

Robert Wiktor Krogstad

Andrespråksinnlæreres reaksjon på avvik i norsk

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

Masteroppgave i lektorutdanning i språkfag med nordisk

Veileder: Kristin Melum Eide

Medveileder: Mila Vulchanova

Mai 2023

Robert Wiktor Krogstad

Andrespråksinnlæreres reaksjon på avvik i norsk

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

Masteroppgave i lektorutdanning i språkfag med nordisk
Veileder: Kristin Melum Eide
Medveileder: Mila Vulchanova
Mai 2023

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



Kunnskap for en bedre verden

Sammendrag

Alle setningene under i 1A-D er eksempler på eksperimentets stimuli. Setning 1A har sekvensen plassert i en kontekst der den morfologiske verbformen, infinitiv, passer inn i setningen. I de neste tre setningene fører ordsekvensen til avvik, siden konteksten i 1B krever et verb i presens, kontekst 1C krever et verb i perfektum partisipp, 1D krever et verb i preteritum.

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.

Når nordmenn leser slike setninger fører de morfologiske avvikene til en annerledes øyeatferd, altså leseatferd, som kan måles ved hjelp av blikksporing. Nordmenn stopper opp og bruker lengre tid på å lese ved slike avvik i teksten. Dette ble påvist i studien til Vinichenko (2021).¹

I min masteroppgave utfører jeg den samme typen studie som Vinichenko (2021), men denne gangen med en testgruppe som ikke har norsk som førstespråk. Øyeatferden til voksne andrespråksinnlærere skiller seg fra den hos dem som har norsk som førstespråk, men det er også interessante likheter i noe at øyeatferden også. I min studie brukte jeg samme metode og samme stimuli som Vinichenko (2021), fordi mine resultater skulle sammenlignes mest mulig direkte med resultatene fra Vinichenkos studie.

Testsetningene inneholdt altså morfologiske avvik der avvikene besto i enten finittavvik (1B), tempusavvik (1C) eller avvik med både finittet og tempus (1D). I tillegg leste informantene kontrollsetninger uten noen avvik, for å kunne sammenligne øyeatferden fra setninger uten avvik med øyeatferden i testsetningene med de ulike avvikene.

Problemstillingen baserer seg på Eides (2009) teori om at finittet og tempus er to adskilte trekk, og hypotesen både i Vinichenko (2021) og i min studie er derfor at finitt og tempus også *oppfattes* som to adskilte trekk. Det vil så resultere i forskjellige lesemønstre. Som nevnt viser konklusjonene i Vinichenko (2021) at finitt- og tempusavvik faktisk fører til ulik øyeatferd når nordmenn leser setninger med slike avvik. Problemstillingen til mitt prosjekt handler om hvorvidt andrespråksinnlærere av norsk prosesserer de verbmorfologiske avvikene på en måte som ligner på nordmenns, førstespråksinnlærere av norsk, selv om mine informanter har lært norsk i voksen alder. Hovedhypotesen var at andrespråksinnlærere vil vise samme type atferd som nordmenn.

Resultatene stemte delvis eller i liten grad med hypotesen. Bare ved måling av en spesifikk type øyeatferd, fant vi en statistisk signifikant forskjell mellom kontrollstimuli og stimuli med tempusavvik. En mulig forklaring på dette mønsteret er at avvik i tempusbøying har større semantisk betydning, mens avvik i finittet først og fremst gir seg syntaktiske utslag. I diskusjonen av funnene i min studie tar jeg utgangspunkt i andrespråksteoriene *Full Transfer*, *Full Access*-teorien (FTFA) og *Failed Functional Feature Hypothesis* (FFFH). Selv om det trengs mer forskning og mer data, er resultatene og analysen i denne studien mest kompatibel med FFFH, som tilsier at formelle (ikke tolkbare) trekk er vanskeligere å tilegne seg i voksen alder dersom du ikke har dette

¹ Dette er en masteroppgave som ble utført på samme språklab som denne studien: *Nordmenns reaksjoner på avvik i innlærernorsk: En eye-tracking undersøkelse av effekter av finitt- og tempusavvik på språkprosessering.*

trekket i ditt morsmål. FTFA trekker i retning av at det ikke burde finnes noen påvisbar forskjell mellom tilegnelsen av formelle (syntaktiske) og semantiske trekk, i motsetning til det mine funn tyder på.

Abstract

All sentences below in 1A-D are examples of this experiment's stimuli. In sentence 1A the sequence is placed in a context where the morphological verb form, infinitive, fits the sentence. The three other sentences lead to anomalies. This is because the context in 1B demands a verb in present, 1C demands verb in perfectum partisipp, and 1D demands a verb in past tense.

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.

When a Norwegian reads sentences like these (1A-D), the morphological anomalies lead to a different eye behaviour, i.e. reading behaviour, which can be measured with eye tracking. Norwegians stop and spend more time reading when these anomalies occur. This was shown in the study done by Vinichenko (2021)²

I did the same type of study as Vinichenko (2021) in my master thesis, but this time with another test group, which don't have Norwegian as first language. The eye behaviour of the adult second language learners is different from those who have Norwegian as first language, but there are also interesting similarities with the eye behaviour as well. The same method and the same stimuli from Vinichenko (2021) were used in this study, because my results could be compared directly with results from Vinichenko's study.

The test sentences contained morphological anomalies where they either contained a finite anomaly (1B), a tempus anomaly (1C) or both finite and tempus anomaly (1D). In addition, the informants read control sentences without anomalies, so the eye behaviour from sentences without anomalies could be compared to the test sentences with anomalies.

The research question is based Eide's (2009) theory that finiteness and tense are to separate features and hypothesis, both in Vinichenko's (2021) and my study, are therefore that finiteness and tempus is perceived as two distinct features. Which means it will result in different reading patterns. As mentioned, the conclusion in Vinichenko (2021) show that finite and tempus anomalies lead to different eye behaviour when Norwegians read different sentences with these anomalies. The research question for my project is about whether or not second language learners of Norwegian process the morphological anomalies in verbs in some way like Norwegians (first language learners of Norwegian), even though my informants learned Norwegian as adults. The main hypothesis was that second language learners will show the same type of behaviour as Norwegians.

The results partially or to a small degree fit the hypothesis. We found just one measurement of a specific type of eye behaviour where we found a statically significant difference between the control stimuli and the test stimuli with tempus anomaly. An explanation to this pattern is that anomalies in tempus conjugation have more semantic meaning, while finite anomalies first and foremost have syntactic consequences. In the

² This is a master thesis which was done in the same language lab as this study: *Nordmenns reaksjoner på avvik i innlærernorsk: En eye-tracking undersøkelse av effekter av finitt- og tempusavvik på språkprosessering*. The title may be translated to: *Norwegians reactions to anomalies in learner Norwegian: An eye-tracking study of effects finite and tempus anomalies on language processing*.

discussion of the results of my study the theoretical foundation is the second language theories *Full Transfer, Full Access-* (FTFA) and *Failed Functional Feature Hypothesis* (FFFH). Even though more research and more data are needed, the results and analysis in this study is most compatible with FFFH, which means formal (uninterpretable) features are harder to acquire as an adult if you don't have this feature in your mother tongue. FTFA on the other hand would imply that there shouldn't be a demonstrable difference between formal (syntactic) and semantic features, in contrast to what my results show.

Forord

Da vil jeg takke mine to fantastiske veiledere Kristin Melum Eide og Mila Vulchanova, som har gitt meg muligheten til å ta i bruk språklaboratoriet og utstyret til å gjennomføre det som har vært et svært spennende masterprosjekt for meg. De har virkelig vært gode veiledere som har hjulpet meg med å fullføre prosjektet og med å gjøre meg til en bedre språkviter igjennom hele prosessen.

Enda en takk til Kristin Melum Eide, som har hjulpet til med teoretiske innspill, oppbyggingen av prosjektet og har vært en uvurderlig pådriver for gjennomføringen av masterprosjektet. Veiledningene var preget av varme, ærlighet og et engasjement uten like, som virkelig har vært helt nødvendig, til stor hjelp og er beundringsverdig.

Enda en takk til Mila Vulchanova, som foreleste i emnet *språk og kognisjon* og inspirerte meg til å utforske psykolingvistikken som fagfelt. Som medveilederen min har hun vært god hjelp og støttespiller med det tekniske, statistikken og forskningslitteratur. Kunnskapen om og engasjementet for feltet, som hun har, er høyst imponerende og har vært verdifull og berikende for meg.

Takk til medstudenten min Amanda Isaksen for å hjelpe til med rekruttering av informanter, lære meg å bruke blikksporingkameraet og å gjennomføre eksperimentene.

Takk til informantene for å stille opp og gjøre dette prosjektet mulig.

Og sist, men absolutt ikke minst: takk til foreldrene mine, Inger Gudny Krogstad og John Gunnar Krogstad, for å støtte meg underveis, for å vise interesse for masteroppgaven min og for å være leseglade begge to.

Innholdsfortegnelse

SAMMENDRAG	1
ABSTRACT	3
FORORD	5
INNHOLDSFORTEGNELSE	6
TABELLER OG FIGURER	7
1 INNLEDNING	8
1.1 Finitthetsteori	10
1.2 Hvorfor psykolingvistisk perspektiv på andrespråk?	12
1.3 Forskningsfronten	12
2 METODE	14
2.1 Blikksporing i språkvitenskap	14
2.2 Kognisjon og blikksporing	15
2.3 Design av eksperiment	16
2.4 Hva måles under blikksporing?	19
2.5 Utvalg av informanter	20
2.6 Gjennomføring	23
2.7 Utstyr	23
2.8 Mulige feilkilder og motvirkende tiltak	24
2.9 Diskusjon om prosjektets validitet og reliabilitet	24
3 RESULTATER OG STATISTISK ANALYSE	26
3.1 Variabler	26
3.2 Prosedyre for datarensing og analyse	28
3.3 Deskriptiv statistikk	28
3.4 Statistisk analyse – ANOVA og T-test	32
3.4.1 rm-ANOVA	32
3.4.2 T-test – første fikseringslengde	34
3.4.3 Sammenligning av grupper	36
3.5 Kapitteloppsummering	36
4 DISKUSJON	37
4.1 Resultatdiskusjon	37
4.1.1 Hva sier det om øyefatferd og språkprosessering?	38
4.1.2 Teori til diskusjon - Andrespråksteori	39
4.1.3 Analyse med hjelp av andrespråksteori	41
4.2 Metodologisk diskusjon	41
4.3 Videre forskning	42
4.4 Konklusjon	44
5 REFERANSELISTE	45
6 MASTEROPPGAVENS RELEVANS FOR LÆRERYRKET	49

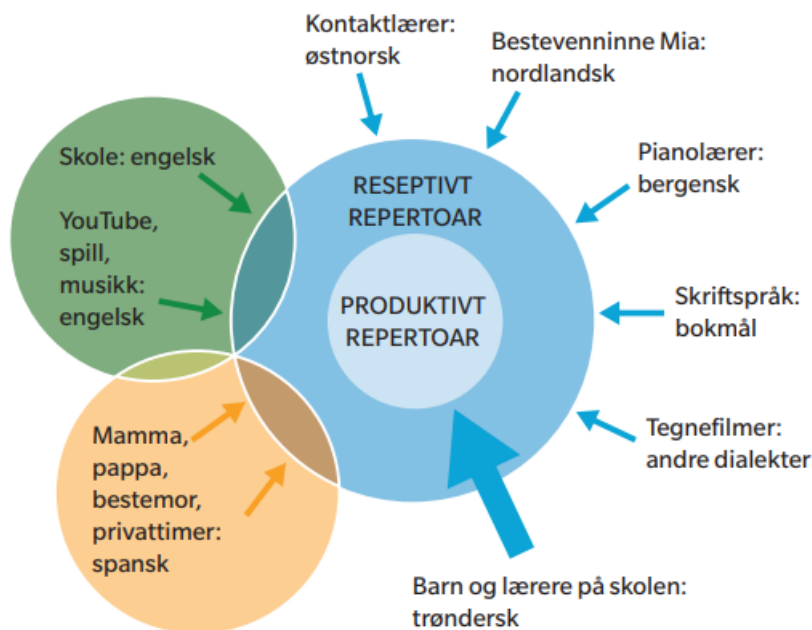
Tabeller og figurer

<i>Tabell 1</i>	<i>10</i>
<i>Tabell 2</i>	<i>11</i>
<i>Tabell 3</i>	<i>22</i>
<i>Tabell 4</i>	<i>26</i>
<i>Tabell 5</i>	<i>29</i>
<i>Tabell 6</i>	<i>29</i>
<i>Tabell 7</i>	<i>30</i>
<i>Tabell 8</i>	<i>30</i>
<i>Tabell 9</i>	<i>33</i>
<i>Tabell 10</i>	<i>35</i>
<i>Tabell 11</i>	<i>43</i>
<i>Figur 1</i>	<i>9</i>
<i>Figur 2</i>	<i>16</i>

1 Innledning

Språket vårt er kanskje det viktigste redskapet vi har. Det brukes til å kommunisere på ulike måter i ulike områder i livene våre, enten det skal skrives en jobbsøknad, en artikkel, en rapport til sjefen, musikk skal nytes eller når vi skal lytte og kommunisere med fremmede, nære og kjære. Med dagens stadig globaliserte og internasjonaliserte verdenssamfunn som vi nå lever i, hvor det blir stadig mer og mer vanlig å flytte over landegrensener og kunne flere språk. Behovet for kunnskap om menneskets evne til å lære seg nye språk blir mer og mer relevant når landegrensene blir utvisket. Selv om det finnes mange ulike, og betydelige, språkvitenskapelige perspektiver, har språkvitenskapen de siste 50 årene vært preget av den chomskyanske eller generative tradisjonen, som har hatt stor innflytelse og også har vært viktig i den utbredte forståelsen av språktilegning. Generativ grammatikk antar typiske at mennesker har en universalgrammatikk i hjernen, som er en slags algoritme som styrer hvordan mennesket lærer og tilegner seg språk. Teorien innebærer også at mennesket kan lære seg språk i barn- og ungdomsalderen lettere enn i voksen alder, fordi universalgrammatikken er bedre utnyttbar hos barn, mer fleksibel og internaliserte grammatiske strukturer og regler formes lettere hos barn og unge (Penfield & Roberts, 1959 i Hyltenstam & Abrahamsson, 2003). Når man blir eldre, blir hjernen mindre plastisk og det reduserer universalgrammatikkens evne til å tilegne seg nye språklige og grammatiske strukturer. Tanken bak er at det er vanskeligere å tilegne seg den optimale språkkompetansen etter en viss alder, etter såkalte kritiske eller sensitive perioder i livet. Innenfor disse sensitive periodene er det enklere å lære seg ulike språklige fenomener som syntaks, fonologi og så videre (Hyltenstam & Abrahamsson, 2003). Et helt sentralt begrep i denne sammenhengen er *kompetanse*, som handler om hvor mye språkbrukeren kan om språket og hvilke mentale representasjoner språkbrukeren har i kognisjonen (Brown et al., 1996, s. 2). Kompetansebegrepet er relevant også i andrespråkstilegning, fordi vi vil finne ut hvor mye andrespråksinnlærerne av norsk faktisk kan lære seg som voksen i sitt nye språk.

Kompetanse har blitt forsøkt målt ved å forske på det produktive systemet i korpusstudier, casestudier, dokumentstudier, intervjuer og lignende hvor man forsker på hva andrespråksinnlærere produserer av språk. Vansken med det er at man i slike studier ofte måler *performanse*, som er faktisk bruken av språket i konkrete situasjoner (Brown et al., 1996, s. 2). Da forsker man ofte på det produktive systemet og repertoaret. I denne studien er det det reseptive systemet vi søker å forstå og forske på. Det repertoaret innebærer alt informantene forstår, men ikke nødvendigvis produserer (Eide, 2022, s. 388). Selv om informantenes prestasjoner i en testsituasjon med blikksporing også kan sies å være påvirket av performansen (som trøtthet, dagsform, oppmerksomhet etc.) er det likevel rimelig å anta at en slik testmetode i større grad nærmer seg informantens underliggende kompetanse, som er det vil vi undersøke. Dette er blant annet fordi øyeatferd som måles ved blikksporing i større grad er raske, ubevisste prosesser.



Figur 1 (Eide, 2022, s. 388): Eksempel på produktivt og reseptivt repertoar hos en seksåring i Trondheim.

Den delen av repertoaret vi undersøker, er tempus- og finittmorfologi i norske verb hos andrespråksinnlærere av norsk. Vi vil finne ut hvordan de reagerer på tempus- og finittavvik i norske verb når norsk ikke er førstespråket deres og de har begynt å lære norsk etter at de kritiske eller sensitive periodene er forbi.

Eksempler på testsetninger. A er kontrollsetning, B har finittavvik, C har tempusavvik og D har både finitt- og tempusavvik. Interessefeltene er markert i klammer

- 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.

Testsetningene inneholdt altså morfologiske avvik der avvikene besto i enten feil i finittthet (1B), feil tempus (1C) eller feil i både finittthet og tempus (1D). I tillegg leste informantene kontrollsetninger, uten noen avvik, for å kunne sammenligne øyeatferden ved slike setninger med den vi kunne observere i testsetningene med avvik.

