

10027

Bacheloroppgave

Hurtig recalibrering av audiovisuelle asynkroniteter i talestimuli

Mai 2022

NTNU

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

Bacheloroppgave

2022



10027

Hurtig rekalkibrering av audiovisuelle asynkroniteter i talestimuli

Bacheloroppgave
Mai 2022

NTNU

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



Kunnskap for en bedre verden

HURTIG REKALIBRERING AV AUDIOVISUELLE ASYNKRONITETER I TALESTIMULI

KANDIDATNUMMER: 10027

PSY2900 BACHELOROPPGAVE I PSYKOLOGI- OPPFATTET SAMTIDIGHET I AUDIOVISUELL TALEPERSEPSJON
TRONDHEIM. 16. MAI 2022

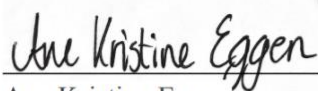
DAWN M. BEHNE

Forord

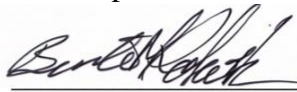
As a starting point for this project, the advisor introduced students to the project's research question and some related issues, together with initial supporting literature. Further literature was identified by the students and shared with the group, and occasionally supplemented by the project advisor.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

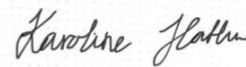
With this basis, each student submits a report (written individually) which has the form and style of a journal article. Students are allowed and encouraged to work together, but the final product must be their own. The report can be in Norwegian or English.



Ane Kristine Eggen
Date: 10.05.2022



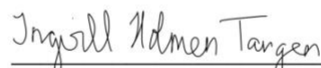
Bente Mari Aakvik
Date: 10.05.2022



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

Samtidig vil jeg takke prosjektets veileder Dawn M. Behne for god oppfølging og veiledning. Jeg vil også rette en takk til Darren Rhodes for hjelpen med datahåndteringen.

Tilslutt vil jeg takke prosjektgruppa for et godt og tett samarbeid underveis i bachelorprosjektet.

Sammendrag

Tidligere forskning tyder på at det forekommer en hurtig recalibrering i henhold til tidligere presentert stimuli. Denne studien undersøkte om SOA- 1 ville påvirke VLT, ALT eller PSS. Hver parameter ble sammenlignet på tvers av tre grupper med synkron, audioledende og videoledende stimuli etter en «simultaneity judgement task» (SJ- oppgave). Stimulusen som ble presentert i eksperimentet var audio/ video- justeringer av ordlyden /ba/. Tidligere forskningsresultater tyder på at datamateriale fra en SJ- oppgave vil representeres mer nøyaktig ved å tilpasse to uavhengige kurver enn én Gaussisk kurve som ofte brukes i slik forskning. Derfor ble PSS ekstrahert ved hjelp av begge kurvene for å kunne se forskjellen på de to. En enveis repeated measures ANOVA viste at tersklene for å oppdage både audioledende og videoledende stimuli flyttet seg mot den ledende sansen. Samsvarende funn viste at også PSS skiftet mot den ledende sansen. Resultatene tyder også på at vinduet på opplevd synkronitet er asymmetrisk, hvor mennesket tillater mer videoledende stimuli i opplevelsen av synkroner hendelser. Resultatene gir implikasjoner på at s- kurver tilpasser datamaterialet bedre enn en Gaussisk kurve, men at helningen på kurvene derfor blir ulik, slik at krysningspunktet for PSS kan variere og føre til mer ulike funn enn ved Gaussisk kurve.

Hurtig rekalkibrering av audiovisuelle asynkroniteter i talestimuli

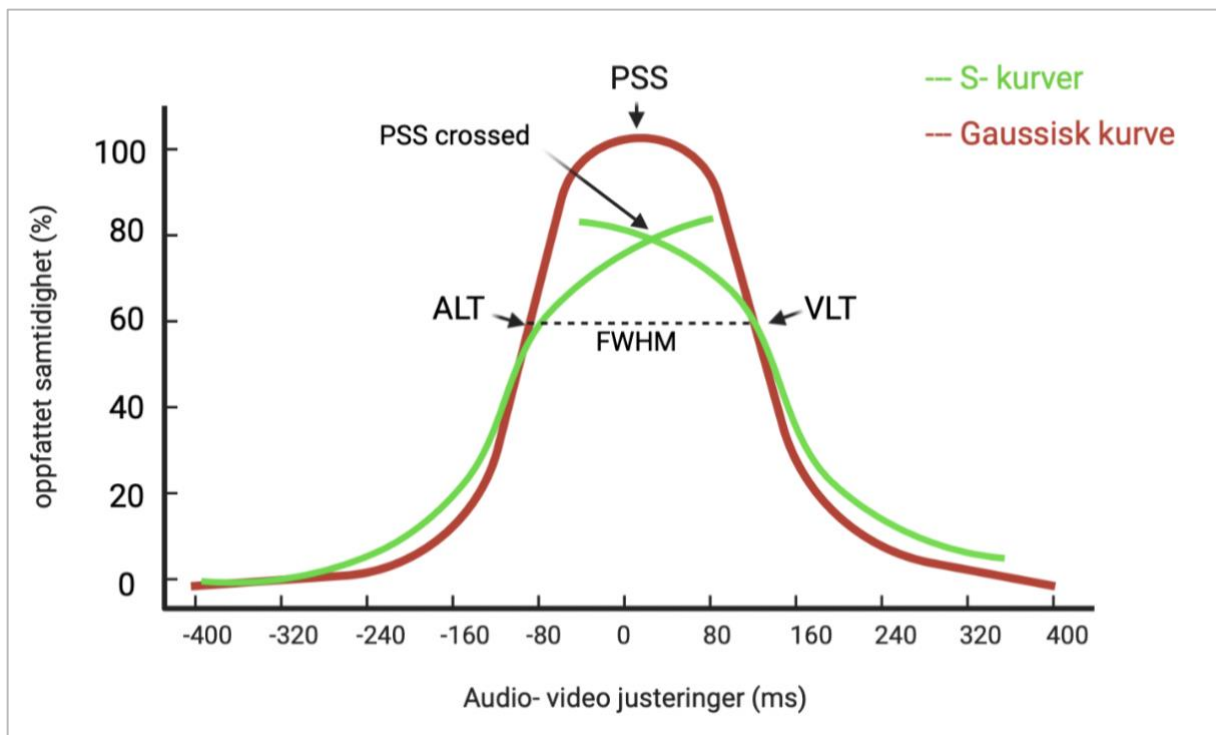
Hver dag blir vi presentert for utallige mengder sensorisk input fra flere ulike kilder. Dette gjør at sansene som stimuleres, sammen må avgjøre om de sensoriske hendelsene hører sammen og dermed sammenkoble stimuliene til en enkelt sanseoppfatning. I en slik prosess kan stimuliene komme fra samme kilde og samtidig være asynkrone i prosesseringen i hjernen, men likevel bli oppfattet som synkrone fordi de er sterkt sammenkoblet til samme hendelse (Vroomen & Keetels, 2010). En slik antagelse fra oppfatterens perspektiv, kalles for «unity assumption» og er svært vanlig å oppleve ved for eksempel talestimuli, fordi mennesket er så vant til at den auditive og visuelle stimulusen tilhører en og samme hendelse, at det dermed oppfattes som synkront tross den fysiske tidsforskjellen (Vroomen & Keetels, 2010, p. 876). Slike egenskaper på tvers av ulike sanser krever at den tidsmessige nærheten må være tilstede for at individet skal oppfatte at stimuliene kom fra ett og samme objekt eller hendelse (Vroomen & Keetels, 2010). Forskning viser at det er forskjell i hvor fort lyd og audio prosesseres i luft, samtidig som det også forekommer en forskjell i hastigheten på prosesseringen i hjernen (Sugita, 2003).

