

Kandidatnummer: 10098

# Rekalibrering og oppfattet samtidighet av audiovisuelle stimuli i SJ-oppgaver

BA08 - Oppfattet samtidighet i audiovisuell  
talepersepsjon

Bacheloroppgave i PSY2900 Bacheloroppgave i psykologi

Veileder: Dawn M. Behne

Mai 2022



Kandidatnummer: 10098

# **Rekalibrering og oppfattet samtidighet av audiovisuelle stimuli i SJ-oppgaver**

BA08 - Oppfattet samtidighet i audiovisuell  
talepersepsjon

Bacheloroppgave i PSY2900 Bacheloroppgave i psykologi  
Veileder: Dawn M. Behne  
Mai 2022

Norges teknisk-naturvitenskapelige universitet  
Fakultet for samfunns- og utdanningsvitenskap  
Institutt for psykologi



Kunnskap for en bedre verden




## Foreword

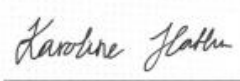
As a starting point for this project, the advisor introduced students to the project's research question and some related issues, together with initial supporting literature. Further literature was identified by the students and shared with the group, and occasionally supplemented by the project advisor. 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.

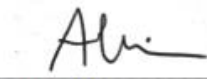
The group had regular seminars, discussions, and close supervision throughout the semester, as well as optional feedback on writing. Students worked as a group to carry out all phases of the project. Literature and materials related to the experiment were stored on a wiki, shared by everyone on the project.


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

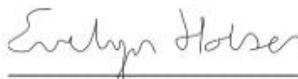
  
Karoline Hatlen  
Date: 11.05.2022


  
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

## **Acknowledgements**

Jeg vil takke mine medstudenter for et hyggelig semester sammen i bachelorgruppa. Det har vært til stor hjelp å ha en samlet gruppa som har jobbet godt sammen for å fullføre dette prosjektet. En stor takk til Dawn M. Behne for de utallige timene hun har lagt ned for at dette prosjektet gikk såpass knirkefritt som det gjorde. Vil også takke Darren Rhodes og Peter Svensson for deres arbeid med datamaterialet vårt.

Studiet trengte ikke å bli registrert av NSD.

## Sammendrag

Denne studien har sett på hvordan forandringer i SOA vil kunne føre til endringer i oppfatningen av audiovisuelle talestimuli gjennom flere eksponeringer av forskjellige SOA-nivåer. Tidligere forskning har sett på hvordan PSS påvirkes av recalibrering gjennom bayesisk perspektiv, der det blir en kumulativ effekt av flere stimuli, eller en temporal recalibrering der det er utelukkende den nyligste stimulusen som har påvirkningskraft (Roseboom, 2019; Van der Burg et al., 2013). Fokuset til denne studien var å se på hvordan ALT og VLT vil forandres gjennom en temporal recalibrering på en SJ-oppgave, og hvordan dette hang sammen med forandringer i FWHM (Full Width Half Maximum) og PSS som er funnet i Roseboom (2019) og Van der Burg et al. (2013). «Simultaneity judgement task» (SJ) med talestimuli (/ba) ble brukt for å måle oppfattelsen av samtidighet. To S-kurver blir satt sammen til en gaussisk kurv. Denne grafiske framstillingen av dataen har støtte fra tidligere forskning som antyder at denne framstillingen av dataen kan være mer presis enn gaussiske kurver vanligvis er i slike studier. Resultatene fra dette studiet gir også implikasjoner på at S-kurver kan skape mer nøyaktighet for målinger i SJ-oppgaver. Funnene i dette studiet tyder på en negativ ettervirkning for video-ledende stimuli, en mindre konservativ ALT enn forventet og viser viktigheten av hvordan man måler PSS.

## **Rekalibrering og oppfattet samtidighet av audiovisuelle stimuli i SJ-oppgaver**

Mennesker oppfatter verden rundt seg gjennom persepsjonen man skaper gjennom sansene. Hver for seg kan sansene gi deg bruddstykker av informasjon, men gjennom å integrere disse bruddstykkene skapes et helhetlig bilde. Multisensorisk persepsjon er denne samhandlingen mellom sansene. Sansene samhandler gjennom synkronisering og integrering, hvor hjernen må oppfatte synkroniteten til de ulike stimuliene for deretter å integrere synkrone sammen til multisensorisk persepsjon (Keetels & Vroomen, 2012). Dette er en forenkling av hva som kan tenkes å skje i multisensorisk persepsjon.

### **Persepsjon**

Å forstå hvordan hjernen utfører multisensorisk persepsjon kan være komplisert. «The binding problem» forsøker å forklare vanskelighetene som oppstår gjennom multisensorisk persepsjon, men har ikke en gitt modell som forklarer dette fenomenet (Feldman, 2013). Hjernen og den intersensoriske integrasjonen bruker mer temporal enn spatial samvariasjon (Bushara et al., 2001). Ofte er det snakk om en temporal sammenfatning/synkronitet mellom ulike hjerneområder som skaper sensorisk multimodalitet. En tidligere studie har funnet en nevralt aktivering i insula som styrer temporal audiovisuell (AV) interaksjon i kortikal prosessering (Bushara et al., 2001). Temporal samhandling er en del av kryssmodalitet, men løser ikke «the binding problem». «Temporal binding» sier ikke noe om hvordan man kan bruke temporal synkronitet, men den sier at «temporal binding» brukes (Keetels & Vroomen, 2012). «Temporal binding» fastsetter bare at sannsynligheten for temporal synkronitet vil øke hvis modaliteter har flere like egenskaper (Keetels & Vroomen, 2012). Dette kalles «Unity Assumption», og er en oppfatning mange forskere har av hvorfor temporal synkronitet oppstår. Dersom modalitetene som oppfattes har flere like egenskaper, vil sannsynligheten for at temporal synkronitet oppstår øke.

### **Synkronitet**

Hvor stort må avviket mellom de ulike perseptene være for at man merker at bildet forstyrres? For å måle samtidighet for perseptene bruker man «stimulus onset asynchronies» (SOA) (Keetels & Vroomen, 2012). Oppfattelsen av samtidighet, som en følge av SOA-nivåer, vil variere fra person til person og er derfor ikke noen universell verdi for hvor samtidighet starter eller slutter. Skillet mellom synkron eller asynkron kalles for «Just-noticeable difference» (JND), og tidligere studier antyder at verdiene ligger fra 25 ms til 80 ms



forsinkelse (Vroomen & Keetels, 2012). Disse verdiene vil variere avhengig av hvilke sanser som oppfatter stimuli, der audiovisuelle stimuli har en JND på 25-50 ms (Keetels & Vroomen, 2005).

Oppfattelsen av samtidighet kan også sees på gjennom «temporal binding window» (TBW). TBW er et teoretisk konsept som viser til tidsintervallet der individer integrerer og binder sammen ulike modaliteters/sansers stimuli (Wallace & Stevenson, 2014). For å regne ut dette vinduet kan man bruke den tekniske fremgangsmåten, «Full Width Half Maximum» (FWHM). Da finner man avstanden mellom «audiovisual treshold» (ALT) og «audiovisual treshold» (VLT), og dette blir i prinsipp samme tidsintervallet som for TBW. Selve vinduet kan variere, men verdiene for TBW er ofte større enn ved JND (Keetels & Vroomen, 2012). Vinduet for oppfatning av synkronitet vil påvirkes av stimuli (Kaganovich & Schumaker, 2016), alder (Noel et al., 2016), oppmerksomhet (Keetels & Vroomen, 2012) og nevrologisk tilstand (Stevenson et al., 2014). Tidligere forskning viser til at visuell-auditiv (VA) stimuli, det vil si stimuli der visuell kommer før auditiv, fører til et større behov for temporal segregering enn AV-stimuli vil behøve for at asynkroniteten vil oppfattes (Bushara et al., 2001).