Med teknologisk utvikling og fremgang har det skjedd flere metodologiske fremskritt, som gir oss helt nye metoder som kan brukes i språkforskning og kan hjelpe oss å besvare dette prosjektets problemstilling ved hjelp av blikksporing («eye-tracking»). I generativ tradisjon har en svært utbredt metode vært grammatikalitets- og akseptabilitetsvurderinger, som har vært nyttige for sitt formål, men slike introspektive metoder er bare en måte å måle kompetanse på. Denne metoden har dessuten blitt mye kritisert, fordi vurderinger ved introspeksjon ikke nødvendigvis har vært ansett som veldig pålitelige.³ Derfor er metodetriangulering, hvor man tar i bruk flere metoder,

³ Men se Sprouse & Almeida (2013)

nødvendig for å forstå språkprosessering ut ifra flere perspektiver. Dette gjør er det aktuelt å bruke blikksporing, som i større grad er blitt forbundet med kognitiv psykologi og språkvitenskap, fordi vi kan måle øyets språk- eller leseatferd. I motsetning til spørreskjema eller akseptabilitetsvurderinger er øyeatferd antatt å være direkte knyttet til hjernen og vil kunne gi innsikt som nærmer seg andrespråksinnlæreres kompetanse uten at resultatene blir tilslørt av bevisste vurderinger. Da kan det forskes på om informanter registrerer tempus- og finittavvik ut ifra øyeatferd ved lesing i stedet for å be dem vurdere avvikene selv. Med generativ grammatikk som paradigme og teoretisk grunnmur, har vi et godt utgangspunkt for å diskutere om blikksporingsdataene kan si noe om andrespråksinnlærere sin underliggende kompetanse i norsk.

1.1 Finittetsteori

Morfologi er formlære og handler blant annet om hvordan ord tar form i språk (Booij, 2012, s. 24). Prosjektet handler om hvorvidt andrespråksinnlærere av norsk oppfatter og prosesserer tempus- og finittetsmorfologi i norske verb, hvilket betyr at det handler om hvordan informantene reagerer på verbenes form, som påvirkes av tempus- og finittetsbøying. Ut ifra Eides finittetsteori (2012; 2009) tas det utgangspunkt i at verbmorfologien markerer bøying i tempus enten i fortid eller ikke-fortidsform og ved finittet markeres det i finitt og ikke-finitt form. Norske verb har flere nyanser og funksjoner enn disse to verbbøyingene, men disse er de sentrale tempusformene. Fellesnevneren for fortellende setninger er at de må inneholde et finitt verb for å være grammatisk korrekt. Derfor er finittet et sentralt grammatisk trekk i norsk verbmorfologi og syntaks, fordi det bare er et finitt verb i setninger, det må stemme overens med andre verb og være plassert korrekt i setningen, altså andre plass i setningen. Se tabell 1 for en oversikt:

	+Finitt	-Finitt
+Fortid	Preteritum	Perfektum partisipp
-Fortid	Presens	Infinitiv

Tabell 1

Testsetningene i denne studien forholder seg bare til finitte setninger hvor mangel på finitt verb regnes som ugrammatisk. Se på eksemplene under.

(1a) *Jeg spiller fotball*

(1b) *Jeg liker å spille fotball*

I (1a) og (1b) er det verdt å merke at verbet *spiller-spille* har ulik bøying i finitt og ikke-finitt form, slik at distinksjonen blir dermed tydeliggjort morfologisk sett. På engelsk er ikke distinksjonen morfologisk til stede hvis vi oversetter setningene i (1).

(2a) *I play football*

(2b) *I like to play football*

(2c) *I played football*

(2d) *I have played football*

I eksemplene på engelsk har verbet *play-play, played-played* lik bøying både som finitt og ikke-finitt verb i motsetning til norsk, som eksemplifiserer hvordan finittet er mer markert og sterkere til stede i norsk verbmorfologi enn i engelsk verbmorfologi. Det kan

dermed være et argument for at finittet er noe eget av egen størrelse og kanskje er en «parameter» i generativ forstand. På den andre siden kan man se at uregelrette verb i engelsk skiller mellom finitt og ikke-finitt i fortidsformene preteritum og perfektum partisipp morfologisk sett.

(3a) *I see the cat*

(3b) *I will see the cat*

(3c) *I saw the cat*

(3d) *I have seen the cat*

Norsk oversettelse:

(4a) *Jeg ser katten*

(4b) *Jeg vil se katten*

(4c) *Jeg så katten*

(4d) *Jeg har sett katten*

Selv om engelsk har en del uregelmessige verb som skiller mellom finitte og ikke finitte verb morfologisk, gjelder ikke dette det produktive systemet som nye verb i språket bøyes etter.

Derimot på norsk kan man se at det er ulik finitthetsmorfologi i både sterke, svake og uregelmessige verb, og alle nye verb som kommer inn i språket. Det vil si at finittet er mer produktivt i norsk. Hvis en ser på tabellen under (se fotnote for forklaring), kan man se en oversikt over bøyning av norske verb som er både sterke og svake. Legg merke til at det er fire ulike bøyinger på norsk i både sterke og svake verb, mens det bare er forskjell på finitt og ikke-finitt bøyning på engelsk i svake, produktive verb. Samtidig er det viktig å påpeke at de sterke verbene i engelsk er frekvente, selv om de bøyes etter en ikke produktiv regel, men tabellen kan illustrere hvordan finitthetsmorfologi er mer produktivt i norsk.

	+Finitt		-Finitt	
+Fortid	Preteritum		Perfektum Partisipp	
	Sterke verb	Svake verb	Sterke verb	Svake verb
	sang («sang»)	likte («liked»)	sunget («sung»)	likt («liked»)
-Fortid	Presens		Infinitiv	
	Sterke verb	Svake verb	Sterke verb	Svake verb
	synger («sing»)	liker («like»)	syng («sing»)	like («like»)

Tabell 2 (Eide, 2012, s. 64)⁴.

Ut ifra eksemplene kan en se at produktiv finitthetsmorfologi er mer utbredt i norske produktive verbklasser sammenlignet med engelske verb. I eksempelvis asiatiske språk som vietnamesisk, thai og kinesiske språk er det ikke verbmorfologi i det hele tatt

⁴ Tabellen er hentet fra kapitlet til Eide om tempus og aspekt i norsk i (Eide, 2012, s. 64), men oversettelsene av de norske verbene til engelsk er lagt til for å illustrere forskjellen mellom norsk og engelsk finitthetsbøyning.

hverken for tempus eller finitthet, slik at det teoretisk sett kan gjøre det vanskeligere for innlærere med disse språkene som førstespråk å lære seg tempus- og finitthetsmorfologi, spesielt kan vi anta at dette gjelder for finitthet. Grunnen til det er at tempus i større grad har en semantisk-konseptuell basis enn finitthet, som i større grad er et rent syntaaktisk trekk (Eide, 2016). De verbmorfologiske fenomenene som er blitt presentert i dette kapitlet er det grammatiske teorigrunnet for stimulisetningene som er brukt i eksperimentene (se innledningsvis eller i vedlegg i kap. 7.3 for eksempler). I blikksproingen ble det undersøkt om deltakerne reagerer mest på tempusavvik, på finittavvik hver for seg, eller om det er en større forskjell når begge typer avvik opptrer i samme setning.

1.2 Hvorfor psykolingvistisk perspektiv på andrespråk?

Forskningsprosjektet baserer seg på psykometrisk, kvantitativ empiri på 9 deltakere, der målet er å bruke blikksproing til å si noe om andrespråksinnlæreres evne til å lære seg nye grammatiske strukturer etter de kritiske og sensitive periodene for språktilegning. Hvorfor bruke denne metoden for innhenting av empiri? Hvorfor ikke bruke kvalitative metoder for gjennom å tolke andrespråksinnlæreres tekster i dokumenter eller korpus? Fordelen med blikksproing er at den samler inn øyets leseatferd i sanntid og ikke baserer seg på informantenes subjektive vurderinger om hva de mener kjennes grammatisk eller ikke. Grunnet for metoden blikksproing brukt i språkforskning er antakelsen om at det er en forbindelse mellom øyet og sinnet, slik at dataene kan brukes til å tolke hvordan språket prosesseres under lesing. Ulikt mange andre metoder undersøker denne metoden det reseptive systemet og ikke det produktive systemet. Mens andre metoder typisk i større grad må forholde seg til performanse, kan man med denne metoden komme nærmere den underliggende kompetansen, riktignok også målt via performanse i form av øyefatferd i det reseptive systemet, men som likevel måler raske, ubevisste prosesser. I stedet for å be informantene om å produsere ulike setninger selv for å vise hva de kan produsere, målspråklig eller ikke, vil vi kunne måle hvordan de mottar tekst med morfologiske avvik og om de oppdager avvikene via sitt reseptive system. Med hjelp av blikksproingskameraet gjennomføres eksperimentene på en gruppe informanter, som kan gi resultater med svar på om norskkompetansen til innlærere av norsk som andrespråk. Målet med metoden er dermed å bruke det psykolingvistiske perspektivet til å samle inn empiri, som kan svare på teoretiske spørsmål som ofte stilles i generativ grammatikk. Disse to språkvitenskapelige perspektivene kan sammen gi interessant innsikt i hvordan førstespråket kan påvirke kompetansen i målspråket. Denne metoden, blikksproing, blir gjort mer detaljert rede for i metodekapitlet.

1.3 Forskningsfronten

Et felt i utvikling de siste tre tiårene innen språkvitenskapen, er andrespråksforskning, som handler om hvordan vi evner å lære oss flere språk igjennom livet. Det økte behovet for kunnskap om andrespråk på grunn av økt globalisering og flerspråklig, gjør at vi trenger mer kunnskap om å lære seg flere språk. Det åpner også for interessante teoretiske problemstillinger om tilegning av språk. En av de grunnleggende antakelsene i ulike andrespråksteorier er at noen språklige trekk er enklere å tilegne seg enn andre. For eksempel semantikk hvor oversettelsen av et felles konsept som *hest* fra et språk til et annet kan være relativt enkelt å tilegne seg. I sterk motsetning til dette har vi syntaktiske strukturer og operasjoner, som er vanskeligere å bruke på en målspråklig

måte. I generativ andrespråksforskning kan det skilles mellom to store familier av teorier om graden av tilgang til universalgrammatikk i voksen alder: *Full transfer, full aksess* (FTFA) og *Failed functional feature hypothesis* (FFFH), som begge baserer seg på prinsipp- og parameter-teori (chomskyansk generativ grammatikk). FTFA sier at universalgrammatikk er tilgjengelig og man kan lære seg nye prinsipper også i voksen alder, mens FFFH mener det er noen funksjonelle trekk som ikke kan tilegnes. Disse teoriene blir brukt og ytterligere gjort rede for i diskusjonsdelen (kap. 4)

Denne studien baserer seg på og er en oppfølgingsstudie til Vinchenkos masteroppgave fra 2021, som fant ut at nordmenn med norsk som førstespråk bruker ulik mengde tid når de leser setninger med grammatisk verbmorfologi sammenlignet med ugrammatisk verbmorfologi. Dessuten var det forskjell mellom norske verb med tempusavvik sammenlignet med finittavvik. Det betyr at når nordmenn leste norske setninger, var øyeatferden ulik mellom de uliketyperne av morfologiske avvik, og dette tyder på de oppfatter tempus og finittethet som ulike trekk. Vinichenko (2021) presenterer altså evidens for at de norske informantene har ulik atferd ved finitthetsavvik og tempusavvik, da informantene brukte lengre tid på å prosessere verb med tempus- eller finittavvik. Det tyder blant annet på høyere kognitiv belastning når det skjer avvik, spesielt tempusavvik (Vinichenko, 2021, s. 87). Problemstillingen for mitt masterprosjekt er altså om voksne andrespråksinnlærere av norsk har kompetansen til å oppfatte finitt- og tempusavvik, hver for seg, og dessuten som to distinkte trekk.

Problemstillingen er delt opp i to hypoteser med tre underhypoteser:

Alternativ hypotese: Finitt og tempus oppfattes som to adskilte trekk. Avvik i finitt- og tempusmorfologi oppfattes som to distinkte typer avvik, som resulterer i forskjellig lesemønstre.

Nullhypotese: Finitt- og tempustrekkene oppfattes ikke som to adskilte trekk; finitt- og tempusavvikene påvirker derfor øyebevegelse på samme måte.

Underhypotesene handler om spesifikk øyeatferd, det vil si regresjoner og fikseringer. De ulike øyeatferdene blir definert og forklart ytterligere i kapittel 3.1. Foreløpig kan vi si at både økt antallet fikseringer og økt antall regresjoner handler om at informanten vil bruke lengre tid i testsetninger enn i kontrollsetninger.

Underhypoteser:

- 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: lenge 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.
- Både tempus- og finittavvik: lengre og flere fikseringer og flere regresjoner enn både i kontrollsetninger og i setninger med bare en type brudd.

2 Metode

Metoden er helt grunnleggende for enhver vitenskapelig undersøkelse og de metodologiske valgene må være nøye gjennomtenkt, slik at ulike feilkilder påvirker resultatene i minst mulig grad. Forståelsen av metodens innvirkning på studiens kvalitet er derfor helt nødvendig. I dette kapitlet brukes begrepet som indre og ytre validitet og reliabilitet. Validitet handler om det man måler er det man måler, mens reliabilitet handler om gjentatte målinger ville gitt samme resultat, altså om målingene er pålitelige (Ringdal, 2018, s. 103). Indre validitet innebefatter om årsaksslutningen med rimelig sikkerhet at X er årsak til Y, mens den ytre validitet er når resultatene kan generaliseres til andre kontekster (Ringdal, 2018, s. 136). Et siste begrep som blir brukt er økologisk validitet og det betyr at det som blir observert i laboratoriet kan generaliseres til «naturlig» atferd i verden (Schmuckler, 2001). I dette kapitlet blir studiens metode gjort rede for, forklart og begrunnet. Det tar for seg blant annet den vitenskapsfilosofiske bakgrunnen for hypotetisk-deduktiv metode, blikksporing i språkvitenskap, det eksperimentelle designet, gjennomføring og mulige feilkilder.

Metodedesignet til denne studien baserer seg på hypotetisk-deduktiv metode, som sees som *vitenskapens metode*, der hypotesen først blir formulert, deretter utlede observasjon, teste hypotesen og se om hypotesen kan bekreftes eller forkastes (Haraldsen, 2020, s. 9). Med det vil mange hevde falsifisering er viktigere for vitenskapelige undersøkelser enn bekreftelse (2020, s. 13). Hvis vi ikke klarer å falsifisere hypotesen, betyr det at den ikke er funnet usann og da tyder det gjerne på at hypotesen kan ha rot i virkeligheten. Når hypotesen testes, er det nyttig å skille mellom nullhypotesen og den alternative hypotesen. I stedet for å finne en bekreftelse på den alternative hypotesen, kan nullhypotesen falsifiseres, som er det motsatte eller negasjon av den alternative hypotesen. Hvis nullhypotesen blir falsifisert, kan vi utelukke at det ikke er en sammenheng mellom to variabler. Dette reflekterer et vanlig, popperiansk vitenskapssyn (Haraldsen, 2020, s. 21). I undersøkelsen er det helt nødvendig å utelukke flest mulige forstyrrende faktorer, som kan påvirke resultatet. Vi mennesker er ikke pålitelige og er ikke gode til å bekrefte eller falsifisere hypoteser ut ifra egne observasjoner (Haraldsen, 2020, s. 8). Dette baserer seg på Poppers vitenskapsteori, som går ut på at det bare er mulig å falsifisere en hypotese i motsetning til å bekrefte en hypotese. Derfor, og fordi mennesker kan være mer og mindre upålitelige i møte med empiri, er falsifikasjon en mer pålitelig fremgangsmåte i møte mellom hypotese og empiri. Når man har en hypotese (e.i. alternativ hypotese) er det også en nullhypotese, som sier den alternative hypotesen ikke er sann. Hvis nullhypotesen blir falsifisert og den alternative hypotesen ikke falsifisert, er sannsynligheten for at den alternative hypotesen er sann høyere (Haraldsen, 2020, s. 21)

2.1 Blikksporing i språkvitenskap

Blikksporing er en (kvantitativ) metode som nettopp er egnet til et hypotetisk-deduktivt rammeverk. Blikksporing måler øyebevegelser på en skjerm, i dette tilfellet. I språkforskning har blikksporing blitt blant annet brukt til å få innsikt i leseprosesser, og på grunn av sine metodologiske fordeler som høy grad av objektivitet, vært en «gullstandard» i psykologisk og kognitiv leseforskning (Rayner, 2009, s. 1474 i Conklin et al., 2018, s. 6). Metoden muliggjør det å observere den visuelle atferden hos språkbrukere når de leser. Blikksporing har for eksempel blitt brukt til å avdekke at semantisk prosessering ser ut til være en essensiell komponent i leseoppgaver i RAN («Rapid Automated Reading») (Chau et al., 2021), noe som tjerner til å vise nytten av blikksporing i leseforskning. Med det er blikksporing dessuten blitt interessant for

andrespråksforskning, fordi metoden åpner for å se hvordan språkprosessering skjer og språkkompetanse brukes i sanntid (Roberts og Sivanova-Chanturia, 2013). Blikksporing er et nyttig verktøy for å forstå hvordan språkprosessering skjer hos andrespråksbrukere, i dette tilfellet norsk som andrespråk, og med det kunne få innblikk i de psykologiske språkprosesseringene under lesing og få mer forståelse av menneskets tilegnelsesevne av andrespråk. Forskere har ofte vært interessert i S2-innlæreres performanse og deres sensitivitet til tvetydigheter og grammatikaliteter (Roberts og Sivanova-Chanturia, 2013, s. 214). Metoden er derfor særdeles relevant for mitt eksperiment, fordi det forsker nettopp på innlæreres leseperformanse av visse testsetninger, det undersøker forhold i det reseptive systemet og hvordan informantene reagerer på grammatikalitet (og mangel på sådan) i sanntid. I dette prosjektets eksperiment er formålet spesifikt å forstå om andrespråksinnlærere av norsk har kompetansen til å registrere tempus- og finittavvik i verb når de leser fortløpende.

