavvik: lengre og flere fikseringer og flere regresjoner enn både i kontrollsetninger og i setninger med finittavvik. Mindre og kortere fikseringer og færre regresjoner enn i setninger med både tempus- og finittavvik

- Tempus- pluss finittavvik: lengre og flere fikseringer og flere regresjoner enn både i kontrollsetninger og i setninger med bare én type brudd.

### 6.1 Finitt- vs. tempusavvik

Mange psyko- og nevrologvistiske studier har vist at forskjellige typer av språklige avvik forårsaker forskjellige reaksjoner (forskjell i øyebevegelser, hjernens elektriske aktivitet, osv.). Mesteparten av slike studier har fokusert på forskjeller mellom syntaktiske og semantiske avvik, og det finnes ganske få studier som sammenligner reaksjoner på forskjellige typer av (morfo)syntaktiske avvik. Likevel finnes det noen studier som har kartlagt forskjeller i reaksjoner på ulike typer av morfologiske avvik: for eksempel oppdaget Barber og Carreiras (2005) forskjeller i hjerteresponser på avvik i genus- og numerusmorfologi i spansk, mens Arnold (2019) påviste forskjeller i øyebevegelser ved avvik i numerus-, person-, og tempusmorfologi i spanske verb. Derfor forventet jeg å se forskjellige reaksjoner på avvik i finitt- og tempusmorfologi, dersom disse faktisk er adskilte kategorier, slik som Eide (2009) hevder.

Resultater for de tidlige målingene (*first fixation duration* og *gaze duration*) i det verbale interessefeltet viste en signifikant forskjell mellom ugrammatikalske setninger og kontrollsetninger, men ingen forskjell mellom finitt- og tempusavvik. Dette indikerer at selv om deltakerne la merke til avvikene veldig fort, hadde begge typer av avvik samme effekt på tidlig språkprosessering. Dette funnet er i samsvar med tidligere studier, som viser at tidlige målinger er sensitive for at setningen inneholder avvik, men ikke nødvendigvis viser forskjeller mellom ulike typer av avvik (se f.eks. Ni et al. (1998)).

Derimot var det en signifikant forskjell mellom finitt- og tempus-betingelser i nesten alle de sene målingene i det verbale interessefeltet, noe som indikerer at disse avvikene ikke er akkurat like lett å kompensere: verb med avvik i tempusmorfologi er tilsynelatende vanskeligere å integrere i konteksten. Signifikant forskjell i målingene *dwelling time* og *fixation count* viser at testverb med tempusavvik totalt tiltrakk seg mer oppmerksomhet enn testverb med finittavvik, mens *regression path duration*-målingen gjenspeiler tiden brukt på det første forsøket å kompensere avviket. Interessant nok var det ingen forskjell mellom de eksperimentelle betingelsene i *regressions out*-målingen: dette betyr at sannsynligheten for en regresjon var det samme for tempus- og finittavvik, men hvis regresjonen faktisk oppstod, var den lengre for

setningene med tempusavvik, selv om setningene med finittavvik som regel hadde mer informasjon tilgjengelig før testverbet på grunn av pseudokoordineringsverb, noe som kunne ha ført til lengre regresjoner. Disse forskjellene mellom regresjonsmålingene er i tråd med resultatene for de tidlige målingene: type av avvik hadde ikke betydning for tidlig språkprosessering (lik varighet på første fiksering og lik sannsynlighet for regresjon), men påvirket senere prosessering (varigheten på regresjonen og den totale lesetiden, altså hvor lang tid deltakere trenger for å analysere setningen på nytt og integrere det avvikende verbet i konteksten).

De veldig sene målingene, altså *spillover*-effekter som oppstår i det postverbale interessefeltet, viste en signifikant interaksjon mellom finitt- og tempusbetingelsene i *regression path duration* og *regressions out*-målingene, og en nær-signifikant interaksjon i *fixation count*. Det er ikke overraskende at det ikke var signifikant interaksjon mellom de eksperimentelle betingelsene i *fixation count* og *dwell time*: selv om den totale lesetiden for de tre betingelsene var ulik, nådde ikke effekten signifikans trolig på grunn av forskjeller i det postverbale interessefeltet i forskjellige datasett, samt forskjeller i deltakernes individuelle lesemønster. Disse forskjellene var en sannsynlig kilde til tilfeldig varians i dataene, noe som gjør det vanskeligere å oppdage effekten. I tillegg tiltrekker ikke det grammatikalske postverbale interessefeltet seg like mye oppmerksomhet som det ugrammatikalske testverbet, noe som gjør at regresjonsmålingene kan være en bedre representasjon av avvikprosessering på dette stadiet i lesing. En sammenligning av *regressions out*-målingen for det verbale og det postverbale interessefeltet viser at informantene bestemte seg for å vende tilbake i teksten for å kompensere avviket ganske sent, altså etter at blikket deres hadde beveget seg bort fra avviket (sannsynlighet for regresjon var mer pålitelig i det postverbale feltet), selv om de oppfattet testverbet som umiddelbart avvikende (ifølge resultater for *first fixation duration*-målingen i det verbale interessefeltet). Sene regresjoner kan også være en mulig forklaring på fravær av effekten i de tidlige målingene for testbetingelsene i det postverbale feltet: større behov for å vende tilbake i teksten fra det postverbale feltet kunne ha ført til kortere førstegangslesetid.

En viktig generalisering som kommer frem fra disse resultatene er at finitt- og tempusavvik fører til forskjellige reaksjoner, hvor finittavvik medfører kortere lesetid og færre regresjoner enn tempusavvik. Disse resultatene er altså forenelige med begge deler av den alternative hypotesen og støtter påstanden om at finittet og tempus er to adskilte kategorier, samt at finittet er et mindre fremtredende trekk enn tempus.

### **Kan morfologiske ulikheter forklare disse resultatene?**

Det er mulig å mistenke at kortere lesetid i setninger med finittavvik i motsetning til andre eksperimentelle betingelser er forårsaket av forskjeller i bøyingsmorfologi: *synge* og *synger* ligner mer på hverandre visuelt enn *synge* og *sang* eller *synge* og *sunget*, ettersom de sistnevnte parene inneholder forskjeller i både bøyingsaffikset og rotvokalen (ablaut). Som følge av dette kan man være fristet til å konkludere at resultatene i denne studien er forårsaket av en større visuell forskjell mellom det forventede verbet og det verbet deltakerne faktisk fikk se i betingelser T og F+T enn i betingelse F. Imidlertid stemmer ikke denne konklusjonen, ettersom den attesterte effekten var til stedet også for svake verb, som har like komplisert bøyingsmorfologi i alle tempusformer (*å passe – passer – passet – har passet*).

### **6.2 Finitt- pluss tempusavvik**

Ikke alle prediksjonene ble bekreftet av resultatene av denne studien: betingelse F+T, som inneholdt et dobbelt (finitt- pluss tempus) avvik, medførte ikke betydelig lengre fikseringer eller flere regresjoner enn finitt- eller tempus-betingelse i noen av målingene. Gjennomsnittene for setningene med dobbelt-avvik var høyere enn for setningene med finittavvik, men lavere enn gjennomsnittene for setningene med tempusavvik. I tillegg var det ikke noen signifikant interaksjon mellom dobbelt-avvik og de enkle avvikene, noe som gjør det vanskelig å trekke konklusjoner om hvordan effekten produsert av dobbelt-avviket står i forhold til effektene produserte av de enkle avvikene, altså om dobbelt-avviket er mer likt finitt- eller tempusavviket.

Dette funnet kan umiddelbart se ut til å kaste tvil over antagelsen om at finittet og tempus er to adskilte trekk: for det første, hvis finitt- og tempusavvik faktisk har opphav i to adskilte kategorier, synes det at en kombinasjon av disse bør produsere en kombinert effekt. For det andre, hvis ulike resultater for betingelser F og T faktisk ble forårsaket av forskjeller mellom finittet og tempus, bør reaksjoner på dobbelt-avviket (som også inneholder et tempusavvik) være signifikant forskjellig fra reaksjoner på finittavviket. En rekke tidligere studier kan imidlertid bidra til å gjøre denne tvilen mindre relevant: Hahne og Jescheniak (2001), Hagoort (2003) og Friederici et al. (2004) er blant de studiene som sammenlignet effekter av enkle og doble avvik på språkprosessering og attesterte at en kombinasjon av to avvik ikke alltid fører til en additiv effekt. I disse studiene var det semantiske og syntaktiske avvik som ble sammenlignet, og funnene ble derfor tolket som resultat av «syntaks-før-semantikk»-



hypotesen, som foreslår at i språkprosessering må den syntaktiske strukturen til en setning bygges opp før semantisk informasjon kan integreres. Forfatterne av disse studiene hevder altså at oppdagelse av et syntaktisk avvik i en tidlig fase av språkprosessering forhindrer videre integrasjon av semantisk informasjon, og at det semantiske avviket derfor ikke bidrar med ytterligere effekt. Fordi stimuli med semantiske avvik brukt i disse studiene var grammatikalske setninger som inneholdt et svært usannsynlig ord (f.eks. *The honest umbrella was in the garage*), kan ikke disse resultatene automatisk generaliseres til en kombinasjon av to morfologiske avvik (finitthet og tempus), selv om tempusavviket kan regnes som et semantisk avvik ifølge noen perspektiver (se f.eks. Baggio (2008); Bos et al. (2013)). Resultater av disse studiene er imidlertid en viktig indikasjon på at to distinkte trekk som produserer forskjellige reaksjoner på egen hånd ikke nødvendigvis forårsaker en større effekt i kombinasjon.