Kompleksiteten til en stimuli vil også påvirke TBW ved at mer kompleks stimuli trenger mer temporal segregering enn mindre kompleks stimuli (Kaganovich & Schumaker, 2016). Kompleks stimuli som språk skiller seg fra andre typer audiovisuelle stimuli ved at hjernen bruker andre hjerneområder enn andre audiovisuelle stimuli gjør (Noesselt et al., 2005). Dette fører til at dette studiet, som bruker språk som stimuli, skiller seg fra andre studier som bruker for eksempel flash/beep eller andre mindre kompleks stimuli. Tidligere forskning viser til at TBW for språk stimuli ved McGurk-effekten kan gå opp til 300 ms asynkronitet ved video-ledende stimuli, eller 80 ms for audio-ledende stimuli (Vatakis & Spence, 2010; Dixon & Spitz 1980). Her ser man også at TBW for språk stimuli er mer sensitiv for audio-ledende enn den er for video-ledende (Vatakis & Spence, 2010)

Alder vil spille en rolle for TBW ved at størrelsen av vinduet vil variere etter hvor man er i livsløpet, hvorav barn ofte oppfatter stimuli som synkront når det egentlig er asynkront stimuli (Kaganovich & Schumaker, 2016). Eldre mennesker vil oppleve en gradvis svekkelse av sansene, der sansene ikke vil være like skarpe som tidligere i livet. Særlig hørselen blir

svekket med alderen, men det finnes en motvirkning til dette (Begau et al., 2021). Erfaring kan delvis motvirke den nedsatte perseptive evnen, ved multimodale stimuli (Begau et al., 2021). Et eksempel på dette er et cocktail-party scenario, hvor det er mange stimuli fra flere kilder. Eldre mennesker sliter ofte mer i slike situasjoner enn yngre voksne, men dette vil som oftest gjelde dersom den tiltenkte stimuli kommer fra et menneske i bevegelse (Begau et al., 2021). Eldre mennesker og yngre mennesker har ganske lik evne til å skille og gruppere stimuli i et cocktail-party scenario, men dette forutsetter at personen som de skal oppfatte sensoriske stimuli fra står i ro og ikke er i bevegelse (Begau et al., 2021). Hvis tiltenkt stimuli skifter person eller avstand i rommet oppstår det ofte vanskeligheter for eldre mennesker. De sliter med å opprettholde samme nivå av adskillelse for samsvarende stimuli, slik yngre mennesker klarer. Erfaring vil hjelpe med å kunne peile seg inn på en stimulus, dersom det er en familiær stemme og stimulusen har en konstant avstand (Begau et al., 2021). Dette kan være en grunn til funnene som viser at alder ikke gir signifikante utslag for hvordan man oppfatter og filtrer støy i cocktail-party scenario.

Man har lenge visst at oppmerksomhet spiller en rolle for hastigheten av oppfattelsen til en stimulus. Edward B. Titchener forklarte fenomenet allerede i 1908, og er kalt «law of prior entry» (Keetels & Vroomen, 2012). «Law of prior entry» sier at hvis to stimuli blir presentert samtidig vil den stimulusen som er i fokus være den som blir oppfattet først (Apa dictionary, 2022). Forskning med «temporal order judgments» (TOJ) (Spence et al., 2001) og elektroencefalografi (EEG) (Hillyard & Munte, 1984; Eimer & Schröger, 1998) viser at hastigheten av oppfattelsen til en stimulus øker med oppmerksomhet. Påvirkningen av oppmerksomheten vil også variere avhengig av hvilken sans som brukes. Berøring førte til et større nødvendig skifte for å skape synkronitet enn det gjorde for visuell stimulus (Spence et al., 2001). Dette hører til dels sammen med erfaring der man kan øke presisjon av synkronitets oppfatninger gjennom å vite hva man skal ha oppmerksomhet på. Fallgruven er kjent ved at man får en «respons bias» der stimulusen i fokus vil bli oppfattet først i synkronitetstester (Keetels & Vroomen, 2012). Boenke et al. (2009) mente at denne biasen oppstår som følge av selektiv oppmerksomhet ville forminskes dersom lengden på de gitte stimuli ble forlenget. De mente at dersom en temporal rekalkibrering oppstår, er behovet der for en lengre stimuli på en modalitet/sans for å kunne føre til et positivt eller negativt skifte i oppfattelsen rekkefølgen av en multimodal stimulus.

En deltakers nevrologiske tilstand vil også ha en innvirkning på resultatene av synkronitetstester. Tidligere forskning tyder på en sammenheng mellom et forstørret AV vindu for AV stimuli og nevrologiske tilstander som dysleksi (Hairston et al., 2005; Virsu et al., 2003) og schizofreni (de Gelder et al., 2003). Samme gjelder for individer på autisme-spekteret, som ofte har sensoriske hindringer som påvirker språkoppfattelse og synkronitetsoppfattelse (Stevenson et al., 2014), hvorav TBW blir større sammenlignet med andre individer på samme alderstrinn. Større TBW fører til mindre presis «temporal binding» som vil forverres etter økende kompleksitet av stimuli. «Temporal binding» blir mer upresis ved at de oftere oppfatter asynkrone stimuli som synkron (Foss-Feig et al., 2010). Tidligere forskning viser at individer på autisme-spekteret ikke bare får større TBW med kompleks stimuli som språk, men også med enklere stimuli som flash/beep (Stevenson et al., 2014). Dette viser til at deres sosiale ulikheter, som ofte oppstår på grunn av språkproblemer, kan ha grunnlag i deres kognitive evne til å oppfatte stimuli som synkront eller asynkront. Man kan dermed se at menneskers evner til å oppfatte synkronitet er en del av den sosiale tilpasningen for individer. Likheter finner man i case-studien av en deltaker som går under aliaset PH. Han opplevde at auditive stimuli kom før visuelle stimuli da han så på tv en dag. Case-studien viste at han hadde lesjoner i basal ganglia og pons som førte til visuelt lag (Freeman et al., 2013). Han måtte ha auditivt lag på rundt 200 ms for at han skulle oppfatte audiovisuelle stimuli som synkron. Dette er et unikt tilfelle, men viser til at nevrologiske ulikheter kan føre til ekstreme resultater i TOJ tester.

### **Avstandens rolle i oppfattet samtidighet**

Gjennom å se på JND vil man se at sansene oppfatter synkronitet forskjellig, hvorav avstand og intensitet vil ha påvirkning på hastigheten til stimuli (Begau et al., 2021). Avstand fra stimuli vil være meget relevant for hvor fort man kan oppfatte en stimulus (Van der Stoep, 2020). Her kommer «horizon of simultaneity» inn. Dette er avstanden (0-10 meter) hvor lyd vil ankomme øret før audio når retina. Går man over avstanden som befinner seg innenfor «horizon of simultaneity» vil man ofte oppfatte visuell stimulus man ser før auditiv stimulus (Keetels & Vroomen, 2012). Dette kommer av hastigheten til de ulike stimuliene. Gjennom luft har lyd en hastighet på 330 ms, mens lys har en hastighet på 300 000 000 ms (Spence & Squire, 2003). Denne forskjellen i hastighet blir noe jevnet ut ved avstander på 10 meter med at auditive stimuli blir prosessert raskere enn visuelle stimuli. «Horizon of simultaneity» begrenses til 10 meter, men hvis man ser på synkronitet så vil avstanden kunne variere desto mer (Poppel et al., 1990). Dermed kan man oppfatte synkronitet på langt større avstander enn

10 meter. Her kommer erfaringen inn som et hjelpemiddel for å oppfatte synkronitet på lengre avstander. Visuelle språk stimuli, som munnbevegelse, kan hjelpe med oppfattelsen av auditive språk stimuli (Begau et al., 2021). Dette kommer av språklig erfaring, der erfaring til individer kan gjøre at visuelle stimuli kan predikere visuelle stimuli. Dette vil føre til en økning i prosesseringshastighet av auditive stimuli som et resultat av oppmerksomheten. Den økte hastigheten for auditive stimuli vil kunne føre til at man oppfatter en stimuli som synkron selv om avstanden er betydelig lengre enn «horizon of simultaneity». Annen forskning som tar for seg påvirkningen av avstand og intensitet antyder et behov for forskjellige TBW-er avhengig av avstanden mellom stimuli og mottaker (Van der Stoep, 2020; Noel, Serino et al. 2019). En mulig inndeling er mellom «near space» og «far space» hvor man vanligvis skiller mellom disse rundt 130 ms (Noel et al., 2019).

### **Alder**

Alm & Behne (2013) kom fram til at sensitiviteten for audio-ledende stimuli er mer ømfintlig for middelaldrene og eldre voksne enn for yngre voksne. De foreslo at dette kom av erfaringsbasert læring av bindingen mellom audiovisuelle stimuli. Når hørselen blir svekket med alderen vil personer i større grad bruke visuelle stimuli for å predikere auditive stimuli. For mennesker er det ikke bare selve hørselen som påvirkes av alder, men også den kognitive prosesseringshastigheten svekkes (Alm & Behne, 2013). De fant også en signifikant forskjell mellom unge og middelaldrende voksne på ALT, hvor de middelaldrende voksne hadde en mer konservativ ALT. Unge voksne hadde en mer «ekstrem» ALT verdi sammenlignet med middelaldrende voksne, hvor ALT var lengre fra fysisk synkron for unge voksne.