Den oppfattede samtidigheten av for eksempel asynkront audiovisuelt talestimuli, oppfattes som synkron fordi vi mennesker hele tiden tilpasser oss omgivelsene og ønsker at verden skal fremstå som en helhet. I den sammenheng er begrepet rekalkibrering relevant og kan beskrives som grunnlaget for at vi kan oppfatte asynkrone multisensoriske stimuli som synkrone (Van Der Burg et al., 2015). Altså ved asynkrone stimuli er hjernen fleksibel og har kapasitet til å justere eller rekalkibrere persepsjonen av de asynkrone multisensoriske stimuliene slik at de deretter vil kunne oppfattes som mer synkrone (Fujisaki et al., 2004; Van Der Burg et al., 2015; Vroomen & Keetels, 2010). Tidligere forskning har funnet at evnen til rekalkibrering av asynkrone stimuli både kan være avhengige av stimuli som blir presentert like før, så godt som langtids erfaringer (Fujisaki et al., 2004; Roseboom, 2019; Van Der Burg et al., 2013). Et eksempel på dette er at vi er vant til at audio kommer etter video på grunn av at bilde prosesseres med en raskere hastighet enn lyd i luft, noe som dermed fører til en lavere sensitivitet for videoledende stimuli (Alm & Behne, 2013; Sugita, 2003).

På grunn av funnene i den tidligere forskningen vil det derfor være nyttig å undersøke om den tidligere presenterte stimulusen, altså den tidligere «stimulus onset asynchrony» (SOA-1), kan ha en påvirkning på videoledende terskel (VLT), audioledende terskel (ALT) og «point of subjective simultaneity» (PSS), som er punktet hvor flest opplever et stimuli som synkront. ALT er punktet på en Gaussisk kurve hvor 50% opplever synkront stimuli og 50% opplever audioledende (AL) asynkront stimuli, mens VLT er punktet på kurven hvor halvparten opplever synkront stimuli eller videoledende (VL) asynkront stimuli. PSS er toppen på den Gaussiske kurven, hvor det er mest sannsynlig at de fleste opplever synkront stimuli (se figur 1).

Figur 1

Grafisk fremstilling av parameterne ALT, VLT, PSS og FWHM.



Notat. En grafisk fremstilling av de ulike parameterne ALT, VLT, FWHM og PSS, samt de to ulike måtene å tilpasse kurver til datamaterialet i s- kurver og Gaussisk kurve.

Audiovisuelle stimuli og tale

Audiovisuelle (AV) stimuli kan bli presentert på mange forskjellige måter. Det kan gå fra de mer komplekse talestimulene, til simple stimuli som for eksempel flash og beep. Vatakis og Spence (2006) foreslo at jo enklere stimuliene var, desto lettere kunne asynkronitet i oppfattede stimuli oppdages. For eksempel ville flash/ beep stimuli være enklere å skille enn mer kompleks stimuli, som tale. Talestimuli vil også være forskjellig fra flash/ beep stimuli på grunn av at det er i forandring og har en konstant variasjon (Vatakis & Spence, 2006). I tillegg presenteres mengder av talestimuli hver dag, både i virkeligheten og over nett eller skjerm. Slike tilfeller kan gjøre noe med erfaringen vi har med tanke på hvordan disse stimuliene blir persipert. Et flash/ beep stimuli, som et simpelt stimuli, vil ha høy sensitivitet til asynkronitet, mens stimuli som tale kan ha blitt påvirket erfaringen vi har. Videosamtaler over nett, hvor det stadig er en forsinkelse i enten video eller audio, kan være et eksempel på dette. Påvirkningene vi blir utsatt for i hverdagen, kan være med å justere sensitiviteten vi har for å oppdage asynkrone stimuli.

I forbindelse med audiovisuell synkronitet knyttet til talestimuli, vil de to stimuliene være tett knyttet på grunn av at mennesket bruker den visuelle stimulusen til å predikere det auditive signalet ved å lese på leppene til den som snakker (Alm & Behne, 2013). Slike prediksjoner hvor den ene stimulusen blir brukt for å forvente den andre, finner en ikke i like stor grad i flash/ beep stimuli fordi slike stimuli ikke er like sterkt sammenkoblet. Tross denne mulige forskjellen i oppfattelsen av flash/ beep, kontra oppfattelsen av talestimuli, er mye av de samme resultatene funnet på tvers av talestimuli og enkle flash/ beep stimuli. Yarrow et al. (2011) fant rekalkibrering ved at hvis beep kommer etter flash, vil PSS skifte i samme retning, slik at beepet som kommer etter flash neste gang vil oppfattes som mer samtidig. I tillegg støttet Van Der Burg et al. (2013) antakelsen om at PSS skifter mot den ledende sansen. Samtidig ble det funnet at PSS også skiftet mot den ledende sansen i den mer komplekse talestimulusen, hvor skiftet var større for eldre voksne sammenlignet med unge voksne (Alm & Behne, 2013). Denne forskjellen viser igjen til betydningen erfaring har på persepsjonen av AV synkronitet.

Et annet eksempel på opplevelsen av synkronitet er at hjernen tar lydhastighet i betraktning når samtidigheten til et lysglimt og et lydsignal skal avgjøres (Sugita, 2003). I et eksperiment ble et LED lys plassert 40 meter fra personen, som hadde hodetelefoner hvor det auditive signalet ble presentert. Funnene i denne studien indikerte at det tok omtrent 120 ms for lyden å reise 40 meter, mens terskelen for at deltakeren skulle oppdage en lydforsinkelse lå på en forsinkelse på 106 ms (Sugita, 2003). Et slikt funn kan vise tilbake på hvordan både

erfaringer og prosesseringshastigheter kan være med å påvirke oppfattelsen av synkronitet. At oppfattelsen av synkronitet ved lydforsinkelse har et større vindu enn for forsinkelse i video (Cecere et al., 2016), kan speiles tilbake på at det i virkeligheten tar lengre tid for lyd å reise enn det tar for lys. Hjernen tar altså denne forsinkelsen i betraktning når den prosesserer informasjonen fra modalitetene, ved at vinduet endres avhengig av avstanden til den synlige lydkilden og dermed avgjør synkroniteten til stimuliene (Sugita, 2003).

