

Temporal recalibrering i audiovisuell persepsjon

Oppfattet samtidighet i audiovisuell talepersepsjon

Navn/kandidatnummer: 10001

PSY2900 Bacheloroppgave i psykologi

NTNU

Fakultet for samfunns og utdanningsvitenskap, institutt for psykologi

30.05.2022

Trondheim

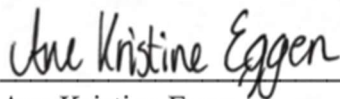
Veileder: Dawn M. Behne

## Foreword

As a starting point for this project, the advisor introduced students to the project's research question and some related issues, together with initial supporting literature. Further literature was identified by the students and shared with the group, and occasionally supplemented by the project advisor. Hypotheses were formulated by the students with supervision, based on the research question and issues presented. Students had the possibility to focus on one or all of the hypotheses in their reports. The experiment was created by the advisor. The students carried out all phases of data collection for the experiment. Data handling was arranged by the advisor and students participated in the process. Statistical analyses and their interpretation were discussed as a group. Students have had the datafile and could run additional/alternative analyses if they chose.

The group had regular seminars, discussions, and close supervision throughout the semester, as well as optional feedback on writing. Students worked as a group to carry out all phases of the project. Literature and materials related to the experiment were stored on a wiki, shared by everyone on the project.

With this basis, each student submits a report (written individually) which has the form and style of a journal article. Students are allowed and encouraged to work together, but the final product must be their own. The report can be in Norwegian or English.



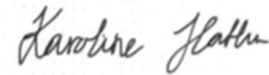
Ane Kristine Eggen

Date: 10.05.2022



Bente Mari Aakvik

Date: 10.05.2022




Karoline Hatlen

Date: 11.05.2022



Angus Wilson

Date: 10.05.2022



Ingvill Holmen Tangen

Date: 10.05.2022



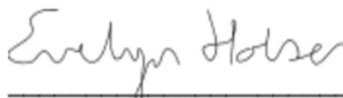
Vegard Dahn

Date: 10.05.2022



Astrid Brøvig Silde

Date: 10.05.2022



Evelyn Holsen

Date: 10.05.2022



Advisor: Dawn M. Behne

Date: 10.05.2022



Benjamin Bornø

Date: 10.05.2022



Linda Marie Leirvik

Date: 11.05.2022



Thea Nordstrøm

Date: 10.05.2022

## Abstract

Tidligere forskning tyder på at det foreligger en temporal rekalkibreringseffekt for synkronitet i audiovisuell persepsjon. Denne effekten er blitt vist å kunne foregå hurtig etter eksponering for kun et enkelt stimuli.

Denne studien tar sikte på å undersøke effekter av rekalkibrering i audiovisuell persepsjon ved å kartlegge hvordan parameterne ALT, VLT og PSS varierer med graden av synkronitet for umiddelbart foregående stimuli i en SJ-oppgave. Studien fant at hurtig temporal rekalkibrering foregikk ved å demonstrere hvordan parameterne ALT, VLT og PSS varierer med varierende grad av audiovisuell asynkronitet for det umiddelbart foregående stimuli. I overensstemmelse med tidligere forskning viste rekalkibreringen seg å foregå i form av negative ettereffekter, der synkronitetsbedømmelse ble forskjøvet mot den spesifikke forsinkelsen.

Rekalkibreringseffekten som ble funnet, viste seg å være større ved videoledende stimuli enn ved audioledende stimuli, noe som gir implikasjoner på at rekalkibrering foregår asymmetrisk. Resultatene støtter opp under tidligere forskning som har foreslått at temporal audiovisuell integrasjon foregår ved ulike mekanismer.

## Temporal rekalkibrering i audiovisuell persepsjon

Persepsjon innebærer oppfattelse av stimuli i sanseorganer og den påfølgende formidlingen og bearbeidelsen av sanseinntrykkene. Hendelser i miljøet oppfattes i mange tilfeller av mer enn ett sansesystem, og prosessering av flere ulike sanseinntrykk fra ulike stimuli er nødvendig for å konstruere en helhetlig forståelse av miljøet en befinner seg i (Stein & Meredith, 1993). Forskning har demonstrert menneskers evne til å vurdere forløp i hendelser på tvers av ulike sanser (Spence et al., 2001), og behandlingen av multisensorisk informasjon fra ulike modaliteter virker å avhenge av amodale egenskaper ved sanseinformasjonen (Keetels & Vroomen, 2012). Sansestimuli fra syn og hørsel er, i større grad enn stimuli fra de kjemiske sansene, i stand til å levere spatial og temporal informasjon om hendelser i miljøet. Stimuliene fra de to modalitetene deler amodale egenskaper, og interaksjon mellom sansesystemene kan resultere i mer nøyaktig og helhetlig persepsjon (Bulkin & Groh, 2006).

Lys og lyd beveger seg med ulik hastighet. For at de auditive og visuelle aspektene ved en hendelse skal kunne oppfattes temporalt nøyaktig, er det en forutsetning at det eksisterer nevrologiske mekanismer som korrigerer oppfatningen av stimuli fra ulike modaliteter. Hjernens behandler auditive og visuelle stimuli på ulike måter, og følgende også med ulik prosesseringshastighet (Harris et al., 2010). De nevralt sammenhengene som danner forutsetning for multimodal interaksjon er ikke fullstendig kartlagt, men hos pattedyr antas aktivitet i bestemte områder i mellomhjernen å ha en sentral rolle i prosessen (King & Palmer, 1985; Meredith et al., 1987). I et samspill med både kortikale og subkortikale strukturer virker Superior Colliculus, en struktur i mellomhjernen der visuell, auditorisk og somatosensorisk informasjon integreres i forkant av motoriske reaksjoner, å spille en viktig rolle i temporal integrering av informasjon (Engel et al., 1999).

Meredith et al. fant at aktivitet i Superior Colliculus økte drastisk når auditive og visuelle stimuli forekom i temporal og spatial nærhet (Meredith et al., 1987). I en studie som undersøkte nevralt responser i superior colliculus hos marsvin ved eksponering for visuelle og auditive stimuli, observerte King og Palmer at aktivitet hos ulike nevroner var avhengig av amodale egenskaper ved stimuliene. Ved auditive og visuelle stimuli presentert separat, viste noen av nevronene seg å respondere kun til en modalitet, andre til begge modalitetene, og andre igjen til ingen av modalitetene. Ved presentasjon av bimodale stimuli fremviste mange av nevronene ulike former for «kryss-modal» interaksjon. Interaksjonen viste seg å være

avhengig av den relative temporale og spatiale posisjonen til stimuliene (King & Palmer, 1985).

Interaksjonen mellom auditive og visuelle stimuli er blitt foreslått å foregå avhengig av ulike forhold, der de amodale egenskapene ved stimuli er felles. Ved spatial integrasjon antas stimuli som kommer fra samme posisjon å ha samme opphav, og ved erfaringsmessig integrasjon antas stimuli som tidligere har vært assosiert med hverandre å høre sammen (Bulkin & Groh, 2006). En tredje form for multimodal interaksjon, temporal integrasjon, skyldes at ulike stimuli ankommer hjernen samtidig, noe som resulterer i at stimuliene antas å skyldes samme hendelse; stimuli kan sies å ha «temporalt sammentreff» (Radeau, 1994). Integrasjon av multisensorisk informasjon avhenger av en viss toleranse for asynkronitet mellom stimuliene fra de ulike modalitetene. Tidsspennet for der audiovisuell stimuli oppfattes for å være synkront ved temporal integrasjon, kalles «the temporal binding window» (det temporale integrasjonsvinduet) (Colonius & Diederich, 2004).