En annen studie, utført av Rossi et al. (2005), undersøkte prosessering av to typer av morfosyntaktiske avvik (hver for seg og i kombinasjon) og oppdaget et mønster i sen språkprosessering som lignet på resultater av min studie: dobbelt-avvik forårsaket en effekt som var større enn effekten av ett av avvikene, men mindre enn effekten av det andre. Rossi og kollegaer forklarte dette mønsteret med syntaktiske forskjeller mellom stimulisetninger i de ulike betingelsene, nærmere bestemt forskjell i distanse mellom testverbet og det ordet som skapte en forventning av en spesifikk verbform i deltakere. Muligens var resultatene av min studie også forårsaket av syntaktiske forskjeller i konteksten, ettersom betingelse T inneholdt et sammensatt tempus (presens perfektum), mens betingelser F og F+T inneholdt enkle tempora (presens og preteritum). Sammensatte (perifrastiske) tempora har en forskjellig syntaktisk struktur (Julien, 2001), og Håkansson (2001) har påvist at disse tempora er vanskeligere å tilegne seg, samt at barn med språkspesifikke vansker ofte sliter kun med perifrastiske, men ikke enkle tempora. Det finnes imidlertid lite psykolingvistisk evidens for at perifrastiske tempora er vanskeligere å prosessere enn enkle tempora. Tvert imot viste Grisot og Blochowiak (2017) at enkle og perifrastiske tempora i fransk er prosessert på samme måte, mens Roberts og Liszka (2013) fant evidens for at engelske språkbrukere prosesserer avvik i presens perfektum fortere enn i preteritum. Den viktigste konklusjonen her er at til tross for det uventede resultatet for dobbelt-avviket i Rossis (2005) studie, er det ukontroversielt at de to typene av avvik som ble undersøkt i den studien (bruk av feil ordklasse og tempusfeil) er distinkte fra hverandre. Resultatet oppdaget i min studie for betingelse F+T motstrider altså ikke den

alternativ hypotesen om at finitthet og tempus er to adskilte trekk, selv om et slikt resultat heller ikke støtter denne hypotesen.

Det fremstår som vanskelig å trekke konklusjoner om mulige kilder for denne uforventede effekten i betingelse F+T, først og fremst fordi det eksisterer svært få studier jeg kan sammenligne mine resultater med. Samarbeid mellom teoretisk lingvistikk og eksperimentell psyko-/nevrolingvistikk har vært tilfeldig og relativt nødtørftig: formelle teorier om språk blir ofte utviklet uten å bruke evidens fra eksperimentelle studier, mens psyko- og nevrolingvistiske studier ofte foretrekker å fokusere på kognitive prosesser og kun tar i bruk utvalgte brokker av språkvitenskapelige teorier i sitt eksperimentdesign og tolkning av resultater (Mancini, 2018). Dessuten har tidligere, delvis relevante, psykolingvistiske studier basert på generativ syntaks gjort helt andre antagelser om infinitive verbformer og/eller finitthet enn jeg gjør i denne studien: det finnes tilsynelatende ingen annen psykolingvistisk studie som har undersøkt finittavvik uavhengig fra kongruens- og tempusavvik, mens studier som undersøker tempusavvik som regel inkluderer ikke infinitivsformer. Tradisjonelt sett regnes infinitiver nemlig som verbformer uten tempus. Dette er dermed en risiko som rammer alle pionerstudier av denne typen, der eksperimenter settes opp for å teste nye formalingvistiske teorier.

Til slutt vil jeg understreke at selv om resultatet for betingelse F+T kan synes vanskelig å forklare og isolert ikke støtter hypotesen om finitthet og tempus som distinkte kategorier, er det viktig å huske at «absence of evidence is not evidence of absence». Med dette mener jeg at dette uforventede resultatet ikke kan regnes som bevis på at finitthet og tempus *ikke* er to distinkte trekk, dette resultatet er kun et tegn på at videre forskning er nødvendig for å korroborere eller falsifisere denne teorien. I tillegg til det faktum at kildene til denne effekten er uklare, var ikke-signifikante forskjeller mellom betingelser F og F+T til stede, og effekter som ikke er signifikante statistisk sett kan godt være signifikante ute i verden. Det er også mulig at med nøyere kontrollerte stimuli og et større informantutvalg kunne denne effekten ha oppnådd statistisk signifikans (Alderson, 2004). Det er derfor viktig å tolke resultater av signifikanstester veldig forsiktig for å unngå feilslutninger.

### 6.3 Videre forskning

Videre forskning er definitivt nødvendig for å entydiggjøre resultatene av denne studien, både for å finne mulige kilder til effekten produsert av betingelse F+T, og for

å være sikker på at resultatene for betingelser F og T faktisk var forårsaket av forskjeller mellom trekkene finitthet og tempus. Da må andre grunner med rimelighet kunne utelukkes, som for eksempel syntaktiske forskjeller mellom stimulisetningene. Selve den totale mangelen på psykolingvistisk forskning på finitthet uavhengig av tempus tilsier klart at videre studier er nødvendig for å finne ytterligere kognitiv evidens som kan kaste lys over disse to kategoriens natur – og slik enten støtte eller falsifisere Eides (2009) teori.

Det naturlige neste steg i videre forskning på dette temaet er derfor å sammenligne andre typer sammenblandinger av verbformer på tvers av finitt- og tempusdistinksjoner. Omfanget av denne studien tillot meg bare å undersøke én type av sammenblanding (bruk av infinitivsformen i stedet for presens/preteritum/partisipp), ettersom en undersøkelse av andre typer av sammenblandinger ville ha ført til et for stort antall stimulisetninger og en meget komplisert analysemodell. Forskning på andre typer av sammenblandinger (f.eks. bruk av presens i stedet for infinitiv/preteritum/partisipp) er nødvendig fordi den kan bidra til å verifisere om forskjeller mellom finitt- og tempusbetingelsene oppdaget i denne studien var forårsaket av forskjeller mellom tempus og finitthet eller av andre ting, som syntaktiske forskjeller mellom stimulisetningene. Det er også interessant å se hvilke resultater som blir produsert av betingelse F+T i andre typer av sammenblandinger.

En fremtidig undersøkelse av andre typer av sammenblandinger kan også bidra til generaliserbarheten av resultatene av denne studien: for eksempel kan det være at avvik som inneholder infinitivsformer prosesseres forskjellig fra avvik som inneholder andre verbformer, ettersom infinitiv er verbets grunnform. Slabakova (2016, s. 191) påstår at infinitivsformer er «uspesifiserte», selv om de ofte inneholder bøyingsaffiks akkurat som andre, «spesifiserte» verbformer, og at dette gjør infinitivsformer lettere tilgjengelig i det mentale leksikonet. I tillegg er bruk av infinitivsformen (*bare infinitive*) i stedet for en annen verbform et vanlig tverrspråklig fenomen i barnespråk, som Wexler (1994) refererer til som en *Optional Infinitive*-fase i førstespråkstilegnelse. Wagley et al. (2019) viste at voksne språkbrukere er mer følsomme for *developmental errors*, altså de feilene som barn gjør når de tilegner seg førstespråket sitt (som f.eks. bruk av infinitiv i stedet for andre verbformer), enn mot *non-developmental errors*, altså feilene som ikke oppstår i førstespråkstilegnelse. Dette tyder på at erstatning av den forventede verbformen med infinitiv kan være vanskeligere å prosessere enn avvik der andre verbformer brukes.

Til slutt er det viktig å ta i betraktning at i noen norske dialekter brukes infinitivsformen av verb i både infinitiv- og presens-kontekster (Sjekkeland, 2005). Dette kan i motsetning til Wagleys studie føre til at bruk av infinitivsformen i stedet for presensform ikke fremstår like avstikkende som andre typer sammenblandinger i norsk og derfor prosesseres fortere.

Videre forskning kan også dra nytte av undersøkelse av potensielle kvalitative forskjeller mellom finitt- og tempusavvik. Denne studien har kun undersøkt den kvantitative forskjellen, altså antall regresjoner og varigheten på fikseringer produsert av hvert avvik, som gjenspeiler hvor vanskelig eller kognitivt krevende hvert avvik var å prosessere. Kvalitative forskjeller mellom finitt- og tempusavvik kan tyde på at *parser* (mekanismen som foretar setningens syntaktiske analyse) benytter forskjellige strategier for å kompensere finitt- og tempusavvikene. Dette kan forventes som følge av tidligere diskuterte semantiske forskjeller mellom finitthet og tempus, ettersom syntaktiske og semantiske avvik produserer kvalitativt forskjellige lesemønstre (Ni et al., 1998). Kvalitativ analyse av eye-tracking-data innebærer analyse av hele setninger inndelt i flere interessefelter.

Elektroencefalografi (EEG) kan også vise hvorvidt det finnes kvalitative (eller kvantitative) forskjeller mellom reaksjoner på finitt- og tempusavvik. I tillegg kan EEG-data gi innsikt i hvorvidt finitthet og tempus har adskilte mentale representasjoner: funksjonelt distinkte nevrokognitive prosesser produserer kvalitativt ulike hjerneresponser (*event-related potentials* eller *ERPs*), som varierer med hensyn til latens, polaritet, og spatial distribusjon (Frisch et al., 2004). Slik evidens synes nødvendig for å støtte Eides teori fra et nevrokognitivt perspektiv, ettersom atferdsdata (øyebevegelser) ikke gir direkte tilgang til kognitive og nevronale prosesser som underligger prosessering av språklige avvik. Morfologiske avvik forårsaker som regel to distinkte ERP-komponenter: *left anterior negativity* (LAN) og *late centro-parietal positivity* (P600), hvor den første komponenten indikerer oppdagelse av et morfosyntaktisk avvik, mens den andre gjenspeiler syntaktisk reanalyse (Rossi et al., 2005). Det finnes imidlertid en del variasjon i ERP-mønster oppdaget i tidligere studier av tempusavvik (se Baggio (2008) for en oversikt). Dette er ikke overraskende, ettersom det ifølge Bos et al. (2013) finnes en forskjell mellom tempusavvik (i.e. avvik i verbmorfologi) og avvik i temporal referanse (i.e. semantisk uoverensstemmelse mellom tempuset uttrykt av verbformen og tempuset uttrykt av det temporale adverbet). Så vidt jeg vet finnes det ikke noen ERP-studier som har undersøkt finittavvik uavhengig av tempus- og kongruensavvik, noe som gjør det umulig på

nåværende tidspunkt å sammenligne ERP-komponentene produsert av finitt- og tempusavvik.