Det kan også tas i betraktning at både lys og lyd har begge kort vei fra sanseorganet til sin primære korteks, i motsetning til for eksempel den taktile banen som kan ha en lang vei opp til prosesseringsområdet i hjernen hvis for eksempel stimulusen kommer i tåen til personen. Variasjon i banen informasjonen må reise før den blir prosessert og oppfattet, kan dermed påvirke oppfattelsen av synkronitet. Derfor vil man kunne si at audiovisuelle stimuli har en større nærhet fra sanseorganene til prosesseringsområdene i hjernen, i motsetning til for eksempel visuell- taktile stimuli. Denne forskjellen gjør at det er større sjans for at det forekommer en asynkronitet på grunn av prosesseringshastighet ved den visuell- taktile stimulusen enn ved den audiovisuelle. For eksempel vil det ta 30 ms lengre tid for taktile stimuli fra hånden å prosesseres enn visuell prosessering (Harrar & Harris, 2008). I tillegg viser tidligere forskning at taktile stimuli kombinert med enten auditive eller visuelle har mindre påvirkning på den opplevde synkroniteten (Harrar & Harris, 2008). Likevel er det en forskjell i hastigheten lyd og lys prosesseres i hjernen, hvor det tar omtrent 40 ms lengre tid for lys enn lyd (Alm & Behne, 2013). Disse forskjellene i prosesseringshastighet utjevner hverandre når personen som oppfatter stimuliene står 10 meter fra hendelsen, på grunn av forsinkelsen lyd har i luft. Dette blir kalt «horizon of simultaneity». Derfor vil audioledende stimuli ofte bare bli oppfattet ved hendelser nærmere enn ti meters avstand (Alm & Behne, 2013; Sugita, 2003), mens på avstander over 15 meter vil lys komme før lyd (Pöppel et al., 1990). På grunn av at oppfattelsen av AL asynkroniteter ofte forekommer på en distanse under 10 meter, vil de være lettere å predikere og dermed lettere å lære, noe som gjør at audioledende terskel blir mer et mer stabilt mål ved erfaringer og læring (Alm & Behne, 2013).

Rekalibrering og persepsjon av asynkrone audiovisuelle stimuli

Rekalibrering er et viktig aspekt når det kommer til oppfattelsen av asynkrone stimuli som synkrone. Det kan forklares som det som gjør det mulig at et asynkront stimuli faktisk kan oppleves som synkront. Det har tidligere blitt diskutert hvorvidt tidligere stimuli eller erfaringer kan påvirke den opplevde samtidigheten (PSS) i audiovisuelle stimuli. Det finnes to

modeller som prøver å forklare hvordan rekalkibrering fungerer. Den ene Bayesianske modellen forklarer hvordan tidligere erfaringer kan påvirke oppfattelsen av audiovisuell synkronitet (Vilares & Kording, 2011). De foreslår at når mennesket blir presentert for et stimuli, vil individet tillegge tidligere kunnskap (prior) når stimulusen blir persipert. Vilares og Kording (2011) refererer og til «cue combinations» når det kommer til at tidligere kunnskap blir brukt for å tolke stimulusen. For eksempel viser Arnold et al. (2010) at audio og video kombinert vil føre til en større sensitivitet til språkinnhold relativt til bare sensoriske signal fra ett av sanseorganene. Altså hjernen kombinerer det auditive og visuelle signalet og kobler dermed tidligere presenterte stimuli eller erfaringer til for å øke sensitiviteten og forståelsen av for eksempel hva som blir sagt (Arnold et al., 2010). Samtidig vil oppfattelsen av samtidigheten til de audiovisuelle signalene bli oppfattet som mer synkron når de opptrer som cues som ofte hører sammen, i henhold til tidligere erfaringer ved at lepper beveger seg samtidig som ordlyder kommer ut. Dette støttes også av Navarra et al. (2005) sin antakelse om et temporalt synkronitetsvindu hvor vi kan tilpasse oss asynkroniteter, særlig for korrelerte multisensoriske input som ved at leppene stemmer overens med lyd.

I tillegg kan man se på den andre modellen som forklarer hvordan tidligere presentert stimuli kan påvirke den oppfattede samtidigheten i audiovisuelle talestimuli. Denne går ut på at stimulusen som kommer rett før den presenterte stimulusen, kan påvirke den opplevde synkroniteten. Dette skiller seg fra den Bayesianske modellen ved at den fokuserer eksplisitt på hvordan stimulusen som blir presentert like før har en påvirkning, i stedet for at erfaringer som individ kunne ha tilegnet seg for lenge siden kan ha betydning. Roseboom (2019) brukte en modell som ble kalt $n-1$ SOA, eller SOA-1 slik det blir brukt i denne artikkelen. SOA-1 betyr med andre ord asynkroniteten som ble presentert like før den påfølgende stimulusen. «The point of subjective simultaneity» (PSS), altså punktet hvor sannsynligheten er størst for at stimulusen oppfattes som samtidig, skiftes mot visuelt ledende stimuli ved audiovisuelle stimuli (se figur 1). Denne skjevheten kan være en effekt av preferansen for videoledende stimuli på grunn av den naturlige tidsmessige hastigheten i informasjonsprosesseringen, som gjør at individet er vant til at lys kommer før lyd (Keetels & Vroomen, 2012). Flere tidligere studier som har brukt audiovisuelle stimuli, har fått resultater som gir en PSS som heller mot visuelt ledende stimuli (Alm & Behne, 2013; Roseboom, 2019; Van Der Burg et al., 2015). Fujisaki et al. (2004) fant at adaptasjonen for tidsforsinkelsen i stimuliene som ble presentert var skiftet mot tidsforsinkelsen til de tidligere adapterte stimuliene og var konsistente med antakelsen om at hjernen rekalkibrerer i henhold til asynkrone audiovisuelle input. I tillegg viser Van Der Burg et al. (2015) og Yarrow et al. (2011) at PSS vil skifte mot den ledende

sansen. Altså hvis audio leder video, vil PSS være mer audioledende, mens hvis video kommer før audio, vil PSS være mer videoledende.

Når det kommer til VLT og ALT har det blitt funnet forskjeller i forbindelse med i hvilken grad mennesket rekalibrerer etter den foregående stimulusen. Forskning viser at ALT vil være mer stabil ved alderen, hvor erfaringen med VL stimuli gjør at individet får mindre toleranse for AL stimuli (Alm & Behne, 2013). I tillegg til at VLT og ALT endrer seg forskjellig, vil sensitiviteten for AL stimuli være høyere enn for VL stimuli, noe som vil påvirke terskelen for å detektere AL og VL asynkroniteter (Cecere et al., 2016). Forskjellen mellom sensitiviteten vedrørende de to tersklene gjør at synkronitetsvinduet, altså vinduet mellom ALT og VLT hvor man opplever stimuli som synkront, vil være asymmetrisk (Cecere et al., 2016). Det vil si at det er enklere å detektere audioledende asynkroniteter enn videoledende (Alm & Behne, 2013). Forskning viser at synkronitetsvinduet også blir bredere ved eksponering av asynkron audiovisuell stimuli av for eksempel tale (Navarra et al., 2005). Et bredere synkronitetsvindu betyr at det tillates mer asynkroniteter i opplevelsen av synkronitet, noe som tyder på at mennesket rekalibrerer ved eksponering asynkron AV stimuli.