Det temporale integrasjonsvinduet har vist seg å være smalere og ikke formbart for audioledende stimuli, i motsetning til å være både videre og trenbart for visuelt ledende stimuli (Powers et al., 2009). Resultater fra persepsjonseksperimenter, der deltagerne skal avgjøre hvorvidt auditivt og visuelt stimuli er synkroner, «synchrony judgement task» (SJ-oppgave), viser at punktet som oppfattes som maksimalt synkront, «point of subjective simultaneity» (PSS), som regel befinner seg mer mot videoledende side; det temporale integrasjonsvinduet kan sies å være forskjøvet mot videoledende side (Keetels & Vroomen, 2012). Dette innebærer at audiovisuelt stimuli som oppfattes som synkront i realiteten er visuelt ledende. Keetels & Vroomen angir en mulig forklaring på fenomenet som at auditive stimuli prosesseres hurtigere enn visuelle stimuli, noe som medfører at visuelle stimuli må presenteres tidligere for å kompensere for forsinkelsen. En annen mulig forklaring for fenomenet baserer seg på at naturlige forekommende audiovisuelle stimuli oftere er videoledende siden lys beveger seg hurtigere enn lyd (2012).

Cecere et al. fant ved et eksperiment med SJ-oppgaver, basert på designet etter treningsparadigmet til Powers et al. (2009), at trening av synkronitetsbedømmelse der en modalitet var ledende, enten audioledende eller videoledende, ikke påvirket oppfatning av synkronitet for stimuli der den andre modaliteten var ledende. Dette kan tyde på at audiovisuell temporal binding for audioledende og videoledende stimuli er uavhengige av hverandre og mulig styrt av ulike mekanismer (Cecere et al., 2016). Alm og Behne fant at personer mellom 50 og 60 år viste lavere toleranse for audioledende stimuli i en SJ-oppgave,

og at PSS ble forskjøvet mer hos disse deltagerne enn hos yngre deltagere. En mulig forklaring på funnene ble foreslått å være at naturlig forekommende audioledende stimuli er mer forutsigbart enn naturlig forekommende videoledende stimuli, noe som kan gjøre terskelverdien for når et audiovisuelt stimuli oppfattes som audioledende (ALT) sårbar for audiovisuell erfaring. I studien fantes ikke audiovisuell erfaring å ha noen slik effekt på terskelen for når et audiovisuelt stimuli oppfattes som videoledende (VLT) (Alm & Behne, 2013).

Tidsmessig oppfatning av audiovisuelle hendelser har vist seg å være fleksibel ved at det foregår en tilpasning i hva som oppfattes som synkront, ved en effekt kalt temporal rekalkibrering (Fujisaki et al., 2004; Vroomen et al., 2004). Både Fujiaki et al. og Vroomen et al. demonstrerte igjennom to studier i 2004, uavhengige a hverandre, at eksponering for audiovisuelt stimuli med en bestemt tidsmessig forsinkelse mellom visuell og auditiv stimulus, over flere minutter, resulterte i en forskyvning i bedømmelse av synkronitet mot den bestemte forsinkelsen. Fujisaki et al. fant ved SJ-eksperimenter at PSS for stimuli som etterfulgte audioledende stimuli ble mer audioledende, og likedan ble PSS som etterfulgte videledende stimuli også mer videoledende(2004). Vroomen et al. viste den samme effekten på PSS med bruk av en «temporal order judgement» (TOJ)- oppgave, der deltagerne bedømmer rekkefølgen på de presenterte modalitetene (2004). Liknende effekt, der rekalkibreringen er demonstrert som en tilpasningsprosess som foregår over tid, er senere blitt funnet i flere studier (Hanson et al., 2008; Keetels & Vroomen, 2007; Navarra et al., 2007; Navarra et al., 2005; Roseboom & Arnold, 2011).

Temporal rekalkibrering er blitt observert å kunne foregå uten behov for at det foregår adaptasjon over tid (Van der Burg et al., 2013; Van der Burg & Goodbourn, 2015). I et eksperiment med SJ-oppgaver fant Van der Burg et al. at rekalkibrering ved asynkron multisensorisk stimuli også kan forekomme etter eksponering av kun et enkelt stimuli. Her ble synkronitetsbedømmelse ved en «stimulus onset asynchrony» (SOA) påvirket av det umiddelbart foregående SOA (SOA-1). I studien argumenteres dette for å tyde på at mekanismen bak rekalkibrering foregår på sensorisk nivå og ikke inngår i høyere kognitiv prosessering (2013). I en liknende studie av Van der Burg og Goodbourn ble denne hurtige formen for rekalkibrering demonstrert å foregå til tross for at deltagerne var klare over at de ulike modalitetene åpenbart hadde forskjellige opphav, noe som igjen kan argumenteres for å tyde på at hurtig rekalkibrering foregår på sensorisk nivå (2015). Van der burg et al. demonstrerte i studien fra 2013 at videoledende stimuli resulterte i større rekalkibreringseffekt

enn audioledende stimuli. Rekalibreringseffekten har også vist seg å være asymmetrisk ved at den foregår over et videre tidsspenn ved videoledende SOA-1 enn ved audioledende SOA-1 (Van der Burg et al., 2013).

Sensorisk rekalibrering har vist seg å kunne produsere to ulike effekter på oppfatningen av et stimuli; man kan skille mellom negative ettereffekter og positive ettereffekter. Repetert eksponering av et bestemt multimodalt stimuli resulterer i negative «ettereffekter», der ettereffekten innbefatter at SOA oppfattes å være mindre likt de foregående stimuliene. Ved enkle eksponeringer av et multimodalt stimuli vil neste stimuli resultere i positive ettereffekter der SOA vil oppfattes å likne mer på SOA-1 (Roseboom, 2019). For audiovisuell temporal persepsjon har derimot både repeterte eksponeringer over tid og enkle eksponeringer, henholdsvis resulterende i effektene omtalt som temporal og hurtig temporal -rekalibrering, vist seg å gi negative ettereffekter i SJ-oppgaver. Roseboom fant at eksponering for enkle audiovisuelle stimuli, resulterte i negative ettereffekter i en SJ-oppgave, men positive ettereffekter i to andre persepsjonsoppgaver: «magnitude estimation judgement» og «temporal order judgement», der deltagerne henholdsvis anslår graden av synkronitet for SOA, og bedømmer rekkefølgen på auditivt og visuelt stimuli. Roseboom foreslår at disse funnene må sees for å støtte opp under en forståelse av audiovisuell persepsjon der negative ettereffekter skyldes mekanismer som foregår på sensorisk nivå, og der positive ettereffekter oppstår som en konsekvens av høyere kognitive prosesser (Roseboom, 2019).

Det er blitt rettet kritikk mot kriteriene som typisk er blitt lagt til grunn for å tolke resultater i studier som tar for seg temporal persepsjon. Parametere som ALT, VLT og PSS som anvendes i forskning, ekstraheres avhengig av kurver som tilvirkes utfra datamaterialet. Responsene fra SJ og TOJ -oppgaver er gjerne utgangspunkt for analyse der sentraltendens beregnes og brukes for å representere det PSS. Yarrow bemerker at det ikke er gitt at antagelsen om at hvis PSS ikke sammenfaller med det faktiske punktet for synkronitet så skyldes dette asynkrone stimuli, stemmer. Effekten kan like gjerne skyldes variasjon i kriteriene som ligger til grunn for å bedømme asynkrone og synkrone stimuli. Yarrow påpeker at bruken av en normalfordelt kurve, som har vært vanlig, kan være problematisk der bruken ikke er begrunnet. Parameteren PSS mister noe av sin validitet ved at den er avhengig av en kurvetilpasning til dataen som ikke er begrunnet. (Yarrow et al., 2011).