## 6.4 Konklusjoner

Denne studien har undersøkt hvorvidt finitthet og tempus oppfattes som to adskilte trekk ved å bruke et eksperimentelt avviksparadigme. Begrunnelsen bak dette designet er at avvik i ulike kategorier er påvist å produsere forskjellige reaksjoner i form av øyebevegelser og hjernerensjoner. Hvis finitthet og tempus faktisk er to adskilte kategorier, bør finitt- og tempusavvik føre til ulike reaksjoner.

Resultater av eye-tracking-eksperimentet viste en statistisk signifikant forskjell mellom reaksjoner på finitt- og tempusavvik (bruk av infinitivsformen i stedet for presens- og partisippform). Prediksjonen om at finitthet og tempus oppfattes som distinkte kategorier, samt at finittavvik er lettere å prosessere enn tempusavvik synes til å være bekreftet av disse resultatene (lengre lesetid og flere regresjoner for tempus- enn for finittbetingelse). Prediksjonen om at en kombinasjon av disse to avvikene er vanskeligere å prosessere enn hvert avvik alene ble imidlertid ikke bekreftet, ettersom setninger med dobbelt-avviket ikke fikk betydelig lengre lesetid enn finitt- eller tempusavvik. Grunnen for dette resultatet er ikke helt klart, og det virker ikke mulig å bestemme hva dette resultatet betyr for teorien om finitthet og tempus basert utelukkende på denne studien. Videre forskning i form av eye-tracking eller EEG-studier kreves for å klargjøre dette resultatet. I tillegg er det viktig å huske at denne studien undersøkte kun et begrenset antall av mulige avvik, og at andre typer av finitt- og tempusavvik må undersøkes før sikre konklusjoner om finitthet og tempus kan trekkes.

Det er på sin plass å være forsiktig i slutningene før videre forskning gjennomføres for å etablere hvorvidt mønster observert i denne studien også gjelder andre typer av finitt- og tempusavvik. Med dette i tankene viser resultatene av denne studien et kvantitativt forskjellig mønster av øyebevegelser til finitt- og tempusavvik, noe som tyder på at kompensering av tempusavvik fører til høyere kognitiv belastning sammenlignet med finittavvik. Dataene produsert av dette eksperimentet indikerer at nullhypotesen i denne studien kan forkastes; hypotesen om at finitthet og tempus er to distinkte trekk er altså korroborert av disse resultatene.

## 6.5 Chapter summary: Discussion

The results of this study have shown a difference in native Norwegian speakers' reactions to different types of anomalies in verbal morphology. The data obtained in the current eye-tracking experiment have shown that the participants recovered significantly quicker from the finiteness violations than from the violations of tense. However, the prediction concerning the combined violation condition was not borne out: this condition did not vary significantly from either one of the pure violation conditions.

The observed differences in reactions to finiteness and tense anomalies are consistent with the claim that finiteness and tense are two distinct features. In addition, longer reading times produced by the tense condition corroborate the idea that tense is a more salient feature than finiteness. However, the results produced by the combined violation condition cannot easily be explained, and neither corroborate nor contradict the original hypothesis. A number of other studies that have investigated cognitive effects of combining two violations have also failed to observe an additive effect. Some studies (Hahne & Jescheniak, 2001) have attributed that result to the idea that syntax precedes semantics in sentence processing, and that detection of a syntactic error at an early stage blocks semantic integration processes. Other studies (Rossi et al., 2005) have claimed that differences in syntactic structure between conditions were driving effects. It is unclear whether the sources of the effect produced by the combined condition in this study are similar to those suggested by Hahne and Jescheniak or Rossi and colleagues, but these studies are an important indication that two distinct violations that produce different reactions on their own do not necessarily result in an additive effect when combined. Therefore, I assume that the results for the combined violation condition do not contradict the idea that the patterns observed in the finiteness and tense conditions were caused by the difference between the finiteness and tense features.

To summarize, the results of this study show interesting differences in eye-movements when reading sentences with different types of morphosyntactic violations, albeit the findings concerning the nature of the two functional features investigated in this study – finiteness and tense – remain inconclusive. Further research is certainly necessary to clarify the effects discovered in this study, as well as to confirm or falsify Eide's proposal, which states that finiteness and tense are two distinct functional categories. Possible avenues for future research on this topic

include further eye-tracking experiments investigating other types of finiteness and tense violations, i.e. other variations of verb form “mix-ups”, as well as EEG experiments, in order to supplement the behavioural data obtained in this study with neurological data, which could shed some light onto the representation of these two functional categories in the brain and the nature of syntactic reanalysis processes triggered by the finiteness and tense anomalies.





# Referanser

- Abdi, H. (2007). The Bonferonni and Šidák Corrections for Multiple Comparisons. I N. Salkind (Red.), *Encyclopedia of Measurement and Statistics*. SAGE Publications, Inc.
- Alderson, P. (2004). Absence of evidence is not evidence of absence. *BMJ (Clinical research ed.)*, 328(7438), 476-477. <https://doi.org/10.1136/bmj.328.7438.476>
- Allen, E. & Seaman, C. (2007). Likert Scales and Data Analyses. *Quality Progress*, 64-65.
- Anderson, J. R., Bothell, D. & Douglass, S. (2004). Eye movements do not reflect retrieval processes: Limits of the eye-mind hypothesis. *Psychological Science*, 15(4), 225-231. <https://doi.org/10.1111/j.0956-7976.2004.00656.x>
- Arnold, E. W. (2019). *Using Eye Tracking to Examine Working Memory and Verbal Feature Processing in Spanish* [Brigham Young University]. *ASK-korpuset*. <http://clarino.uib.no/ask/ask>
- Baddeley, A. D., Thomson, N. & Buchanan, M. (1975). Word length and the structure of short-term memory. *Journal of Verbal Learning and Verbal Behavior*, 14(6), 575-589. [https://doi.org/https://doi.org/10.1016/S0022-5371\(75\)80045-4](https://doi.org/https://doi.org/10.1016/S0022-5371(75)80045-4)
- Baggio, G. (2008). Processing Temporal Constraints: An ERP Study [<https://doi.org/10.1111/j.1467-9922.2008.00460.x>]. *Language Learning*, 58(s1), 35-55. <https://doi.org/https://doi.org/10.1111/j.1467-9922.2008.00460.x>
- Barber, H. & Carreiras, M. (2005). Grammatical Gender and Number Agreement in Spanish: An ERP Comparison. *Journal of Cognitive Neuroscience*, 17(1), 137-153. <https://doi.org/10.1162/0898929052880101>
- Barron, J. (2000). The morphosyntactic correlates of finiteness. Proceedings of the LFG00 Conference, University of California, Berkeley.
- Bokmålsordboka*. Språkrådet og Universitetet i Bergen. <http://ordbok.uib.no>.
- Bos, L., Dragoy, O., Stowe, L. & Bastiaanse, R. (2013). Time reference teased apart from tense: Thinking beyond the present. *Journal of Neurolinguistics*, 26, 283-297. <https://doi.org/10.1016/j.jneuroling.2012.10.001>

- Braze, D., Shankweiler, D., Ni, W. & Palumbo, L. C. (2002). Readers' Eye Movements Distinguish Anomalies of Form and Content. *Journal of Psycholinguistic Research*, 31(1), 25-44. <https://doi.org/10.1023/A:1014324220455>
- Carter, B. T. & Luke, S. G. (2020). Best practices in eye tracking research. *Int J Psychophysiol*, 155, 49-62. <https://doi.org/10.1016/j.ijpsycho.2020.05.010>
- Chamorro, G., Sorace, A. & Sturt, P. (2016). What is the source of L1 attrition? The effect of recent L1 re-exposure on Spanish speakers under L1 attrition. *Bilingualism: Language and Cognition*, 19(3), 520-532.
- Chomsky, N. (1965). *Aspects of the Theory of Syntax* (50. utg.). The MIT Press. <http://www.jstor.org/stable/j.ctt17kk81z>
- Chomsky, N. (1995). *The Minimalist Program*. The MIT Press. <https://doi.org/10.2307/j.ctt17kk8xd>
- Clahsen, H., Balkhair, L., Schutter, J.-S. & Cunnings, I. (2013). The time course of morphological processing in a second language. *Second Language Research*, 29(1), 7-31.
- Clifton Jr, C., Staub, A. & Rayner, K. (2007). Eye movements in reading words and sentences. I *Eye movements: A window on mind and brain*. (s. 341-371). Elsevier. <https://doi.org/10.1016/B978-008044980-7/50017-3>
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. L. Erlbaum Associates.
- Comrie, B. (1985). *Tense*. Cambridge University Press. <https://doi.org/DOI:10.1017/CBO9781139165815>
- Conklin, K. & Pellicer-Sánchez, A. (2016). Using eye-tracking in applied linguistics and second language research. *Second Language Research*, 32(3), 453-467. <https://doi.org/10.1177/0267658316637401>
- Conklin, K., Pellicer-Sánchez, A. & Carrol, G. (2018). *Eye-tracking: A Guide for Applied Linguistics Research*. Cambridge University Press.
- Cooper, R. M. (1974). The control of eye fixation by the meaning of spoken language: A new methodology for the real-time investigation of speech perception, memory, and language processing. *Cognitive Psychology*, 6(1), 84-107. [https://doi.org/https://doi.org/10.1016/0010-0285\(74\)90005-X](https://doi.org/https://doi.org/10.1016/0010-0285(74)90005-X)
- Cowper, E. (2002). *Finiteness*. University of Toronto.
- Cowper, E. (2016). Finiteness and Pseudofiniteness. I K. M. Eide (Red.), *Finiteness Matters: On finiteness-related phenomena in natural languages*. John Benjamins. <https://doi.org/10.1075/la.231.02cow>
- Crawford, T. J., Smith, E. S. & Berry, D. M. (2017). Eye Gaze and Aging: Selective and Combined Effects of Working Memory and Inhibitory Control [Original