Van Der Burg et al. (2015) fant at hurtig rekalibrering fant sted ved presentasjon av asynkron SOA- 1. Funnene indikerte at deltakerne rekalibrerte, altså justerte opplevelsen av synkronitet i samsvar med asynkroniteten som ble presentert like før. De fant at PSS ble mer videoledende ved presentasjon av asynkron VL stimuli sammenlignet med AL asynkrone stimuli. Samsvarende funn viser at endring i erfaring skjer hurtig i etterkant av et asynkront stimuli (Roseboom, 2019). I følge Roseboom (2019) vil hurtig eksponering påvirke avgjørelser som øker perseptuell stabilitet, mens langtidserfaringer endrer egenskapen til sansene for å øke perseptuell presisjon.

SJ- oppgaver og bruk av tilpassede kurver

I tidligere forskning er det blitt brukt flere ulike oppgaver for å måle opplevd synkronitet. De mest brukte metodene er TOJ (temporal order judgement) oppgaven og SJ (simultaneity judgement) oppgaven. En studie fant at TOJ oppgaven var mer sensitiv til den spatiale separeringen av stimuli, mens SJ oppgaven var mer sensitiv til den temporale separeringen av to stimuli (Vatakis et al., 2008). SJ oppgaven, som vil bli brukt i denne forskningen, vil inneholde presentasjoner av AV stimuli med ulik grad av synkronitet, audioledende, eller visuelt ledende stimuli, hvor deltaker skal trykke på knapper ettersom de mener at stimuliene er synkron eller ikke. Det er for å kunne finne ut hvordan sensitiviteten til

deltakerne er til asynkrone stimuli, samtidig som man finner ut om stimuliene som blir presentert i forkant kan ha en mulig påvirkning på svaret. Etter en slik oppgave er det vanlig å plassere resultatene i en Gaussisk kurve, som er en normalfordeling av svarresponsene. En fordel med SJ oppgavene er at de kan gi et mer reliabelt estimat på PSS, fordi den ikke er like utsatt for responsbias. Likevel er en mulig responsbias ved SJ oppgaver, at deltakerne antar på forhånd at to stimuli hører til den samme multisensoriske hendelsen, og dermed trykker på knappen med «synkron». I TOJ (temporal order judgement- task) sitt tilfelle ville det ikke være sannsynlig at stimulusen har en bias på deltakerne med tanke på at svarresponsen bikker mot enten «audio først» eller «visjon først» (Vatakis et al., 2008).

I etterkant av en SJ-oppgave vil dataen kunne få tilpassede kurver for videre analyser. To av kurvene som blir brukt knyttet til slike oppgaver er Gaussisk kurve og s- kurver. Yarrow et al. (2011) forklarer at ved bruk av to s- kurver, vil stimulusens samtidighet bli definert av forskjellen mellom de to kumulative kurvene. I stedet for å se på PSS på den Gaussiske kurven, som kun er én parameter, får dermed to uavhengige s- kurver vise den opplevde samtidigheten ved å se på vinduet mellom to parameter, ALT og VLT, altså omfanget mellom de to kriteriene. Yarrow et al. (2011) mener at den oftest brukte Gaussiske kurven, kun er brukt for enkelthets skyld.

Studiets formål

I tidligere forskning har det vist seg at PSS har et relativt tydelig mønster, mens ALT og VLT har mindre litteratur som viser til tydelige svar. I tillegg er det tidligere brukt to forskjellige kurvetilpasninger til datamaterialet tilhørende SJ- oppgaver. I denne studien vil det derfor bli kjørt en ANOVA analyse på om PSS, ALT og VLT endrer seg i henhold til SOA-1. Basert på de tidligere funnene til Roseboom (2019) vil det forventes at det vil være en signifikant økning i den absolutte verdien til PSS for videoledende SOA-1, sammenlignet med når audio kommer før video og synkron SOA-1. I tillegg vil denne hypotesen også støtte seg på Van Der Burg et al. (2015) sine funn om at PSS vil skiftes mot den ledende sansen. I den aktuelle studien vil det bli brukt kurvetilpasninger ved å benytte både Gaussisk og s- kurver for å se på forskjellen dette fører med seg.

Forskning viser også at det er en forskjell i hvordan VLT og ALT endres i forhold til SOA- 1 (Alm & Behne, 2013). Funnene tyder på at VLT vil skiftes mot å bli mer videoledende ved VL SOA-1, altså terskelen for å oppleve VL asynkroniteter øker. ALT har vist seg å være mer stabil på grunn av læring og en større sensitivitet for AL stimuli, noe som kan skyldes erfaringer med VL stimuli (Alm & Behne, 2013) og på grunn av verdens

asynkroniteter og prosesseringshastigheter både i hjernen og i luft (Alm & Behne, 2013; Sugita, 2003). På grunnlag av disse funnene vil det forventes en mer videoledende VLT-verdi ved videoledende SOA-1, sammenlignet med synkron SOA-1. Samtidig forventes det ikke noen forskjell i ALT-verdien ved synkron SOA-1, sammenlignet med asynkron SOA-1.

Metode

Design

Norske studenters responser på SJ- oppgaver basert på AV justeringer av stavelsen /ba/, ble brukt for å måle hvilken grad den tidligere stimulusen påvirket det neste. Det var 21 AV- justeringer som ble presentert fra 400 ms audioledende (AL), til 400 ms videoledende (VL), eller synkront. De 21 AV- justeringene ble presentert 21 ganger, som til sammen utgjorde 441 unike trails. Etter eksperimentet ble datamaterialet plassert inn i en Gaussisk kurve hvor det ble ekstrahert ut tre parametere; VLT (video lead treshold), ALT (audio lead treshold) og PSS (point of subjective simultaneity).

Deltakere

Utvalget i studien besto av 30 frivillige studenter med norsk morsmål, som alle hadde en alder på mellom 20 til 28 år ($M = 23.03$, $SD = 1.63$), hvor 21 (70%) var kvinner, åtte (27%) var menn og en (3%) person hadde ikke oppgitt kjønn. Tre personer ble ekskludert fra eksperimentet da de ikke oppfylte kravene for studiet. Basert på en poweranalyse for estimering av utvalgsstørrelse gjort ved bruk av SPSS, viste det at $N = 6$ var nok til en power på 99%, hvis det ble forventet medium effekt på signifikanskriteriet $\alpha = .05$ og power = .95. Det betyr at $N = 30$ er mer enn nok for å teste studiens hypotese. Individene måtte gjennom en rekke pre- tester før de kunne utføre selve eksperimentet. De fikk derfor først utdelt et samtykkeskjema som måtte skrives under på. Deretter bekreftet de sitt morsmål og alder gjennom et spørreskjema. Deltakerne bekreftet sin hendthet gjennom en hendthetstest (Oldfield, 1970), hvor de som var høyrehendte ble med i eksperimentet. For å fastsette deltakernes syn, gikk alle gjennom en Snellen synstest hvor deltakere med en binokulær synsskarphet på 20/ 25 eller bedre kunne inkluderes i eksperimentet (Watt, 2003). Dette tilsvarte å kunne lese gjennom linje syv på Snellen- testen. En hørselstest i form av en audiometriprosedyre ble gjennomført for å avgjøre hvem som oppfylte eksperimentets hørselskrav og kunne delta. Deltakere med en hørsel på under 15 dB og nedover, fra 250 til 4000 Hz ble inkludert. En øyedominstest ble utført, uten at resultatet kunne utgjøre et grunnlag for ekskludering fra eksperimentet.