## Studiets hensikt

Denne studien tar sikte på å undersøke effekter av recalibrering i audiovisuell persepsjon ved å kartlegge hvordan parameterne ALT, VLT og PSS varierer med graden av synkronitet for umiddelbart foregående stimuli i en SJ-oppgave. Temporal recalibrering er blitt vist å forekomme hurtig ved at eksponering for et enkelt audiovisuelt stimuli resulterer i effekt på synkronitetsbedømmelsen av neste stimuli (Van der Burg et al., 2013). Dersom en slik hurtig recalibreringseffekt foreligger, forventes det å observere at parameterne ALT, VLT og PSS, avhenger av graden av asynkronitet i SOA-1. Videre har hurtig temporal recalibrering blitt observert å resultere i negative ettereffekter, der et stimuli vurderes som ulikt foregående stimuli, for SJ-oppgaver (Roseboom, 2019). Ved negativ recalibrering i eksperimentet, forventes parameterne, ALT, VLT og PSS, å finnes å være mer videoledende med mer videoledende SOA-1, og mer audioledende ved audioledende SOA-1. En slik effekt vil stemme overens med en forståelse av audiovisuell persepsjon der ulike former for audiovisuell integrasjon eksisterer og må forstås for å foregå ved ulike mekanismer (2019).

Forskning tyder på at videoledende stimuli resulterer i videoledende forskyving av PSS og resulterer i større recalibreringseffekt sammenliknet med audioledende stimuli. (Alm & Behne, 2013; Van der Burg et al., 2013) På bakgrunn av dette forventes det å observere en recalibreringseffekt der PSS finnes å være mer videoledende ved videoledende SOA-1 enn ved synkron og audioledende SOA-1. På liknende vis forventes VLT å være mer videoledende ved videoledende SOA-1 enn ved synkron SOA-1. Van der burg et al. fant at det var en asymmetri i recalibreringseffekten, ved at recalibrering foregikk over et videre tidsspenn for videoledende SOA-1 enn ved audioledende SOA-1 (Van der Burg et al., 2013). Dersom man legger til grunn en sterk asymmetri i en eventuell recalibreringseffekt, forventes det i ytterste konsekvens å ikke observere noen forskjell i ALT på tvers av SOA-1.



## Metode

### Design

Med hensikt å undersøke effekten erfaring fra umiddelbart foregående audiovisuelt stimuli har på oppfattelsen av synkronitet for påfølgende audiovisuelt stimuli, ble det gjennomført et persepsjonseksperiment med gjennomføringer av en SJ-oppgave. Ordlyden /ba/, uttalt av en kvinne, utgjorde auditivt stimulus, og ble presentert sammen med en video av kvinnen som uttalte ordlyden, som utgjorde visuell stimulus. Oppgaven bestod i at deltagerne ved eksperimentet avgjorde hvorvidt auditiv stimulus og visuell stimulus fremstod som synkront eller ikke, i en rekke audiovisuelle stimuli med varierende grad av audiovisuell synkronitet. Utfra responsdataen ble terskelverdier for parameterne ALT, VLT og PSS ekstrahert. Effekten det foregående presenterte audiovisuelle stimuluset (SOA-1) hadde på avgjørelse av synkronitet for det påfølgende stimuluset (SOA), ble undersøkt med et «within subjects»-design der uavhengig «within subjects» variabel var SOA-1, delt i tre nivåer, og avhengige variabler var parameterne ALT, VLT, og PSS.

### Deltagere

Trettien personer i alderen 20 til 28 år, med norsk som morsmål, ble rekruttert til deltagelse i studien. Rekrutteringen foregikk ved Norges Tekniske og Naturvitenskapelige Universitet (NTNU), i Trondheim i Norge. Deltagelse i eksperimentet forutsatte til at deltagerne innfridde forutbestemte kriterier satt til synsskarphet, hørsel og hånddominans. Disse kriteriene ble kontrollert for ved gjennomføring av pretester forut for eksperimentgjennomføring.

Deltagernes hånddominans ble kartlagt med bruk av Edinburgh handedness Inventory (Oldfield, 1971). Deltagerne besvarte spørsmål i et skjema, der de redegjorde for hvilken hånd de foretrakk å bruke ved gjennomføring av ulike oppgaver, og en lateralitetskvotient for hver deltager ble beregnet ut fra svarene. En forutsetning for deltagelse i eksperimentet var at deltageren ble vurdert som høyrehendt. Deltagernes hørsel ble vurdert ved en hørselsundersøkelse gjennomført i et lydisolert kammer etter modell av standard audiometrisk

prosedyre fra British Society of Audiology (2018). Kun deltagere som viste en terskel for oppfattelse av lyd under 20 dB, ved frekvenser fra og med 250 til og med 4000 Hz, på begge ører, ble inkludert i eksperimentet. Deltagernes synsskarphet ble undersøkt og vurdert utfra resultater ved gjennomføring av en Snellen-test (Tsui & Patel, 2020), gjennomført på en iMac med 21,5 tommer skjerm med oppløsning 1920 x 1080 piksler og grafikkort av typen ATI Radeon HD 5670 512 MG graphics. Størrelsen på Snellen-testen var 13,9 cm x 9,8 cm. Deltagelse i eksperimentet forutsatte en synsskarphet på 20/25. Som en del av pretestingen ble deltagerne også undersøkt for øyedominans, men resultatene fra undersøkelsen ble ikke vurdert som kriterier for ekskludering fra eksperimentet.

En deltager ble ekskludert forut for gjennomføring av persepsjonseksperimentet som følge av ikke innfridde krav satt til synsskarphet. Det selekterte utvalget ( $N = 30$ ) bestod av 8 menn, 21 kvinner og 1 deltager av ikke oppgitt kjønn, i alderen 20 - 28 år ( $M = 23.03$ ,  $SD = 1.63$ ). Etter gjennomføring av eksperimentet ble data fra ytterligere 1 deltager ekskludert som følge av at avgitte responser fra SJ-oppgaven ikke ville være anvendbare i senere dataanalyse (se reformatering). Samtlige deltagere skrev under på et samtykkeskjema i forkant av eksperimentet.

## Materiale

Stimuliene anvendt i persepsjonseksperimentet ble produsert ved talelaben ved psykologisk institutt, NTNU, opprinnelig til bruk i en studie av Magnus Alm og Dawn Behne (2013). Den audiovisuelle stimulien som ble anvendt i SJ-oppgaven ble produsert ut fra et opptak av en kvinne med norsk østlandsdialekt som uttalte ordlyden /ba/. Opptaket ble gjort i et lydisolert rom og kvinnen ble instruert til å uttale stavelsen med en flat intonasjon og å unngå ansiktsbevegelser. Den opprinnelige studien opptakene var tilsiktet å produsere stimuli for, inkluderte også stavelserne /da/ og /ga/. Hver av stavelserne ble repetert ti ganger, der 1 gjennomføring for hver stavelse ble selektert for eksperimentell bruk.

Videopptak ble gjort med et kamera av typen PDWF800 Sony Professional XDCAM HD422 Camcorder, plassert 2 meter unna kvinnen. Videofilene som ble produsert på MPEG-4 format hadde en bildefrekvens på 30 fps, og oppløsning på 1920 x 1200 piksler. To stk. Røde NT1-A mikrofoner, plassert i knehøyde foran kvinnen, tok opp lyd. Den ene mikrofonen var koblet til kameraet, den andre til en datamaskin av typen Apple Macintosh G5 igjennom et lydkort av typen RME FIREFACE 400. Datamaskinen tok opp lyd fra de to

kildene med opptaksfrekvens på 48 kHz, i programmet Praat version 5.1 (Boersma and Weenink, 2009).