### **Rekalibrering av audiovisuell talepersepsjon**

Det er ikke bare forskjellene i de nåværende stimuli som man må ta hensyn til, men også hvordan tidligere stimuli spiller en rolle for oppfatningen. Rekalibrering er tendensen til å minske forskjellene mellom objekter eller hendelser som vanligvis hører sammen (Keetels & Vroomen, 2012). For audiovisuelle språk stimuli vil dette være å minske forskjellen mellom språk og bevegelse på munnen. Van der Burg et al. (2013) argumenterer for at rekalibrering foregår nesten uten umiddelbart, dermed hurtig rekalibrering. Denne hurtige rekalibreringen er en sensorisk effekt, og ikke en kompleks kognitiv prosess (Van der Burg et al., 2013). Effekten av rekalibrering vil gradvis minke ved nye stimuli som presenteres, men den vil ikke svekkes av tiden mellom stimuli (Machulla et al., 2012). Hurtig rekalibrering deles ofte opp i

to oppfatninger, bayesisk og temporal. I den bayesiske tilnærmingen til rekalkibrering tenker man at alle eksponeringene man har hatt for stimuli vil påvirke oppfatningen av den nåværende stimulusen. Her finner man positive ettervirkninger; tidligere stimuli vil bli oppfattet mer lik den nåværende stimulusen, vil være assimilierende (Roseboom, 2019). Dersom flere konsekvente meget audio-ledende stimuli presenteres, vil den neste stimulusen oppfattes som mer audio-ledende enn den egentlig er (Roseboom, 2019). Disse positive ettervirkningene finner man på høyere kognitive nivåer som tyder på at denne typen tilnærming til rekalkibrering er å foretrekke for mer kompleks stimuli som TOJ og «magnitudo judgments» (MJ) (Roseboom, 2019).

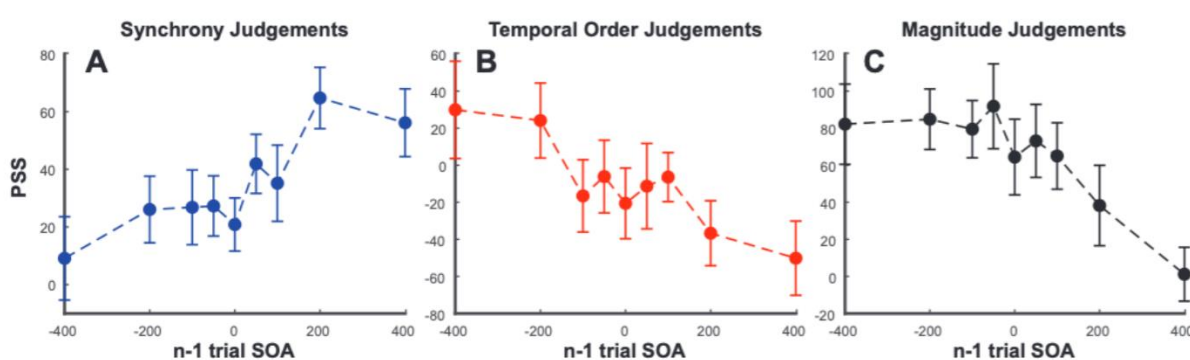


Figure 2. Average point of subjective synchrony (PSS) for 18 participants, estimated from (A) synchrony judgements, (B) temporal order judgements, and (C) magnitude judgements, depending on  $n - 1$  trial audiovisual stimulus-onset asynchrony (SOA). Error bars depict standard error of the mean. See the online article for the color version of this figure.

*Notat.* Bildet er hentet fra Roseboom (2019) og viser funnene i studiet. Modellene viser ettervirkningene av rekalkibreringen for SJ, TOJ og MJ-oppgaver. For SJ-oppgavene ser man en negativ ettervirkning som en følge av rekalkibreringen, mens den får en positiv ettervirkning for TOJ og MJ-oppgavene.

Temporal rekalkibrering er et resultat av hjernens tendens til å justere subjektiv synkronitets oppfattelse, gjennom et skifte i sensorisk prosessering, for å hindre forskjeller mellom multimodale stimuli (Fujisaki et al., 2004). Temporal rekalkibrering fører til negative ettervirkninger for SJ, TOJ og MJ-oppgaver (Roseboom, 2019). Negative ettervirkninger er når en stimulus oppfattes som mindre lik foregående stimuli (Roseboom, 2019). For eksempel så vil en ny stimulus bli oppfattet som mindre audio-ledende hvis SOA-1 var audio-ledende for de foregående stimuli (Roseboom, 2019). Negative ettervirkninger er knyttet opp mot sensorisk adaptasjon der det foregår en rekalkibrering av den sensoriske oppfatningen av

informasjonen. Temporal rekalkibrering foregår dermed på et sensorisk nivå, og ikke på et komplekst kognitivt nivå slik som bayesisk rekalkibrering (Roseboom, 2019).

### **S-kurver og valg av PSS**

Valget av kurvetilpasning er et tema som er blitt mer utforsket de siste årene. Atferds studier har tidligere brukt en normalfordelt kurve, også kjent som en gaussisk kurve (Yarrow et al., 2011). Yarrow et al. kom fram til at den gaussiske kurven har en svakhet med at den ikke brukes på grunn av kompatibiliteten med funnene i studier, men brukes heller som en standard for atferds studier. Dette fører til at PSSaverage som man finner på den gaussiske kurven kan mangle nøyaktigheten som en målemetode for det høyeste punktet av synkronitet (Yarrow et al., 2011). Av den grunn kan det være gunstigere å bruke to S-kurver istedenfor en gaussisk kurve for SJ-oppgaver (Yarrow et al., 2011).

Bruken av to S-kurver åpner muligheten for en ny tilnærming til PSS. Den utbredte måten å måle PSS på er gjennom å ta FWHM og dele den på to, for å finne midtpunktet (Keetels & Vroomen, 2012). Dette vil teoretisk sett være det høyeste punktet for oppfattelse av samtidighet. Om dette er den beste måten å gjøre det i praksis på er et annet spørsmål. En annen måte å finne PSS på, er å velge å bruke PSS som en verdi av skjæringspunktet mellom de to S-kurvene. Verdien av PSS vil ikke bestemmes av FWHM (ALT + VLT), som kan føre til et mer nøyaktig mål for PSS siden ALT og VLT ikke har en direkte sammenheng med akkurat hvor oppfattelsen for synkronitet er mest nøyaktig.

### **Studiets formål**

Studiet fokuserer på hvordan endringer i SOA vil kunne føre til endringer i oppfatningen av synkronitet i audiovisuelle talestimuli som en følge av rekalkibrering. Tidligere forskning har sett på hvordan PSS påvirkes av rekalkibrering gjennom bayesisk perspektiv, en kumulativ effekt av flere stimuli, eller en temporal rekalkibrering der den nyligste stimulusen har tilnærmet all påvirkningskraft (Roseboom, 2019; Van der Burg et al., 2013). Fokuset er på hvordan ALT og VLT påvirkes av temporal rekalkibrering av en SJ-oppgave, sammen med forandringer i FWHM og PSS som er funnet i Roseboom (2019) og Van der Burg et al. (2013). «Simultaneity judgement task» (SJ) med talestimuli (/ba) blir brukt for å måle

oppfattelsen av samtidighet, der valget av SJ-oppgaver kommer av at studiet prioriterer den umiddelbare oppfatningen av talestimulien som presenteres. Med dette vil man finne den sensoriske oppfatningen av synkronitet med å begrense tiden personen har til å svare. I likhet med funnen i Van der Burg et al. (2013) så forventer studiet at når SOA-1 er video-ledende blir PSS mer video-ledende sammenlignet med synkron og audio-ledende SOA-1.. Studiet forventer også at når SOA-1 er video-ledende blir VLT mer video-ledende sammenlignet med synkron SOA-1. Dette kommer av at studiet forventer en negativ ettervirkning i SJ-oppgaver i likhet med andre atferds studier også har funnet (Roseboom, 2019). Til slutt forventes ingen forskjell mellom ALT ved SOA-1 som er audio-ledende og synkron SOA-1. Dette henger sammen med tidligere funn som tyder på en mer konservative ALT (Alm & Behne, 2013; Vatakis et al., 2007; Vatakis & Spence, 2010).