Materiale

I forkant av eksperimentet ble det produsert et audiovisuelt opptak som skulle presenteres for deltakerne, slik at deres sensitivitet for asynkroniteter kunne bli målt (Alm & Behne, 2013). Det audiovisuelle opptaket som ble brukt, var av en ung kvinne med norsk morsmål, hvor alle smykker og briller som kunne være forstyrrende ble tatt vekk (Alm & Behne, 2013). Opptaket ble utført i et lydisolert rom, med en grå monoton vegg i bakgrunn for å unngå støy som kunne ha en innvirkning på opptaket og opplevelsen av lyd kvalitet. Kvinnen ble også bedt om å snakke uten særlig intonasjon og med minimalt med blinking og ansiktsuttrykk (Alm & Behne, 2013). Kvinnen satt to meter foran et PDWF800 Sony Professional XDCAM HD422 Camcorder kamera hvor to mikrofoner i typen Røde NT1-A var plassert utenfor kameraets rekkevidde. Den ene mikrofonen var koblet til kameraet og lyden fra den andre gikk via et RME FIREFACE 400 til en Apple Macintosh G5 datamaskin, hvor lyden deretter ble tatt opp ved bruk av Praat versjon 5.1 (Boersma & Weenink, 2009). Stavelsen /ba/ ble brukt fordi det er en stavelse hvor artikuleringen vises tydelig på leppene, i motsetning til en /ga/, hvor lyden blir produsert bak i ganen, noe som gjør at /ba/ er et bedre alternativ i forbindelse med AV synkronitetsmåling (Alm & Behne, 2013). /Ba/ ble justert med 21 forskjellige «stimulus onset» asynkroniteter (SOA) fra 400 ms audioledende stimuli, til 400 ms videoledende stimuli. Ti av stimuliene var audioledende, ett var synkront og 10 var videoledende.

Prosedyre

Eksperimentet ble utført i NTNUs talelaboratorium på Dragvoll i Trondheim. Det aktuelle studiet tok utgangspunkt i et eksperiment som bestod av flere SJ oppgaver. Eksperimentet ble delt opp i tre deler på grunn av at eksperimentets kvalitet i presenteringen av stimuliene skulle være optimalt. De 30 deltakerne som ble kvalifisert etter eksperimentets kriterier, ble plassert i en stol med fire ben for å sikre at avstanden til skjermen forble som den skulle. Videoen av stimuliene som ble brukt, ble presentert ved hjelp av SuperLab, på en iMac- skjerm (27 tommer, 5120 x 2880 piksler, AMD Radeon R9 M295X 4GB graphics) som var plassert omtrent 70 cm fra øynene til deltakerne. De visuelle stimuliene av stavelsen /ba/ ble presentert på skjermen, mens de auditive stimuliene ble presentert gjennom AKG K271 stereolukkede dynamiske øreomsluttende studio- hodetelefoner på 68 ± 1 dBA. SuperLab versjon 2 ble brukt for å presentere stimuliene i tre blokker av trials, med i alt 450 trials. Den første blokken hadde 144 stimuli, den andre på 153 stimuli og den siste blokken hadde 153 stimuli. Det ble flere stimuli tilsammen fordi de ni stimuliene som ble gitt like før pausene

måtte bli presentert på nytt etter pausen for å kunne måle påvirkningen den fikk på den påfølgende stimulusen.

Hver blokk hadde randomiserte stimuli for alle deltakerne, slik at alle de 21 stimuliene ble presentert én gang ved siden av hver og en av de andre AV-justeringene. MatLab (R2021b) ble brukt for randomiseringen av de 21 stimuliene. Dette ble gjort for å kunne måle hvilken påvirkning den foregående stimulusen hadde på den neste. Det som skjedde i en trail, var at et stimuli med en AV-justering ble presentert for deltaker som skulle avgjøre om stimulusen var synkron eller ikke, altså en SJ-oppgave (simultaneity judgement task). Ved å se på responsen tilhørende den neste presenterte stimulusen ble det målt om den foregående AV-justeringen endret oppfattelsen av synkroniteten eller rekaliserte på grunn av SOA-1. Deltakerne fikk beskjed om å gi respons ved bruk av en Cedrus RB responsboks (730/740 modell) som var utstyrt med en knapp for opplevd synkronitet og en for opplevd asynkronitet. Boksene varierte med å ha synkron-knappen til høyre eller til venstre på boksen, for å unngå mulige responsbiaser ved responderingen av stimuliene knyttet til at for eksempel alle deltakerne var høyrehendte. I tillegg var det en blå knapp på boksen som deltakerne kunne bruke for å trykke seg videre etter instruksjoner i eksperimentet. Eksperimentet tok i alt omtrent 60 til 70 minutter, inkludert pretester, hvor utføringen av selve eksperimentet tok omtrent 20 til 30 minutter.