Videoopptaket, med tilhørende lydopptak fra mikrofonen koblet til kameraet, ble delt i 1400 ms lange videofiler som hver omfattet en uttalt stavelse, med programmet AVID Media Composer 3.5. Audiofilene med lydopptak fra den eksterne mikrofonen ble på liknende måte delt opp i segmenter som hver omfattet en stavelse, med programmet Praat version 5.1 (Boersma & Weenink, 2009). Etter en systematisk seleksjonsprosess, der videofilene ble vurdert etter kriterier relatert til auditiv og visuell kvalitet, opp mot tilsiktet bruk, ble en segmentert videofil selektert. Auditive stimuli /ba/, i lydfilen tilhørende det selekterte videosegmentet, hadde en varighet på 404 ms. Lydopptaket fra den eksterne mikrofonen ble synkronisert med lydopptaket fra mikrofonen tilkoblet kameraet i programmet Logic Pro 8.0.2.

Audiovisuelt stimuli ble konstruert ved at lydopptaket fra mikrofonen tilkoblet kameraet ble erstattet med det korresponderende lydopptaket fra den eksterne mikrofonen for det selekterte videoklippet i AVID Media Composer. Det auditive segmentet ble justert trinnvis relativt til det visuelle segmentet for å konstruere asynkrone audiovisuelle stimuli der begynnelsen på auditivt stimuli differerte med 40 ms. Eksperimentet i studien til Alm og Behne (2013) anvendte stimuli fra og med 440 ms audioledende SOA til og med 440 ms videoledende SOA. Spennet for audiovisuelle stimuli er i det her aktuelle eksperimentet er redusert til å omfatte fra og med 400 ms audioledende SOA til og med 400 ms videoledende SOA. Totalt 21 ulike audiovisuelle stimuli, der 10 var auditivt ledende, 1 var synkron og 10 var visuelt ledende ble tatt i bruk i persepsjonseksperimentet.

## Prosedyre

Persepsjonseksperimentet ble gjennomført i talelaben ved psykologisk institutt, NTNU. Opptil to deltagere gjennomførte i samme rom samtidig, da på hver sin side av rommet og vendt vekk fra hverandre, for å begrense distraksjoner i størst mulig grad. Deltagerne fikk benytte seg av synskorrigerende briller eller linser dersom de brukte dette til vanlig. Deltagerne ble plassert i stoler slik at avstanden til dataskjermen var 70 cm. Eksperimentet ble gjennomført på en iMac med en 27 tommer stor monitor med oppløsning på 5120 x 2880 piksler og grafikkort av typen AMD Radeon R9 M295X 4GB. Lysstyrken på skjermen var justert til nivå 12/16. Lyd ble presentert igjennom studiohodetelefoner av typen

AKG K273. Programmet Superlab 6.2 ble brukt til å presentere SJ-oppgaven og å samle responsdata. Responsene fra SJ-oppgaven ble registrert med responsbokser av typene Cedrus RB-730 og Cedrus RB-740. Deltagerne ble instruert til å bruke pekefinger fra hver hånd for å trykke på to ulike knapper markert «synk» for synkrone stimuli, og «asynk» for asynkrone stimuli.

SJ-oppgaven bestod av 3 blokker som hver omfattet 2 pauser, med henholdsvis 144, 153 og 154 trials. De 21 ulike avpassingene av audiovisuelle stimuli ble arrangert (randomisert i matlab med et «inhouse»-script) slik at hvert audiovisuelle stimuli, presentert som SOA, etterfulgte samtlige andre stimuli, presentert som SOA-1, en enkelt gang i løpet av en hel eksperimentgjennomføring. Dette utgjorde til sammen 441 ulike kombinasjoner av SOA og SOA-1 som gav grunnlag for datamateriale. Første SOA etter hvert avbrekk hadde ingen umiddelbart foregående SOA-1 og responsen fra SJ-oppgaven for dette audiovisuelle stimuliet inngikk dermed ikke i datainnsamlingen. Etterfølgende hvert avbrekk ble derfor det siste presenterte SOA fra før avbrekket repetert som første stimuli, for å inngå som SOA-1 til neste SOA. En eksperimentgjennomføring bestod dermed til sammen av 450 AV-stimuli, der 441 av disse bidro til data som inngikk i datainnsamling. Deltagerne fikk også innledende presentert 4 eksempelstimuli for å bli kjent med oppgaven. Hver enkelt deltager gjennomførte en egen versjon av eksperimentet der kombinasjonen av SOA og SOA-1 var randomisert på en unik måte. Randomiseringen av stimuliene, og organiseringen av de 3 delene med 2 pauser hver, ble foretatt i matlab\_R2021\_b ved bruk av to ulike script som hver ble produsert for disse formålene.

## Resultater

### Reformatering

Responsene fra SJ-oppgaven i Superlab ble samlet inn i sin randomiserte form som rådata, fordelt på 3 datafiler for hver enkelt deltager. Deltagerresponsene ble undersøkt for gjennomgående avvik fra normalen, med hensikt å kartlegge anvendbarheten av dataen. 1 deltager måtte ekskluderes etter senere omformatering av data som følge av urimelige responser, der det ikke ville være mulig å ekstrahere ut brukbare terskelverdier for parameterne som inngikk i videre analyse. Utvalget som dataanalysen tok utgangspunkt i bestod dermed av responsene til 29 deltagere. De 3 datafilene, som hver omfattet 1/3 av responsene til en deltager, ble samlet til en enkelt fil per deltager, og omformatert i et «inhouse» Matlab script. Responsdataen for hver enkelt SOA ble derandomisert og arrangert etter grad av faktisk audiovisuell synkronitet for SOA-1. Synkront stimuli tilsvarer nullpunktet for asynkront stimuli, definert som 0 ms videoledende stimuli, dette innebærer at audioledende stimuli benevnes som videoledende stimuli med negative måltall.

Graden av faktisk audiovisuell synkronitet for SOA-1 ble delt inn i grupperingene audioledende SOA-1, tilnærmet synkron SOA-1 og videoledende SOA-1, som hver spente over et 240 ms intervall fordelt på 7 AV-stimuli, henholdsvis: [-400 ms, -160 ms]; [-120 ms, 120 ms]; og [160 ms, 400 ms]. Innenfor hver gruppering ble prosentvis synkronitet for deltagerresponsene beregnet for hver enkelt SOA, til sammen 21 SOA per deltager, for å fremstille to Sigmoid-kurver. Disse ble benyttet til å ekstrahere terskelverdier for parameterne ALT, VLT, og PSS innenfor hver SOA-1 gruppering. To ulike verdier for PSS ble ekstrahert: PSS-average, som representerer midtpunktet på en gausskurve; og PSS-cross, som representerer krysningspunktet mellom to sigmoidkurver, uavhengige av hverandre. PSS-cross ble konstruert ved at dataen for ALT og VLT ble tilpasset sigmoidkurvene. Etter reformatering og ekstraksjon av terskelverdier for de ulike parameterne ble datamateriale fra hver enkelt deltager samlet i en enkelt Excel-fil der de var gjenstand for videre dataanalyse.