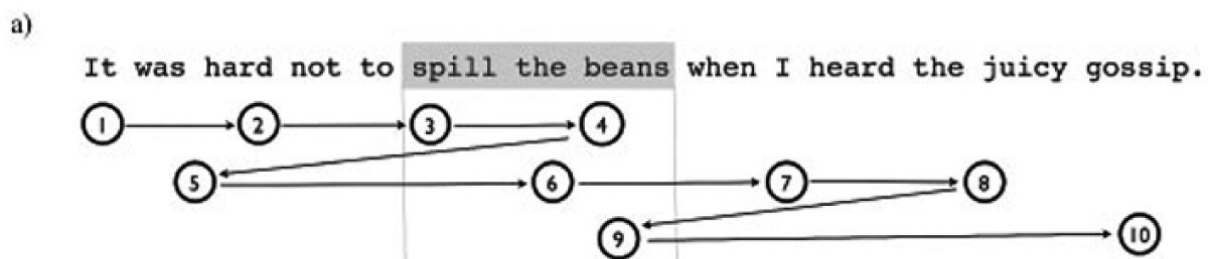
2.2 Kognisjon og blikksporing

Hva er det teoretiske grunnlaget for metoden blikksporing brukt i språkvitenskap? Den grunnleggende antakelsen er *øye-sinn-hypotesen* som antar at varigheten på fikseringene tilsvarer innsatsen som er påkrevd i prosesseringen av setningene og ordene (Just og Carpenter, 1980; Godfroid, 2020, s. 88 i Vinichenko, 2021, s. 33). I motsetning til lytting kan man ved lesing selv begrense hastigheten på inntak av informasjon, hoppe over deler eller fokusere på enkelte ord, og ut ifra atferdsmønstret til leserne kan man få innsikt i forståelsesprosessene. Med blikksporing har man for eksempel funnet ut at den gjennomsnittligere leseren leser 200 ord i minuttet og at nesten alle innholdsord («content words») blir fiksert på, mens fyllord og funksjonsord i mindre grad blir fiksert på (Just & Carpenter, 1980).

Samtidig ser det ut til å være en diskrepans mellom øye-sinn-hypotesen og kognitiv teori om arbeidsminne, formulert av blant andre Baddeley og Hitch (1974). Arbeidsminne kan defineres som komponenter som prosesserer et sett med sammenkoblede og interagerende informasjon, der disse komponentene opprettholder den informasjonen i korttidsminne lenge nok til å aktivt manipulere de lagrede informasjonenhetene (Becker & Morris, 1999). Bakgrunnen for diskrepansen er hvordan arbeidsminnekapasiteten er begrenset og at ikke all informasjon rekkes å bli prosessert. Tunge prosesseringskrav kan minske mengden informasjon som blir lagret og vedlikeholdt, fordi kognisjonen genererer konkurrerende strukturer eller aktivt hindrer vedlikeholdet av inntatt informasjon. At folks arbeidsminne er begrenset, er universelt, men mengden kapasitet og hvordan informasjonen blir prosessert er forbundet med store individuelle variasjoner i leseforståelsesprestasjoner, noe som gjør at språkbrukere med høy arbeidsminnekapasitet kan integrere flere elementer av teksten samtidig (Just og Carpenter, 1980). Til tross for at det er individuelle forskjeller i arbeidsminnekapasitet og man dermed ikke kan garantere en direkte avlesbar, universell øye-sinn-forbindelse slik den er formulert i den basale hypotesen, er hypotesen fortsatt en nødvendig antakelse for at blikksporing skal være en valid metode. Blant annet derfor er det sentralt at overordnede mønstre observeres med hjelp av statistikk for å besvare problemstillingen, slik det da også gjøres i denne undersøkelsen

Gamle påstander som at man kan «se inni sinnet» på informanten, kan skape et feil bilde av hva blikksporing faktisk kan si oss, fordi man kommer med sterke antakelser om at det som blir fiksert, blir prosessert eller vurdert og at tiden som brukes til å fikse på elementer direkte og utvetydig reflekterer tiden det tar å prosessere informasjonen. Fordelen er likevel at blikksporeren observerer øyenes faktiske leseatferd og at

opptakene med blikksporing inneholder mye data. Dette gir oss muligheten til å kvantifisere hva som skjer når en leser møter et ord. Dette gjør man ved å lage ROI'er («region of interest»), som også kalles *interessefelt*. Når interessefeltet er definert, måles øyeatferden i relasjon til dette interessefeltet. Da skiller man mellom tidlige og sene målinger («early measures» og «late measures»), der den førstnevnte kategorien av målinger innebærer blant annet første fikseringslengde og lesetid for første passering, som er målinger under automatisk prosessering og startfasen av prosesseringen. Sene målinger består av gjenlesing, lesetid andre gang, total lesetid og fikseringsantall, som skal reflektere strategisk prosessering, tilbakeblikk («revisits»), omanalysering («reanalysis»), etter antakelsen på grunn av prosesseringsvansker. Mindre frekvente og lengre regresjoner tyder på en delvis mislykket umiddelbar forståelse av teksten (Conklin og Pellicer-Sánchez, 2016). Det betyr at det ikke nødvendigvis er en direkte sammenheng mellom prosessering og tiden øyet hviler på interessefeltet, men at det er en rimelig antakelse at det er en grad av sammenheng og det er mulig å analysere øyeatferden for å finne ut hva som inngår i spørsmål om kognitiv prosessering.



Figur 2⁵

2.3 Design av eksperiment

Som nevnt er det eksperimentelle designet basert på et tidligere masterprosjekt, selv om utvalget informanter er annerledes, et utvalg som representerer en annen populasjon. Det er et metodologisk poeng at dette eksperimentet blir gjennomført så likt som overhodet mulig, slik at resultatene fra eksperimentene kan sammenlignes direkte (Eddington, 2008; Hüttermann et al., 2013). Derfor blir den samme stimulilista som Vinichenko (2021) brukte i sitt eksperiment også brukt i dette eksperimentet. Stimuliene består av forskjellige setninger som informantene skal lese, og etter hver setning blir informantene stilt forståelsesspørsmål («comprehension questions») for å sikre at setningene faktisk blir lest, prosessert og forstått (Conklin et al., 2018, s. 35).

Informantene blir presentert for 83 setninger hver, hvor de tre første er utformet for at informantene skal øve og bli vant til formatet, 40 av setningene er faktiske testsetninger, mens de resterende 40 er kamuflasjesetninger, slik at deltakerne ikke skal oppdage eller vite med sikkerhet hva som faktisk testes. Kamuflasjesetningene er andre setninger med andre grammatiske avvik for å kamuflere målet med eksperimentet (Vinichenko, 2021, s. 48). I eksperimentet til Vinichenko lignet kamuflasjesetningene testsetningene i syntaktisk struktur, men kamuflasjesetningene hadde ikke verbmorfologiske avvik. I stedet var halvparten av kamuflasjesetningene grammatikalske og den andre halvdelene inneholdt avvik i genus i bestemte artikler og dessuten V2-avvik, det vil si avvikende

⁵ Hentet fra Conklin og Pellicer-Sánchez (2016)

plassering av det finitte verbet. (2021, s. 48). Testsetningene er stimuliene informantene blir testet for, og hver deltaker blir presentert for 10 kontrollsetninger, 10 setninger med finittavvik, 10 setninger med tempusavvik og 10 setninger med både finitt- og tempusavvik i hvert eksperiment. Disse utgjør da de ulike betingelsene testpersonene blir eksponert for og vil finne effekten av.

I analysen skilles det mellom fire betingelser:

- (i) Kontrollsetninger (betingelse K) der verbet i interessefeltet brukes på en grammatisk måte. Se eksempel 1-A
- (ii) Setninger med finittavvik (betingelse F), der infinitivformen brukes i stedet for presensformen, som bryter med finitthetsdistinksjonen. Se eksempel 1-B
- (iii) Setninger med tempusavvik (betingelse T), der infinitivformen brukes i stedet for partisippformen, som bryter med tempusdistinksjonen. Se eksempel 1-C
- (iv) Setninger med både finitt- og tempusavvik (betingelse F+T), der infinitivformen brukes i stedet for preteritumsformen, som er brudd på både finitt- og tempusdistinksjonen. Se eksempel 1-D⁶

Eksempler på testsetninger. A er kontrollsetning, B har finittavvik, C har tempusavvik og D har både finitt- og tempusavvik. Interessefeltene er markert i klammer

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.

Interessefeltene i eksperimentet ble definert på forhånd. I testsetningene er det to interessefelt: verbale interessefelt og post-verbale interessefelt, der verbet er i det verbale interessefeltet og nomenet etter verbet er i det post-verbale interessefeltet.

I disse eksemplene på testsetninger med kontrollsetning og de tre ulike avvikene (1A-D), er verbet *passe* det verbale interessefeltet og nomenet *barna* det post-verbale interessefeltet.

Grunnen til at det post-verbale interessefeltet også blir observert, er at det kan skje en «spillover»-effekt, hvor det kan oppstå en forsinket effekt av morfosyntaktiske avvik rett etter det primære interessefeltet. Derfor er det et interessefelt til bak verbet for å fange opp effektene, som ikke blir fanget opp i det første, verbale interessefeltet (Conklin og Pellicer-Sánchez, 2016 i Vinichenko, 2021, s. 54). Det er heller ikke nødvendig med flere interessefelt enn to, fordi effekten av morfosyntaktiske avvik er umiddelbare og kortvarige. Det vil si at regresjoner og lengre lesetid i skjer i det verbale interessefeltet (Braze et al., 2002 i Vinichenko, 2021, s. 54) utenom mulige «spillover»-effekter, som blir fanget opp i det post-verbale feltet.

2.3.1 Kritisk innvending til stimulilisten

I dette prosjektets eksperiment brukes den samme stimulilisten som i Vinichenkos(2021) studie . I eksperimentet brukes 40 testsetninger med 4 ulike betingelser á 10 per betingelse (A-D) (se vedlegg 7.3). Det tas her forbehold om at lengre fikseringer kan forårsakes av rare og unaturlige formuleringer eller merkelig innhold, som igjen bidrar til

⁶ For å se resten av testsetningene, se vedlegg i 7.3.

mer støy i dataene. Det gjør empirien mindre reliabel og vanskeligere å finne effekt av stimuliene (Vinichenko, 2021, s. 46). Derfor er også en rimelig innvending at setningene for mitt prosjekt i større grad kunne blitt tilpasset andrespråksinnlærere, slik at ikke de møter setninger de synes er merkelige eller unaturlig på grunn av komplisert norsk. I tillegg vil bruk av pseudokoordineringsverb, for å motvirke støyen som kommer av forflytting av det post-verbale interessefeltet, potensielt være et problem for andrespråksinnlærere. Grunnen til det er at pseudokoordineringsverb kan fremstå som overflødige og for uformelle i skriftlig språk. Det samme gjelder topikalisering eller tematisering, som også kan virke overflødig, rare eller unaturlige for informanten. Det kan bidra til uønsket tilfeldig variasjon. Stimuliene bør være enkle og logiske, slik at bare den eksperimentelle variasjonen påvirker språkprosesseringen (Vinichenko, 2021, s. 46). Det er målet ideelt sett, men det er en hårfin balanse mellom enkle, logiske setninger og ikke for åpenbare, samtidig som man også skal lage setninger som inneholder akkurat de betingelsene som skal testes. Til tross for at testsetningene kan forårsake noe støy, noe som potensielt kan bidra til å redusere reliabiliteten, viser resultatene fra Vinichenko (2021) at det er fullt mulig å få signifikante, betydningsfulle resultater fra eksperimentet, noe som tilsier at eksperimentet likevel kan fortelle oss noe vesentlig om prosessering av verbmorfologiske avvik.

2.3.2 Mer om begrunnelsen for bruk av samme stimuliliste

I det forrige avsnittet gikk vi igjennom noen kritiske innvendinger mot stimulilisten med tanke på reliabilitet. Hvorfor bruke den samme stimulilisten brukes når den potensielt kunne ha vært vesentlig forbedret? Dette prosjektet, i motsetning til Vinichenkos studie, undersøker andrespråksinnlærere av norsk, men hypotesene og eksperimentet er likt. Altså representerer utvalget i denne studien en annen populasjon enn Vinichenkos utvalg, men ellers er designet til eksperimentene helt like. Hvis dataene fra Vinichenkos eksperiment sammenlignes med dataene fra denne undersøkelsen med nøyaktig likt design, med unntak av populasjonen utvalget representerer, blir dette utvalget den eksperimentelle manipulasjonen for denne studien. Da kan effekten av to utvalg, som representerer ulike populasjoner, sammenlignes, men da bør de utsettes for akkurat de samme betingelsene. Hvis stimulilisten hadde blitt endret, ville det minsket det direkte sammenligningsgrunnlaget. Det er nærliggende å tenke at utvalget fra det tidligere eksperimentet er kontrollgruppen til dette eksperimentets utvalg, fordi kontrollgruppen da representerer nordmenn med norsk som førstespråk. Siden Vinichenko fant signifikante resultater med disse testsetningene på nordmenn, er det rimelig å bruke dem igjen, spesielt når formålet er at Vinichenkos data skal kunne sammenlignes med dataene fra denne studien.

2.4 Hva måles under blikksporing?

For å svare på problemstillingen, må vi finne ut som andrespråksinnlærere oppfatter en distinksjon mellom tempus og finittet i norske verb. Med -blikksporing kan vi kartlegge dette ved å se om informantene statistisk sett bruker lengre tid på å prosessere verb med ulike morfologiske avvik, gitt at mer forsinkende øyeatferd på et interessefelt rimeligvis kan antas å bety at det brukes lengre tid på å prosessere det grammatiske trekket på grunn av avviket. Når informantene eksponeres for de eksperimentelle betingelsene og blikket til informantene spores, kan det måles om det er en statistisk signifikant distinksjon i øyeatferd mellom de ulike betingelsene, samt sammenligne det med grammatikalske kontrollsetninger. Blant annet derfor skiller det mellom underhypoteser og nullhypotese og alternativ hypotese, der underhypotesene handler om spesifikk øyeatferd og de to sistnevnte hypotesene kan gi svar på problemstillingen.

Når det kommer til blikksporing av lesing, handler det om å observere i hovedsak tre fenomener: *fikseringer* («fixations»), *sakkader* («saccades») og *regresjoner* («regressions») (Conklin og Pellicer-Sánchez, 2016). Det man vil undersøke i setningen som verb i dette eksperimentet, defineres som *interessefelt* i setningen. Data om hvordan øyeatferden i interessefeltene blir samlet inn og skiller fra andre blikksporingsdata, som også blir samlet inn. Det som måles er hvordan informantene fikserer og flytter blikket på de definerte interessefeltene, hvor blikket lander og hvor lenge det dveler ved hvert interessefelt (Vinichenko, s.27 og 33-34). Siden sakkader er så raske øyebevegelser, antas det at informasjon ikke tilegnes under sakkader (Liversedge, et al., 1998; Rayner, 1998; 2009 i Roberts og Sivanova-Chanturia, 2013, s. 214). Derfor måles ikke sakkader i seg selv, men i hovedsak fikseringer og regresjoner. Regresjoner på visse måter like sakkader i og med at det innebærer raske øyebevegelser. Likevel er regresjoner interessante, de innebærer forflytting av blikket tilbake til interessefeltet. Dette regnes vanligvis som tegn på en forsinkelse i prosesseringen eller vansker med å tilegne seg informasjonen. I denne studien blir (1) *første fikseringslengde* («first fixation duration»), (2) *blikklengde* («gaze duration»), (3) *regresjonsstielengde* («Regression path duration»), (4) *dveletid* («Dwell time»), (5) *fikseringsantall* («fixation count») og (6) *regresjon ut* («regression out») målt i både verbalt og postverbalt interessefelt. I motsetning til de andre målingene, er «regresjon ut» en binær variabel (Field, 2013, s. 8), som måler sannsynligheten for at en sakkade forlater det aktuelle interessefeltet tilbake til et lavere et før det senere interessefeltet ble fiksert eller ikke. I den målingen måler man sannsynligheten for at blikket vendes tilbake fra et gitt interessefelt under førstegangslæsning. Målet er derfor i prosentvis sannsynlighet for at denne øyeatferden skjer. Det er de spesifikke målene som blir gjort av blikksporeren, og de utgjør avhengige variabler, mens de ulike betingelsene (setning K, T, F og F+T) er uavhengige variabler (Vinichenko, 2021, s. 56-57). Det vil si at det antas at de uavhengige variablene (verbmorfologiske avvikene) vil føre til en effekt på de avhengige variablene (målingene av øyeatferd) (Field, 2013, s. 7).

2.5 Utvalg av informanter

Utvalget består av 9 voksne andrespråksinnlærere av norsk, og vi søkte primært etter informanter som behersker engelsk, tysk, kinesisk, vietnamesisk eller thailandsk, fordi disse språkene varierer sterkt med hensyn til verbmorfologi. Noen av språkene har for eksempel markert finitthetsmorfologi og noen ikke, andre har verken tempus eller finitthet som bøyingskategori. Informantene skulle ideelt sett ha et av disse språkene som førstespråk og være mellom 18 og 40 år. Optimalt sett burde antallet informanter være høyest mulig, fordi standardfeil som kan forekomme med denne metoden primært motvirkes ved å øke antall deltakere i eksperimentene. Det betyr at et høyere antall i seg selv regnes for å føre til mer reliable data (Conklin et al., 2018, s. 72). Dette har vært in mente i vurderingen når deltakere med andre førstespråk enn nevnt ovenfor ble inkludert. Den nedre aldersgrensen på 18 år er satt ut fra at eksperimentet undersøker voksne innlærere av norsk. Som nevnt innledningsvis tyder mye tidligere forskning på at alder har mye å si for hvordan tilegning av et andrespråk skjer. Puberteten ser ut til å være et relativt etablert brytningspunkt (Johnson & Newport, 1989), selv om det er ulike antakelser om antall og varighet av kritiske og sensitive perioder for språktilegning. Det er derfor nødvendig for undersøkelsens indre validitet at informantene er over 18 år gamle og at de har lært norsk etter fylte 18 år, fordi premisset er at de er ferdige med disse kritiske og sensitive periodene for språktilegning når de lærer norsk. Alderen til informantene er derfor relevant for undersøkelsens validitet, fordi alder har en effekt på språktilegning og det ville kunne påvirke resultatene om de lærte norsk før fylte 18 år. I denne typen eksperiment er den øvre aldersgrensen for informantene i utvalget vanligvis satt til 40 år, på grunn av potensielle fysiske og kognitive endringer. Generell aldring kan allerede i dette alderssjiktet påvirke blikkets bevegelser, særlig raskhet og fluktuering («smooth-pursuit»). Hvis man derimot ikke skal benytte seg av målinger fra akkurat disse fenomenene, blir denne øvre aldersvariabelen mindre viktig (Salvi et al, 2006; Crawford et al., 2017 i Vinichenko, 2021, s. 48).