- Research]. *Frontiers in Human Neuroscience*, 11(563).  
<https://doi.org/10.3389/fnhum.2017.00563>
- Dodge, R. & Cline, T. S. (1901). The angle velocity of eye movements. *Psychological Review*, 8(2), 145-157. <https://doi.org/10.1037/h0076100>
- Domínguez, L., Hicks, G. & Slabakova, R. (2019). Terminology choice in generative acquisition research: The case of “incomplete acquisition” in heritage language grammars. *Studies in Second Language Acquisition*, 41(2), 241-255. <https://doi.org/10.1017/S0272263119000160>
- Dowiasch, S., Marx, S., Einhäuser, W. & Bremmer, F. (2015). Effects of aging on eye movements in the real world [Original Research]. *Frontiers in Human Neuroscience*, 9(46). <https://doi.org/10.3389/fnhum.2015.00046>
- Durlak, J. A. (2009). How to Select, Calculate, and Interpret Effect Sizes. *Journal of Pediatric Psychology*, 34(9), 917-928. <https://doi.org/10.1093/jpepsy/jsp004>
- Dussias, P. E. & Sagarra, N. (2007). The effect of exposure on syntactic parsing in Spanish-English bilinguals. *Bilingualism*, 10(1), 101.
- Eide, K. M. (2009). Finiteness: The haves and the have-nots. IJ. H. Artemis Alexiadou, Thomas McFadden, Justin Nuger, Florian Schäfer (Red.), *Advances in Comparative Germanic Syntax*. John Benjamins Publishing Company. <https://doi.org/10.1075/la.141>
- Eide, K. M. (2012). Om tempus og aspekt i norsk. I H.-O. Enger, J. T. Faarlund & K. I. Vannebo (Red.), *Grammatikk, bruk og norm* (1. utg., s. 57-78). Novus forlag.
- Eide, K. M. (2016). Finiteness, inflection, and the syntax your morphology can afford: On finiteness-related phenomena in natural languages. I (s. 121-168). <https://doi.org/10.1075/la.231.05eid>
- Eide, K. M. & Hjelde, A. (2018). Om verbplassering og verbmorfologi i amerikanorsk. *Maal og Minne*, 110(1). <http://ojs.novus.no/index.php/MOM/article/view/1529>
- Eisikovits, E. (1987). Variation in the lexical verb in inner-Sydney English. *Australian Journal of Linguistics*, 7(1), 1-24. <https://doi.org/10.1080/07268608708599371>
- EyeLink 1000 User Manual*. (2009). [Bruksanvisning]. SR Research Ltd., Mississauga, Ontario, Canada.
- Faarlund, J. T. (1987). Om beskriving og forklaring i lingvistisk teori. *NLT - Norsk lingvistisk tidsskrift*, 1, 13-28.
- Faarlund, J. T., Lie, S. & Vannebo, K. I. (1997). *Norsk referansegrammatikk*. Universitetsforl.
- Field, A. (2013). *Discovering Statistics using IBM SPSS Statistics* (4. utg.). Sage Publications Ltd.

- Foppolo, F. & Staub, A. (2020). The puzzle of number agreement with disjunction. *Cognition*, 198, 104161. <https://doi.org/10.1016/j.cognition.2019.104161>
- Forker, D. (2011). Finiteness in Hinuq. *Linguistic Discovery*, 9(2). <https://doi.org/10.1349/PS1.1537-0852.A.391>
- Frekvensordlister for norsk talespråk og skriftspråk*. Tekstlaboratoriet, ILN, Universitetet i Oslo. <http://www.tekstlab.uio.no/frekvensordlister/>
- Friederici, A. D., Gunter, T. C., Hahne, A. & Mauth, K. (2004). The relative timing of syntactic and semantic processes in sentence comprehension. *NeuroReport*, 15(1). [https://journals.lww.com/neuroreport/Fulltext/2004/01190/The\\_relative\\_timing\\_of\\_syntactic\\_and\\_semantic.32.aspx](https://journals.lww.com/neuroreport/Fulltext/2004/01190/The_relative_timing_of_syntactic_and_semantic.32.aspx)
- Frisch, S., Hahne, A. & Friederici, A. (2004). Word category and verb-argument structure information in the dynamics of parsing. *Cognition*, 91, 191-219. <https://doi.org/10.1016/j.cognition.2003.09.009>
- Frisson, S., Bélanger, N. N. & Rayner, K. (2014). Phonological and orthographic overlap effects in fast and masked priming. *Quarterly journal of experimental psychology* (2006), 67(9), 1742-1767. <https://doi.org/10.1080/17470218.2013.869614>
- Ghasemi, A. & Zahediasl, S. (2012). Normality Tests for Statistical Analysis: A Guide for Non-Statisticians. *International Journal of Endocrinology and Metabolism*, 10(2), 486-489. <https://doi.org/10.5812/ijem.3505>
- Gibbs, R. W. (2006). Introspection and cognitive linguistics: Should we trust our own intuitions? *Annual Review of Cognitive Linguistics*, 4, 135-151. <https://doi.org/10.1075/arcl.4.06gib>
- Givón, T. (1990). *Syntax: A functional-typological introduction* (Bd. II). John Benjamins.
- Glahn, E., Håkansson, G., Hammarberg, B., Holmen, A., Hvenekilde, A. & Lund, K. (2001). Processability in Scandinavian Second Language Acquisition. *Studies in Second Language Acquisition*, 23(3), 389-416. <https://doi.org/10.1017/s0272263101003047>
- Godfrey-Smith, P. (2021). *Theory and Reality: An Introduction to the Philosophy of Science, Second Edition*. University of Chicago Press. <https://books.google.no/books?id=muv1zQEACAAJ>
- Godfroid, A. (2020). *Eye Tracking in Second Language Acquisition and Bilingualism: A Research Synthesis and Methodological Guide* [Book]. Routledge. <http://search.ebscohost.com/login.aspx?direct=true&db=nlebk&AN=1761991&site=ehost-live>

- Goldberg, J. & Wichansky, A. (2003). *Eye tracking in usability evaluation: A practitioner's guide*.
- Gollücke, V. (2009). *Eye-Tracking. Grundlagen, Technologien und Anwendungsgebiete*. GRIN Verlag.
- Gonzalez-Marquez, M. (2007). *Methods in Cognitive Linguistics* [Book]. John Benjamins Publishing Co.  
<http://search.ebscohost.com/login.aspx?direct=true&db=e000xww&AN=229841&site=ehost-live>
- Grano, T. (2017). Finiteness contrasts without Tense? A view from Mandarin Chinese. *Journal of East Asian Linguistics*, 26(3), 259-299.  
<https://doi.org/10.1007/s10831-017-9159-8>
- Gretsch, P. & Perdue, C. (2007). Finiteness in first and second language acquisition. I I. Nikolaeva (Red.), *Finiteness*. Oxford University Press.
- Grisot, C. & Blochowiak, J. (2017). Temporal connectives and verbal tenses as processing instructions. *Pragmatics & Cognition*, 24(3), 404-440.  
<https://doi.org/10.1075/pc.17009.gri>
- Hagoort, P. (2003). Interplay between Syntax and Semantics during Sentence Comprehension: ERP Effects of Combining Syntactic and Semantic Violations. *Journal of Cognitive Neuroscience*, 15(6), 883-899.  
<https://doi.org/10.1162/089892903322370807>
- Hahne, A. & Jescheniak, J. D. (2001). What's left if the Jabberwock gets the semantics? An ERP investigation into semantic and syntactic processes during auditory sentence comprehension. *Brain Res Cogn Brain Res*, 11(2), 199-212.  
[https://doi.org/10.1016/s0926-6410\(00\)00071-9](https://doi.org/10.1016/s0926-6410(00)00071-9)
- Håkansson, G. (2001). Tense morphology and verb-second in Swedish L1 children, L2 children and children with SLI. *Bilingualism: Language and Cognition*, 4, 85-99. <https://doi.org/10.1017/S1366728901000141>
- Heyman, T., Rensbergen, B., Storms, G., Hutchison, K. & De Deyne, S. (2014). The Influence of Working Memory Load on Semantic Priming. *Journal of experimental psychology. Learning, memory, and cognition*, 41.  
<https://doi.org/10.1037/xlm0000050>
- Holmberg, A. & Platzack, C. (1995). *The Role of Inflection in Scandinavian Syntax*. Oxford University Press.  
<http://search.ebscohost.com/login.aspx?direct=true&db=e000xww&AN=143747&site=ehost-live>
- Holmqvist, K., Nyström, M., Andersson, R., Dewhurst, R., Jarodzka, H. & van de Weijer, J. (2011). *Eye Tracking: A Comprehensive Guide To Methods And Measures*.