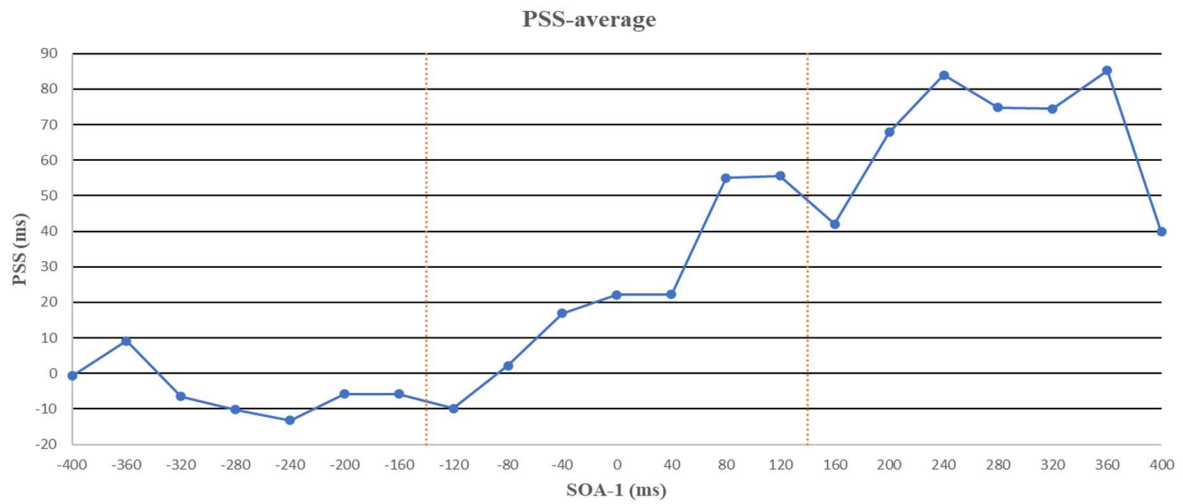
### Analyse

Det samlede datamateriale med de ekstraherte terskelverdiene for de ulike parameterne ble analysert i IBM statistics 27. Hensikten med analysen var å undersøke eventuelle effekter av rekalkibrering i audiovisuell persepsjon ved å se hvordan parameterne ALT, VLT og PSS

varierte med ulike SOA-1. Et sekundært mål med analysen var å undersøke validiteten i definisjonen av parameteren PSS i forskning på audiovisuell persepsjon, derav ble to mål for PSS inkludert i dataanalysen. Det ble gjennomført fire separate enveis repeated measure ANOVA analyser for å sammenlikne effekten de ulike SOA-1 grupperingene hadde på terskelverdiene for parameterne ALT, VLT og PSS. Den uavhengige «within-subjects»-variabelen SOA-1 var delt i tre nivåer: audioledende SOA-1; tilnærmet synkron SOA-1; og videoledende SOA-1. De avhengige variablene var parameterne ALT, VLT, PSS-average og PSS cross.

Figur 1

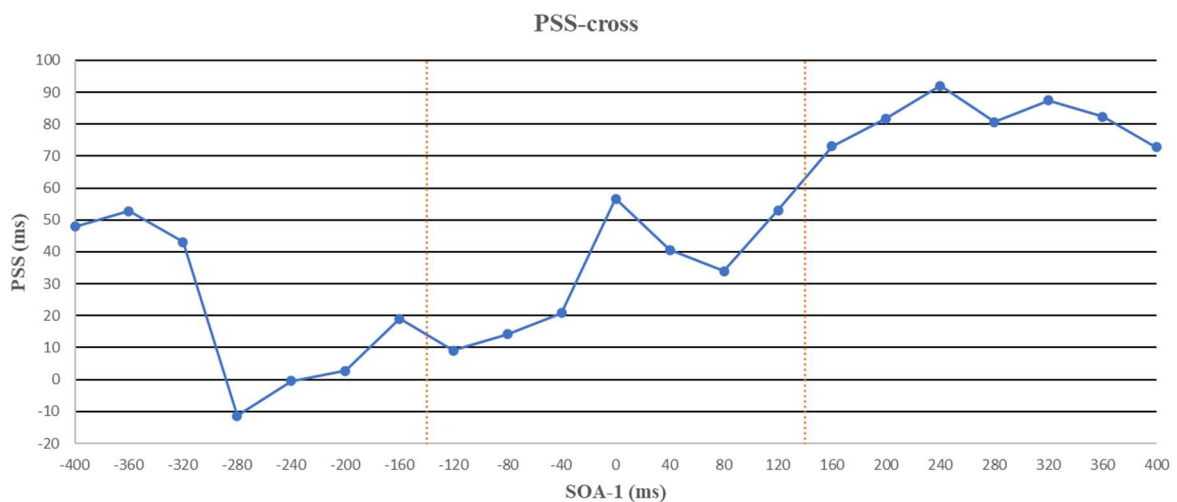
Tendens for PSS-average ved SOA-1



*Notat.* Gjennomsnittsverdier basert på hele utvalget, N = 29, for PSS-average ved SOA-1 målt i ms videolevende stimuli. Loddrette havstrukne linjer ved -140 ms videolevende SOA og 140 ms videolevende SOA markerer skille mellom «within subjects»-grupperingene.

Figur 2

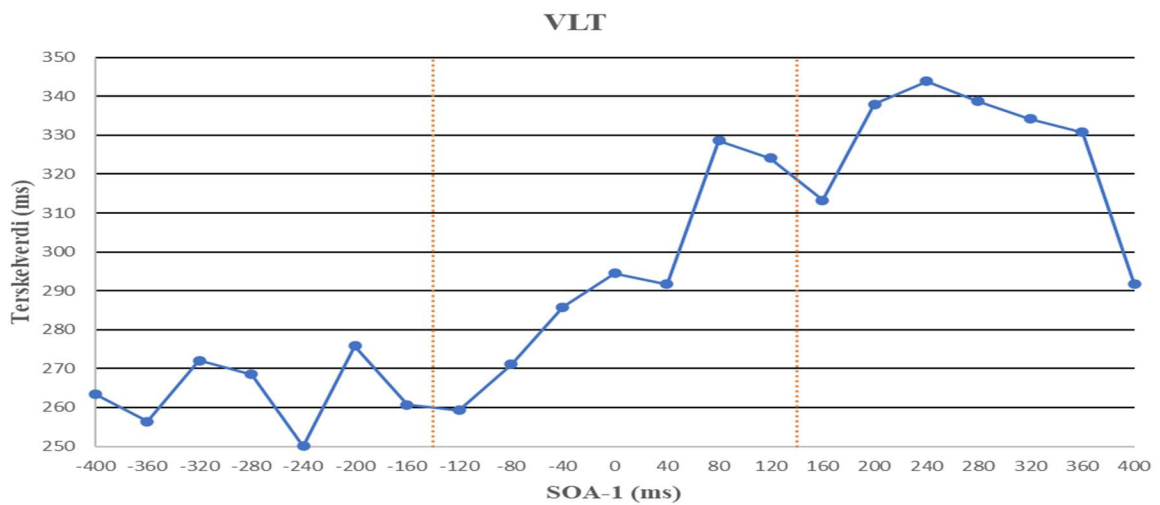
Tendens for PSS-Cross ved SOA-1



*Notat.* Gjennomsnittsverdier basert på hele utvalget, N = 29, for PSS-cross ved SOA-1 målt i ms videolevende stimuli. Loddrette havstrukne linjer ved -140 ms videolevende SOA og 140 ms videolevende SOA markerer skille mellom «within subjects»-grupperingene.