Generelt kan vi altså si at aldersutvalget er et spørsmål om reliabilitet, fordi variabler som fysiske og kognitive endringer kan påvirke dataene fra eldre informanter og gi uoversiktlige utslag, sammenlignet med yngre informanter. Det ser likevel ikke ut til å være avgjørende for dataene som brukes for å besvare problemstillingen i denne studien. Derfor valgte vi å inkludere også en deltaker som er over 40 år, først og fremst for å øke antallet testpersoner, men også fordi en kan anta at det blir vanskeligere å lære seg nye språk desto eldre man er, og at dette øker sannsynligheten for utslag i eksperimentet. I tillegg, hvis de scorer høyt på andre viktige variabler, som hvilket førstespråk de har, har vi latt dette trumfe risikoen for at dataene blir påvirket av forstyrrende, mer irrelevante variabler.

Et annet kriterium for utvalget er informantenes kompetansenivå i norsk, der målet er at de skulle være på C-nivå. Et så høyt kompetansenivå øker sannsynligheten for at mellomspråket deres har nådd potensialet for å bli så målspråkligt som mulig. Vi endte likevel opp med å inkludere informanter også på B1- og B2-nivå i tillegg til de på C-nivå, fordi det ikke var gjennomførbart å rekruttere nok informanter til prosjektet med så strenge utvalgskriterier, mange ellers godt kvalifiserte kandidater ville ha blitt ekskludert på grunn av et litt for lavt kompetansenivå i norsk. Derfor ble det valgt å inkludere deltakere også på B1- og B2-nivå for å få nok deltakere, noe som i sum også styrker reliabiliteten og den ytre validiteten til resultatene. Som nevnt over er det essensielt å motvirke standardfeil for at resultatene skal være mest mulig signifikante. I tillegg ble også lesevansker og diagnostisert ADHD et eksklusjonskriterium, fordi det ville påvirket

resultatene i for stor grad og redusere resultatenes reliabilitet. Hvis informant(e) har dysleksi vil det påvirke resultatene (Rayner, 1998, s. 22). Dette er en vanlig eksklusjonsårsak i blikksporingsstudier; når man prøver å finne ut hva som er en normaltendens som kan generaliseres til en typisk populasjon, kan ikke utvalget inkludere deltakere med dysleksi eller ADHD.

Et viktig kriterium for at en gitt informants data skulle kunne brukes i analysen var at informantene forsto det som sto i setningene, fordi vi ville at informantene skulle være kompetente nok til å forstå setningene. I eksperimentene ble de underveis spurt ja-/nei-spørsmål om setningenes innhold. Antallet riktige svar på ja-nej spørsmålene ble brukt for å vurdere om deltakerne skulle inkluderes, fordi dette var en indikasjon på om deltakeren forsto setningene. Nøyaktighetsterskelen for forkasting av deltakere ble satt på 80% i prosjektet til Vinichenko, hvor alle deltakerne i eksperimentet hennes fikk 85% riktige svar, og derfor ble ingen informanter forkastet (Vinichenko, 2021, s. 55). I dette prosjektet ble det lagt en nøyaktighetsterskel på 70%. Noe annet enn at andrespråksgruppen har færre korrekte svar enn nordmenn var ikke å forvente. Av de ni informantene som gjensto etter utsiling var det ikke alle som klarte 80%, men alle klarte 70%. Siden utvalget var lite i utgangspunktet, valgte jeg derfor å inkludere alle informantene. At alle svarte over 70% korrekt, tyder i det minste på at de prosesserer og forstår setningene godt nok til at det ikke er tilfeldig om de svarer korrekt eller ikke. Igjen hadde det vært ideelt at nøyaktighetsterskelen var høyere, men dataene ville etter mitt syn blitt mindre reliable hvis det var færre deltakere.

Til slutt kan det hende at vi likevel må ekskludere data fra enkeltinformanter, for eksempel fordi de er utsatt for sporingstap («track loss»). Sporingstap skjer enten på grunn av blinking, sminke, enkelte kontaktlinsetyper, kalibreringsvansker eller biologiske særegenheter hos informanten (Vinichenko, 2021, s. 54). Dette kan først kontrolleres når eksperimentene er over, og først da kan det vurderes hvorvidt dataene fra enkelte informanter skal forbli en del av utvalget som kan analyseres. Det skjer ikke så sjelden at data må tas ut, enten på grunn av at enkelte informanter viser seg å likevel ikke passer til kriteriene som er satt for utvalget, på grunn av sporingstap eller av andre, uforutsette årsaker. I vårt eksperiment endte vi med 9 testpersoner.

Tabellen under viser en oversikt over deltakernes språkbakgrunn med hvilket språk de kan, når de begynte å lære språkene (AoO), hvor lenge de har lært norsk, kompetansenivå i norsk ut ifra A1-C2-klassifikasjonen (CEFR) og ortografisk bakgrunn (hvilket alfabet de har brukt i sin primære skolegang).

Bakgrunnsinformasjon om hver deltaker

ID	Alder	Språkbakgrunn	Læretid	Kompetansenivå	Ortografisk språkbakgrunn
1	29	Engelsk, 0 år Kinesisk, 14 år Finsk, 18 år Portugisisk, 20 år Norsk, 28 år	1 år	B2	Det latinske alfabet Kinesisk skrifttegn
2	28	Russisk, 0 år Engelsk, 12 år Spansk, 18 år Norsk, 27 år	1 år	B1	Det kyriliske alfabet Det latinske alfabet
3	25	Urdu 0 år Engelsk 4 år Norsk 25 år	>1 år	B2	Urdualfabetet Det latinske alfabet
4	32	Gresk, 0 år Albansk, 0 år Engelsk, 4 år Norsk, 31 år	1 år	B1	Det latinske alfabet
5	26	Hindi, 0 år Engelsk, 5 år Norsk, 25 år	1 år	A2	Devanagari skrift Det latinske alfabetet
6	26	Tysk 0 år Italiensk 6 år Engelsk 11 år Fransk 14 år Spansk 23 år Norsk 25 år	1 år	B1/B2	Det latinske alfabet
7	25	Afrikaans, 0 år Engelsk, 1 år Norsk, 23 år	2 år	B2	Det latinske alfabet
8	25	Russisk, 0 år Engelsk, 6 år Norsk, 23 år	2 år	B1	Det kyriliske alfabet Det latinske alfabet
9	45	Kinesisk, 0 år Engelsk, 13 år Norsk, 38 år	7 år	B2	Kinesisk skrifttegn Det latinske alfabet

Tabell 3

2.6 Gjennomføring

Eksperimentet må gjøres klart ved at datamaskinene skrus på med blikksporingskameraet og programvaren gjøres klart. Før undersøkelsen starter, får deltakeren både informasjonsskrivet og samtykkeskjemaet for både dette prosjektet og en annen masterstudents prosjekt, fordi begge prosjektene skal bruke de samme dataene, til to ulike formål. Disse dokumentene er sendt på e-post i god tid slik at deltakeren får tid til å lese dem på forhånd. På samtykkeskjemaet står det en «subjektkode», som knytter enkeltpersonene til de spesifikke dataene, slik at dataene knyttet til hver deltaker kan etterspores i tilfelle noen trekker seg og vil ha dataene sine slettet. Når det formelle er unnagjort, blir det forklart hvordan eksperimentet blir gjennomført. De får instruksjoner om hvordan de kan stille inn stolen og stativet etter egne preferanser, slik at de sitter mest mulig komfortabelt, noe som er viktig for deltakernes velvære og for at hodet skal være mest mulig stabilt når de leser. Så blir prosedyren for undersøkelsen gitt skriftlig på skjermen og forklart muntlig. Deltakerne blir informert om at de bør unngå blinking, og at det er lagt opp til pauser underveis. Før eksperimentet starter, blir kameraet kalibrert for å tilpasse kameraet til hver deltaker, og så valideres denne kalibreringen ved å undersøke nøyaktigheten til blikksporingen.

Et annet viktig aspekt ved gjennomføringene av eksperimentene er at forskeren legger til rette for at informantene får de best mulige forutsetningene for å konsentrere seg, slik at undersøkelsen blir mest mulig reliabel. Da kan vi minimere risikoen for at resultatene blir påvirket av forstyrrende variabler («confounding variables»), som er eksterne faktorer for undersøkelsen resultatene kan påvirkes av (Field, 2013, s. 15). Målet er å isolere effekten av de uavhengige variablene på de avhengige variablene. Et annet ansvar forskeren har, er å ikke påvirke resultatene gjennom eksempelvis å si på forhånd hva som testes i så detaljert grad at det påvirker leseatferden til informantene. Derfor ble det ikke nevnt at undersøkelsen fokuserer på verb, fordi da ville kanskje informantene rette blikket i større grad mot verb for å finne ut hva som undersøkes. Et annet tiltak for å motvirke feilkilden om overdrevent metalingvistisk fokus er å kamuflere testsetningene ved å legge inn setninger med andre typer avvik, som ikke er avvikene vi undersøker. 40 av de 83 setningene som blir presentert for deltakerne, er såkalte kamuflasjesetninger.

2.7 Utstyr

I Vinichenko (2021) brukes *EyeLink 1000* for å spore blikket til leseren, fordi den har høyere samplingfrekvens enn andre blikksporingskamera, og høy samplingfrekvens gjør dataene mer reliable når blikkmønster i små interesseområder undersøkes (Vinichenko, 2021, s. 49). Ulempen er at *EyeLink 1000* ikke er mobil, i sammenligning med andre kamera med lavere samplingfrekvens. Siden forskningsprosjekts design er ment å være så lik Vinichenkos design utenom populasjonen utvalget studiene underøsker, brukes også det samme blikksporingskameraet med de samme innstillingene for å gi et godt sammenligningsgrunnlag mellom førstespråksgruppen og andrespråksgruppen. *EyeLink 1000s* samplingfrekvens er 1000Hz og romopløsning på 30' bueminutt, som gjør at den fanger opp korte og raske øyebevegelse med høy presisjon. Eksperimentet ble programmert i «Experiment Builder»-programvaren (SR Research Ltd., Toronto, Canada) av Vinichenko (2021). Leseren bruker også hakestøtte for å redusere hodebevegelser (s. 49). Det er nyttig når man samler inn data fra små områder og som dermed krever høy presisjon. Dataenes kvalitet determineres av nøyaktighet («accuracy») og presisjon, der nøyaktighet trengs i studier av små interessefelt (Conklin et al., 2018, s. 23). Hvis interessefeltet er lite, bør samplingsfrekvensen være over 250Hz (Conklin et al. 2018, s. 31). Samplingfrekvensen til kameraet er da høy for å samle mest mulig data, og

hakestøtte blir brukt for at dataene ikke blir påvirket av hodebevegelser, noe som ville ha gjort dataene mindre nøyaktige og reliable. Blikksporingseksperimentet bør gjennomføres et sted som er stille, ideelt sett lydtett, for å minimere distraherende lyder utenfra. I tillegg bør rommet være godt belyst, men ikke på en slik måte at sollys eller kunstig lys fører til refleksjon på dataskjermen eller blikksporeren (Conklin et al., 2018, s. 74). Eksperimentene i vår studie har ikke blitt gjennomført i et lydtett rom, men det har vært godt belyst uten direkte sollys eller annet lys som førte til refleksjon på skjerm eller blikksporeren. Og det har heller ikke vært distraherende lyder utenfra som forstyrret deltakerne.

2.8 Mulige feilkilder og motvirkende tiltak

Det er ulike muligheter til feilkilder som kan oppstå i løpet av eksperimentene, som at forskeren gjør feil. Det kan være at kameraet blir feilkalibrert eller er så dominerende i rommet at det fører informantenes oppmerksomhet bort fra eksperimentet, eller mangel på justering av sporingstap. Disse potensielle feilkildene vil kunne påvirke resultatene og gjøre dem mindre reliable. Derfor bør disse mulige feilkildene motvirkes. Vi søkte i størst mulig grad å sørge for at kalibrering og andre tilretteleggende forhold ble så grundig utført som mulig, men det var likevel ett eksperiment hvor kameraet var litt unøyaktig. I tillegg ble skjermen brukt som en slags skillevegg mellom informant og forsker for å gjøre forskeren mindre synlig. Det viktigste tiltaket ble nevnt i 4.4, og bør nevnes her igjen: Det er at man kan motvirke generell statistisk støy, også kalt standardfeil, ved å ha et større utvalg, fordi standardfeil blir motvirket av et større utvalg (Conklin et al., 2018, s. 72). Det må dermed vurderes hvor mye tid og ressurser man har til å gjennomføre mange eksperimenter, men desto flere deltakere jo mer reliable blir resultatene. På den andre siden er det ikke alltid mulig å gjennomføre så mange eksperimenter som er ideelt eller ønskelig, så derfor er det også viktig å gjennomføre eksperimentene på en kvalitetssikret måte og god måte, slik at (1) målingene blir gjort riktig og (2) reduserer mengden støy som minsker prosjektets reliabilitet. Hvis gjennomføringen ikke er god nok og det er for mange forstyrrende variabler, har det ikke noe å si hvor stort utvalget er. Da vil man ende opp med upålitelige data uansett, bare i større mengde. Vi sørget derfor for å minimere disse feilkildene i størst mulig grad, for at resultatene skal både være valide og reliable nok til at det er et sammenligningsgrunnlag mellom resultatene fra dette utvalget og Vinichenko (2021) sitt utvalg.

2.9 Diskusjon om prosjektets validitet og reliabilitet

I dette prosjektet elisiterer vi data fra innlærere av norsk som andrespråk for å kunne bedømme deres evne til å legge merke til morfologiske tempus- og finittavvik i norske verb når de leser norske setninger. Problemstillingen er dermed om de har kompetansen til å oppfatte finitt- og tempusavvik, både kombinert og som distinkte trekk i voksen alder. Metodens validitet bygger på øye-sinn-hypotesen ved å anta at det er en sammenheng mellom eksempelvis fikseringenes varighet og innsatsen prosesseringene krever. Med det måles informantenes øyeatferd, og målene er blant annet hvor lang tid de fikserer, antall fikseringer, regresjoner i både verbale og post-verbale interessefelt. Gitt øye-sinn-hypotesen er antakelsen at de vil bruke lengre tid på å prosessere verbet hvis de oppdager avvik i språket, og det vil dermed gi utslag på disse målingene i form av lengre tid eller flere fikseringer. Siden det er vist tidligere at nordmenn bruker lengre tid på ugrammatikalske verbformer enn grammatikalske verbformer, tyder det på at grammatikalske setninger leses og prosesseres raskere. Det styrker evidensen for at metoden er valid. På den andre siden kan det for eksempel brukes lengre tid i

interessefeltet av ulike årsaker som at setningene blir unaturlige og rare som nevnt i 4.7.1, fordi setningene kan bli rare når man lager setninger som tvinger frem spesifikke skrivefeil. De skal heller ikke være så rare at de tar opp oppmerksomheten til leseren. Hvis deltakerne skal bli oppmerksomme på noe, bør det i så fall være avvikene man undersøker. Men gjennomføres eksperimentet på nok deltakere, vil det kunne få statistisk utslag og gi innsikt om informantene som gruppe legger merke til tempus- og finittavvik i norske verb på en måte som ligner eller avviker fra atferden til de med norsk som førstespråk.

En annen fordel er at blikksporing er relativt økologisk valid med tanke på hvor mye, detaljert og nyansert data kameraet samler inn, men *self-paced reading* (SPR) er også et metodealternativ for leseforskning.⁷ Det ville muligens vært enklere å rekruttere flere informanter til et slikt prosjekt, siden det kan gjøres hjemme digitalt. Et større utvalg ville kanskje gjort undersøkelsen mer reliabel, men da ville man ikke kunne måle den faktiske øyeatferden når deltakerne leser. SPR har mindre «økologisk validitet», fordi det er unormalt å lese som man gjør i slike undersøkelser hvor man leser et ord om gangen. I tillegg er dataene mindre nyanserte og reliable enn blikksporing. Blikksporing er også en mer valid metode i henhold til problemstillingen når kanskje det viktigste metodiske poenget er at metoden skulle skille seg minst mulig fra Vinichenko (2021) sin metode utenom populasjonen utvalget representerer, slik at gruppene kan sammenlignes ut ifra dataene. Bruken av SPR ville dermed gjøre sammenligningsgrunnlaget mindre reliabelt, fordi SPR måler leseatferd så ulikt og sammenligningene ville derfor vært mindre valide og reliable. Da ville vi fått mindre data fra hver deltaker og undersøkelsen ville ikke kunne si noe om øyeatferd når de leser. Etter mitt syn er den store fordel til blikksporing både mer indre valid og økologisk valid sammenlignet med SPR utenom at informantene må lese setningene med haken på et stativ i en laboratoriesetting, fordi det lar dem lese setningene naturlig og det er mange flere mål av faktisk øyeatferd. Ulempen er at det muligens er vanskeligere å rekruttere flere informanter til et blikksporingseksperiment, fordi oppsettet er ikke mobilt og derfor må informanter komme til laboratoriet, som øker terskelen for å delta. Det igjen går utover utvalgets størrelse, som igjen går utover reliabiliteten.