Resultat

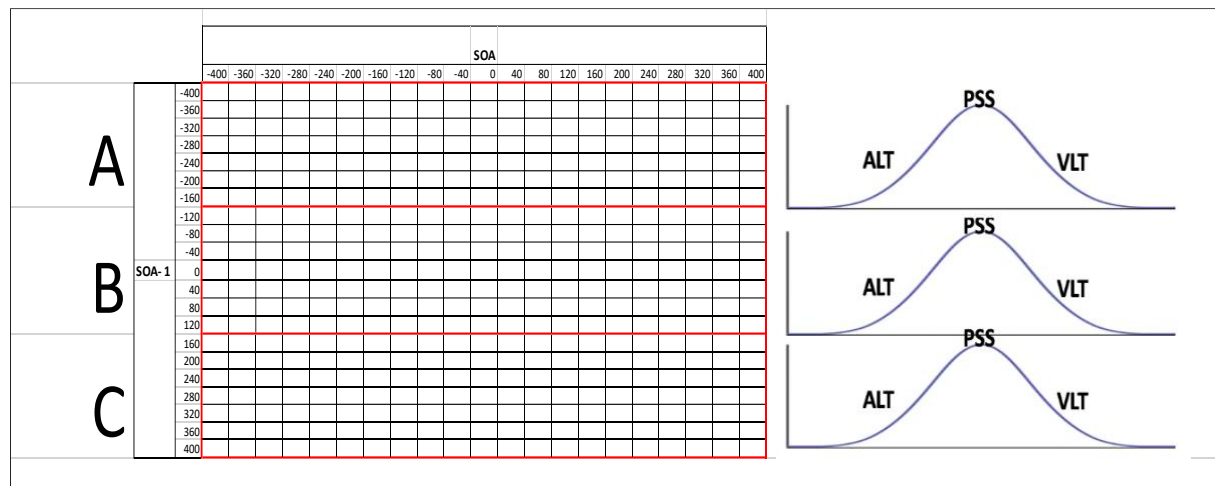
Datahåndtering

Etter eksperimentet var gjennomført, ble SOA-ene derandomisert basert på SOA-1. Det ble gjort dataekstraksjon for hver deltaker i MatLab, hvor prosentandelen synkrone responser ble kalkulert for SOA på tvers av alle SOA-1. To s-kurver ble tilpasset prosentandelen synkrone responser (y) for SOA (x). Det ble valgt s-kurver for å tilpasse dataen best mulig slik at ALT og VLT kunne forbli uavhengig av hverandre. Deretter ble ALT ekstrahert på den ene s-kurven og VLT på den andre, noe som utgjorde grunnlaget for å inkludere eller ekskludere deltakerne. Grunnen til at det ble satt kriterier på dataen til hver deltaker var at det måtte fremgå terskler for AL og VL. I denne prosessen ble en av deltakernes responser ekskludert fra studien grunnet svært usannsynlige svar som skilte seg sterkt fra normalen og ikke ga tydelige VL og AL terskler. Dette førte til at utvalget ble $N = 29$. Basert på ALT og VLT ble PSS og FWHM (full width of half maximum) kalkulert. FWHM ble kalkulert ut ifra avstanden mellom ALT og VLT, mens PSS ble funnet ved å se på krysningspunktet til de to s-kurvene.

Den endelige datafilen med alle deltakernes responser knyttet til de 21 SOA- 1 ble delt inn i grupper som tilsvarte AL- stimuli (-400 ms til -160 ms), VL stimuli (160 ms til 400 ms) og synkrone (-120 ms til 120 ms). Audioledende stimuli var gruppe A, de omtrent synkrone stimuliene ble til gruppe B, og de siste syv videoledende stimuliene ble gruppe C (se figur 2). Deretter kunne en ANOVA bli kjørt på alle de tre parameterne ALT, PSS og VLT på tvers av de tre gruppene for å se om det forelå en forskjell mellom VL, AL eller synkron SOA-1.

Figur 2

En illustrasjon over inndelingen av de 21 SOA- 1 inn i tre grupper.



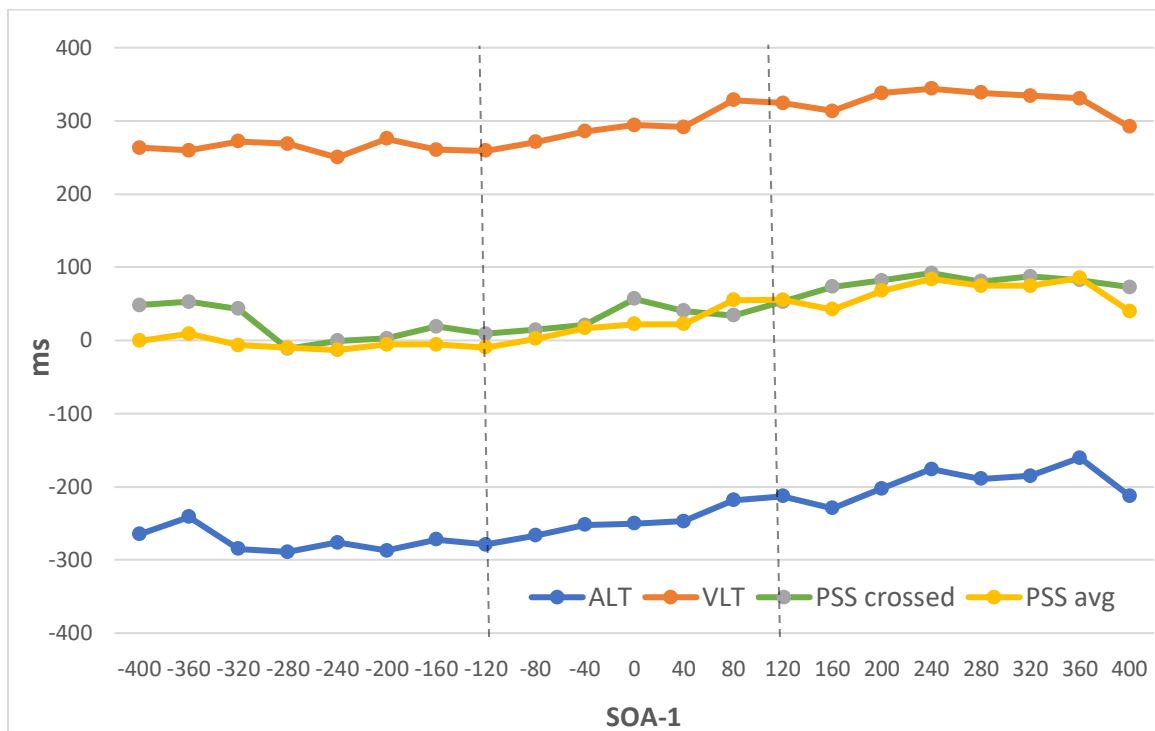
Notat. Dette viser hvordan de ulike SOA- 1 for alle deltakerne ble inndelt i tre grupper (A, B og C). De første syv SOA-1 (-400 ms til -160 ms) ble kalt gruppe A, som representerte de audioledende stimuliene. Gruppe B besto av de neste syv SOA- 1 (-120 ms til 120 ms) som utgjorde de synkrone stimuliene. Til slutt ble de syv siste SOA- 1 (160 ms til 400ms) plassert i gruppe C, som tilsvarte de videoledende stimuliene.

Statistiske funn

Målet med denne studien var å finne ut om den videoledende, audioledende eller synkrone stimulusen som ble presentert før den neste, hadde en effekt på ALT, VLT og PSS (se Figur 3). Datamaterialet fra SJ- oppgavene ble analysert med IBM SPSS Statistics Data Editor versjon 27. Fire enveis Repeated Measures ANOVA ble utført for parameterne ALT, VLT, PSS avg og PSS cross som avhengig variabler mellom hver av de tre gruppene av SOA-1 som de uavhengige variablene.

Figur 3

En oversikt over endring i ALT, VLT, PSS i forhold til forandret SOA-1.