Figur 3

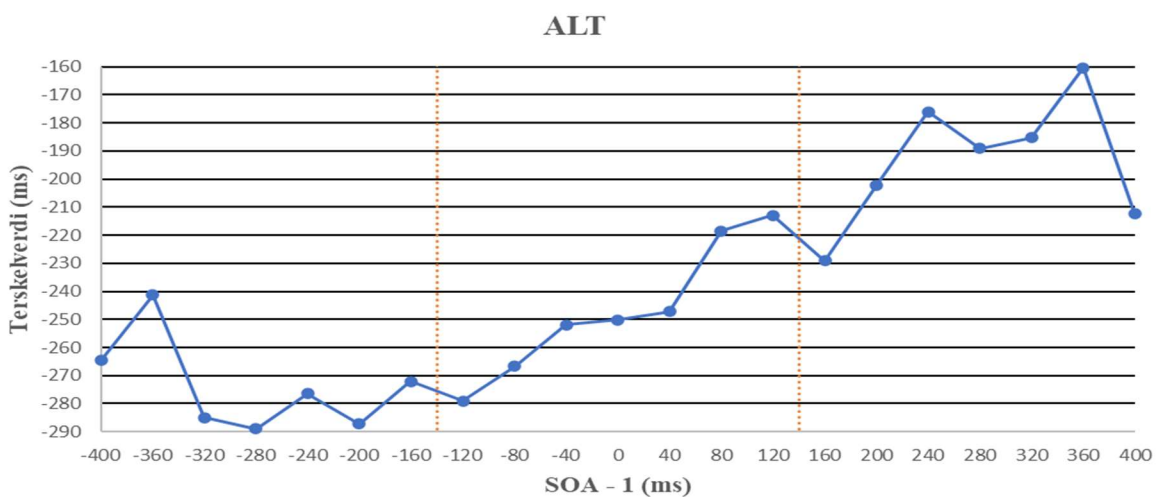
*Tendens for VLT ved SOA-1*



*Notat.* Gjennomsnittlige terskler basert på hele utvalget, N = 29, for VLT ved SOA-1 målt i ms videoledende stimuli. Loddrette havstrukne linjer ved -140 ms videoledende SOA og 140 ms videoledende SOA markerer skille mellom «within subjects»-grupperingene.

Figur 4

*Tendens for ALT ved SOA-1*



*Notat.* Gjennomsnittlige terskler basert på hele utvalget, N = 29, for VLT ved SOA-1 målt i ms videoledende stimuli. Loddrette havstrukne linjer ved -140 ms videoledende SOA og 140 ms videoledende SOA markerer skille mellom «within subjects»-grupperingene.



### PSS-average

Mauchley's test for sphericity indikerte at antagelsen om sfærisitet mellom SOA-1 grupperingene var brutt,  $\chi^2(2) = 11.15, p = .004$ . Ved frihetsgrader justert etter Greenhouse Geisser, ( $\epsilon = .75$ ), ble det funnet å være en signifikant forskjell mellom minst to av de ulike SOA-1 grupperingene,  $F(1.49, 41.85) = 22.11, p < .001$ . En post hoc pairwise comparison justert med Bonferroni-korreksjon utført på SOA-1 grupperingene, viste at det forelå signifikant forskjell,  $\Delta M = -18.06, p = .003$ , mellom audioledende SOA-1,  $M = -10$  ms,  $SD = 64$ , og tilnærmet synkron SOA-1,  $M = 8$  ms,  $SD = 54$ . Det ble også funnet å være en signifikant forskjell i PSS-verdi,  $\Delta M = -17.82, p = .001$ , mellom tilnærmet synkron SOA-1,  $M = 8$  ms,  $SD = 54$ , og videoledende SOA-1,  $M = 26$  ms,  $SD = 44$  ms. Det forelå signifikant forskjell,  $\Delta M = -35.89$  ms,  $p < .001$ , i PSS-verdi mellom audioledende SOA-1,  $M = -10$  ms,  $SD = 64$  ms, og videoledende SOA-1,  $M = 26$  ms,  $SD = 44$  ms. Tendensen i resultatene viste at gjennomsnittsverdien for PSS-average gikk fra å være audioledende ved audioledende SOA-1, til å bli mer videoledende med mer videoledende SOA-1 gruppering (Se figur 1.).

### PSS-cross

Mauchley's test for sphericity indikerte at antagelsen om sfærisitet mellom SOA-1 grupperingene bestod,  $\chi^2(2) = 1.37, p = .504$ . Det ble ikke funnet å være noen signifikant forskjell mellom de ulike SOA-1 grupperingene,  $F(2, 56) = 0.92, p = .403$ . En post hoc pairwise comparison justert med Bonferroni-korreksjon utført på SOA-1 grupperingene, viste ingen signifikant forskjell i PSS-verdi,  $\Delta M = -12.01, p = .881$ , mellom audioledende SOA-1,  $M = 21$  ms,  $SD = 69$  ms, og tilnærmet synkron SOA-1,  $M = 33$  ms,  $SD = 53$  ms. Forskjell i PSS-verdi var ikke signifikant,  $\Delta M = -3.51, p = 1.000$ , mellom tilnærmet synkron SOA-1,  $M = 33$  ms,  $SD = 53$  ms, og videoledende SOA-1,  $M = 36$  ms,  $SD = 47$  ms. Forskjell i PSS-verdi var ikke signifikant,  $\Delta M = -15.57, p = .753$ , mellom audioledende SOA-1,  $M = 21$  ms,  $SD = 69$  ms, og videoledende SOA-1,  $M = 36$  ms,  $SD = 47$  ms. Post hoc test ble gjennomført til tross for at det ikke ble funnet noen signifikant effekt av SOA-1 på de avhengige parameterne i en repeated measure ANOVA. Hensikten var å avdekke eventuelle tendenser som kunne bidra til å belyse virkningene av, og vurdere holdbarheten til, de ulike måtene å beregne PSS. Post hoc test viste, en tendens der gjennomsnittsverdien for PSS-cross var mer videoledende med mer videoledende SOA-1 (se figur 2.).

## VLT

Mauchley's test for sphericity indikerte at antagelsen om sfærisitet mellom SOA-1 grupperingene var brutt,  $\chi^2(2) = 13.16, p = .001$ . Ved frihetsgrader justert etter Greenhouse Geisser, ( $\epsilon = .72$ ), ble det funnet å være en signifikant forskjell mellom minst to av de ulike SOA-1 grupperingene,  $F(1.44, 40.41) = 16.77, p < .001$ . En post hoc pairwise comparison justert med Bonferroni-korreksjon utført på SOA-1 grupperingene, viste at det forelå signifikant forskjell i terskelverdi for VLT,  $\Delta M = -13.48, p = .015$ , mellom audioledende SOA-1,  $M = 246$  ms,  $SD = 81$  ms, og tilnærmet synkron SOA-1,  $M = 260$  ms,  $SD = 74$  ms. Det ble funnet å være signifikant forskjell i terskelverdi for VLT,  $\Delta M = -16.61, p = .001$ , mellom tilnærmet synkron SOA-1,  $M = 260$  ms,  $SD = 74$  ms, og videoledende SOA-1,  $M = 277$  ms,  $SD = 68$  ms. Det var signifikant forskjell i terskelverdi for VLT,  $\Delta M = -30.08, p < .001$ , mellom audioledende SOA-1,  $M = 246$  ms,  $SD = 81$  ms, og videoledende SOA-1,  $M = 277$  ms,  $SD = 68$  ms. Tendensen var at gjennomsnittsverdien for VLT var mer videoledende med mer videoledende SOA-1 gruppering (se figur 3.).

## ALT

Mauchley's test for sphericity indikerte at antagelsen om sfærisitet mellom SOA-1 grupperingene var brutt,  $\chi^2(2) = 9.38, p = .009$ . Ved frihetsgrader justert etter Greenhouse Geisser, ( $\epsilon = .77$ ), ble det funnet å være en signifikant forskjell mellom minst to av de ulike SOA-1 grupperingene,  $F(1.55, 43.30) = 15.26, p < .001$ . En post hoc pairwise comparison justert med Bonferroni-korreksjon utført på SOA-1 grupperingene, viste at det forelå signifikant forskjell i ALT-verdi,  $\Delta M = -23.51, p = .013$ , mellom audioledende SOA-1,  $M = -266$  ms,  $SD = 109$  ms, og tilnærmet synkron SOA-1,  $M = -243$  ms,  $SD = 86$  ms. Det ble også funnet å være en signifikant forskjell i terskelverdi for ALT,  $\Delta M = -19.04, p = .008$ , mellom tilnærmet synkron SOA-1,  $M = -243$  ms,  $SD = 86$  ms, og videoledende SOA-1,  $M = -224$  ms,  $SD = 74$  ms. Det ble funnet en signifikant forskjell i ALT verdi,  $\Delta M = -42.55, p < .001$ , mellom audioledende SOA-1,  $M = -266$  ms,  $SD = 109$  ms, og videoledende SOA-1,  $M = -224$  ms,  $SD = 74$  ms. Tendensen fra resultatene var at gjennomsnittsverdien for ALT ble mindre audioledende ved mer videoledende SOA-1 (se figur 4.).