Formålet med studiet er å se om man får lignende resultater som Roseboom (2019) og Van der Burg et al. (2013) om man fokuserer på VLT og ALT i større grad enn PSS. De brukte PSS for å vise signifikante skifter, mens dette studiet skal forklare endringer i ALT og VLT som en følge av temporal recalibrering i SJ-oppgaver.

## *Metode*

### **Design**

SJ-oppgave responser fra unge voksne på oppfattelse av audiovisuelle stimuli som audio-ledende, synkront eller video-ledende, hvor man hører stavelsen /ba/ på forskjellige SOA-nivåer, blir sammenlignet med hverandre i forhold til recalibrering som ble gjort på den forrige stimulusen.

30 norsktalende deltakere i en alder mellom 20 og 30 år deltok i eksperimentet. Kriteriene som ble satt for syn var å kunne utføre Snellen testen og ha et monokulært syn på 20/25. Snellen-testen ble utført på en 21,5-in. iMac med en oppløsning på 1920 x 1080. Selve størrelsen på Snellen testen var på 13,9 cm x 9,8 cm og avstanden fra individet til skjermen var på 70cm. For hørsel var kriteriene å ha en hørselsgrense på 20db på frekvensene 250, 500, 1000, 2000 og 4000 Hz for begge ørene. Selve hørselstesten ble utført i samsvar med en standard ren tone audiometrisk prosedyre i samsvar med «British Society of Audiology» (2004) (Alm & Behne, 2013).

En poweranalyse ble kjørt før eksperimentene startet. Ved et gitt alfa-nivå på .05, tre inndelinger og 147 målepunkter fikk man medium effekt ved et utvalg på seks deltakere. Tidligere studier hadde 20 deltakere (Roseboom, 2019) og 24 deltakere (Van der Burg et al., 2013), derfor var tanken at dette studiet også burde ha tilnærmet likt antall. Studiet endte dermed opp med 29 deltakere som innfrir både kravene fra poweranalysen og tidligere studiers utvalg.

### **Deltakere**

Deltakere som ikke innfridde disse kriteriene, ble ekskludert fra eksperimentet. Tre deltakere ble ekskludert, én på grunn av språkkravet; norsk som morsmål. Én annen ble ekskludert med hensyn til kriteriet av alder (19 år) og til slutt ble en kvinne ekskludert fra eksperimentet på grunn av at hun ikke innfridde kriteriene for syn. Det ene øyet var godkjent synsskarphet, men det andre øyet innfridde ikke kravet om monokulær synsskarphet på 20/25.

Av de 30 deltakerne var 21 av disse kvinner, 8 menn og 1 oppga ikke kjønn. Gjennomsnittsalderen ( $M$ ) var på 23 år blant deltakerne som gjennomførte eksperimentet, mens aldersintervallet var 20 til 28 år. Standardavviket ( $SD$ ) ble tilnærmet lik 2. Studiet rekrutterte deltakere på NTNU Dragvoll gjennom en qr-kode. 42 individer meldte seg på gjennom qr-koden, hvorav 30 endte med å være på eksperimentet.

Flere oppga seg selv som musikere i spørreskjemaet, 10 til sammen. Dette er ikke noe som blir sett på i eksperimentet, men kan være en forklaring for mulige avvik i resultatene (Alm & Behne, 2013). Ingen av deltagerne svarte at de hadde hatt et inntak av alkohol de siste 24 timene, mens én svarte at han/henne hadde en nevrologisk sykdom (ADHD). Dette førte ikke til ekskludering, men muligheten er der for at dette har hatt en innvirkning på denne deltakerens resultater. Et spørsmål var hvor mange timer deltakerne brukte på videospill i løpet av en uke. Grunnen for dette spørsmålet er at individer med utbredt erfaring fra «action» videospill kan ha en bedre evne til å oppfatte og integrere multimodale stimuli som audiovisuelle stimuli (Donohue et al., 2010).



## **Materiell**

Stimuli som ble brukt var stavelen /ba/ på SOA-nivåer mellom -400 og 400 ms. De audiovisuelle opptakene ble gjort på NTNU, Speech Laboratory på Department of Psychology (Alm & Behne, 2013). På opptakene var det en ung kvinne som snakket med en lett gjenkjennelig østlandsdialekt (Oslo). Opptakene minsket distraksjonene ved å fjerne briller eller smykker, ha jevn og flatt toneleie og minimalt med ansiktsmimikk på den unge kvinnen.

Kvinnen i opptaket satt i et sete som var to meter unna kameraet, PDWF800 Sony Professional XDCAM HD422 Camcorder (Tokyo, Japan) (Alm & Behne, 2013). Rommet var lydisolert og to røde NT1-A mikrofoner ble plassert foran henne i knehøyde, hvorav den ene var koblet til kameraet og den andre var koblet til en RME FIREFACE 400 (Haimhausen, Germany). Denne var igjen koblet til en Apple Macintosh G5 datamaskin (Cupertino, CA), hvor to audiokanaler tok opp sekvensen med en «sampling rate» på 48 kHz når man bruker Praat version 5.1 (Alm & Behne, 2013). Resultatet av denne filmingen ble 10 MPEG-4 videofiler med uttalelser av stavelen /b/. Videokvalitet var på 30 fps (frames per second) med en oppløsning på 1920 x 1200 (Alm & Behne, 2013)

I eksperimentet brukte man to 27-in. iMac'er med headsett og en responsboks koblet til. Oppløsningen på iMacene var på 5120 x 2880 og har grafikkort AMD Radeon R9 M295X 4GB. iMac'en som ble brukt til Snellen testen var en 21,5-in., med oppløsning på 1920 x 1080 og grafikkort ATI Radeon HD 5670 512 MB. To forskjellige typer responsbokser ble brukt i eksperimentet, A-S og S-A. Det vil si at en hadde asynkron på venstresiden og synkron på høyresiden, og omvendt for den andre responsboksen. Responsboksene er AD Instruments MLE1347 med 7 knapper totalt, men bare tre av disse var i bruk. Selve modellene het Cedrus RB-730 og Cedrus RB-740, hvor det var tre knapper; asynkron, synkron og en blå knapp for å gå videre i eksperimentet.

## **Prosedyre**

Før eksperimentet startet skulle deltager skrive under samtykkeskjema, svare på et spørreskjema, ta Snellen testen og en hørselstest. Ble disse godkjent gikk de videre til selve eksperimentdelen. Eksperimentet ble gjennomført i Talelaben på instituttet for Psykologi, NTNU Dragvoll. To iMacer ble brukt gjennom studiet, der lysstyrken ble satt til 12 hakk og lydstyrken var på 6 hakk. Selve eksperimentet ble gjennomført i tre deler hvor eksperimentkjørerne måtte skifte filer i Superlab 6-2. mellom hver del. Dette ble gjennomført på grunn av dårlig prosesseringskraft i de to iMacene som ble brukt, følgelig var det et behov for oppdeling i tre deler istedenfor en helhet. For hver oppdeling ble den siste stimulusen gjentatt for å ikke miste den tiltenkte effekten av rekalleringen i pausene. I den første delen av eksperimentet var det 141 stimuli, og 150 i hver av de to andre delene som ble brukt. Til sammen ble 441 stimuli brukt i dette studiet, ikke medregnet de fire eksempel stimuliene i starten av den første delen og de åtte stimuliene som gjentok seg etter pausene. Hver stimulus ble kalt for en «trial» der man skulle trykke «asynk» eller «synk» avhengig av hva man oppfattet som audiovisuell synkroni eller ikke. Test-boksene som ble brukt var forskjellige der én type hadde «asynk» på venstre hånd og «synk» på høyre hånd, mens den andre typen hadde omvendt. Disse ble fordelt slik at det skulle bli så likt som overhodet mulig mellom bruken av de forskjellige, slik at man til en viss grad kan utelukke bias mot en gitt hånd.

Lengden på eksperimentet varierte fra deltaker til deltaker. Noen var raske med å svare, mens andre tok seg bedre tid. Dette ble særlig tydelig dersom man hadde to deltakere inne samtidig for å kjøre eksperimentet. Da måtte den som ble ferdig først vente til den andre ble ferdig før en av testerne kom for å skifte del eller avslutte eksperimentet. Eksperimenttiden var satt til ca. 30 minutter, men tiden var som sagt svært varierende.