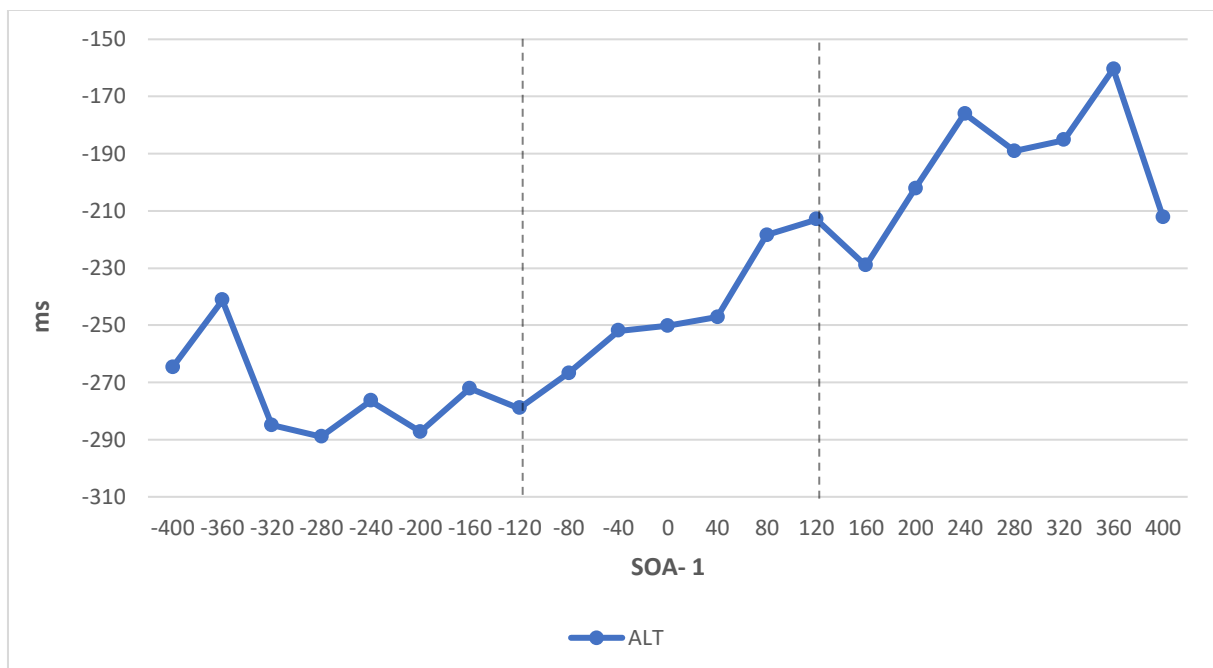


Notat. En oversikt over hvordan alle tre parameterne endrer seg i sammenheng med SOA-1.

SOA- 1 hadde en signifikant hovedeffekt på ALT, $F(1.55, 43.30) = 15.26, p < .001$ (se figur 4). Når SOA- 1 ble mer videoledende, ble ALT mer videoledende. Den største forskjellen lå mellom gruppe A (audioledende SOA- 1) ($M = -266$ ms, $SD = 109$) og gruppe C (videoledende SOA- 1) ($M = -224$ ms, $SD = 74$), $\Delta M = 42.55, p < .001$. Hypotesen antok at det ikke skulle foreligge noen signifikante forskjeller i ALT når SOA- 1 var synkron og SOA- 1 var asynkron. Mellom gruppen for synkron SOA- 1 (gruppe B) ($M = -243$ ms, $SD = 86$) og VL SOA- 1 ble det funnet en signifikant forskjell, $\Delta M = 19.04, p = .008$. For gruppe B og A ble det også funnet en signifikant gruppeforskjell, $\Delta M = 23.51, p = .013$.

Figur 4

Endring i ALT avhengig av forandringen i SOA- 1.

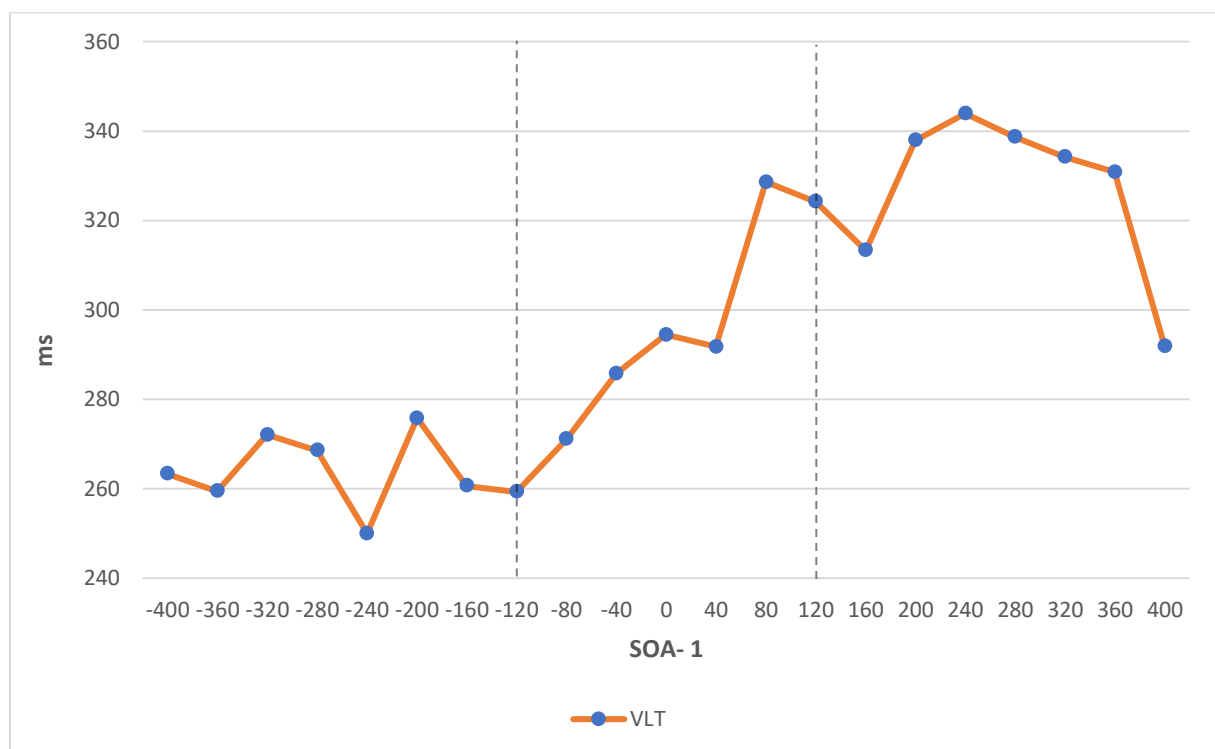


Notat: En grafisk fremstilling av hvordan ALT- verdien endrer seg i sammenheng med de 21 SOA- 1 audio- video justeringene, fra audioledende til venstre (-400 ms) til videoledende til høyre (400 ms).

SOA- 1 hadde en signifikant hovedeffekt på VLT, $F(1.44) = 16.77, p < .001$ (se figur 5). VLT- verdien ble funnet å være signifikant større ved videoledende SOA- 1 (gruppe C) ($M = 277$ ms, $SD = 68$) enn ved synkron SOA- 1 (gruppe B) ($M = 260$ ms, $SD = 74$), $\Delta M = 16.61, p = .001$. Den største forskjellen ble funnet mellom AL SOA- 1 (gruppe A) ($M = 246$ ms, $SD = 81$) og VL SOA- 1, $\Delta M = 30.01, p < .001$.