## Oppsummering

PSS-average ble funnet å være signifikant mer videoledende ved videoledende SOA-1 enn ved både tilnærmet synkron, og audioledende -SOA-1. PSS-cross ble på liknende funnet å være mer videoledende ved videoledende SOA-1 enn ved både tilnærmet synkron, og audioledende -SOA-1, men forskjellene mellom SOA-1 grupperingene var mindre enn for PSS-average, og ikke signifikante. VLT ble funnet å være signifikant mer videoledende ved videoledende SOA-1 enn ved tilnærmet synkron SOA-1. ALT ved tilnærmet synkron SOA-1 varierte signifikant fra både ALT ved audioledende og videoledende -SOA-1. For samtlige av de avhengige parameterne gjaldt det at terskelverdiene var mer videoledende ved mer videoledende SOA-1, og mer audioledende ved mer audioledende SOA-1.

## Drøfting

Tidligere forskning har demonstrert effekten av temporal rekalkibrering i audiovisuell persepsjon (Fujisaki et al., 2004; Vroomen et al., 2004), både der den foregår i form av positive ettereffekter etter repetert eksponering for et bestemt audiovisuelt stimuli, og der den forekommer som negative ettereffekter ved eksponering for enkle stimuli i SJ-oppgaver (Roseboom, 2019; Van der Burg et al., 2013). I denne studien ble effekter av slik hurtig temporal rekalkibrering undersøkt ved å se på hvordan parameterne ALT, VLT og PSS varierte avhengig av hvorvidt SOA-1 var audioledende ([-400 ms, -160 ms]), tilnærmet synkron ([-120 ms, 120 ms]) eller videloedende ([160 ms, 400 ms]). Dersom det foregikk hurtig temporal rekalkibrering i SJ-eksperimentet var det forventet å finne at parameterne varierte med SOA-1 (Van der Burg et al., 2013). Det var videre forventet å finne at rekalkibreringen som eventuelt fant sted, resulterte i negative ettereffekter der parameterne ville finnes å være mer videoledende med mer videoledende SOA-1, og mer audioledende ved mer audioledende SOA-1 (Roseboom, 2019). En slik effekt stemmer overens med en forståelse av audiovisuell persepsjon som innebærer at det eksisterer ulike mekanismer for temporal rekalkibrering avhengig av adaptasjon har foregått over tid eller ikke.

## Hurtig temporal rekalkibrering

ALT viste seg, tross antagelsen om ingen variasjon, å variere med SOA-1,  $F(1.55, 43.30) = 15.26, p < .001$ . Dette vitner om at vurderingene som ble gjort rundt parameteren ble

basert på misantatte forutsetninger. Det temporale integrasjonsvinduet har vist seg å være smalere, og mindre formbart for audioledende SOA-1 (Powers et al., 2009; Van der Burg et al., 2013), men dette medfører ikke i at rekalkibrering ikke foregår for parameteren ALT. Som forventet varierte VLT med graden av synkronitet for SOA-1,  $F(1.44, 40.41) = 16.77, p < .001$ . Det samme gjorde PSS-average definert som midtpunktet på en normalfordelt kurve tilpasset datamaterialet,  $F(1.49, 41.85) = 22.11, p < .001$ . PSS definert som krysningspunktet mellom to, hver for seg uavhengige sigmoid kurver, ble funnet å ikke være signifikant,  $F(2, 56) = 0.92, p = .403$ . Effekten i variasjonen var dessuten relativt minimal relativ til funnene ved PSS-average. Dette kan vitne om at datamaterialet ikke var riktig tilpasset analysen som ble gjort. Alt i alt viser funnene at parameterne ALT, VLT og PSS-average varierer med varierende grad av asynkronitet for SOA-1, noe som støtter opp under hypotesen om at hurtig negativ rekalkibrering foregår ved audiovisuell persepsjon.

Tendensen i resultatene viser at samtlige parametere blir mer videoledende mer mer videoledende SOA-1. Disse funnene samsvarer med tidligere forskning som legger til grunn en forståelse av audiovisuell persepsjon der hurtig temporal rekalkibrering foregår som et resultat av mekanismer som virker på et lavere sensorisk nivå (Roseboom, 2019; Van der Burg et al., 2013; Van der Burg & Goodbourn, 2015). Roseboom foreslår at ettereffekter, slik som funnet i dette eksperimentet, må forstås som å foregå på et sensorisk nivå, og positive ettereffekter å foregå som et resultat av høyere kognitive prosesser (2019). Denne forståelsen går overens med Van der Burg sin antakelse om at hurtig rekalkibrering nødvendigvis er et resultat av sensoriske mekanismer (Van der Burg et al., 2013). Eksistensen av ulike mekanismer for temporal integrering av audiovisuelle stimuli, avhengig av hvorvidt det har foregått tilpasning til stimuli over tid, kan sees å ha likhetstrekk med Cecere sine funn som antyder at persepsjon av visuelt ledende og auditivt ledende stimuli behandles uavhengig av hverandre, noe som igjen kan indikere at det eksisterer ulike mekanismer for temporal integrering av audiovisuell stimuli (Cecere et al., 2016). En konsekvens av cecere sine funn, at audioledende og visuelle stimuli er uavhengige, medfører at en normalfordelt kurve vil være lite hensiktsmessig ved analyse av datamateriale som kommer ut av persepsjonseksperimenter. Dersom det foreligger ulike mekanismer for rekalkibrering, resulterer dette i større usikkerhet rundt kurvetilpasning og parameterfremstilling (Yarrow et al., 2011).

Antagelsen om at PSS ville være mer videoledende ved videoledende SOA-1 enn ved synkron og audioledende SOA-1 ble vist å stemme dersom definisjonen av PSS var basert på

en gaussisk kurve. Der PSS ble definert som krysningspunktet mellom to uavhengige sigmoidkurver ble effekten av SOA-1 funnet å være ikke signifikant. Årsaken til dette følger nødvendigvis av kriteriene som legges til grunn for definisjonene av parameterne. Der PSS-average er basert på en normalfordelt kurve, tilpasset alt datamaterialet under ett, er PSS-cross basert på to sigmoidkurver, uavhengige av hverandre, men avhengige av parameterne ALT og VLT. Datagrunnlaget som effektstørrelse og signifikans baseres på er følgende forskjellig for PSS-cross og PSS-average, noe som i dataanalysen har resultert i vidt forskjellige resultater. Sigmoidkurvene som ble konstruert og tatt i bruk i dataanalysen skulle tentativt innebære at parameterne som avhenger av kurveutformingen hadde større nærhet til virkeligheten. Yarrow argumenterer for at det er for mye usikkerhet knyttet til ekstrahering av paramtere som inngår i forskning på persepsjon (Yarrow et al., 2011). De sprikende resultatene for PSS, avhengig av kurven parameteren baseres på, i denne studien, illustrerer Yarrow sitt poeng.