- Inhoff, A. W. & Rayner, K. (1986). Parafoveal word processing during eye fixations in reading: Effects of word frequency. *Perception & Psychophysics*, 40(6), 431-439. <https://doi.org/10.3758/BF03208203>
- Jacob, R. & Karn, K. (2003). Eye Tracking in Human-Computer Interaction and Usability Research: Ready to Deliver the Promises. I (Bd. 2, s. 573-605). <https://doi.org/10.1016/B978-044451020-4/50031-1>
- Johannessen, J. B. (2003). Innsamling av språklige data: Informanter, introspeksjon og korpus. I J. B. Johannessen (Red.), *På språkjakt*. Unipub forlag.
- Judd, C. H., McAllister, C. N. & Steel, W. M. (1905). General introduction to a series of studies of eye movements by means of kinoscopic photographs. I J. M. Baldwin, H. C. Warren & C. H. Judd (Red.), *Psychological Review, Monograph Supplements*. The Review Publishing Company.
- Julien, M. (2001). The Syntax of Complex Tenses. *The Linguistic Review*, 18, 125-167. <https://doi.org/10.1515/tlir.18.2.125>
- Just, M. A. & Carpenter, P. A. (1980). A theory of reading: From eye fixations to comprehension. *Psychological Review*, 87(4), 329-354. <https://doi.org/10.1037/0033-295X.87.4.329>
- Kjærbæk, L. (2018). *Morfologisk utvikling hos barn med dansk som modersmål*. 15. Nordiske Symposium om Børnesprog, Odense, Danmark.
- Klein, W. (2009). Finiteness, universal grammar, and the language faculty. *Crosslinguistic approaches to the psychology of language: Research in the tradition of Dan Isaac Slobin*, 333-344 (2009).
- Kløve, M. H. & Husby, O. (2008). *Norsk som andrespråk: undervisningsopplegg i uttale*. Abstrakt forlag.
- Knickerbocker, F., Johnson, R., Starr, E., Hall, A., Preti, D., Slate, S. & Altarriba, J. (2019). The time course of processing emotion-laden words during sentence reading: Evidence from eye movements. *Acta Psychologica*, 192, 1-10. <https://doi.org/10.1016/j.actpsy.2018.10.008>
- Kuperman, V. & Van Dyke, J. (2011). Effects of individual differences in verbal skills on eye-movement patterns during sentence reading. *Journal of Memory and Language*, 65, 42-73. <https://doi.org/10.1016/j.jml.2011.03.002>
- Lakens, D. (2013). Calculating and reporting effect sizes to facilitate cumulative science: a practical primer for t-tests and ANOVAs [Review]. *Frontiers in Psychology*, 4(863). <https://doi.org/10.3389/fpsyg.2013.00863>
- Lezak, M. D. (1982). The Problem of Assessing Executive Functions [<https://doi.org/10.1080/00207598208247445>]. *International Journal of Psychology*, 17(1-4), 281-297. <https://doi.org/https://doi.org/10.1080/00207598208247445>

- Libben, G. & Archibald, J. (2019). Morphological theory and second language acquisition. I.
- Lie, S. (2017). *Pseudokoordinering?*
- Liversedge, S. P., Paterson, K. B. & Pickering, M. J. (1998). Eye movements and measures of reading time. I *Eye guidance in reading and scene perception*. (s. 55-75). Elsevier Science Ltd. <https://doi.org/10.1016/B978-008043361-5/50004-3>
- Mai, Y. (2018). Statistical Power Analysis for Oneway ANOVA with Binary or Count Data. I Z. Zhiyong & Y. Ke-Hai (Red.), *Practical Statistical Power Analysis Using Webpower and R* (s. 87-143). Granger, IN: ISDSA Press.
- Mancini, S. (2018). When Grammar and Parsing Agree [Opinion]. *Frontiers in Psychology*, 9(336). <https://doi.org/10.3389/fpsyg.2018.00336>
- Næs, O. (1972). *Norsk grammatikk: Elementære strukturer og syntaks*. Fabritius. <https://books.google.no/books?id=JVhCDQEACAAJ>
- Næss, S. M. (2006). *Kan han snakker norsk? En generativ analyse av ja/nei-spørsmål i norsk som andrespråk*. [Norges teknisk-naturvitenskapelige universitetet (NTNU)].
- Ni, W., Fodor, J. D., Crain, S. & Shankweiler, D. (1998). Anomaly Detection: Eye Movement Patterns. *Journal of Psycholinguistic Research*, 27(5), 515-539. <https://doi.org/10.1023/A:1024996828734>
- Nikolaeva, I. (2007). *Finiteness: Theoretical and Empirical Foundations* [Book]. OUP Oxford. <http://search.ebscohost.com/login.aspx?direct=true&db=nlebk&AN=191604&site=ehost-live>
- Nunan, D. (1992). *Research Methods in Language Learning*. Cambridge University Press.
- Pinker, S. (2000). *Words and Rules: The Ingredients of Language*. HarperCollins. <https://books.google.no/books?id=wxMZ8ssz2iIC>
- Platzack, C. (1996). The Initial Hypothesis of Syntax: A Minimalist Perspective on Language Acquisition and Attrition. I H. Clahsen (Red.), *Generative Perspectives on Language Acquisition: Empirical findings, theoretical considerations and crosslinguistic comparisons*. John Benjamins Publishing Company. <https://doi.org/10.1075/lald.14.15pla>
- Płużyczka, M. (2018). The First Hundred Years: a History of Eye Tracking as a Research Method. *Applied Linguistics Papers*, 4/2018, 101-116. <https://doi.org/10.32612/uw.25449354.2018.4.pp.101-116>

- Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin*, 124(3), 372-422.  
<https://doi.org/10.1037/0033-2909.124.3.372>
- Rayner, K. & Balota, D. A. (1989). Parafoveal Preview and Lexical Access During Eye Fixations in Reading. I W. D. Marslen-Wilson (Red.), *Lexical Representation and Process* (s. 261-290). MIT Press.
- Rayner, K. & Pollatsek, A. (1989). *The psychology of reading*. Prentice-Hall, Inc.
- Rayner, K., Warren, T., Juhasz, B. J. & Liversedge, S. P. (2004). The Effect of Plausibility on Eye Movements in Reading. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 30(6), 1290-1301.  
<https://doi.org/10.1037/0278-7393.30.6.1290>
- Reed, J. B. & Meyer, R. J. (2007). Edmund Burke Huey (1870–1913): A brief life with an enduring legacy. I *Shaping the reading field: The impact of early reading pioneers, scientific research, and progressive ideas* (s. 159-175).
- Roberts, L. & Liszka, S. A. (2013). Processing tense/aspect-agreement violations online in the second language: A self-paced reading study with French and German L2 learners of English. *Second Language Research*, 29(4), 413-439.
- Rossi, S., Gugler, M. F., Hahne, A. & Friederici, A. D. (2005). When word category information encounters morphosyntax: An ERP study. *Neuroscience Letters*, 384(3), 228-233. <https://doi.org/10.1016/j.neulet.2005.04.077>
- Sagarra, N. & Ellis, N. C. (2013). From seeing adverbs to seeing verbal morphology: Language experience and adult acquisition of L2 tense. *Studies in Second Language Acquisition*, 35(2), 261-290.  
<https://doi.org/10.1017/S0272263112000885>
- Salvi, S. M., Akhtar, S. & Currie, Z. (2006). Ageing changes in the eye. *Postgraduate medical journal*, 82(971), 581-587.  
<https://doi.org/10.1136/pgmj.2005.040857>
- Sampson, G. (2002). Regional variation in the English verb qualifier system. *English Language and Linguistics*, 6(1), 17-30.  
<https://doi.org/10.1017/S1360674302001028>
- Simonsen, H. G. & Bjerkan, K. M. (1998). Testing past tense inflection in Norwegian: a diagnostic tool for identifying SLI children? *International Journal of Applied Linguistics*, 8(2), 251-270. <https://doi.org/10.1111/j.1473-4192.1998.tb00132.x>
- Sjekkeland, M. (2005). *Dialektar i Noreg*. Høyskoleforlaget.
- Slabakova, R. (2016). *Second Language Acquisition*. Oxford University Press.  
<https://books.google.no/books?id=XTD0CwAAQBAJ>



- Smythies, J. (1996). A Note on the Concept of the Visual Field in Neurology, Psychology, and Visual Neuroscience. *Perception*, 25(3), 369-371. <https://doi.org/10.1068/p250369>
- Sprouse, J. & Almeida, D. (2013). The empirical status of data in syntax: A reply to Gibson and Fedorenko. *Language and Cognitive Processes*, 28, 222-228. <https://doi.org/10.1080/01690965.2012.703782>
- Sprouse, J. & Almeida, D. (2017). Setting the empirical record straight: Acceptability judgments appear to be reliable, robust, and replicable. *Behavioral and Brain Sciences*, 40, e311, Artikkel e311. <https://doi.org/10.1017/S0140525X17000590>
- Staub, A. (2007). The parser doesn't ignore intransitivity, after all. *Journal of experimental psychology. Learning, memory, and cognition*, 33(3), 550-569. <https://doi.org/10.1037/0278-7393.33.3.550>
- Staub, A. & Rayner, K. (2007). Eye movements and on-line comprehension processes. I M. G. Gaskell (Red.), *The Oxford Handbook of Psycholinguistics* (s. 327-342). Oxford University Press.
- Stowe, L. A., Kaan, E., Sabourin, L. & Taylor, R. C. (2018). The sentence wrap-up dogma. *Cognition*, 176, 232-247. <https://doi.org/10.1016/j.cognition.2018.03.011>
- Sullivan, G. M. & Feinn, R. (2012). Using Effect Size-or Why the P Value Is Not Enough. *Journal of graduate medical education*, 4(3), 279-282. <https://doi.org/10.4300/JGME-D-12-00156.1>
- Svartdal, F. (2015). *Psykologiens forskningsmetoder* (4. utg.). Fagbokforlaget.
- Tsimpli, I. M. & Dimitrakopoulou, M. (2007). The Interpretability Hypothesis: evidence from wh-interrogatives in second language acquisition. *Second Language Research*, 23(2), 215-242. <https://doi.org/10.1177/0267658307076546>
- Utdanningsdirektoratet. (2014). *Undersøkelse av nynorsk som hovedmål*.
- van Peer, W., Hakemulder, F. & Zyngier, S. (2012). *Scientific Methods for the Humanities*. John Benjamins Publishing Company. <https://doi.org/10.1075/lal.13>
- Volkman, F. C. (1986). Human visual suppression. *Vision Research*, 26(9), 1401-1416. [https://doi.org/10.1016/0042-6989\(86\)90164-1](https://doi.org/10.1016/0042-6989(86)90164-1)
- Warren, P. (2012). *Introducing Psycholinguistics*. Cambridge University Press. <https://doi.org/10.1017/CBO9780511978531>
- Warren, T., McConnell, K. & Rayner, K. (2008). Effects of context on eye movements when reading about possible and impossible events. *Journal of experimental*