## ***Resultat***

### **Databehandling:**

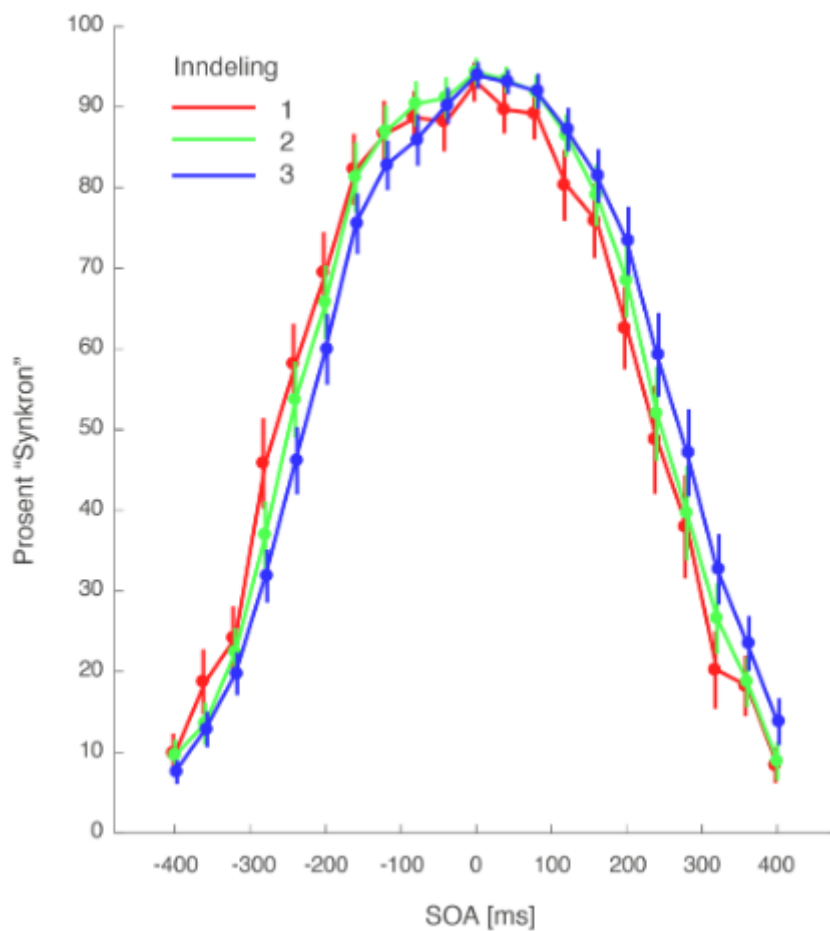
Den gaussiske kurven som blir brukt er to S-kurver som sammen skaper en gaussisk kurve. «Cumulative» gaussisk kurve som brukes i Yarrow et al. (2011) er lik tilnærmingen som dette studiet har til bruken av to S-kurver. Forskjellene er at Yarrow et al. (2011) fokuserer på SJ og TOJ sammen, bruker utelukkende PSSaverage og har ikke talestimuli. De to S-kurvene vil være en kurve for audio-ledende stimuli, og en for visuell-ledende stimuli. Man vil få PSS og

FWHM som hos en vanlig gaussisk kurve, men presisjon i endringer for de to gaussiske kurvene vil bli høyere på grunn av at de to S-kurvene blir uavhengig av hverandre. Det vil også føre til muligheten for en annen type PSS som ikke går på å finne midtpunktet mellom ALT og VLT (PSSaverage). Studiet foreslår en annen tilnærming; PSScrossover som bruker skjæringspunktet mellom de to S-kurvene som PSS-verdi.

Studiet gikk gjennom hver deltager for å sjekke om alle hadde fått alle stimuli. Deretter kombinerte studiet de tre delene (A,B,C) til en fil, hvor de brukte denne samlede filen for «data ekstraksjon» i MatLab, R2021b. Der ble SOA strukturert etter SOA-1, og en prosentandel av synkronitet for hver SOA ble kalkulert. Prosentandelen synkrone responser ble deretter tilpasset en gaussisk kurve per inndeling, hvor ALT og VLT ble ekstrahert. Én deltaker hadde ikke terskelverdier for audio-ledende (AL) eller video-ledende (VL), og ble dermed ekskludert fra resultatene. Studiet ville bruke to S-kurver istedenfor en gaussisk kurve, dermed ble dataen justert til å passe inn i to S-kurver. S-kurvene var også tilpasset prosentandelen synkrone responser per inndeling, men det kan tenkes at den har en større sensitivitet for skifter i ALT og VLT siden AL og VL har hver sin kurve som er uavhengig av hverandre. Gjennom ALT og VLT kalkulerer man for PSSaverage, PSScrossover, FWHM, ALTslope og VLTslope for hver deltaker. Utregningen av «PSSaverage» var lik som den som ble brukt av Yarrow et al. (2011), mens PSScrossover fant man i krysningpunktet mellom de to S-kurvene. Utvalget er  $n = 29$  for alle de ulike inndelingene, hvor hver inndeling vil ha tre grafiske fremstillinger av to S-kurver.

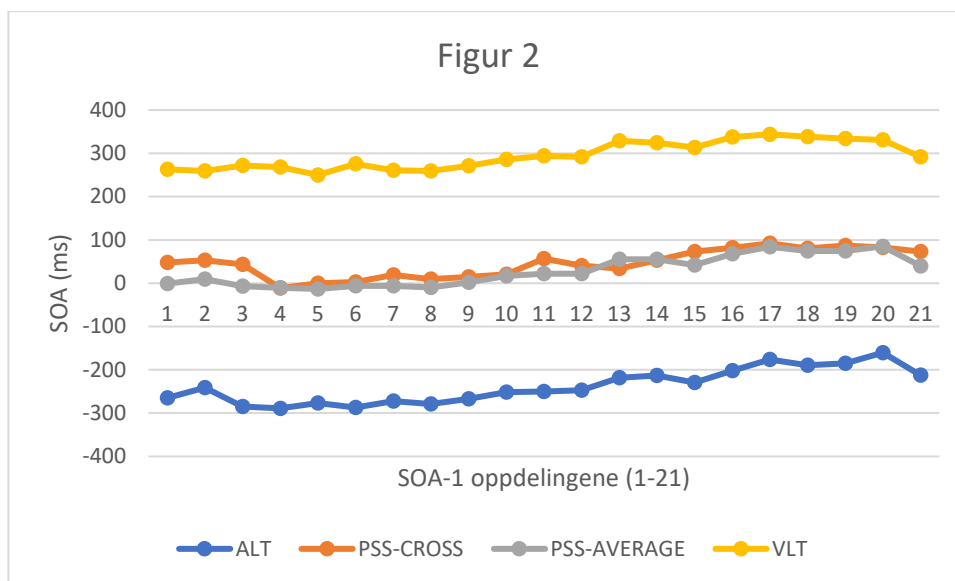
Resultatene i MatLab ble satt inn i et Excel-ark med svarene for de spesifikke deltakerne i prosjektet i et ark, og resultatene av inndelingene i et annet ark. Tre inndelinger (A,B,C) med verdier for ALT, VLT, PSS i krysningpunktet mellom de to S-kurvene og «PSSaverage», hvor SOA-verdiene var fra -400 til 400 ms med 21 SOA-verdier. Det ble dermed 7 av disse i hver inndeling, -400 til -160 ms, -120 til 120 ms og 160 til 400 ms. De ble likt fordelt der tanken var at inndeling A skulle være audio-ledende asynkron, inndeling B skulle være synkron og inndeling C skulle være video-ledende asynkron.

Figur 1

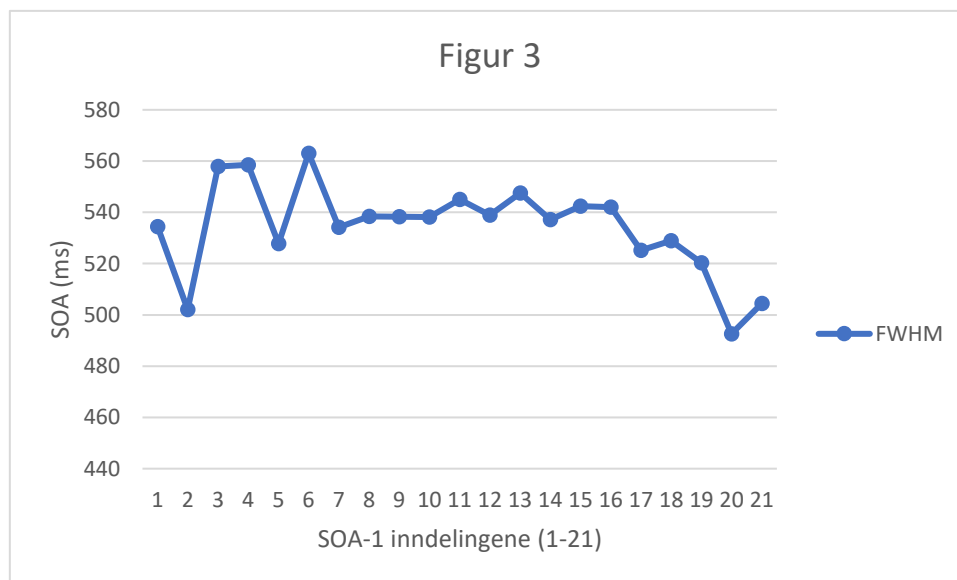


*Notat.* Dette er en grafisk fremstilling av de tre inndelingene der de to S-kurvene er satt sammen til en tilnærmet gaussisk kurve. Inndeling 1 (rød), også kjent som inndeling A, viser en video-ledende SOA-1, inndeling 2/B (grønn) viser en tilnærmet synkron SOA-1, og til slutt så viser inndeling 3/C (blå) en audio-ledende SOA-1. Kurvene er et resultat av de gjennomsnittlige verdiene for de 29 deltakerne.