## Konklusjon

Studien fant at hurtig temporal rekalkibrering foregår i audiovisuell persepsjon ved å demonstrere hvordan parameterne ALT, VLT og PSS varierer med varierende grad av audiovisuell asynkronitet for SOA-1 i en SJ-oppgave. I overensstemmelse med tidligere forskningsfunn viste rekalkibreringen seg å foregå i form av negative ettereffekter, der synkronitetsbedømmelse i SJ-oppgaven ble forskjøvet mot forsinkelsen mellom modalitetene. Rekalkibreringseffekten som ble funnet viste seg å være større ved videoledende stimuli enn ved audioledende stimuli, noe som gir implikasjoner på at rekalkibrering foregår asymmetrisk. Resultatene støtter opp under tidligere forskning som har foreslått at temporal audiovisuell integrasjon foregår ved ulike mekanismer.

## Litteratur

- Redaksjonen for norsk APA-stil. (2021). Norsk APA-manual: En nasjonal standard for norskspråklig APA-stil basert på APA 7th (Versjon 1.7). Unit.  
<https://www.unit.no/tjenester/norsk-apa-referansestil>
- Boersma, P., and Weenink, D. (2009). "PRAAT: Doing phonetics by computer (Version 5.1) (Computer program)," [www.praat.org](http://www.praat.org)
- Alm, M., & Behne, D. (2013). Audio-visual speech experience with age influences perceived audio-visual asynchrony in speech. *The Journal of the Acoustical Society of America*, 134(4), 3001-3010. <https://doi.org/10.1121/1.4820798>
- British Society of Audiology. (2018). Recommended procedure: Pure tone air and bone conduction threshold audiometry with and without masking  
<https://www.thebsa.org.uk/wp-content/uploads/2018/11/OD104-32-Recommended-Procedure-Pure-Tone-Audiometry-August-2018-FINAL.pdf>
- Bulkin, D. A., & Groh, J. M. (2006). Seeing sounds: visual and auditory interactions in the brain. *Curr Opin Neurobiol*, 16(4), 415-419.  
<https://doi.org/10.1016/j.conb.2006.06.008>
- Cecere, R., Gross, J., & Thut, G. (2016). Behavioural evidence for separate mechanisms of audiovisual temporal binding as a function of leading sensory modality. *European Journal of Neuroscience*, 43(12), 1561-1568.  
<https://doi.org/https://doi.org/10.1111/ejn.13242>
- Colonus, H., & Diederich, A. (2004). Multisensory interaction in saccadic reaction time: a time-window-of-integration model. *J Cogn Neurosci*, 16(6), 1000-1009.  
<https://doi.org/10.1162/0898929041502733>
- Engel, A. K., Fries, P., König, P., Brecht, M., & Singer, W. (1999). Temporal binding, binocular rivalry, and consciousness. *Conscious Cogn*, 8(2), 128-151.  
<https://doi.org/10.1006/ccog.1999.0389>
- Fujisaki, W., Shimojo, S., Kashino, M., & Nishida, S. (2004). Recalibration of audiovisual simultaneity. *Nat Neurosci*, 7(7), 773-778. <https://doi.org/10.1038/nm1268>
- Hanson, J. V., Heron, J., & Whitaker, D. (2008). Recalibration of perceived time across sensory modalities. *Exp Brain Res*, 185(2), 347-352. <https://doi.org/10.1007/s00221-008-1282-3>
- Harris, L., Harrar, V., JAeK, P., & Kopinska, A. (2010). Mechanisms of simultaneity constancy. *Space and time in perception and action*, 232-253.
- Keetels, M., & Vroomen, J. (2007). No effect of auditory-visual spatial disparity on temporal recalibration. *Exp Brain Res*, 182(4), 559-565. <https://doi.org/10.1007/s00221-007-1012-2>
- Keetels, M., & Vroomen, J. (2012). Frontiers in Neuroscience Perception of Synchrony between the Senses. In M. M. Murray & M. T. Wallace (Eds.), *The Neural Bases of Multisensory Processes*. CRC Press/Taylor & Francis Copyright © 2012 by Taylor & Francis Group, LLC.
- King, A. J., & Palmer, A. R. (1985). Integration of visual and auditory information in bimodal neurones in the guinea-pig superior colliculus. *Exp Brain Res*, 60(3), 492-500.  
<https://doi.org/10.1007/bf00236934>

- Meredith, M. A., Nemitz, J. W., & Stein, B. E. (1987). Determinants of multisensory integration in superior colliculus neurons. I. Temporal factors. *J Neurosci*, 7(10), 3215-3229. <https://doi.org/10.1523/jneurosci.07-10-03215.1987>
- Navarra, J., Soto-Faraco, S., & Spence, C. (2007). Adaptation to audiotactile asynchrony. *Neurosci Lett*, 413(1), 72-76. <https://doi.org/10.1016/j.neulet.2006.11.027>
- Navarra, J., Vatakis, A., Zampini, M., Soto-Faraco, S., Humphreys, W., & Spence, C. (2005). Exposure to asynchronous audiovisual speech extends the temporal window for audiovisual integration. *Brain Res Cogn Brain Res*, 25(2), 499-507. <https://doi.org/10.1016/j.cogbrainres.2005.07.009>
- Oldfield, R. C. (1971). The assessment and analysis of handedness: the Edinburgh inventory. *Neuropsychologia*, 9(1), 97-113. [https://doi.org/10.1016/0028-3932\(71\)90067-4](https://doi.org/10.1016/0028-3932(71)90067-4)
- Powers, A. R., 3rd, Hillock, A. R., & Wallace, M. T. (2009). Perceptual training narrows the temporal window of multisensory binding. *J Neurosci*, 29(39), 12265-12274. <https://doi.org/10.1523/jneurosci.3501-09.2009>
- Radeau, M. (1994). Auditory-visual spatial interaction and modularity. *Curr Psychol Cogn*, 13(1), 3-51. <https://doi.org/10.1007/bf02686854>
- Roseboom, W. (2019). Serial dependence in timing perception. *Journal of Experimental Psychology: Human Perception and Performance*, 45(1), 100-110. <https://doi.org/10.1037/xhp0000591>
- Roseboom, W., & Arnold, D. H. (2011). Twice upon a time: multiple concurrent temporal recalibrations of audiovisual speech. *Psychol Sci*, 22(7), 872-877. <https://doi.org/10.1177/0956797611413293>
- Spence, C., Shore, D. I., & Klein, R. M. (2001). Multisensory prior entry. *J Exp Psychol Gen*, 130(4), 799-832. <https://doi.org/10.1037//0096-3445.130.4.799>
- Stein, B. E., & Meredith, M. A. (1993). *The merging of the senses*. The MIT press.
- Tsui, E., & Patel, P. (2020). Calculated Decisions: Visual acuity testing (Snellen chart). *Emerg Med Pract*, 22(4), Cd1-cd2.
- Van der Burg, E., Alais, D., & Cass, J. (2013). Rapid recalibration to audiovisual asynchrony. *J Neurosci*, 33(37), 14633-14637. <https://doi.org/10.1523/jneurosci.1182-13.2013>
- Van der Burg, E., & Goodbourn, P. T. (2015). Rapid, generalized adaptation to asynchronous audiovisual speech. *Proc Biol Sci*, 282(1804), 20143083. <https://doi.org/10.1098/rspb.2014.3083>
- Vroomen, J., Keetels, M., de Gelder, B., & Bertelson, P. (2004). Recalibration of temporal order perception by exposure to audio-visual asynchrony. *Brain Res Cogn Brain Res*, 22(1), 32-35. <https://doi.org/10.1016/j.cogbrainres.2004.07.003>
- Yarrow, K., Jahn, N., Durant, S., & Arnold, D. H. (2011). Shifts of criteria or neural timing? The assumptions underlying timing perception studies. *Conscious Cogn*, 20(4), 1518-1531. <https://doi.org/10.1016/j.concog.2011.07.003>