Avslutningsvis for dette kapitlet er det relevant å diskutere om hvorvidt problemstillingen kan besvares ut ifra målene som er valgt ut for oppgaven. De utvalgte målene er knyttet opp mot underhypotesene som spesifikt kommer med prediksjoner på hvordan øyeatferd vil kunne bli påvirket av både tempus- og finittavvik isolert og sammen. Hvis resultatene er signifikante og viser ulike resultater mellom testbetingelsene, vil det da avkrefte nullhypotesen, som gjør den alternative hypotesen mulig. Hvis nullhypotesen ikke avkreftes, vil det si at det ikke er nok evidens for det og begge hypoteser er mulige. Dette skal være valid og reliabelt nok til at problemstillingen kan besvares ut ifra de setningene som presenteres i eksperimentene, men det kan hende det ikke gjelder andre former av setninger med finitt- og tempusavvik, i tillegg til at det kan være individuelle forskjeller. Det er verdt å merke seg at den hypotetisk-deduktive metoden er streng og derfor nyttig når øyeatferd tallfestes og analyseres statistisk, og derfor er det nødvendig å ta høyde for at terskelen for å dra konklusjoner er høy. I tillegg må det legges til at ikke-signifikante resultater fortsatt er betydningsfulle resultater, fordi det likevel sier oss mye om leseatferden til andrespråksinnlærere av norsk. Jeg kommer derfor til å

⁷ Til sammenligning med metoden *self-paced reading* (SPR), som var et reelt alternativ for blikksporing i tilfelle det ikke ble mulig å gjennomføre.

diskutere begge typer av resultat nedenfor, både statistisk signifikante resultater og de som bare viser trender og tendenser som ikke er statistisk signifikante.

3 Resultater og statistisk analyse

I dette kapitlet gjennom går jeg hva som måles, hvordan dataene ble fremstilt, prosjektets sentrale data og hva funnene gir av implikasjoner for dette prosjektets hypoteser. I underkapitlet om variablene blir disse definert og kort forklart; deretter beskriver jeg prosedyren for datarensing og analyse i prosjektet. Videre presenterer jeg den deskriptive statistikken, både fra dette prosjektet og Vinichenkos; og til slutt gjør jeg rede for den statistiske analysen, som innebefatter rm-ANOVA og T-test. Her analyserer jeg også resultatene med Bonferroni-korreksjon. Tabellen under gir en oversikt over definisjoner på statistiske begreper brukt i analysen.

Begrepsoversikt for statistikk

M(ean)	Gjennomsnitt.
SD	Standardavvik sier noe om hvor mye skårene fra observasjonene varierer fra gjennomsnittet.
P-verdi	Sannsynlighet for at observasjonen er sann gitt at nullhypotesen er sann.
T-verdi	Verdien oppgitt i T-testen og uttrykker forholdet mellom systematisk og tilfeldig variasjon.
Cohens d (d_{rm})	Effektstørrelse for repeterte målinger og brukes for T-verdi i denne analysen.
F-ratio	Uttrykker forholdet mellom systematisk og tilfeldig variasjon.
Partiell eta kvadrert	Effektstørrelse brukt for rm-ANOVA i denne analysen.

Tabell 4 (Vinichenko, 2021, s. 5-6)

3.1 Variabler

I dette designet er det seks avhengige variabler. Alle er målinger av ulike øyeatferder; det de avhenger av, er de uavhengige variablene. De uavhengige variablene på sin side er betingelsene i stimulisetningene, altså setninger med finittsavvik, tempusavvik eller begge avvikene samtidig. I tillegg har vi kontrollsetningene som ikke har avvik det hele tatt, som kontrollerer hvordan øyeatferden er ved setninger uten morfologiske avvik. De avhengige variablene utgjør øyeatferd i møte med de uavhengige variablene, som er ulike norske setninger med tempus og finittavvik sammen eller hver for seg. Da er det nødvendig at de seks avhengige variablene blir definert, for å forstå hvordan dataene kan oversettes til faktisk beskrevet øyeatferd. I Vinichenko ble de definert⁸ slik:

- Første fikseringslengde: «varigheten på den første fikseringen i interessefeltet». Det er det tidligste tidspunktet man kan se effekter av ordets egenskaper. Dette måles i millisekunder (Liversedge et al., 1998).

⁸ Disse definisjonene er hentet fra Vinichenko (2021, s. 56-57) med kildene hun oppgir. Noe er skrevet om, men definisjonene er i store trekk tatt ifra hennes masteroppgave.

- Blikklengde er den oppsummerte varigheten til alle fikseringer gjort i et interessefelt under førstegangslesing før blikket forlot interessefeltet til høyre eller venstre. Dette måles i millisekunder (Rayner et al., 2004).
- Dveletid er den oppsummerte varigheten på alle fikseringer i interessefeltet. Dette måles i millisekunder.
- Regresjonsstilengde 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. Dette måles i millisekunder (Rayner et al., 2004).
- Fikseringsantall er antall fikseringer i interessefeltet. Dette måles i antall (Conklin et al., 2018).
- Regresjon ut er sannsynligheten for at blikket vender tilbake i teksten fra et gitt interessefelt under førstegangslesing. Dette måles i sannsynlighet i prosent for at blikket vender tilbake (Rayner og Pollatsek, 1989) (2021, s. 56-57)

Som nevnt i metodekapitlet, er grunnen til at akkurat disse målingene er valgt ut som grunnlag for den foreliggende undersøkelsen, at disse målingene ble brukt av Vinichenko, og de er også de samme som er brukt tidligere, i eksempelvis Ni et al., 1998 og Braze et al., 2002 (2021, s.57). Et annet aspekt ved målingene er å påpeke forskjellen på tidlige og sene målinger, hvor tidlige målinger er første fikseringslengde og blikklengde, mens sene målinger er de fire resterende målingene. De tidlige målingene skal måle tidlig kognitiv prosessering, og de sene målingene skal måle sen kognitiv prosessering, men noen av målingene, som dveletid, er en sum av både tidlige og sene fikseringer (Vinichenko, 2021, s. 56). Hvis målingene viser ulikt resultat mellom de ulike betingelsene, tyder det på ulik kognitiv prosessering. Hvis det brukes lengre tid på de morfologiske avvikene, kan det bety at informanten bruker lengre tid på å prosessere verbet fordi de forventer morfologi som er riktig, gitt konteksten, i tempus og finitthet. Til slutt bør det påpekes at disse variablene måles i både verbale og postverbale felt og resultatene for de forskjellige interessefeltene presenteres hver for seg. Det blir gjort for å skille ut *spillovereffekten* fra det postverbale interessefeltet. Derfor er resultatene presentert i fire tabeller hvor hver tabell inneholder deskriptiv statistikk av de seks målingene fra hvert interessefelt. To av tabellene viser resultatene fra andrespråksgruppen i denne masteroppgaven og de tabeller viser resultatene fra førstespråksgruppen i Vinichenkos masteroppgave.

3.2 Prosedyre for datarensing og analyse

Dataene fra eksperimentene ble lagret i Dataviewer, som er et program for å lagre, fremstille og eksportere blikksporingsdataene (SR Research Ltd., Toronto, Canada). For hver observasjon av hver setning ble det sett over manuelt om det var springstap i dataene i interessefeltene, for å vurdere om hvilke observasjoner som måtte slettes. Deretter ble de dataene hvor fikseringene og sakkadene ble vist over eller under interessefeltene ifølge blikksporingsdataene, justert i den vertikale akse der dette var rimelig. Hvis majoriteten av fikseringene og sakkadene var over eller under hele setningen, ble alle dataene av fikseringene markert og flyttet vertikalt til riktig plassering i interessefeltene for at dataene skulle registreres i interessefeltrapporten. Interessefeltrapporten eksporterer all dataene fra de ulike setningene fra de ulike informantene fra Dataviewer til Excel.

Etter at dataene ble eksportert til Excel, måtte dataene renses, som ville si at dataene ble sortert og fremstilt slik at de ble et datasett som kunne analyseres i SPSS. Som nevnt i kapittel 2.5 ble nøyaktighetsterskelen satt til 70%, og derfor trengte ikke noen av deltakernes data å bli ekskludert. Resten av prosessen innebefattet å slette data fra kamuflesjetninger og øvingssetninger, data fra de ikke-verbale interessefeltene, og å skille dataene mellom de ulike betingelsen fra hverandre. Dermed ble det laget et datasett i SPSS for det verbale interessefeltet og et for det postverbale interessefeltet, slik at disse to kunne analyseres separat. I disse datasettene hadde vi så kolonner med de relevante observasjonene fra hver deltaker, totalt 24 rader (6 variabler ganger 4 betingelser). Dataene er da klare til å bli lagt inn i SPSS og analysert.

3.3 Deskriptiv statistikk

Dette underkapitlet handler om deskriptiv statistikk, som er statistikk som beskriver dataene fra eksperimentet etter at dataene er renses, ubrukelige data tatt ut og dataene er fremstilt for tolkning. I tabellen kan man se gjennomsnitt (M) og standardavvik (SD) for de ulike målingene utenom målingen «regresjon ut», som måles i sannsynlighet i prosent. Å vise den deskriptive statistikken er interessant for å framstille en oversikt over dataene, noe som kan gi innsikt i ulike tendenser som man ikke nødvendigvis får øye på med rm-ANOVA eller T-test. En annen fordel med å vise til standardavvik, for eksempel, er at det kan si noe om dataenes kvalitet, for det viser seg at et høyt standardavvik indikerer også hvor mye «støy» det er i dataene da dette indikeres av mye variasjon (Svartdal, 2015 i Vinichenko, 2021, s. 59). Vertikalt går de avhengige variablene og horisontalt går de ulike betingelsene der K står for kontrollsetning, F for setning med finittavvik, T for setning med tempusavvik og FT for setning med både finitt- og tempusavvik.

3.3.1 Deskriptiv statistikk av verbalt interessefelt

Andrespråkgruppen

	Første fikseringslengde (ms)		Blikklengde (ms)		Dveletid (ms)		Regresjonsstilengde (ms)		Fikseringsantall		Regresjon ut
	M	SD	M	SD	M	SD	M	SD	M	SD	%
K	226	64	317	158	529	470	418	343	2.38	2.09	16
F	247	75	360	161	567	371	459	347	2.37	1.46	18.6
T	270	114	348	216	596	441	477	374	2.31	1.56	13.8
FT	250	74	355	195	531	307	481	510	2.27	1.32	13.9

Tabell 5

Førstespråkgruppen

	Første fikseringslengde (ms)		Blikklengde (ms)		Dveletid (ms)		Regresjonsstilengde (ms)		Fikseringsantall		Regresjon ut
	M	SD	M	SD	M	SD	M	SD	M	SD	%
K	214	62	242	97	296	181	286	206	1.40	0.76	8.4
F	237	84	265	125	376	244	302	192	1.62	0.94	7.1
T	239	96	274	134	437	277	345	260	1.89	1.10	10.7
FT	241	93	276	137	407	252	324	234	1.79	1.00	8.4

Tabell 6

3.3.2 Deskriptiv statistikk av postverbalt interessefelt

Andrespråkgruppen

	Første fikseringslengde (ms)		Blikklengde (ms)		Dveletid (ms)		Regresjonsstilengde (ms)		Fikseringsantall		Regresjon ut
	M	SD	M	SD	M	SD	M	SD	M	SD	%
K	281	64	723	577	962	777	944	892	3.56	2.78	13
F	280	114	582	399	934	707	725	623	3.60	2.53	11
T	267	91	498	462	805	733	696	1057	3.17	2.69	16
FT	259	114	500	487	748	770	737	1435	2.99	2.77	15

Tabell 7

Førstespråkgruppen

	Første fikseringslengde (ms)		Blikklengde (ms)		Dveletid (ms)		Regresjonsstilengde (ms)		Fikseringsantall		Regresjon ut
	M	SD	M	SD	M	SD	M	SD	M	SD	%
K	228	67	351	208	429	273	415	300	1.91	1.14	9.3
F	235	82	336	206	450	311	470	358	1.97	1.24	23.5
T	233	81	331	187	480	310	560	471	2.14	1.26	9.3
FT	234	75	338	200	470	307	495	376	2.09	1.27	27.5

Tabell 8

3.3.3 Sammenligning

For å sammenligne de deskriptive statistikktabellene til Vinichenko og dette prosjektet, er en overordnet trend at utvalget med norsk som andrespråk gjennomsnittlig bruker lengre tid på å lese setningene. Det kan bety at de har behov for lengre tid til å prosessere setningene generelt. Det kan man forvente av andrespråkgruppen, fordi de ikke har like

mye kompetanse i å lese språket som førstespråkgruppen har og dermed trenger mer tid til å prosessere innholdet. Selv om de bruker lengre tid enn førstespråkgruppen, er det fortsatt noen andre trender som viser likheter mellom grupene. Variablene første fikseringslengde, blikklengde, dweetid og regresjonsstilengde i verbalt interessefelt viser gjennomsnittene fra andrespråkgruppen noen like tendenser med førstespråkgruppen som at begge utvalgene bruker gjennomsnittlig lengre tid på betingelsene med finitt- eller tempusavvik enn ved kontrollbetingelsene. På den andre siden er standardavvikene generelt høyere i andrespråkgruppen relativt til gjennomsnittet enn førstespråkgruppen, selv om de overordnet også bruker gjennomsnittlig lengre tid. De viser også en lik trend i standardavvikene mellom betingelsene. Standardavviket for betingelsene med avvik til dweetid i postverbalt interessefelt blir eksempelvis lavere enn kontrollbetingelsene for både andrespråkgruppen og førstespråkgruppen. I tillegg ser man at standardavviket øker ved testbetingelsene i begge gruppene, samtidig som man ser at standardavvikene generelt er høyere i andrespråkgruppen relativt til gjennomsnittet enn førstespråkgruppen. At standardavviket er høyt relativt til gjennomsnittet tyder på større individuell variasjon (Jin og Eide, 2015, s. 317; Busterud, 2014, s. 328). Det vil i så fall regnes som forstyrrende variabler. Individuelle forskjeller kan forklare dette, men det blir diskutert i diskusjonskapitlet (kap. 4). Det er derimot verdt å påpeke at individuelle forskjeller vil få større utslag med utvalg på 9 informanter i en undersøkelse.

Vinichenko påpeker om dataene sine at betingelse FT ikke har de høyeste verdiene av de tre testbetingelsene i noe av interessefeltene, noe som ikke stemmer overens med den alternative hypotesen (2021, s. 59). I andrespråkgruppen er ikke dette tilfellet under alle målinger, som første fikseringslengde, dweetid og fikseringsantall i verbalt interessefelt, men i regresjonsstilengde er denne verdien (FT) høyest. Under blikklengde er FT nest høyest, der F har lengst blikklengde. Det går derfor imot trendene til førstespråkgruppen. Om den deskriptive statistikken fra det post-verbale interessefeltet kommenterer Vinichenko (2021) at det er ulikt mønster der sammenlignet med det verbale interessefeltet, fordi blikklengde («gaze duration») viser lengre lesetid for kontrollbetingelsen enn for noen av testbetingelsene, hvor forklaringen kan være at et «verre» morfologisk avvik fører til at personenes blikk forlater det post-verbale interessefeltet, for å gå tilbake i setningen og «reparere bruddet» (s. 59). Det er også trenden i den deskriptive statistikken fra det post-verbale feltet for andrespråkgruppen. Om det er forårsaket av *spillover* fra avvikene er vanskelig å si, men ser man på resultatene fra kapittel 5.4.1.1, viser disse at det ikke er statistisk signifikant forskjell mellom betingelsene. I diskusjonsdelen (kapittel 6). går jeg nøyere inn på hva som kan være årsaken til dette.

3.4 Statistisk analyse – ANOVA og T-test

For å forstå tallene bak den deskriptive statistikken, kan rm-ANOVA og tohalet T-test brukes for å isolere hvordan hver måling blir påvirket av betingelsene i eksperimentet. Rm-ANOVA identifiserer om det er en stor nok forskjell mellom deltakerne i eksperimentet og om det kommer av den eksperimentelle manipulasjonen. Hvis variansen forårsaket av manipulasjonen er høy sammenlignet med andre faktorer, vil man få en høy F-ratio. F-ratio er en viktig verdi i analysen. Desto mer denne avviker fra tallet 1, desto større sannsynlighet er det for at dataene kan forklares av noe annet enn tilfeldigheter (Field, 2013, s. 549). Når vi får etablert om noen av målingene er statistisk signifikante med ANOVA, kan vi se på sammenhengene mellom de ulike betingelsene. Da brukes en tohalet T-test, for å finne ut hvilke betingelser som viser systematisk variasjon seg imellom. Under blir disse to statistiske analysene forklart og resultatene presentert.