## Presentasjon av resultater:



Figur 2: En visuell presentasjon av de ulike SOA-nivå punktene (1-21), hvor man finner ALT, «PSScross», «PSSaverage» og VLT.



Figur 3: En visuell presentasjon av FWHM etter SOA-1 punktene (1-21). Ble ikke med i figur 2 siden det ville vært vanskeligere å se tendensene i grafen hvis den hadde økt til 700 ms SOA på y-aksen.

Inndelingene (A,B,C) ble deretter satt inn i SPSS versjon 27. Her ble det kjørt fire forskjellige enveis-ANOVA; «Repeated Measures» under «Analyze» på verktøyslinjen, og verktøyet brukes når man skal finne varians innenfor grupper med en uavhengig variabel på mer enn to nivå. De avhengige variablene i studiet er ALT, VLT, PSScrossover, PSSaverage og FWHM, hvor disse hadde tre nivåer per (A,B,C). De uavhengige variablene er de tre inndelingene av SOA-1. Post hoc testen som ble brukt var Bonferroni. Én ble ekskludert etter databehandlingen på grunn av tydelige avvik. 180 S-kurver til sammen, men 6 ble fjernet når en deltager ble ekskludert.

Mauchly's test of Sphericity ble kjørt for å sjekke homogeniteten til utvalget. Mauchly's var ikke signifikant for PSScrossover ( $F(1,55)(p = .504)$ ), og homogeniteten er oppholdt. Derimot var Mauchly's signifikant for ALT ( $p = .004$ ), PSSaverage ( $p = .004$ ), VLT ( $p = .001$ ), og FWHM ( $p = .015$ ), slik at  $f$ -verdien blir tilpasset og  $f$ -verdiene som rapporteres nedenfor er basert på Greenhouse Geiser.

Det ble brukt en enveis-ANOVA for å finne forskjeller innad i de ulike inndelingene i resultatene for ALT, ( $F(1,55) = 15.26, p < .001$ ). En Bonferroni post hoc test viste at den største signifikante negative forskjellen i ALT,  $\Delta M = -42.55, p < .001$ , var mellom inndeling A,  $M = -266.49, SD = 108.54$ , og inndeling C,  $M = -223.94, SD = 74.43$ . Det var også en signifikant negativ forskjell i ALT,  $\Delta M = -23.51, p = .013$ , mellom inndeling A, og inndeling B,  $M = -242.98, SD = 85.95$ . Til slutt var det også en signifikant negativ forskjell i ALT,  $\Delta M = -19.04, p = .008$ , mellom inndeling B og inndeling C.

En enveis-ANOVA ble brukt for å finne forskjeller innad i de ulike inndelingene i resultatene for PSSaverage, ( $F(1,49) = 22.11, p < .001$ ). Bonferroni post hoc test viste også at den største signifikante negative forskjellen i PSSaverage,  $\Delta M = -35.89, p < .001$ , var mellom inndeling A,  $M = 26.30, SD = 43.74$ , og inndeling C,  $M = -9.58, SD = 64.06$ . Det var også en signifikant negativ forskjell,  $\Delta M = -18.06, p = .003$ , mellom inndeling A, og inndeling B,  $M = 8.48, SD = 53.56$ . Til slutt var det også en signifikant negativ forskjell,  $\Delta M = -17.82, p = .001$ , mellom inndeling B og inndeling C.

En enveis-ANOVA ble brukt for å finne forskjeller innad i de ulike inndelingene i resultatene for VLT, ( $F(1,44) = 16.77, p < .001$ ). VLT,  $\Delta M = -30.08, p < .001$ , hadde en signifikant negativ forskjell mellom inndeling A,  $M = 246.46, SD = 81.50$ , og inndeling C,  $M = 276.55, SD = 68.45$ . Det var også en signifikant negativ forskjell i VLT,  $\Delta M = -13.48, p = .015$ , mellom inndeling A og inndeling B,  $M = 259.94, SD = 74.19$ . Til slutt var det også en signifikant negativ forskjell i VLT,  $\Delta M = -16.61, p = .001$ , mellom inndeling B og inndeling C. Det var ikke en signifikant forskjell i PSScrossover mellom de tre forskjellige inndelingene A, B og C. Heller ikke noen signifikante resultater for FWHM.

### *Diskusjon*

#### **PSS**

Den første hypotesen om at SOA-1 er video-ledende blir PSS mer video-ledende sammenlignet med synkron og audio-ledende SOA-1, støttes eller forkastes avhengig av valg av PSS. Hvis man velger å bruke PSScrossover får man ikke-signifikante resultater, mens et valg av PSSaverage vil føre til signifikante resultater. I dette studiet valgte man å bruke PSScrossover som verdi siden man trodde at dette ville være den mest presise måten å måle PSS på. Derfor blir den første hypotesen forkastet. Det skal sies at PSScrossover går mot positiv retning, men forskjellen er ikke stor nok til at dette er signifikant.

En del av studiets formål var å se på hvordan våre resultater vil være sammenlignet med funnene fra Roseboom (2019) og Van der Burg et al. (2013). Grunnlaget for to av hypotesene er nemlig dette, derfor vil det være relevant å snakke om PSSaverage når man sammenligner med disse studiene. Disse forskjellige resultatene, for de ulike måtene å måle PSS på, tyder på at man får forskjellige verdier avhengig av PSS-en som velges. Studiet spekulerte om PSScrossover ville være mer presis siden ALT og VLT ville være mer uavhengig av hverandre som en følge av de to S-kurvene, og dermed være mer presise. Om dette stemmer eller ikke er umulig å konkludere, men i dette studiet så ser man forskjeller mellom PSScrossover og PSSaverage.

#### **VLT**

Den andre hypotesen; når SOA-1 er video-ledende blir VLT mer video-ledende sammenlignet med synkron SOA-1, blir støttet. Her ser man likheter mellom dette studiet og Van der Burg et al. (2013), med at det vil utføres en «rapid temporal recalibration» hos deltakeren.

Roseboom (2019) tok dette videre med «rapid temporal recalibration», hvor det vil oppstå «serial dependence» som fører til negative ettervirkninger for SJ-oppgaver og positive ettervirkninger for TJ- og MJ-oppgaver (Roseboom, 2019). Dette funnet strider mot tidligere oppfatninger innenfor atferds studier som sier at man får positive ettervirkninger for SJ-oppgaver som en følge av «serial dependence» (Roseboom, 2019). Roseboom argumenterer derfor for at «serial dependence» burde bli sett på som en kombinasjon av positive ettervirkninger og beskrevet ved en bayesisk modell. Funnene fra denne studien fant en positiv ettervirkning for SJ-oppgaver, og hadde en større signifikant forskjell for inndeling A; video-ledende SOA-1., enn den hadde for inndeling B; synkron SOA-1. Resultatene støtter derfor hypotesen og Roseboom sin tilnærming til «serial dependence», selv om at han fant en negativ ettervirkning gjennom sine SJ-oppgaver.

## **ALT**

Den siste hypotesen er at det forventes ingen forskjell mellom ALT ved SOA-1 som er audio-ledende og synkron SOA-1. Denne ble forkastet, siden det var en signifikant forskjell mellom alle de tre inndelingene, og dermed en signifikant forskjell mellom audio-ledende og synkron (inndeling A og B). Dette går imot tidligere forskning som Alm & Behne (2013) der resultatene tydet på en ikke-signifikant forskjell for ALT avhengig av SOA-1. Alm & Behne (2013) brukte flere aldersinndelinger og så på forskjeller mellom dem. I dette studiet var alderen til deltakerne 20 til 28 år, unge voksne. Som Alm & Behne (2013) fant ut så vil middelaldrende voksne ha en mer konservativ ALT, enn yngre voksne. Med vårt utvalg av yngre mennesker så var ALT verdien, i likhet med Alm & Behne (2013), mer liberal. Dette kan være noe av grunnen til de signifikante forskjellene for ALT.