- psychology. Learning, memory, and cognition*, 34(4), 1001-1010.  
<https://doi.org/10.1037/0278-7393.34.4.1001>
- Wexler, K. (1994). Optional infinitives, head movement and the economy of derivations. I D. Lightfoot & N. Hornstein (Red.), *Verb movement* (s. 305-363). CUP.
- Wik, M. A. (2014). *Om tempus og finitthet i norsk som andrespråk*. [Norges teknisk-naturvitenskapelige universitet].
- Wooldridge, J. M. (2012). *Introductory Econometrics: A Modern Approach* (5. utg.). South-Western College Publishing.

# Vedlegg A: Stimuliliste

Interessefeltene er markert med [klammeparenteser]. Dette er bare en indikasjon for leseren av denne oppgaven, under eksperimentet leste informantene setninger uten noen markeringer. Tallet angir hvilket stimuliset settninger tilhører, mens bokstaven indikerer betingelsen (A – kontrollbetingelse, B – finittavvik, C – tempusavvik, D – finitt- pluss tempusavvik).

- 1-A. Julie liker å [passe] [barna] til naboen sin.
- 1-B. Hver helg drar Julie bort dit og [passe] [barna] til naboen sin.
- 1-C. Julie har ofte [passe] [barna] til naboen sin.
- 1-D. I går dro Julie bort dit og [passe] [barna] til naboen sin.
- 2-A. Jeg liker å [passe] [søsknene] mine.
- 2-B. Mens mamma er på jobb, er jeg vanligvis hjemme og [passe] [søsknene] mine.
- 2-C. Jeg har [passe] [søsknene] mine hele sommeren.
- 2-D. Sist fredag var jeg hjemme og [passe] [søsknene] mine.
- 3-A. Jonas liker ikke å [passe] [huset] til sin søster.
- 3-B. Hver sommerferie reiser Jonas dit og [passe] [huset] til sin søster.
- 3-C. Jonas har ofte [passe] [huset] til sin søster.
- 3-D. Sist sommerferie reiste Jonas dit og [passe] [huset] til sin søster.
- 4-A. Jeg må [passe] [alle plantene] i kollektivet mitt.
- 4-B. Hver sommer går jeg rundt og [passe] [alle plantene] i kollektivet mitt.
- 4-C. Helt siden jeg flyttet inn, har jeg [passe] [alle plantene] i kollektivet mitt.
- 4-D. I hele fjor sommer gikk jeg rundt og [passe] [alle plantene] i kollektivet mitt.
- 5-A. Fredrik tilbød seg å [passe] [kjæledyrene] til alle i nabolaget gratis.
- 5-B. I sommerferien kommer Fredrik og [passe] [kjæledyrene] til alle i nabolaget gratis.
- 5-C. I to uker nå har Fredrik [passe] [kjæledyrene] til alle i nabolaget gratis.
- 5-D. Forrige vinter kom Fredrik og [passe] [kjæledyrene] til alle i nabolaget gratis.
- 6-A. Ola velger å [passe] [ungene] sine selv.
- 6-B. Ola blir hjemme og [passe] [ungene] sine selv.
- 6-C. Ola har i tre år [passe] [ungene] sine selv.
- 6-D. Fram til i fjor var Ola hjemme og [passe] [ungene] sine selv.

- 7-A. Drømmen min er å [passe] [delfinene] i Sea World.
- 7-B. Hver sommer drar jeg til USA og [passe] [delfinene] i Sea World.
- 7-C. I to år har jeg [passe] [delfinene] i Sea World.
- 7-D. For to år siden dro jeg til USA og [passe] [delfinene] i Sea World.
- 8-A. Katrine må [passe] [mannen] sin døgnet rundt.
- 8-B. Katrine sitter og [passe] [mannen] sin døgnet rundt.
- 8-C. I tre måneder nå har Katrine [passe] [mannen] sin døgnet rundt.
- 8-D. Fram til hans død satt Katrine og [passe] [mannen] sin døgnet rundt.
- 9-A. Martine lovet å [vaske] [hele leiligheten] vår.
- 9-B. I dag kommer Martine og [vaske] [hele leiligheten] vår.
- 9-C. Martine har allerede [vaske] [hele leiligheten] vår.
- 9-D. Mens vi spiste, kom Martine og [vaske] [hele leiligheten] vår.
- 10-A. Det er jeg som skal [vaske] [den store] dansesalen.
- 10-B. Hver helg drar jeg og [vaske] [den store] dansesalen.
- 10-C. Jeg har nettopp [vaske] [den store] dansesalen.
- 10-D. Etter forestillingen dro jeg og [vaske] [den store] dansesalen.
- 11-A. Romkameraten min pleier å [vaske] [håret] sitt med min dyre sjampo.
- 11-B. Hver dag dusjer romkameraten min tidlig og [vaske] [håret] sitt med min dyre sjampo.
- 11-C. Romkameraten min har ofte [vaske] [håret] sitt med min dyre sjampo.
- 11-D. I går dusjet romkameraten min tidlig og [vaske] [håret] sitt med min dyre sjampo.
- 12-A. Hushjelpen vår skal [vaske] [sengetøyet] på den høyeste temperaturen.
- 12-B. Hver søndag kommer hushjelpen vår og [vaske] [sengetøyet] på den høyeste temperaturen.
- 12-C. Hushjelpen vår har aldri før [vaske] [sengetøyet] på den høyeste temperaturen.
- 12-D. Sist uke kom hushjelpen vår og [vaske] [sengetøyet] på den høyeste temperaturen.
- 13-A. Jenny skal [vaske] [bilen sin] før hun drar på jobb.
- 13-B. Hver dag kjører Jenny innom Circle-K og [vaske] [bilen sin] før hun drar på jobb.
- 13-C. Jenny hadde [vaske] [bilen sin] før hun dro på jobb.
- 13-D. I går kjørte Jenny innom Circle-K og [vaske] [bilen sin] før hun dro på jobb.
- 14-A. Alle må [vaske] [kaffetrakteren] etter bruk.
- 14-B. Alle ansatte her holder kjøkkenet rent og [vaske] [kaffetrakteren] etter bruk.
- 14-C. Jeg har alltid [vaske] [kaffetrakteren] etter bruk.
- 14-D. I fjor holdt vi kjøkkenet rent og [vaske] [kaffetrakteren] etter bruk.
- 15-A. Unni hater å [vaske] [ansiktet] med vann og såpe.

- 15-B. Hver morgen pusser Unni tenner og [vaske] [ansiktet] med vann og såpe.
- 15-C. I flere dager har Unni [vaske] [ansiktet] med vann og såpe.
- 15-D. I går kveld pusset Unni tenner og [vaske] [ansiktet] med vann og såpe.
- 16-A. Det er mitt ansvar å [vaske] [hunden] vår.
- 16-B. Etter hver luftetur tar jeg og [vaske] [hunden] vår.
- 16-C. Jeg har allerede [vaske] [hunden] vår.
- 16-D. I går morges tok jeg og [vaske] [hunden] vår.
- 17-A. Camilla får ikke lov til å [gjøre] [lekser] foran TV-en.
- 17-B. Hver kveld sitter Camilla og [gjøre] [lekser] foran TV-en.
- 17-C. Camilla har alltid [gjøre] [lekser] foran TV-en.
- 17-D. Forrige helg satt Camilla og [gjøre] [lekser] foran TV-en.
- 18-A. Hilde ønsker å [gjøre] [livet] lettere for oss alle.
- 18-B. Hver dag handler Hilde mat, og det [gjøre] [livet] lettere for oss alle.
- 18-C. Hilde har allerede [gjøre] [livet] lettere for oss alle.
- 18-D. I går handlet Hilde mat og [gjøre] [livet] lettere for oss alle.
- 19-A. Sjefen kan ikke lenger [gjøre] [endringer] i kontrakten min.
- 19-B. Hvert år utfører sjefen underveisvurdering og [gjøre] [endringer] i kontrakten min.
- 19-C. Sjefen har aldri [gjøre] [endringer] i kontrakten min.
- 19-D. Etter disputten ble sjefen sint og [gjøre] [endringer] i kontrakten min.
- 20-A. Kristine vil [gjøre] [alt mulig] for å hjelpe oss.
- 20-B. Hver dag kommer Kristine og [gjøre] [alt mulig] for å hjelpe oss.
- 20-C. Hele mitt liv har Kristine [gjøre] [alt mulig] for å hjelpe oss.
- 20-D. Da pappa var syk, kom Kristine og [gjøre] [alt mulig] for å hjelpe oss.
- 21-A. Alle de ansatte er redde for å [gjøre] [feil på] jobben.
- 21-B. Hver dag er det noen som slurver og [gjøre] [feil på] jobben.
- 21-C. På et eller annet tidspunkt har alle [gjøre] [feil på] jobben.
- 21-D. Forrige måned var det flere ansatte som slurvet og [gjøre] [feil på] jobben.
- 22-A. Bærekraftig utvikling skal [gjøre] [verden] til et bedre sted.
- 22-B. Hver dag hjelper hun sårbare barn og [gjøre] [verden] til et bedre sted.
- 22-C. Kjærlighet har alltid [gjøre] [verden] til et bedre sted.
- 22-D. Denne kvinnen ga mat til de fattige og [gjøre] [verden] til et bedre sted.
- 23-A. Jeg trenger ikke noen til å [gjøre] [jobben] min for meg.
- 23-B. Av og til kommer hun og [gjøre] [jobben] min for meg.
- 23-C. Ingen har noen gang [gjøre] [jobben] min for meg.
- 23-D. Han var snill og [gjøre] [jobben] min for meg.
- 24-A. Jeg vil ikke [gjøre] [samme] feil om og om igjen.
- 24-B. Hver dag våkner jeg og [gjøre] [samme] feil om og om igjen.