3.4.1 rm-ANOVA

Rm-ANOVA er engelsk forkortelse av *repeated-measures analysis of variance*, som en statistisk metode for å se om det er gjennomsnittlige forskjeller mellom grupper – i denne studien forskjeller mellom stimuli (Field, 2013, s. 870). En ANOVA med gjentatte målinger brukes når formålet er å sammenligne betingelsers gjennomsnittlige forskjell (s. 544). Det kan også brukes til å teste hypoteser. Når en analyse er gjort, får man en F-ratio, som indikerer om de eksperimentelle manipulasjonene hadde en effekt utenom hva som undersøkes (Field, 2013, s. 554), altså forstyrrende variabler. Hvis F-ratioen da er større enn 1, indikerer det at betingelsene hadde en effekt. Gitt antall frihetsgrader er *den kritiske verdien* for F-ratio i dette eksperimentet 2.72 (*Repeated-Measures ANOVA*, u.å.) Kritisk verdi brukes som en terskelverdi for å bestemme om F-ratioen er statistisk signifikant med hensyn til frihetsgradene. F-ratio kommer frem i denne formelen, som er gjennomsnittlig kvadrat av systematisk variasjon (MS_M) og gjennomsnittlig kvadrat av usystematisk variasjon (MS_R) (Field, 2013, s. 554).

$$F = \frac{MS_M}{MS_R}$$

Under er det en oversikt over resultatene av de ulike målingene fra både dette prosjektet og Vinichenko (2021). Der blir det oppgitt F-verdi, P-verdi og partiell eta kvadrert. Partiell eta kvadrert er en effektstørrelse, som gjenspeiler hvor mye av den totale variasjonen er forklart av den uavhengige variabelen (Vinichenko, 2021, s. 63). Effektstørrelser ved 0.01 regnes som liten; 0.06 som moderat og 0.14 som stor effekt (Vinichenko, 2021, s. 64). Denne undersøkelsen analyserer dataene som innengruppefaktor med 4 nivå som i Vinichenko (2021, s. 60), som vil si at man analyserer alle 4 betingelsene for hver måling. I tabellen er statistisk signifikante resultat markert med skyggelagte ruter.

rm-ANOVA - resultater til sammenligning

	Krogstad 2023		Vinichenko 2021	
	Verbalt IF	Postverbalt IF	Verbalt IF	Postverbalt IF
Første fikseringslengde	F (3,234) = 3.846, p < .01, ηp2= 0.047.	F (3,243) = 0.860, p < .462, ηp2= 0.0171.	F (3,117) = 13.03, p < .001, ηp 2 = .250	F (3,117) = 0.43, p < .73, ηp 2 = .011.
Fikseringsantall	F (3,234) = 0.106, p < .95, ηp2= 0.01.	F (3,243) = 0.969, p < .408, ηp2= 0.012.	F (3,117) = 23.81, p < .001, ηp 2 = .379.	F (3,117) = 6.29, p < .001, ηp 2= .139
Dveletid	F (3,234) = 0.677, p < .056, ηp2= 0.009	F (3,243) = 0.361, p < .781, ηp2= 0.004.	F (2.37,92.24) = 36.07, p < .001, ηp 2 = .480	F (3,117) = 3.82, p < .012, ηp 2= .089
Blikklengde	F (3,234) = 1.047, p < .372, ηp2= 0.013.	F (3,243) = 1.025, p < .382, ηp2= 0.012.	F (3,117) = 8.21, p < .001, ηp 2 = .174	F (3,117) = 0.80, p < .49, ηp 2 = .020.
Regresjonsstilengde	F (3,234) = 0.553, p < .0647, ηp2= 0.007.	F (3,243) = 0.924, p < .43, ηp2= 0.011.	F (3,117) = 7.52, p < .001, ηp 2 = .162	F (3,117) = 14.28, p < .001, ηp 2 = .268
Regresjon ut	F (3,234) = 0.195, p < .90, ηp2= 0.002.	F (3,243) = 0.643, p < .588, ηp2= 0.008.	F (3,117) = 1.30, p = .277, ηp 2 = .032.	F (2.28,88.90) = 22.67, p < .001, ηp 2 = .368.

Tabell 9

Til sammenligning er det bare et av de tolv målingene fra denne undersøkelsen som er statistisk signifikant, mens 10 av Vinichenkos målinger er statistisk signifikante. I tillegg er effekten større i målingene fra Vinichenkos prosjekt. Det tyder på at øyeatferden til andrespråksinnlærere av norsk blir mindre påvirket av finitt- og tempusavvik enn nordmenn som har norsk som førstespråk.

3.4.1.1 Statistisk ikke-signifikante resultater

Etter å ha gjennomført rm-ANOVA på både verbalt og postverbalt interessefelt på 6 ulike målinger for ulike øyeatferder, var alle seks målene på begge interessefeltene, utenom første fikseringslengde i det verbale interessefeltet, ikke statistisk signifikant. Det vil si at ingen av de målene kan falsifisere noen av hypotesene ved undersøkelsen. Alle de ikke-signifikante målingene har en P-verdi på over 0,05 eller høyere. Det vil si at sannsynligheten for at resultatene er den samme eller mer ekstrem er for høy ut ifra denne studiens standard. Det vil si at dataene er for lite reliable til at det kan dras sikre

konklusjoner om de resultatene. I tillegg må F-ratioen være over kritisk verdi for å være statistisk signifikant. Målingenes lave F-ratio og lave effektstørrelser på partiell eta kvadrert, viser også at andrespråksgruppen ikke reagerer betraktelig på de morfologiske avvikene. De statistisk ikke-signifikante resultatene åpner derfor opp for at både alternative hypotesen og nullhypotesen er sann. Det vil si at de aller fleste resultatene fra prosjektet tyder på at både H1 og H0 er mulige. Disse resultatene blir diskutert i kapittel 4.

3.4.1.2 Statistisk signifikant resultat

Etter å ha gjennomført rm-ANOVA viser det seg at en av målingene er statistisk signifikant når vi ser på p-verdien og f-ratioen: første fikseringslengde. Resultatet ble $F(3,234) = 3.846$, $p < .01$, $\eta^2 = 0,047$ (se tabell 9). Det betyr at det er en betydelig forskjell mellom gjennomsnittene til betingelsene i denne målingen av det verbale interessefeltet, slik at en nullhypotese kan falsifiseres. Derfor ble det gjennomført en tohalet T-test for å sammenligne gjennomsnittene mellom betingelsene, slik at vi kan se om det er signifikante forskjeller på nettopp denne formen for øyefatferd - første fikseringslengde.

Et problem med analysen er at når det er så mange som 12 målinger, øker det sannsynligheten betraktelig for at man gjør en Type 1-feil, som er å konkludere at det er en sammenheng når det ikke er det (Field, 2013, s. 68). For å motvirke det kan en gjennomføre en Bonferroni-korreksjon:

$$P_{Crit} = \frac{\alpha}{\kappa}$$

Hvor det nye signifikansnivået for p-verdien brukes for å vurdere om resultatet er signifikant, som regnes ut ved å dele det originale signifikansnivået (α) på antall målinger (κ). I denne undersøkelsen er det 12 målinger, så signifikansnivået med Bonferroni-korreksjonen er da 0.00416. Med Bonferroni-korreksjonen kan ikke nullhypotesen falsifiseres, fordi p-verdien er 0.047. Grunnen er at sannsynligheten for å begå en type 1-feil ved å falsifisere nullhypotesen og mene det er en sammenheng er stor. På den andre siden øker det sannsynligheten for å begå en type 2-feil, som er en falsk negativ (Moran, 2003)⁹. Til sammenligning er p-verdiene på de statistisk signifikante resultatene til Vinichenko under terskelen på 0.00416 utenom dweetid i postverbalt felt, og det betyr at sannsynligheten for å konkludere med en falsk positiv i undersøkelsen til Vinichenko er betydelig lavere.

3.4.2 T-test – første fikseringslengde

Siden en av de avhengige variablene ble statistisk signifikant, ble det bare gjort parvise sammenligninger mellom betingelsene av den variabelen (T-test): første fikseringslengde fra det verbale interessefeltet. I motsetning til ANOVA kan T-testen gjøre parvise sammenligninger hvor man sammenligner gjennomsnittet fra eksperimentet med hva man ville forvente mellom populasjonsgjennomsnittet. Hvis nullhypotesen er sann, vil det ikke være noen forskjell mellom populasjonsgjennomsnittene, altså 0 (Field, 2013, s. 368). I et eksperiment vil skåren mellom betingelser være forskjellige, som kan være store eller små. Forskjellen til deltakernes skårer for hver betingelse blir regnet ut og lagt sammen, finner man den totale mengden differanse, som er \bar{D} i formelen under:

⁹ Hensikten med å gjøre en Bonferronis korreksjon er å vise at statistisk signifikans ikke er absolutt og at resultatene fra denne undersøkelsen ikke er entydige.

$$t = \frac{\bar{D} - \mu_D}{s_D / \sqrt{N}}$$

T-testen kalkulerer forholdet mellom den gjennomsnittlige forskjellen mellom betingelsene minus forventet forskjell mellom populasjonsgjennomsnitt og standardavviket delt på standardfeilen. T-testen er dermed en sammenligning av forventet gjennomsnitt og faktisk gjennomsnitt. Denne formelen gir oss en T-verdi, som brukes til å sammenligne gjennomsnittene til to betingelser. Betingelse F, T og F&T ble sammenlignet parvis med kontrollbetingelsen K, for å se hvordan verb med morfologiske avvik differensierte fra verb uten morfologiske avvik. Kritisk verdi for T-testen ved 80 frihetsgrader er 1.990 (Turney, 2022). I tillegg til T-verdi ble også Cohens *d* regnet ut for å finne en effektstørrelse på resultatene. Effekten regnes som liten ved 0.2, moderat ved 0.5 og stor ved 0.8 (Vinichenko, 2021, s. 64). T-testen viser at det er en statistisk signifikant forskjell mellom kontrollbetingelsen og tempus- og finittavvikene adskilt, men ikke for betingelse F&T, fordi både den kritiske verdien og p-verdien til F&T er for høy til å være statistisk signifikant.

Resultat T-test – første fikseringslengde

	KROGSTAD (2023)	VINICHENKO (2021)
F&K	t (78) = 2.004, p < .049, drm = .255	t (39) = -5.12, p < .001, drm = 1.09
T&K	t (78) = 2.995, p < .004, drm = .337	t (39) = -4.65, p < .001, drm = 1.00
FT&K	t (78) = 1.884, p < .063, drm = .212	t (39) = -5.79, p < .001, drm = 1.19

Tabell 10¹⁰

Gitt at resultatene stemmer, kan man si ut ifra T-testen på den avhengige variabelen første fikseringslengde at nullhypotesen er falsifisert, fordi det er evidens for at det er statistisk signifikant forskjell på gjennomsnittene mellom betingelsene sammenlignet med kontrollsetninger. Det vil si at man ikke kan si at øyeatferden er helt lik i møte med tempus- og finittavvik. Det vil i så fall være et bevis imot nullhypotesen. Når det kommer til effektstørrelsen Cohens *d*, er det en liten til moderat effekt av betingelse F og T sammenlignet med kontrollbetingelsen. P-verdiene for de tre parvise sammenligningene er godt under terskelverdien uten og med Bonferroni-korreksjonen fra undersøkelsen til Vinichenko (se tabell 10). Hvis det blir gjort en Bonferroni-korreksjon på T-testen også, vil den nye terskelen være 0.0166. Det medfører at bare den parvise sammenligningen av setninger med tempusavvik og kontrollsetninger, er statistisk signifikant. Det vil i det tilfellet bety at bare tempusavvik påvirker øyeatferden, hvis man vil redusere sannsynligheten for å begå en type 1-feil, som er å konkludere med en falsk positiv. På den andre siden øker det sannsynligheten for å begå en falsk negativ.

¹⁰ Om T-verdien er positivt eller negativt har ikke noe å si for resultatet. Forskjellen er hvilken rekkefølge T-testen er tatt mellom betingelsene.

3.4.3 Sammenligning av grupper

Til sammenligning med resultatene fra Vinichenko (2021) ser vi at det er større forskjell og mer systematisk forskjell mellom betingelsene i førstespråksgruppen enn andrespråksgruppen når en sammenligner T-verdiene deres. Hvis en også ser på Cohens d , er det en tydelig, stor forskjell i effektstørrelse fra andrespråksgruppen, hvor alle effektene er på 1.00 eller høyere i førstespråksgruppen. Det betyr at avvikene har betydelig større effekt på øyeatferd enn for andrespråksgruppen. Dette gjelder ikke bare første fikseringslengde, men det er ofte moderat og stor effekt på T-testen i de andre målingene også hos førstespråksgruppen, som ikke hadde statistisk signifikans i andrespråksgruppen. Hvis man også tar høyde for Bonferronis korreksjon, er resultatene til Vinichenko T-tester for alle betingelsene for første fikseringslengde statistisk signifikant. Til forskjell fra førstespråksgruppa er bare resultatet for betingelsen tempusavvik fra T-testen i andrespråksgruppa statistisk signifikant med Bonferroni-korreksjonen.

3.5 Kapitteloppsummering

I dette kapitlet har resultatene fra både dette prosjektet og masteroppgaven til Vinichenko (2021). Betraktelig flere resultater fra studien til Vinichenko er statistisk signifikant enn i denne studien. Det viser seg at én av totalt tolv målinger med seks variabler på to ulike interessefelt er statistisk signifikant i rm-ANOVA. Hvis den blir kontrollert med Bonferroni-korreksjonen blir målingen ikke-signifikant, men det kan være en falsk negativ. Den variabelen heter første fikseringslengde. Det ble også gjennomført en T-test for å sammenligne de uavhengige variablene (betingelsene med morfologiske avvik) parvis med kontrollbetingelsen. Der viser det seg at det er en statistisk signifikant forskjell både tempus- og finittavvik sammenlignet med kontrollsetninger uten Bonferroni-korreksjonen, men med korreksjonen er bare tempusavvik statistisk signifikant, til forskjell fra førstespråksgruppen hvor alle betingelsene var statistisk signifikant med Bonferroni-korreksjonen.

4 Diskusjon

I dette kapitlet diskuterer jeg resultatene fra min undersøkelse i lys av tidligere funn og etablerte andrespråksteorier. Jeg ser på hva dataene kan si om hypotesene for forskningsprosjektet, metodens styrker og svakheter, hvilke teoretiske implikasjoner empirien muligens har og hva som først og fremst utmerker seg som løse tråder som bør forskes videre på. Delkapittel 4.4 inneholder oppsummering og konklusjon.

Hypotesene for prosjektet ble presentert i kapittel 2.3, men jeg gjentar dem her for enkelhets skyld:

Alternativ hypotese: Finitt og tempus oppfattes som to adskilte trekk. Avvik i finitt- og tempusmorfologi oppfattes som to distinkte typer avvik, som resulterer i forskjellig leseemønstre

Nullhypotese: Finitt- og tempustrekkene oppfattes ikke som to adskilte trekk; finitt- og tempusavvikene påvirker derfor øyebevegelse på samme måte.

Underhypoteser:

- 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.
- Både tempus- og finittavvik: lengre og flere fikseringer og flere regresjoner enn både i kontrollsetninger og i setninger med bare en type brudd.

4.1 Resultatdiskusjon

Den statistiske analysen og sammenligningen med resultatene fra et tidligere masterprosjekt i kapittel 5 viste at bare en måling var statistisk signifikant uten Bonferronis korreksjon i rm-ANOVA. Her viste F-ratioen at det kan være en forskjell på hvordan andrespråksinnlærere reagerer på ulike betingelser. Det vil si at ulike morfologiske avvik i verb kanskje kan påvirke øyeatferden deres i møte med norske setninger. Derfor ble det gjennomført en T-test for å se spesifikt på de ulike testbetingelsene. Men med Bonferronis korreksjon er forskjellen mellom betingelsene ikke store nok til at det er en statistisk signifikant forskjell. Det betyr, som før diskutert, ikke nødvendigvis at resultatene er statistisk ikke-signifikante, fordi Bonferronis korreksjon kan øke sannsynligheten for å begå en type 2-feil, også kalt falsk negativ (Moran, 2003), som tilsier at en ikke skal utelukke at det er stor nok forskjell mellom betingelsene. Hvis det så likevel er signifikante forskjeller vil det i så fall falsifisere nullhypotesen om at tempus og finittavvik ikke oppfattes som adskilte trekk.

Alle andre målinger enn første fikseringslengde på verbalt interessefelt er ikke-signifikante ved rm-ANOVA, hvor F-ratioene og p-verdiene tilsa at det ikke er statistisk signifikante forskjeller mellom gruppene forårsaket av manipulasjonene til og med uten Bonferronis korreksjon. Hvis vi ser på den deskriptive statistikken, er det en generell trend at det er lengre fikseringer, flere regresjoner og fikseringsantall ved testbetingelsene enn kontrollbetingelsen, som tyder på en systematikk i dataene, men det er vanskelig å vite hvordan denne systematikken best kan forklares. Resultatene fra rm-ANOVA beskriver bare mønstrene som eksisterer i datasettet, det blir opp til oss som

forskere å tolke hva vi ser. Mangelen på statistisk signifikans hos de aller fleste målingene viser at andrespråkgruppen er mindre sensitiv overfor testbetingelsene sammenlignet med førstespråkgruppen, fordi førstespråkgruppen har både flere statistisk signifikante resultater og tydelig større effektstørrelse, altså partiell eta kvadrat og Cohens d , på alle målinger med statistisk signifikans. Likevel, selv om resultatene fra våre målinger i denne undersøkelsen ikke er statistisk signifikante, betyr det ikke at alle har de samme atferdsmønstrene på individnivå. Andrespråkgruppen samlet sett har bare ett signifikant resultat med rm -ANOVA av totalt 12 målinger, hvor effektstørrelsen til og med er liten til moderat, mens førstespråkgruppen sin effektstørrelse er fem ganger større enn førstespråkgruppens effektstørrelse. Derfor er dataene tydelige på at det er en stor forskjell på gruppenes sensitivitet til verbmorfologiske avvik i sin helhet, men det er også øyeatferd som kan være evidens for at de reagerer på og prosesserer avvikene.