## **ALT vs VLT**

Studiets resultater kan tyde på at ALT og VLT påvirkes nesten like mye av SOA-1. Det kan være flere grunner for dette. En av disse kan være inndelingene som ble brukt i studiet. Studiet har en lik fordeling mellom inndelingene med 7 i hver, men man er ikke sikkert at



dette var den beste måten å gjøre dette på. Funnene tyder på at ALT påvirkes i større grad nært fysisk synkron på den audio-ledende enn lengre fra fysisk synkron. Med dette menes det at helningen til ALT grafen er brattere nærmere SOA-1 som er fysisk synkron enn ved asynkron SOA-1 på audio-ledende side (se figur 2). En annen inndeling kan dermed ha ført til andre resultater. Grunnen til dette kan være tendensen til å ha en mer konservativ ALT. ALT for dette studiet er mer liberalt enn tidligere studier, men den brattere helningen innenfor inndeling B på audio-ledende SOA-1 den rigide tendensen mennesker har for oppfattelse for delay av video-ledende stimuli (Alm & Behne, 2013).

Denne forskjellen i sensitivitet vil også gjelde for VLT. VLT vil være mer fleksibel enn ALT (Alm & Behne, 2013), og vil dermed bli mer påvirket av forskjellige faktorer som for eksempel. Læringseffekten burde minimeres gjennom randomiseringen som er forskjellige for alle de ulike deltakerne, men fleksibiliteten i seg selv vil kunne føre til variasjoner i resultatene av VLT. Fleksibiliteten til VLT og den rigide ALT er det som skaper en PSS som ofte er på den video-ledende siden.

### **TBW og TWI**

Temporal binding window (TBW) er ikke noe dette studiet har sett direkte på, men den er en viktig del av studier på samtidighet for AV stimuli. For å se på TBW så vil studiet bruke resultater fra FWHM som ble presentert i resultatene. Det var ingen signifikante forskjeller mellom de ulike inndelingene innad i FWHM, noe som var et overraskende resultat. Når man har språk stimuli så vil TBW forskyves kraftig mot video-ledende stimuli (Noesselt et al., 2005). I dette studiet fikk man ikke denne tydelige forskyvningen siden ALT var mindre konservativ enn forventet. Dette kan være et resultat av at målingene for video-ledende og audio-ledende stimuli blir mer nøyaktig på grunn av økt sensitivitet. Den konservative ALT som tidligere studier har funnet gjennom å se på forskyvningen til PSS kan være et resultat av forskjellig metodikk. Resultatene for PSSaverage viser et signifikant skifte til video-ledende SOA-1 som også gjenspeiler tidligere funn, men PSScrossover og FWHM har ikke signifikante forskjeller mellom inndelingene.

Flere studier (Van der Stoep et al., 2020; Kaganovich & Schumaker, 2016) ser også på hvordan avstandene fra stimuli burde føre til en inndeling av TBW inn i to deler; «far space» og «near space». Dette gjenspeiles i forskningen som viser til forskyvingen av TBW mot video-ledende stimuli. Mennesker er mer vant til å høre delay av audio stimuli enn delay av video stimuli (Keetels & Vroomen, 2012). Studiet til Van der Stoep et al. (2020) antyder at TBW er større for «near space» enn for «far space». Lignende resultater finner man i dette studiet gjennom resultatene for FWHM (se figur 3). I figur 3 ser man tendensen til at FWHM er mindre for SOA-1 punktene 1-7 og 14-21, med andre ord SOA-1 som enten er meget video-ledende eller meget audio-ledende. Til tross for ingen signifikante forskjeller mellom de ulike inndelingene (A,B,C), ser man at resultatene for dette studiet antyder det samme som Van der Stoep et al. (2020). Argumentasjonen om å skille TBW i «far space og «near space» kan dermed være noe som er relevant for fremtidige studier.

## **Metode**

Hvor studiet i stor grad skiller fra tidligere studier er valgene som er gjort innenfor metoden. Valget av S-kurver istedenfor gaussisk kurv, nytt forslag til måling av PSS (crossover) og fokuset på ALT og VLT over PSS er aspekter som skiller seg fra tidligere lignende studier (Roseboom et al., 2019; Van der Burg et al., 2013; Van der Burg et al., 2015; Alm & Behne, 2013). Studiet tok utgangspunkt i mye av metoden hos Alm & Behne (2013) med lignende stimuli, utstyr, lokasjon og redigering. Videoen som ble vist hadde en nøytral, hvit, bakgrunn for å unngå utilsiktede effekter som en følge av farge på bakgrunn (Sung-En et al., 2020). Studiet har bare en stimuli /ba/, hvor andre studier ofte har flere stimuli for å ligne mer McGurk-effekten. Fokuset er derfor ikke i like stor grad på språk eller lignende aspekter, men heller på rekalkibrering utelukkende gjennom SJ-oppgaver. Derfor er dette en ganske så spisset tilnærming og metode for å se på oppfatningen av samtidighet i audiovisuell persepsjon.

Studiet testet også «mixed effects model» i SPSS som nå er blitt mer og mer brukt for ANOVA. Resultatene var tilnærmet like som ved å kjøre en «repeated measueres» analyse for ANOVA, så studiet kom fram til å bruke denne metoden i studiet.

## **Videre forskning**

Dette studiet fokuserte på hvordan recalibrering påvirker oppfattelsen av synkronitet i AV stimuli på SJ-oppgaver. Funnene tyder på «rapid temporal recalibration» som støtter lignende funn fra Roseboom (2019) og Van der Burg et al. (2013). I tillegg til dette ble det brukt S-kurver istedenfor den gaussiske kurve i den grafiske framstillingen av resultatene, og både PSS i krysningspunktet mellom grafene og en gjennomsnittlig PSS-verdi ble brukt. Gjennom dette ville studiet se om PSS er en god måleverdi for samtidighet i AV talepersepsjon. Resultatene antyder et behov for mer forskning på bruken av S-kurver som et alternativ til den gaussiske kurven, i fremtidige studier.

Noe annet som videre forskning kan se på er om helningen av kurvene (stigningstallet til punktene) er en verdifull målemetode eller verdi. Hvor er det at de største forandringene i verdier som ALT, VLT og PSS skjer? Dette kan være noe som kan gi en forklaring på hvorfor dette studie fikk overaskende resultater der ALT og VLT nesten hadde like mye forandring innenfor egne inndelinger. ALT er ofte mer sensitiv for forandringer i SOA/SOA-1 slik at det kan bli en effekt av PSS skifte i forhold til krysning eller gjennomsnitt på grunn av dette.

## **Konklusjon**

Dette studiet har sett på recalibrering og oppfattet samtidighet av audiovisuelle stimuli i SJ-oppgaver. Hypotesene som ble presentert gjenspeiler tidligere lignende forskning innenfor feltet; Roseboom (2019), Van der Burg et al. (2013) og Alm & Behne (2013). Yarrow et al. (2011) oppfattelse av svakhetene til den tradisjonelle gaussiske kurven (PSSaverage) ble tatt høyde for, der studiet endte med å bruke to S-kurver og en ny målemetode for PSS (PSScrossover). Resultatene tyder på en negativ ettervirkning for SJ-oppgaver som støtter tidligere forskning. Noe uventet var funnene for ALT-verdier og PSS. ALT var mindre konservativ enn forventet noe som førte til mindre forskyvning av TBW mot video-ledende SOA. Funnene viste også viktigheten av valget av PSS, siden studiet fikk signifikante forskjeller innad inndelingene for PSSaverage, men ikke for PSScrossover. Resultatene tyder på at nøyaktigheten som øker ved valget av S-kurver og PSScrossover kan føre til ulike resultater sammenlignet med tidligere forskning som så på PSSaverage og brukte en gaussisk kurve.

**Referanseliste:**

Alm, M. & Behne, D. (2013). Audio-visual speech experience with age influences perceived audio-visual asynchrony in speech. *The Journal of Acoustical Society of America*, 134, 3001.

DOI: <https://doi.org/10.1121/1.4820798>

APA Dictionary of Psychology (2021, 23. mars). *Law of prior entry*.