- 24-C. Hele livet mitt har jeg [gjøre] [samme] feil om og om igjen.
- 24-D. I alle mine forhold var jeg dum og [gjøre] [samme] feil om og om igjen.
- 25-A. Silje insisterer på å [bruke] [god tid] på å sminke seg.
- 25-B. Hver morgen sitter Silje og [bruke] [god tid] på å sminke seg.
- 25-C. Silje har alltid [bruke] [god tid] på å sminke seg.
- 25-D. Fredag kveld satt Silje og [bruke] [god tid] på å sminke seg.
- 26-A. Ane liker å [bruke] [pene vinflasker] som lysestaker.
- 26-B. For å pynte rommet sitt, kjøper Ane blomster og [bruke] [pene vinflasker] som lysestaker.
- 26-C. I mange år har Ane [bruke] [pene vinflasker] som lysestaker.
- 26-D. Til middagen i går kjøpte Ane nye servietter og [bruke] [pene vinflasker] som lysestaker.
- 27-A. Kristine liker ikke å [bruke] [solkrem] i ansiktet.
- 27-B. Hver dag tar Kristine vare på huden sin og [bruke] [solkrem] i ansiktet.
- 27-C. Kristine har aldri [bruke] [solkrem] i ansiktet.
- 27-D. Da Kristine bodde i Spania, satt hun mye i sola og [bruke] [solkrem] i ansiktet.
- 28-A. Mamma begynte å [bruke] [mitt gamle] soverom som atelier.
- 28-B. Når jeg ikke er hjemme, kommer mamma og [bruke] [mitt gamle] soverom som atelier.
- 28-C. De siste årene har mamma [bruke] [mitt gamle] soverom som atelier.
- 28-D. I fjor sommer kom mamma og [bruke] [mitt gamle] soverom som atelier.
- 29-A. Politiet vurderer å [bruke] [sporhunder] i jakten på terrorister.
- 29-B. Nå for tida tenker politiet nytt og [bruke] [sporhunder] i jakten på terrorister.
- 29-C. I ukesvis har politiet [bruke] [sporhunder] i jakten på terrorister.
- 29-D. I fjor tenkte politiet nytt og [bruke] [sporhunder] i jakten på terrorister.
- 30-A. Håndverkerne lover å ikke [bruke] [feil maling].
- 30-B. Mange håndverkere slurver og [bruke] [feil maling].
- 30-C. Håndverkerne har to ganger allerede [bruke] [feil maling].
- 30-D. I går slurvet håndverkerne og [bruke] [feil maling].
- 31-A. Jeg foretrekker å [bruke] [linser] i stedet for briller.
- 31-B. På jobb leser jeg mye og [bruke] [linser] i stedet for briller.
- 31-C. I noen måneder har jeg [bruke] [linser] i stedet for briller.
- 31-D. I går ville jeg prøve noe nytt og [bruke] [linser] i stedet for briller.
- 32-A. Eva velger å [bruke] [alle lommepengene] sine på julegaver.
- 32-B. Hvert år drar Eva hjem og [bruke] [alle lommepengene] sine på julegaver.
- 32-C. I flere år har Eva [bruke] [alle lommepengene] sine på julegaver.
- 32-D. I fjor dro Eva hjem og [bruke] [alle lommepengene] sine på julegaver.
- 33-A. Korister må alltid [synge] [rent og] harmonisk.

- 33-B. På prøvene står korister og [synge] [rent og] harmonisk.
- 33-C. Våre korister har alltid [synge] [rent og] harmonisk.
- 33-D. I går kveld sto korister på scenen og [synge] [rent og] harmonisk.
- 34-A. Jeg liker å [synge] [karaoke] med venner.
- 34-B. Hver lørdag kveld drar jeg på byen og [synge] [karaoke] med venner.
- 34-C. Helt siden jeg var tiåring har jeg [synge] [karaoke] med venner.
- 34-D. Forrige helg dro jeg på byen og [synge] [karaoke] med venner.
- 35-A. Det er alltid koselig å [synge] [julesanger] på julaften.
- 35-B. Hele familien min samler seg og [synge] [julesanger] på julaften.
- 35-C. Helt siden jeg var liten har hele familien min [synge] [julesanger] på julaften.
- 35-D. I fjor samlet hele familien min seg og [synge] [julesanger] på julaften.
- 36-A. En ekte prins må kjempe mot dragen og [synge] [ballader] til prinsessen.
- 36-B. I alle eventyr kjemper prinsen mot dragen og [synge] [ballader] til prinsessen.
- 36-C. Denne prinsen har aldri [synge] [ballader] til prinsessen.
- 36-D. I eventyret kjempet prinsen mot dragen og [synge] [ballader] til prinsessen.
- 37-A. Halvard er redd for å [synge] [soloen] under forestillingen.
- 37-B. Hvert år kommer Halvard og [synge] [soloen] under forestillingen.
- 37-C. Så lenge jeg kan huske har Halvard [synge] [soloen] under forestillingene.
- 37-D. I fjor kom Halvard og [synge] [soloen] under forestillingen.
- 38-A. Bandet nekter å [synge] [bare rockesanger] på turnéen.
- 38-B. I år drar bandet til Asia og [synge] [bare rockesanger] på turnéen.
- 38-C. Så langt har bandet [synge] [bare rockesanger] på turnéen.
- 38-D. I fjor dro bandet til Asia og [synge] [bare rockesanger] på turnéen.
- 39-A. Jeg vil gjerne lære å [synge] [opera] profesjonelt.
- 39-B. Venninna mi opptrer på teater og [synge] [opera] profesjonelt.
- 39-C. I nesten ti år har hun [synge] [opera] profesjonelt.
- 39-D. Før ulykken opptrådte hun på teater og [synge] [opera] profesjonelt.
- 40-A. Mannen min liker å [synge] [gamle] folkesanger.
- 40-B. Hver morgen steker mannen min egg og [synge] [gamle] folkesanger.
- 40-C. I to timer nå har mannen min [synge] [gamle] folkesanger.
- 40-D. Etter konserten i går kom mannen min hjem og [synge] [gamle] folkesanger.





## Vedlegg B: Parvise sammenligninger

### I. Verbalt interessefelt

#### *i. First fixation duration*

	<i>M</i>	<i>SD</i>
A: kontrollbetingelse	5.33	.083
B: finittavvik	5.42	.079
C: tempusavvik	5.41	.083
D: fin. + tempusavvik	5.43	.082

I	J	Mean Difference (I-J)	Std. Error	Sig.	95% CI for Difference	
					Lower Bound	Upper Bound
A	B	<b>-.088*</b>	.017	.000	-.136	-.041
	C	<b>-.083*</b>	.018	.000	-.133	-.034
	D	<b>-.098*</b>	.017	.000	-.145	-.051
B	A	<b>.088*</b>	.017	.000	.041	.136
	C	.005	.019	1.000	-.047	.058
	D	-.009	.016	.992	-.053	.034
C	A	<b>.083*</b>	.018	.000	.034	.133
	B	-.005	.019	1.000	-.058	.047
	D	-.015	.020	.976	-.069	.040
D	A	<b>.098*</b>	.017	.000	.051	.145
	B	.009	.016	.992	-.034	.053
	C	.015	.020	.976	-.040	.069

ii. *Gaze duration*

	<i>M</i>	<i>SD</i>
A: kontrollbetingelse	5.43	.087
B: finittavvik	5.50	.100
C: tempusavvik	5.52	.107
D: fin. + tempusavvik	5.53	.114

I	J	Mean Difference (I-J)	Std. Error	Sig.	95% CI for Difference	
					Lower Bound	Upper Bound
A	B	-.070*	.021	.011	-.129	-.012
	C	-.091*	.020	.000	-.147	-.035
	D	-.102*	.023	.000	-.166	-.039
B	A	.070*	.021	.011	.012	.129
	C	-.021	.022	.931	-.082	.041
	D	-.032	.022	.620	-.092	.028
C	A	.091*	.020	.000	.035	.147
	B	.021	.022	.931	-.041	.082
	D	-.011	.028	.999	-.088	.065
D	A	.102*	.023	.000	.039	.166
	B	.032	.022	.620	-.028	.092
	C	.011	.028	.999	-.065	.088

iii. Regression path duration

	<i>M</i>	<i>SD</i>
A: kontrollbetingelse	5.53	.120
B: finittavvik	5.58	.120
C: tempusavvik	5.66	.151
D: fin. + tempusavvik	5.62	.164