Uten Bonferronis korreksjon var en variabel statistisk signifikant: første fikseringslengde. Derfor ble T-testen gjort på denne målingen, mellom betingelsene for den variabelen. I T-testen kom det frem at både tempus- og finittavvik hver for seg gir statistisk signifikante utslag på øyeatferd, men ikke når begge avvikene er i setningen. Samtidig ble det gjort en Bonferronis korreksjon på denne målingen også. Etter korreksjonen var fortsatt resultatet for betingelsen med tempusavvik statistisk signifikant, men finittavviket var ikke lenger signifikant med korreksjonen. Det betyr at ved finittavvik har vi mer usikre utslag og det kan være at det vi har er en falsk positiv eller negativ. Dette gjelder uansett med eller uten korreksjonen og uansett om vi konkluderer med effekt eller mangel på effekt. Det T-testen viser, er at andrespråksinnlærere reagerer på tempusavvik med Bonferronis korreksjon, som esikrere videns for at andrespråkgruppen reagerer på tempusavvik på første fikseringslengde. Det er på den andre siden en klar mulighet for at det signifikante resultatet for finittavvik er en falsk positiv når man ser at p -verdien er 0.49 uten korreksjonen, men det er fortsatt interessant at finittavvik har en effekt, selv om den er liten til moderat. Det at tempusavvik og finittavvik fører til ulike øyeatferd falsifiserer og forkaster nullhypotesen og øker sannsynligheten for at den alternative hypotesen er korrekt. Det ser derfor ut som at utvalget som helhet oppfatter tempus og finittethet som to adskilte trekk, eller at avvik som angår finittethet ikke oppfattes av informanter, men det er likevel påfallende at finittethet har noe effekt ett eller annet sted i gruppen, noe som dermed gir utslag i den deskriptive statistikken. Det ser altså ut til at andrespråkgruppen i større grad legger merke til og er mer sensitive for tempusavvik i større grad enn finittavvik. Det er vanskelig å si når marginene er så små og resultatene tvetydige. Men tendensene går i samme retning, og mønstrene kan ses som svake avskygninger av de mønstrene Vinichenko (2021) påviste for førstespråkgruppen.

4.1.1 Hva sier det om øyeatferd og språkprosessering?

Innledningsvis ble det nevnt at med blikksporing måles øyets performanse i det reseptive systemet, i tillegg til at øye-sinn-hypotesen ble forklart i kapittel 4.2 hvor antakelsen er at varigheten på fikseringene tilsvarer innsatsen prosesseringen av ordene krever. Det er evidens som peker mot at andrespråksinnlærere av norsk prosesserer tempusavvik ved første fikseringslengde og legger merke til tempusavvikene til en viss grad. Vi kan falsifisere nullhypotesen om at finitt- og tempusavvik ikke oppleves som to adskilte trekk, rett og slett fordi avvikene ikke påvirker øyeatferden til gruppen på samme måte. Det sier derimot ikke noe om *i hvor stor grad* verbmorfologien prosesseres. Effektstørrelsen er liten til moderat sammenlignet med førstespråkgruppen, som i så fall vil bety at andrespråkgruppen så vidt legger merke til finittavvik, om de i det hele tatt gjør det. Samtidig er det nødvendig å påpeke hvordan den deskriptive statistikken gir et

annet inntrykk enn rm-ANOVA, fordi testbetingelsene tydeligvis påvirker gjennomsnittene på de ulike målingene. Siden det ble funnet ut at andrespråksgruppen er mye mindre sensitiv til finitt- og tempusavvik i norsk, kan det antas at de prosesserer morfologiske avvik i norske verb i mye mindre grad. Det er usannsynlig at andrespråksgruppen prosesserer de morfologiske avvikene raskere enn førstespråkgruppen, fordi forskningen er klar på at det er vanskeligere å lære et nytt språk som voksen etter en viss alder enn det er å lære det som barn (Hyltenstam & Abrahamsson, 2003), selv om det er unntak hvor noen lett lærer seg nye språk og mange av dem (Jouravlev et al., 2021). Hensikten med denne studien er å se på andrespråksinnlærere som gruppe sammenlignet med førstespråksinnlærere – ikke se på hvordan enkeltindivider er bedre til å lære nye språk enn andre.

Første fikseringslengde, som er det ene statistisk signifikante resultatet, er en tidlig måling og måler varigheten til den første fikseringen. At tempusavvik fører til at den første fikseringen er noe lengre, kan bety på at andrespråksinnlærere av norsk merker at tempus er feil tidlig i prosesseringen ved første øyekast. Det at ikke noen av de andre målingene gir utslag, har også implikasjoner for øyeatferd og språkprosessering. Siden senere målinger gir innsikt i mer bevisste språkprosesseringer som omanalysering, kan en anta at det ikke skjer senere, mer bevisste språkprosesseringer av tempusavvik hos andrespråksgruppen, og at finittavvik ikke prosesseres i det hele tatt. Det tyder på at de ikke «reparerer» eller i betydelig grad blir bevisste på morfologiske tempus- eller finittavvik i det hele tatt, selv om de i tidlig prosessering ved første fiksering muligens legger merke til tempusavvik.

Tempusbøying i verb har mer semantisk betydning en finitthetsbøying, og det kan ha noe å si for andrespråksinnlæreres prosessering av morfologien. Teorier om sensitive perioder for språklæring viser til at myelinisering av nevronene i hjernen fører til modning, som igjen gjør det vanskeligere å lære seg fonologiske og syntaktiske språkssystemer desto eldre man blir, mens modningen ikke har like stor effekt på semantiske og pragmatiske systemer (Pulvermüller & Schumann, 1994). Når man blir eldre, blir det vanskeligere å lære seg syntaktiske systemer og ikke like vanskelig å lære seg semantiske systemer, og det kan i så fall forklare hvorfor de voksne andrespråksinnlærerne av norsk legger merke til tempusavvikene mens finitthetsavvik ikke blir lagt merke til. Det blir diskutert mer i detalj nedenfor ut ifra etablert andrespråksteori.

4.1.2 Teori til diskusjon - Andrespråksteori

Paradigmet undersøkelsen tar utgangspunkt i, er generativ grammatikk, som antar at mennesket har en universalgrammatikk. Det går ut på at mennesket er født og pre-programmert til å lære seg ethvert språk med enhver naturlig forekommende grammatisk struktur så lenge man blir eksponert for dem som barn. Prinsipp- og parameterteorien går ut på at man har universelle grammatiske prinsipper som utgjør vår medfødte språkevne, altså universalgrammatikk, mens spesifikke språk har spesifikke prinsipper og parametere, som blir tilegnet og internalisert når man eksponeres for språk, og universalgrammatikken blir da i mindre grad tilgjengelig.

Hvert språk har sine prinsipper og parametre, og kan også kalles språktrekk, hvor noen språk har like eller ulike språktrekk. Antakelsen går ut på at med alderen «modnes» grammatikken der disse grammatiske prinsippene eller parameterne blir låst fast til en verdi som en gammel bryter blir låst fast, opp eller ned, metaforisk sett (Eide og Busterud, 2015, s. 31-33). Det betyr at det blir vanskeligere for voksne å lære seg nye

språktrekk. Disse parameterne er vesentlig enklere å lære seg som barn enn som voksen og det er argumenter for *kritiske* eller *sensitive perioder* for språktilegning, blant annet fordi det er eksempler på barn som ikke har hatt tilgang på språk i relevante aldersspenn, som fikk vansker med å lære seg språk. Det er usikkert når disse periodene starter og slutter, men det er relativt stor enighet om det er en grense i 6-7-årsalderen og puberteten (Eide og Busterud, 2015, s. 38-40). Med dette paradigmet i bunn, finnes det også spesifikke andrespråksteorier om menneskets evne til å tilegne seg nye språk med nye grammatiske strukturer.

I generativ andrespråkforskning skilles det mellom to hovedhypoteser: trenings- og modningshypotesen. Modningshypotesen går ut på at mennesket tidlig har en enormt god evne til å lære seg språk og at den evnen forsvinner med alderen, mens treningshypotesen mener denne evnen kan trenes, aktiveres eller vedlikeholdes igjennom livet (Johnson & Newport, 1989 i Eide og Busterud, 2015, s. 42). Treningshypotesen forutsetter dermed at universalgrammatikken er tilgjengelig. Da er teoriene *Failed Functional Features Hypothesis* (FFFH) og *Full Transfer, Full Aksess* (FTFA) relevante. Begge teoriene tar utgangspunktet i det generative paradigmet om universalgrammatikk med prinsipper og parametere.

4.1.1.1 Failed Functional Feature Hypothesis

FFFH ble først presentert i studien til Hawkins og Chan (1997) der konklusjonen var at UG-parameterne ikke kunne få ny verdi i innlæringen av andrespråk (Jin, Eide og Busterud, 2015, s. 245). Teorien forklarer andrespråksinnlæreres mangel på målspråklighet med at de UG-parameterene som angår funksjonelle trekk ikke kan justeres i voksen alder hvis trekket ikke allerede finnes i førstespråket. Her skilles det mellom tolkbare og utolkbare trekk. Ikke-funksjonelle trekk eller tolkbare trekk på regnes i større grad som semantiske og kan konseptualiseres i større grad (Jin, Eide og Busterud, 2015, s. 243-244). Her forenkler jeg en hel del for framstillingens skyld, siden semantikk og funksjonelle trekk ikke nødvendigvis er adskilte, uavhengige språkfenomen (Jin, Eide og Busterud, 2015, s. 243; Wik, 2014, s. 20). Det er noe usikkert om finittet er et tolkbart eller utolkbart, men det er forstå det som et utolkbart trekk (Wik, 2014, s. 21-23), hvor noen mener det er et syntaktisk trekk (Cowper, 2016), og selv om det er omdiskutert, blir finittetsbøying i norsk omtalt som et morfologisk fenomen, som har flere syntaktiske egenskaper enn tempusbøying, se Wik (2014). Finittetsbøying blir dermed ikke definert av alle som et utolkbart trekk, men blir likevel ofte forstått som et trekk som har mindre konseptuelt innhold enn tempus. Tempusbøyingene vi har forsket på i denne undersøkelsen skiller mellom fortid og ikke-fortid, som har et klart semantisk tilsnitt, og skillet mellom finitt og ikke-finitt bøying, som er syntaktisk basert. Uansett er fortsatt teorien anvendbar i analysen av empirien.

4.1.1.2 Full Transfer, Full Aksess

FTFA-hypotesen går ut på at UG-grammatikken er fullstendig tilgjengelig og kan endres til og med i andrespråkstilegning. Med tilstrekkelig eksponering fra andrespråket, vil det kunne påvirke parametrene slik at disse kan endres til å passe til det nye språket. Det betyr ikke at innlæreren har en «tabula rasa» der førstespråket ikke har noe å si, fordi jamfør teorien vil andrespråksinnlæreren uansett ta utgangspunkt i parameterverdiene fra førstespråket (Jin, Eide og Busterud, 2015, s. 243), men til forskjell fra FFFH er denne hypotesen ikke deterministisk om hva andrespråksinnlæreren kan tilegne seg. Hypotesen ble formulert for første gang av Schwartz og Sprouse (1996) ut ifra en longitudinell studie de gjennomførte der en tyrker tilegnet seg mer målspråklige mønstre på tysk, som ikke fantes i tyrkisk (Jin, Eide og Busterud, 2015, s. 259-262).

4.1.3 Analyse med hjelp av andrespråksteori

For å forstå de empiriske resultatene i undersøkelsen ut ifra disse teoriene er det nødvendig å trekke inn språkbakgrunnene til deltakerne. Siden dette er et leseeksperiment, kan ortografisk språkbakgrunn være en relevant forklaring. Det er 4 av 9 i utvalget som ikke har det latinske alfabetet som utgangspunkt, og det kan være en mulig forklaring på det generelt er lite sensitivitet for tempus- og finittavvik i de ulike målingene, men at testbetingelsene fortsatt har høyere gjennomsnitt enn kontrollbetingelsen og har noen trender som ligner førstespråkgruppen. Det er også et poeng å se på hvilke talespråk de har som førstespråk, for det er blitt aktivt rekruttert informanter med språk som ikke har like utstrakt finitthetsmorfologi som norsk. Det er eksempelvis flere med engelsk som førstespråk, som har mindre markert finitthetsmorfologi og har mer markert tempusmorfologi. Det kan forklare hvorfor avvik i tempusmorfologi blir lagt merke til, mens finitthetsbøying blir oversett. Det ville gi mening ut ifra FFFH, fordi finitthet er mindre tolkbart enn tempus og er derfor vanskeligere å lære som voksen, ifølge denne teorien.

En annen dimensjon ved innlæring av norsk er hvor lang tid har blitt eksponert for og lært seg norsk. Det er påpekt at for mange andrespråksinnlærere mangler bøyingsmorfologi i stor grad i tidlige faser av innlæring av målspråket (Åfarli, 2015, s. 68). 6 av 9 deltakere har hatt et år eller mindre på å lære norsk, og det kan forklare hvorfor andrespråkgruppen er mindre sensitiv til verbmorfologiske avvik enn førstespråkgruppen. Hvis det forklarer mangelen på sensitiviteten for verbmorfologiske avvik, kan dette også forstås i en FTFA-ramme, fordi informantene da ville trenge mer tid for å bli eksponert for nok norske språkstimuli, spesifikt verb. Med det ville det i tillegg vært interessant å vite mer om hvor mye deres lesekompetanse spesifikt på norsk som er medvirkende til resultatene. Det er naturlig å tenke seg at også denne mulige variabelen påvirker sensitivitet for verbmorfologiske avvik når de leser setningene i undersøkelsen. Det er vanskelig å si noe om hvor mye lesestimuli informantene våre har blitt utsatt for, men det at nesten hele gruppen bare har lært norsk i et år, kan bety at de ikke har fått nok «trening» i å lese, noe som bidrar til at de heller ikke har nok trening i å oppdage norske verb og legge merke til avvik i dem.

4.2 Metodologisk diskusjon

Resultatene er et godt utgangspunkt for å diskutere undersøkelsens validitet og spesielt reliabilitet. Vi vet hvor mange informanter som ble inkludert, bakgrunnen til hver enkelt og hva resultatene er. I dette underkapitlet diskuterer jeg resultatene i lys av metodens styrker og svakheter. Det at resultatene fra de fleste målingene hverken hadde statistisk signifikans eller betydelige effekter kan forklares i et teoretisk andrespråkparadigme hvor det til en viss grad gir mening at øyeatferden til andrespråkgruppen ikke er lik førstespråkgruppen. Det er ikke bare forventet ut ifra FFFH, men også FTFA, fordi begge teoriene forventer at voksne innlærere er mindre sensitive overfor stimulien enn de som har norsk som førstespråk. Siden funnene støtter opp under teori og tidligere funn, tyder det på at dataene er reliable.

Tilsynelatende ser det ut til å være en diskrepans mellom resultatene av rm-ANOVA'ene og den deskriptive statistikken, fordi det er generelt høyere gjennomsnitt på testbetingelsene enn kontrollbetingelsen. Det generelle høyere gjennomsnittet hos alle betingelser tyder jo på at avvikene har en effekt, men det er ikke nok systematikk til å være signifikant i en rm-ANOVA. Hva kan i så fall forklare at testbetingelsene har høyere gjennomsnitt? En mulig forklaring er at det er individuelle, muligens store, forskjeller, som forklarer at gjennomsnittene er høyere. Det er et lite utvalg, med bare 9 deltakere,

som derfor kan påvirke gjennomsnittene. Hvis utvalget hadde vært større, er det mulig at gjennomsnittene ville jevnet seg ut eller at rm-ANOVA hadde gitt flere målinger med statistisk signifikans. Det vil si at resultater med et større utvalg kunne sett annerledes ut og ført til andre konklusjoner. Når utvalget er så lite, kan individuelle forskjeller dra gjennomsnittet opp, og det er ut ifra det målet vanskelig å si om «alle» la litt merke til avvikene eller om en la merke til alle sammen og dermed alene dro opp hele gjennomsnittet. Med et utvalg av denne størrelsesorden kan resultatene lettere bli påvirket av individuelle forskjeller. I så faller det vært interessant å studere nærmere hvilke individuelle forskjeller som kan ha påvirket andrespråksinnlæreres kompetanse i det nye språket.

Når det er et så lite og lingvistisk sett spredt utvalg kan det være forstyrrende variabler som forklarer variasjonen. Det kan være individuelle forskjeller i kompetanse, hvor lenge de har lært norsk, hvilke språk de kan fra før, hvor mange språk de kan og hva slags ortografi språkene de kan fra før har. Dette er diskutert over, og dette illustrerer dermed utfordringen med å forske på en heterogen gruppe som om den var en mer homogen gruppe. Ut ifra den tilgjengelige informasjonen om informantene, individuelle forskjeller forklares av førstespråket, ortografien de er kjent med, hvor lang tid de har hatt på å lære norsk og kompetansenivå (se tabell 3). Siden utvalget er lite, med 9 informanter i en kvantitativ undersøkelse, vil individuelle forskjeller få større utslag. Dette er da potensielle forstyrrende variabler i undersøkelsen, som kan ha påvirket resultatene for undersøkelsen. Det viser at andrespråksgruppen er en heterogen gruppe, som ser ut til å bli påvirket av ulike typer av relevante (men forstyrrende) variabler i større grad enn førstespråkgruppen, og det viser dette prosjektets begrensning, fordi det eksperimentelle designet ikke er lagt opp til å slå fast konklusjoner om hvordan spesifikk språkbakgrunn påvirker leseferden av norsk. Det er riktignok nødvendig å påpeke og diskutere. Mangelen på statistisk signifikans i de fleste målingene hadde muligens sett annerledes ut med et større utvalg, og et større utvalg ville gjort studien mer reliabel, fordi individuelle forskjeller påvirker resultatene i større grad i små utvalg som i denne studien.

Et annet metodologisk tankekors er at problemstillingen og dette prosjektets formål er å finne generaliserbare funn om et spesifikt morfologisk fenomen i en liten, mangfoldig gruppe, som kommer fra vidt forskjellige språkbakgrunner. Hvor man på den ene siden vil ha flest mulig informanter, da dette er en kvantitativ metode, som krever en god del tid og tilstedeværelse og ikke minst fysisk oppmøte, og på den andre siden vil man også filtrere ut informanter med for lavt kompetansenivå, mindre relevante førstespråk (for eksempel svensk), og ulike lesevaner eller ADHD. Kriteriene for utvalget er derfor krevende, fordi det både er tidkrevende å rekruttere mange informanter fra en så liten populasjon til datainnsamling, som krever fysisk oppmøte på universitetet, samtidig som det klart er ønskelig å ha så mange informanter som mulig.