<https://dictionary.apa.org/law-of-prior-entry>

Begau, A., Klatt, L.-I., Wascher, E., Schneider, D. & Getzmann, S. (2021). Do congruent lip movement facilitate speech processing in a dynamic audiovisual multi-talker scenario? An ERP study with older and younger adults. *Behavioural Brain Research*, 412, 113436. DOI:

<https://doi.org/10.1016/j.bbr.2021.113436>

Boenke, L. T., Deliano, M. & Ohl, F. W. (2009). Stimulus duration influences perceived simultaneity in audiovisual temporal-order judgement. *Experimental Brain Research*, 198, 233-244. DOI: 10.1007/s00221-009-1917-z

Bushara, K. O., Grafman, J. & Hallett, H. (2001). Neural Correlates of Auditory-Visual Stimulus Onset Asynchrony Detection. *Journal of Neuroscience*, 21(1), 300-304. DOI:

<https://doi.org/10.1523/JNEUROSCI.21-01-00300.2001>

de Gelder, B., Vroomen, J., Annen, L., Masthof, E. & Hodiament, P. (2003). Audio-visual integration in schizophrenia. *Schizophrenia Research*, 59(2-3), 211-218. DOI:

[https://doi.org/10.1016/S0920-9964\(01\)00344-9](https://doi.org/10.1016/S0920-9964(01)00344-9)

Dixon, N. F. & Spitz, L. (1980). The detection of auditory visual desynchrony. *Perception*, 9(6), 719-721. DOI: <https://doi.org/10.1068/p090719>

Donohue, S. E., Woldorf, M. G. & Mitroff, S. R. (2010). Video game players show more precise multisensory temporal processing abilities. *Attention, Perception, & Psychophysics*, 72, 1120-1129. DOI: <https://doi.org/10.3758/APP.72.4.1120>

Eimer, M. & Schröger, E. (1998). ERP effects of intermodal attention and cross-modal links in spatial attention. *Psychophysiology*, 35(3), 313-327. DOI: <https://doi.org/10.1017/s004857729897086x>

Feldman, J. (2013). The neural binding problem(s). *Cognitive Neurodynamics*, 7, 1-11. DOI: <https://doi.org/10.1007/s11571-012-9219-8>

Freeman, E. D., Ipser, A., Palmbaha, A., Paunoiu, D., Brown, P., Lambert C., Leff, A. & Driver, J. (2013). Sight and sound out of synch: Fragmentation and renormalisation of audiovisual integration and subjective timing. *Cortex*, 49(10), 2875-2887. DOI: <https://doi.org/10.1016/j.cortex.2013.03.006>

Fujisaki, W., Shimojo, S., Kashino, M. & Nishida, S. (2004). Recalibration of audiovisual simultaneity. *Nature Neuroscience*, 7, 773-778. DOI: <https://doi.org/10.1038/nn1268>

Hairston, W. D., Burdette, J. H., Flowers, D. L., Wood, F. B. & Wallace, M. T. (2005). Altered temporal profile of visual–auditory multisensory interactions in dyslexia. *Exp Brain Res* 166, 474–480. DOI: <https://doi.org/10.1007/s00221-005-2387-6>

Hillyard, S. A. & Münte, T. F. (1984). Selective attention to color and location: an analysis with event-related brain potentials. *Perception & Psychophysics*, 36(2), 185-198: DOI: <https://doi.org/10.3758/bf03202679>

Kaganovich, N. & Schumaker, J. (2016). Electrophysiological correlates of individual differences in perception of audiovisual temporal asynchrony. *Neuropsychologia*, 86, 199-130. DOI: <https://doi.org/10.1016/j.neuropsychologia.2016.04.015>

Keetels, M., & Vroomen, J. (2012). Chapter 9: Perception of synchrony between the senses. *The neural bases of multisensory processes* (s. 147- 177). Taylor & Francis Group.

Keetels, M. & Vroomen, J. (2005). The role of spatial disparity and hemifields in audio-visual temporal order judgments. *Experimental Brain Research*, 167, 635-640. DOI: <https://doi.org/10.1007/s00221-005-0067-1>

Machulla, T.-K., Di Luca, M., Froehlich, E. & Ernst, M. O. (2012). Multisensory simultaneity recalibration: storage of the aftereffect in the absence of counterevidence. *Experimntal Brain Research*, 217, 89-97. DOI: <https://doi.org/10.1007/s00221-011-2976-5>

Noel, J.-P., Serino, A. & Wallace M. T. (2019). Increased Neural Strength and Reliability to Audiovisual Stimuli at the Boundary of Peripersonal Space. *Journal of Cognitive Neuroscience*, 31(8), 1155-1172. DOI: [https://doi.org/10.1162/jocn\\_a\\_01334](https://doi.org/10.1162/jocn_a_01334)

Noel, J.-P., Chatelle, C., Perdikis, S., Jöhr, J., Da Silva, M. L., Ryvlin, P., De Lucia, M., Diserens, K. & Serino, A. (2019). Peri-personal space encoding in patients with disorders of consciousness and cognitive-motor dissociation. *NeuroImage: Clinical*, 24, 101940. DOI: <https://doi.org/10.1016/j.nicl.2019.101940>

Pöppel, E., Schill, K. & von Steinbüchel, N. Sensory integration within temporally neutral systems states: A hypothesis. *Naturwissenschaften* 77, 89–91. <https://doi.org/10.1007/BF01131783>

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-referansetil>

Roseboom, W. (2019). Serial Dependence in Timing Perception. *Journal of Experimental Psychology: Human Perception and Performance*, 45(1), 100-110. DOI:

<http://dx.doi.org/10.1037/xhp0000591>

Sung-En, C., Yi-Chuan, C., Akiko, M., Wakayo, Y., Kuang-Tsu, S., Sei-ichi, T. & Su-Ling, Y. (2020). The modulation of background color on perceiving audiovisual simultaneity.

*Vision Research*, 172, 1-10. DOI: <https://doi.org/10.1016/j.visres.2020.04.009>

Van der Burg, E., Alais, D., & Cass, J. (2013). Rapid Recalibration to Audiovisual Asynchrony. *Journal of Neuroscience*, 33(37), 14633-14637. DOI:

<https://doi.org/10.1523/JNEUROSCI.1182-13.2013>

Van der Burg, E., Orchard-Mills, E. & Alais, D. (2015). Rapid temporal recalibration is unique to audiovisual stimuli. *Experimental Brain Research*, 233, 53-59. DOI:

<https://doi.org/10.1007/s00221-014-4085-8>

Van der Stoep, N., Colonius, H., Noel, J.-P., Wallace, M. T. & Diederich, A. (2020). Audiovisual integration in depth: Modeling the effect of distance and stimulus effectiveness using the TWIN model. *Journal of Mathematical Psychology*, 99, 102443. DOI:

<https://doi.org/10.1016/j.jmp.2020.102443>

Vatakis, A., Navarra, J., Soto-faraco, S. & Spence, C. (2008). Audiovisual temporal adaption of speech: temporal order versus simultaneity judgements. *Experimental Brain Research*, 185,

521-529. DOI: <https://doi.org/10.1007/s00221-007-1168-9>

Vatakis, A. & Spence, C. (2010). Audiovisual Temporal Integration for Complex Speech, Object-Action, Animal Call, and Musical Stimuli. *Multisensory Object Perception in the Primate Brain*, 95-121. DOI: [https://doi.org/10.1007/978-1-4419-5615-6\\_7](https://doi.org/10.1007/978-1-4419-5615-6_7)

Virsu, V., Lahti-Nuuttila, P. & Laasonen, M. (2003). Crossmodal temporal processing acuity impairment aggravates with age in developmental dyslexia. *Neuroscience Letters*, 336(3), 151-154. DOI: [https://doi.org/10.1016/S0304-3940\(02\)01253-3](https://doi.org/10.1016/S0304-3940(02)01253-3)

Wallace, M. T. & Stevenson, R. A. (2014). The Construct of the Multisensory Temporal Binding Window and its Dysregulation in Developmental Disabilities. *Neuropsychologia*, 64, 105-123. DOI: <https://dx.doi.org/10.1016%2Fj.neuropsychologia.2014.08.005>

Yarrow, K., Jahn, N., Durant, S. & Arnold, D. H. (2011). Shifts of criteria or neural timing? The assumptions underlying timing perception studies. *Consciousness and Cognition*, 20(4), 1518-1531. DOI: <https://doi.org/10.1016/j.concog.2011.07.003>