I	J	Mean Difference (I-J)	Std. Error	Sig.	95% CI for Difference	
					Lower Bound	Upper Bound
A	B	-.054	.026	.236	-.126	.018
	C	-.137*	.029	.000	-.219	-.056
	D	-.095*	.032	.031	-.185	-.006
B	A	.054	.026	.236	-.018	.126
	C	-.083*	.025	.013	-.154	-.013
	D	-.042	.029	.666	-.123	.040
C	A	.137*	.029	.000	.056	.219
	B	.083*	.025	.013	.013	.154
	D	.042	.038	.849	-.062	.146
D	A	.095*	.032	.031	.006	.185
	B	.042	.029	.666	-.040	.123
	C	-.042	.038	.849	-.146	.062

iv. Dwell time

	<i>M</i>	<i>SD</i>
A: kontrollbetingelse	5.58	.095
B: finittavvik	5.77	.153
C: tempusavvik	5.91	.179
D: fin. + tempusavvik	5.84	.184

I	J	Mean Difference (I-J)	Std. Error	Sig.	95% CI for Difference	
					Lower Bound	Upper Bound
A	B	-.199*	.028	.000	-.276	-.123
	C	-.334*	.028	.000	-.411	-.258
	D	-.266*	.031	.000	-.352	-.181
B	A	.199*	.028	.000	.123	.276
	C	-.135*	.034	.002	-.229	-.041
	D	-.067	.037	.383	-.170	.036
C	A	.334*	.028	.000	.258	.411
	B	.135*	.034	.002	.041	.229
	D	.068	.044	.565	-.054	.189
D	A	.266*	.031	.000	.181	.352
	B	.067	.037	.383	-.036	.170
	C	-.068	.044	.565	-.189	.054

v. *Regressions out*

	<i>M</i>	<i>SD</i>
A: kontrollbetingelse	.08	.075
B: finittavvik	.07	.068
C: tempusavvik	.11	.086
D: fin. + tempusavvik	.08	.082

I	J	Mean Difference (I-J)	Std. Error	Sig.	95% CI for Difference	
					Lower Bound	Upper Bound
A	B	.013	.018	.976	-.036	.062
	C	-.022	.019	.822	-.074	.030
	D	.002	.019	1.000	-.050	.053
B	A	-.013	.018	.976	-.062	.036
	C	-.035	.017	.249	-.082	.012
	D	-.011	.015	.973	-.052	.030
C	A	.022	.019	.822	-.030	.074
	B	.035	.017	.249	-.012	.082
	D	.024	.021	.854	-.036	.083
D	A	-.002	.019	1.000	-.053	.050
	B	.011	.015	.973	-.030	.052
	C	-.024	.021	.854	-.083	.036

vi. *Fixation count*

	<i>M</i>	<i>SD</i>
A: kontrollbetingelse	.24	.106
B: finittavvik	.36	.141
C: tempusavvik	.50	.146
D: fin. + tempusavvik	.43	.165

I	J	Mean Difference (I-J)	Std. Error	Sig.	95% CI for Difference	
					Lower Bound	Upper Bound
A	B	-.115*	.028	.001	-.193	-.037
	C	-.254*	.027	.000	-.330	-.178
	D	-.183*	.029	.000	-.263	-.102
B	A	.115*	.028	.001	.037	.193
	C	-.139*	.030	.000	-.222	-.056
	D	-.068	.035	.308	-.164	.029
C	A	.254*	.027	.000	.178	.330
	B	.139*	.030	.000	.056	.222
	D	.071	.037	.325	-.032	.175
D	A	.183*	.029	.000	.102	.263
	B	.068	.035	.308	-.029	.164
	C	-.071	.037	.325	-.175	.032

## II. Postverbalt intereseffelt

### *i. First fixation duration*

	<i>M</i>	<i>SD</i>
A: kontrollbetingelse	5.39	.099
B: finittavvik	5.41	.094
C: tempusavvik	5.40	.093
D: fin. + tempusavvik	5.41	.084

I	J	Mean Difference (I-J)	Std. Error	Sig.	95% CI for Difference	
					Lower Bound	Upper Bound
A	B	-.018	.020	.941	-.073	.037
	C	-.010	.017	.995	-.057	.038
	D	-.019	.018	.883	-.067	.030
B	A	.018	.020	.941	-.037	.073
	C	.008	.019	.998	-.043	.060
	D	-.001	.019	1.000	-.052	.051
C	A	.010	.017	.995	-.038	.057
	B	-.008	.019	.998	-.060	.043
	D	-.009	.020	.998	-.064	.046
D	A	.019	.018	.883	-.030	.067
	B	.001	.019	1.000	-.051	.052
	C	.009	.020	.998	-.046	.064

ii. Gaze duration

	<i>M</i>	<i>SD</i>
A: kontrollbetingelse	5.71	.307
B: finittavvik	5.68	.274
C: tempusavvik	5.67	.253
D: fin. + tempusavvik	5.68	.267

I	J	Mean Difference (I-J)	Std. Error	Sig.	95% CI for Difference	
					Lower Bound	Upper Bound
A	B	.033	.037	.940	-.070	.137
	C	.045	.029	.554	-.035	.125
	D	.035	.025	.668	-.034	.104
B	A	-.033	.037	.940	-.137	.070
	C	.012	.031	.999	-.075	.098
	D	.002	.033	1.000	-.089	.092
C	A	-.045	.029	.554	-.125	.035
	B	-.012	.031	.999	-.098	.075
	D	-.010	.030	1.000	-.094	.074
D	A	-.035	.025	.668	-.104	.034
	B	-.002	.033	1.000	-.092	.089
	C	.010	.030	1.000	-.074	.094



iii. *Regression path duration*

	<i>M</i>	<i>SD</i>
A: kontrollbetingelse	5.83	.361
B: finittavvik	5.94	.317
C: tempusavvik	6.07	.322
D: fin. + tempusavvik	5.99	.348

I	J	Mean Difference (I-J)	Std. Error	Sig.	95% CI for Difference	
					Lower Bound	Upper Bound
A	B	-.116*	.033	.008	-.209	-.023
	C	-.243*	.038	.000	-.350	-.137
	D	-.160*	.035	.000	-.257	-.063
B	A	.116*	.033	.008	.023	.209
	C	-.127*	.033	.002	-.218	-.037
	D	-.044	.038	.835	-.151	.063
C	A	.243*	.038	.000	.137	.350
	B	.127*	.033	.002	.037	.218
	D	.084	.048	.423	-.049	.216
D	A	.160*	.035	.000	.063	.257
	B	.044	.038	.835	-.063	.151
	C	-.084	.048	.423	-.216	.049

iv. Dwell time

	<i>M</i>	<i>SD</i>
A: kontrollbetingelse	5.89	.329
B: finittavvik	5.93	.328
C: tempusavvik	5.99	.303
D: fin. + tempusavvik	5.97	.336

I	J	Mean Difference (I-J)	Std. Error	Sig.	95% CI for Difference	
					Lower Bound	Upper Bound
A	B	-.045	.032	.666	-.132	.043
	C	<b>-.106*</b>	<b>.032</b>	<b>.014</b>	<b>-.195</b>	<b>-.016</b>
	D	<b>-.084*</b>	<b>.030</b>	<b>.043</b>	<b>-.166</b>	<b>-.002</b>
B	A	.045	.032	.666	-.043	.132
	C	-.061	.035	.418	-.157	.035
	D	-.039	.034	.823	-.133	.054
C	A	<b>.106*</b>	<b>.032</b>	<b>.014</b>	<b>.016</b>	<b>.195</b>
	B	.061	.035	.418	-.035	.157
	D	.022	.039	.995	-.086	.129
D	A	<b>.084*</b>	<b>.030</b>	<b>.043</b>	<b>.002</b>	<b>.166</b>
	B	.039	.034	.823	-.054	.133
	C	-.022	.039	.995	-.129	.086

v. *Regressions out*

	<i>M</i>	<i>SD</i>
A: kontrollbetingelse	.09	.096
B: finittavvik	.23	.135
C: tempusavvik	.32	.142
D: fin. + tempusavvik	.27	.171

I	J	Mean Difference (I-J)	Std. Error	Sig.	95% CI for Difference	
					Lower Bound	Upper Bound
A	B	-.142*	.028	.000	-.221	-.063
	C	-.224*	.023	.000	-.287	-.162
	D	-.178*	.025	.000	-.246	-.110
B	A	.142*	.028	.000	.063	.221
	C	-.082*	.025	.013	-.152	-.013
	D	-.036	.034	.874	-.130	.058
C	A	.224*	.023	.000	.162	.287
	B	.082*	.025	.013	.013	.152
	D	.046	.036	.740	-.052	.145
D	A	.178*	.025	.000	.110	.246
	B	.036	.034	.874	-.058	.130
	C	-.046	.036	.740	-.145	.052

vi. *Fixation count*

	<i>M</i>	<i>SD</i>
A: kontrollbetingelse	.49	.314
B: finittavvik	.53	.306
C: tempusavvik	.61	.279
D: fin. + tempusavvik	.58	.316

I	J	Mean Difference (I-J)	Std. Error	Sig.	95% CI for Difference	
					Lower Bound	Upper Bound
A	B	-.039	.024	.523	-.105	.028
	C	<b>-.114*</b>	<b>.029</b>	<b>.002</b>	<b>-.194</b>	<b>-.034</b>
	D	<b>-.084*</b>	<b>.026</b>	<b>.013</b>	<b>-.155</b>	<b>-.013</b>
B	A	.039	.024	.523	-.028	.105
	C	-.076	.028	.063	-.154	.003
	D	-.046	.029	.534	-.125	.034
C	A	<b>.114*</b>	<b>.029</b>	<b>.002</b>	<b>.034</b>	<b>.194</b>
	B	.076	.028	.063	-.003	.154
	D	.030	.034	.943	-.063	.123
D	A	<b>.084*</b>	<b>.026</b>	<b>.013</b>	<b>.013</b>	<b>.155</b>
	B	.046	.029	.534	-.034	.125
	C	-.030	.034	.943	-.123	.063