4.3 Videre forskning

For ungdom og voksne krever det å mestre finitthetskategorien en mer analytisk og bevisst tilnærming, mens barn kan assimilere inputtmønstrene uten den samme analytiske tilnærmingen til andrespråksinnlærere (Dimroth, 2008). Siden finitthet i seg selv ikke blir undervist om på norskkurset til deltakerne spesifikt, kan det bety at utvalget med andrespråksinnlærere ikke har fått den metaspråklige kunnskapen til å tilnærme seg finitthetsmorfologien analytisk til at de kan legge merke til finittavvikene. Det hadde dermed vært interessant å vite om det å gi andrespråksinnlærerne av norsk den nødvendige metaspråklige kunnskapen, for eksempel via en intervensjon, ville ha

påvirket øyeatferden, for det ville også kunne si noe om øye-sinn-hypotesen. Det ville i så fall vært interessant om metaspråklig kunnskap via en intervensjon kan føre til faktisk atferdsmessig endring for blikket som regnes som mindre kontrollerbart enn andre språkatferder som skriving. Derfor ville det også vært interessant å se på hvordan regelmessig lesing av norsk og mengden lesing av norsk påvirker andrespråksinnlærere av norsk sin evne til å legge merke til morfologiske avvik i blikksporingseksperiment. Det hadde vært et interessant både ut ifra et teoretisk og didaktisk perspektiv.

Siden resultatene viser at andrespråksinnlærere av norsk er mye mindre sensitive overfor morfologiske avvik, sammenlignet med de med norsk som førstespråk, hadde det også vært interessant å gi dem andre typer av elisiteringstester, grammatikalitetsvurderinger, skriveoppgaver hvor de må produsere verbene i norsk eller andre metoder hvor man tester andrespråksinnlæreres kompetanse om tempus- og finittmorfologi, for å se hvordan det henger sammen med deres lavere sensitivitet til tempus- og finittmorfologi. Da kunne vi funnet ut om det er en sammenheng mellom reseptive systemet gjennom øyeatferd og det produktive systemet.

Det er også behov for at dette blir forsket på av et større utvalg og hvor man også kan se på innvirkningen det har å kunne flere språk, samt hvordan spesifikke førstespråk og andrespråk kan påvirke norsk som tredjespråk. Som nevnt tidligere et problem med å forske på andrespråksinnlærere av norsk at det er mange kriterier som skal oppfylles for at andrespråksinnlærere kan inkluderes i undersøkelsen, og da igjen blir det enda færre som faktisk deltar i eksperimentet. I tillegg til det blir det et reliabilitetsspørsmål når det er så variert hvilken språkbakgrunn informantene har, fordi det kanskje er store forskjellige utfall ut ifra språkbakgrunn, hvor lenge man har lært norsk, førstespråket, ortografien og så videre.

En annen fordel med en større undersøkelse er om de tendensene som vi kan se i den deskriptive statistikken ville gjenspeiles som signifikante mønstre hos andrespråksinnlæreres øyeatferd ei en større gruppe, fordi det uansett er påfallende hvordan informantene bruker gjennomsnittlig lengre tid på flere målinger når det både er finitt- og tempusavvik. Da dette er kanskje det største spørsmålet jeg undrer mest over. Hva forklaringen på dette er, kan vi vanskelig si med sikkerhet, men det er mulig å spekulere. Etter å ha analysert dataene fra enkelte informanter med like språktrekk i førstespråket, informant 7 og 6 med henholdsvis afrikaans og tysk, og sammenlignet disse med informantene 1 og 9, med engelsk og kinesisk som førstespråk, finner vi noe interessant. Afrikaans og tysk har mer markert finittbøying enn engelsk og kinesisk, og her viste det seg at gjennomsnittene for reaksjoner på finittavvik var betydelig høyere hos de som hadde afrikans og tysk som førstespråk, og disse målingene skilte seg markant fra informantene med engelsk og kinesisk som førstespråk.

Tabell med gjennomsnitt mellom informanter

	Afrikaans&tysk (informant 7&6)	Engelsk&kinesisk (informant 1&9)
K	223,50	266,36
F	242,26	256,62
T	250,50	260,52
F&T	237,62	256,08

Tabell 11 – viser gjennomsnitt fra første fikseringslengde for disse fire informantene

I tabell 11 kan man se sammenlagte gjennomsnitt av to informanter med to ulike språk. Ut ifra FFFH skal det være vanskeligere å tilegne seg et mer syntaktisk eller funksjonelt trekk som finitthet hvis man ikke har det i førstespråket. Derfor ble det sett på om dataene fra informanter førstespråk som har mer markert verbmorfologi, spesifikt finitthet, viser en ulik trend sammenlignet med de med mindre markert verbmorfologi. I tabellen kan man se at de med afrikaans og tysk som førstespråk viser den trenden vi også finner i Vinichenko (2021) hos nordmenn, men den trenden vises ikke i det hele tatt hos de med engelsk og kinesisk. De bruker til og med gjennomsnittlig kortere tid når det er avvik enn uten avvik. Det viser for det første hvordan individuelle forskjeller kan påvirke trendene vi ser, og for det andre at transfer fra førstespråket muligens kan bidra til å forklare forskjellene. Dette er forskningsspørsmål jeg svært gjerne skulle ha tid til å forfølge, men på grunn av det begrensede formatet masteroppgaven gir, er det ikke anledning til å studere alle variablene disse for informantene enkeltvis. Det ville ha krevd å kjøre alle testene for statistisk signifikans for hver av betingelsene for hver informant enkeltvis, og det var det simpelthen ikke tid til. Likevel er det interessante tendenser som kan observeres bare ut fra de fire informantenes data i tabellen over.

4.4 Konklusjon

I denne studien har jeg undersøkt om andrespråksinnlærere av norsk har kompetansen til å oppfatte morfologiske tempus- og finittavvik og om de oppfattes som distinkte trekk av voksne innlærere av norsk. Med bruk av blikksporing som metode har jeg gjennom denne studien forsøkt å svare på problemstillingen.

Det ble funnet et statistisk signifikant for tempusavvik i første fikseringslengde hvor sannsynligheten for en falsk positiv er liten med Bonferronis korreksjon, mens finitthet bare var signifikant uten korreksjonen. Det er derfor evidens for at andrespråkgruppen oppfattet tempusavvik og muligens finitthet, men det gjelder bare en måling av seks variabler på ett av to interessefelt. Overordnet sett og sammenlignet med førstespråkgruppen er andrespråkgruppen mye mindre sensitive til verbmorfologiske avvik enn førstespråkgruppen, som i det hele gir støtte til Failed Functional Feature Hypothesis, gitt en tilleggsantakelse om at finitthet er et funksjonelt trekk. Avvik i den mest semantiske bøyningen tempus, ser ut til å vekke en påvisbar reaksjon, i motsetning til finitthetsbøyning, som har syntaktiske egenskaper.

Leseren vil kanskje konkludere med at etter denne studien sitter man igjen med flere spørsmål enn man hadde før. Det er fortsatt usikkert hvordan resultatene hadde sett ut med et større utvalg, hvis andrespråksinnlærerne hadde hatt mer tid til å lære seg norsk, eller lært hatt en intervensjon hvor man underviste spesifikt om distinksjonen mellom tempus og finitthet. Nullhypotesen kan derimot forkastes, men det er fortsatt tvil om finitthet faktisk ble oppfattet og også i hvor stor grad tempusmorfologi oppfattes. Det er likevel ikke tvil om at voksne andrespråksinnlærere av norsk, som gruppe betraktet, er mye mindre sensitive til verbmorfologiske tempus- og finittavvik enn voksne informanter med norsk som sitt førstespråk.

5 Referanseliste

- Baddeley, A. D., & Hitch, G. J. (1974). Working memory. In G. H. Bower (Ed.), *The psychology of learning and motivation* Vol. 8 (pp. 47–89). New York: Academic Press.
- Becker, J. T., & Morris, R. G. (1999). Working Memory(s). *Brain and Cognition*, 41(1), 1–8. <https://doi.org/10.1006/brcg.1998.1092>
- Booij, G. (2012). *The Grammar of Words: An Introduction to Linguistic Morphology*. Oxford University Press.
- 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>
- Brown, G., Malmkjaer, K., & Williams, J. (1996). *Performance and Competence in Second Language Acquisition*. Cambridge University Press.
- Busterud, G. (2014). *Anaforiske bindingskonstruksjoner i norsk som andrespråk* [Doktorgradsavhandling]. Norges teknisk-naturvitenskapelige universitet.
- 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
- Cowper, E. (2016). Finiteness and Pseudofiniteness. I K. M. Eide (Red.), *Linguistik Aktuell/Linguistics Today* (Bd. 231, s. 47–78). John Benjamins Publishing Company. <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 93 Research]. *Frontiers in Human Neuroscience*, 11(563). <https://doi.org/10.3389/fnhum.2017.00563>
- Dimroth, C. (2008). Age Effects on the Process of L2 Acquisition? Evidence From the Acquisition of Negation and Finiteness in L2 German. *Language Learning*, 58(1), 117–150. <https://doi.org/10.1111/j.1467-9922.2007.00436.x>
- Eddington, D. (2008). Linguistics and the scientific method. *Southwest Journal of Linguistics*, 27.

- Eide, K. (2009). *Tense, finiteness and the survive principle. Temporal chains in a crash-proof grammar.* (s. 91–132). <https://doi.org/10.1075/la.144.05eid>
- Eide, K. M. (2016). Finiteness, inflection, and the syntax your morphology can afford. I *Finiteness Matters. On finiteness-related phenomena in natural languages.* (s. s.121-170).
- Eide, K. M. (2022). *Språket som superkraft.* Fagbokforlaget
- Eide og Busterud (2015). «Kompetanse og performanse: kunnskap og produksjon». I Eide, K.M (red.) *Norsk andrespråkssyntaks.* Oslo: Novus forlag
- Field, A. (2013). *Discovering Statistics Using IBM SPSS Statistics.* SAGE.
- Godfroid, A. (2020). *Eye Tracking in Second Language Acquisition and Bilingualism : A Research Synthesis and Methodological Guide [Bok].* Routledge.
- Haraldsen, F. (2020) "Hva er vitenskap?" NTNU
- Hawkins, & Yuet-hung Chan, C. (1997). The partial availability of universal grammar in secondlanguage acquisition: the 'failed functional features hypothesis' *Second Language Research, 13(3), 187–226.* <https://doi.org/10.1191/026765897671476153>
- Hyltenstam, K., & Abrahamsson, N. (2003). Maturational Constraints in SLA. I *The Handbook of Second Language Acquisition* (s. 538–588). John Wiley & Sons, Ltd. <https://doi.org/10.1002/9780470756492.ch17>
- Hüttermann, S., Memmert, D., Simons, D. J., & Bock, O. (2013). Fixation Strategy Influences the Ability to Focus Attention on Two Spatially Separate Objects. *PLoS ONE, 8(6), e65673.* <https://doi.org/10.1371/journal.pone.0065673>
- Jin,F. og Eide, K.M. (2015) «Tilegnelse av nominale kategorier» I Eide, K.M (red.) *Norsk andrespråkssyntaks.* Oslo: Novus forlag
- Jin,F., Eide, K.M. og Busterud,G. (2015) «Nyere andrespråksteorier» I Eide, K.M (red.) *Norsk andrespråkssyntaks.* Oslo: Novus forlag
- Johnson, J. S., & Newport, E. L. (1989). Critical period effects in second language learning: The influence of maturational state on the acquisition of English as a second language. *Cognitive Psychology, 21(1), 60–99.* [https://doi.org/10.1016/0010-0285\(89\)90003-0](https://doi.org/10.1016/0010-0285(89)90003-0)

- Jouravlev, O., Mineroff, Z., Blank, I. A., & Fedorenko, E. (2021). The Small and Efficient Language Network of Polyglots and Hyper-polyglots. *Cerebral Cortex*, *31*(1), 62–76. <https://doi.org/10.1093/cercor/bhaa205>
- Just, M. A., & Carpenter, P. A. (1980). A theory of reading: From eye fixations to comprehension. *Psychological Review*, *87*, 329–354. <https://doi.org/10.1037/0033-295X.87.4.329>
- Liversedge, S. P., Paterson, K. B., & Pickering, M. J. (1998). Chapter 3—Eye Movements and Measures of Reading Time. I G. Underwood (Red.), *Eye Guidance in Reading and Scene Perception* (s. 55–75). Elsevier Science Ltd. <https://doi.org/10.1016/B978-008043361-5/50004-3>
- Moran. (2003). Arguments for rejecting the sequential Bonferroni in ecological studies. *Oikos*, *100*(2), 403–405. <https://doi.org/10.1034/j.1600-0706.2003.12010.x>
- 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:102499682873>
- Penfield, W., & Roberts, L. (1959). *Speech and brain mechanisms*. Princeton University Press.
- Pulvermüller, F., & Schumann, J. H. (1994). Neurobiological Mechanisms of Language Acquisition. *Language Learning*, *44*(4), 681–734. <https://doi.org/10.1111/j.1467-1770.1994.tb00635.x>
- Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological Bulletin*, *124*(3), 372-422. <https://doi.org/10.1037/0033-2909.124.3.37277>
- Rayner. (2009). Eye movements and attention in reading, scene perception, and visual search. *Quarterly Journal of Experimental Psychology (2006)*, *62*(8), 1457–1506. <https://doi.org/10.1080/17470210902816461>
- 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>

- Repeated-Measures ANOVA*. (u.å.). Hentet 4. mai 2023, fra
<http://www.statisticslectures.com/topics/repeatedmeasuresanova/>
- Ringdal, K. (2018). *Enhet og mangfold: samfunnsvitenskapelig forskning og kvantitativ metode* (4.utg., p. 562). Fagbokforlaget.
- Roberts. L., & Siyanova-Chanturia, A. (2013). Using eye-tracking to investigate topics in L2 acquisition and L2 processing. *Studies in Second Language Acquisition*, 35, 213–235. doi:10.1017/S0272263112000861
- 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>
- Schmuckler, M. A. (2001). What Is Ecological Validity? A Dimensional Analysis. *Infancy*, 2(4), 419–436. https://doi.org/10.1207/S15327078IN0204_02
- Schwartz, B., & Sprouse, R. (1996). L2 cognitive states and the full transfer/full access hypothesis. *Second Language Research*, 12, 4072. <http://dx.doi.org/10.1177/026765839601200103>
- 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.
- Svartdal, F. (2015). *Psykologiens forskningsmetoder* (4. utg.). Fagbokforlaget.
<https://doi.org/10.1080/01690965.2012.703782>
- Turney, S. (2022, april 29). *Student's t Table (Free Download) | Guide & Examples*. Scribbr. <https://www.scribbr.com/statistics/students-t-table/>
- Vinichenko (2021). *Nordmenns reaksjoner på avvik i innlærernorsk: En eye-tracking undersøkelse av effekter av finitt- og tempusavvik på språkprosessering* [Masteroppgave]. Norges teknisk-naturvitenskapelige universitet
- Wik (2014). *Om tempus og finitthet i norsk som andrespråk* [Masteroppgave]. Norges teknisk-naturvitenskapelige universitet
- Åfarli, T. A. (2015). «UG-basert andrespråksteori og syntaktisk struktur». I Eide, K.M (red.) *Norsk andrespråkssyntaks*. Oslo: Novus forlag

6 Masteroppgavens relevans for læreryrket

I LK20 er det også læreplaner for minoritetsspråklige og de med kort botid. Denne masteroppgaven har undersøkt hvordan andrespråksinnlærere av norsk prosesserer norske verbmorfologiske avvik. Det handler også om hvordan det er å lære seg et nytt språk i et nytt land. I læreplanen sies det at:

Ved å bruke det norske språket aktivt innlemmes elever med kort botid i Norge i kultur og samfunnsliv. Faget skal styrke elevenes evne til kritisk tenkning og ruste dem til å delta i demokratiske prosesser og til å gi uttrykk for tanker, synspunkter og ideer i egne tekster. Det skal også forberede dem på et arbeidsliv som stiller krav til variert kompetanse i lesing, skriving og muntlig kommunikasjon. (Utdanningsdirektoratet, 2020, s.2)

Derfor er det høyst relevant å forske på andrespråksinnlærere av norsk, fordi det gir innsikt i hvordan det er å lære seg nytt språk. Det å kunne lese og forstå det som blir skrevet er nødvendig i et tekstsamfunn i stadig utvikling hvor man må kunne prosessere store mengder tekst i hverdagen, og da er det relevant å forstå hvordan lesing av norsk som andrespråk.

I tillegg til den generelle læringen av språket er det også interessant å ha forsket spesifikt på den grunnleggende ferdigheten lesing ut ifra dette perspektivet. «Å kunne lese i faget norsk for språklige minoriteter med kort botid i Norge er å skape mening i et bredt utvalg av egnede tekster fra nåtid og fortid» (Utdanningsdirektoratet, 2020, s. 4). Lesing og forståelse av verb, som er en helt sentral ordklasse i enhver setning, og å kunne prosessere verbmorfologien er sentralt for å kunne skape og forstå meningen i tekster i nåtid og fortid. Som lærer har man «et særlig ansvar for opplæringen i å kunne lese på norsk» (Utdanningsdirektoratet, 2020, s. 4), og da er det høyst relevant å ha sittet seg inn i hvordan andrespråksinnlærere av norsk prosesserer informasjonen de får fra verb når de leser norske setninger.

Referanse

Utdanningsdirektoratet. (2020). Læreplan i norsk for språklige minoriteter med kort botid i Norge - videregående opplæring (NOR09-04). *Læreplanverket for Kunnskapsløftet 2020*. <https://data.udir.no/kl06/v201906/laereplaner-lk20/NOR09-04.pdf?lang=nob>