Figur 5

Endring i VLT avhengig av forandringen i SOA- 1.



Notat: En grafisk fremstilling av hvordan VLT- verdien endrer seg i sammenheng med de 21 SOA- 1 audio- video justeringene, fra audioledende til venstre (-400 ms) til videoledende til høyre (400 ms).

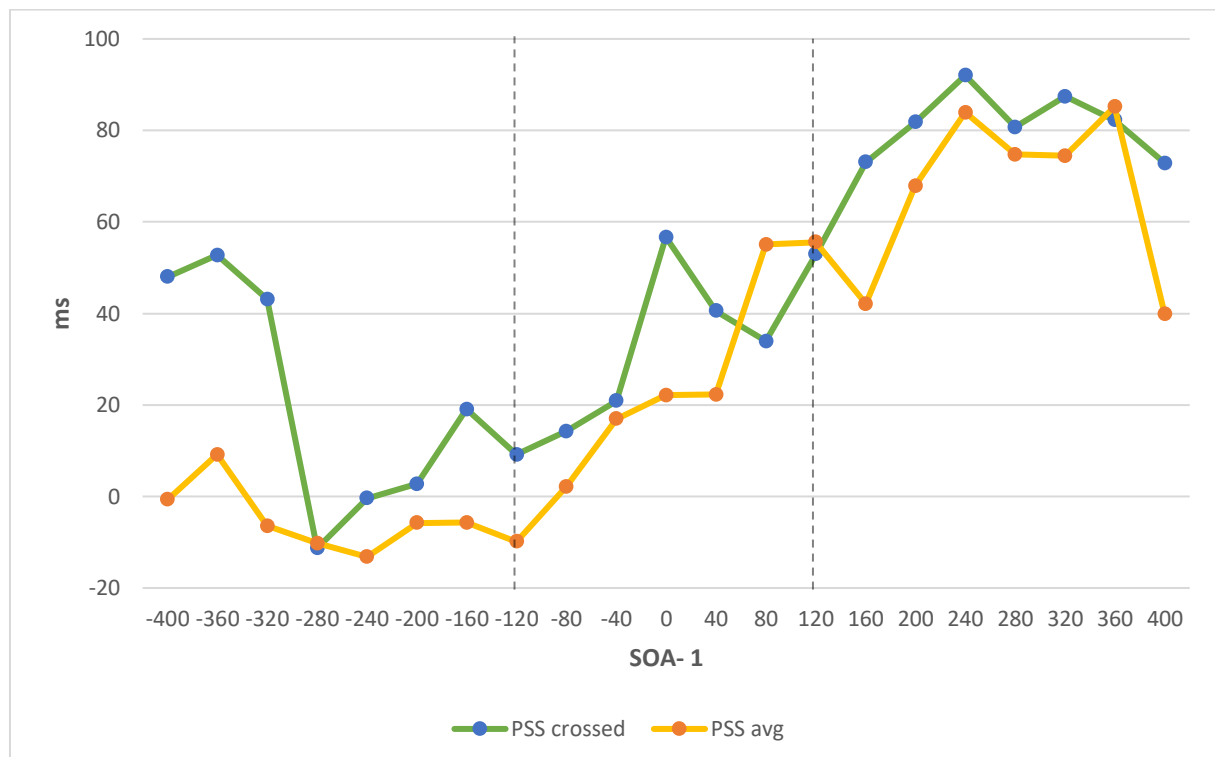
For å ekstrahere PSS fra deltakernes responser, ble det brukt to metoder (se figur 1). Den ene måten å finne PSS på, var å bruke en Gaussisk kurve hvor først avstanden mellom ALT og VLT, «full width at half Maximum» (FWHM), ble funnet for deretter å dele avstanden på to slik at tallet viser midtpunktet mellom ALT og VLT på toppen av den Gaussiske kurven. Den andre metoden for å ekstrahere PSS var ved bruk av å tilpasse to s- kurver til datamaterialet, og deretter finne krysningspunktet for de to kurvene. I hypotesen ble

det antatt at PSS skulle være mer videoledende når SOA- 1 var videoledende, sammenlignet med synkron og audioledende SOA- 1. Den uavhengige variabelen SOA- 1 hadde en signifikant hovedeffekt på PSS ekstrahert fra Gaussisk kurve, $F(1.49) = 22.11$ $p < .001$, hvor den største gruppeforskjellen forelå mellom AL SOA-1 (gruppe A) ($M = -10$ ms, $SD = 64$) og VL SOA- 1 (gruppe C) ($M = 26$ ms, $SD = 44$), $\Delta M = 35.89$, $p < .001$. Det ble også funnet en signifikant forskjell mellom gruppe B ($M = 8$ ms, $SD = 54$) og C, $\Delta M = 17.82$, $p = .001$ (se figur 6).

For PSS ekstrahert ved bruk av tilpassede s- kurver, ble det ikke funnet noen signifikant hovedeffekt av SOA- 1, $F(2) = 0.92$, $p = .403$. Det ble heller ikke funnet noen signifikante gruppeforskjeller mellom gruppe A ($M = 21$ ms, $SD = 69$) og C ($M = 36$ ms, $SD = 47$), $\Delta M = 15.57$, $p = .753$, eller B ($M = 33$ ms, $SD = 53$) og C, $\Delta M = 3.51$, $p = 1.000$.

Figur 6

Endring i PSS crossed og avg. avhengig av forandringen i SOA- 1.



Notat: En sammenligning av de to målene av PSS i henhold til hvordan de endrer seg i sammenheng med de 21 SOA- 1 audio- video justeringene, fra audioledende til venstre (-400 ms) til videoledende til høyre (400 ms).

Analysen av ALT, VLT og PSS på tvers av de tre inndelingene/ gruppene av SOA- 1 viste at det var signifikante forskjeller mellom gruppene, altså en endring i verdiene på grunnlag av forandringen i SOA- 1, for alle parameterne. Endringen var konsistent med at når SOA- 1 ble mer videoledende, ble også parameterne skjøvet mer mot VL sammenlignet med audioledende SOA- 1. Et oppsiktsvekkende funn var at PSS cross ikke viste signifikante forskjeller mellom gruppene av SOA- 1. Dette viser at kurvevalget for PSS vil utgjøre en effekt på resultatene (se figur 6).

Diskusjon

Tidligere forskning på oppfattelsen av synkronitet i audiovisuell talepersepsjon foreslår at det forekommer en hurtig recalibrering som gjør at den foregående stimulusen kan påvirke oppfattelsen av synkroniteten til den påfølgende. I hypotesen ble det antatt at den tidligere stimulusen, altså SOA- 1, hadde en påvirkning på oppfattelsen av den påfølgende stimulusen. Det ble antatt at SOA- 1 ville påvirke parameterne VLT og PSS, mens det ble forventet at ALT ikke skiftet. Det ble funnet en økning i VLT- verdien når SOA- 1 ble mer videoledende sammenlignet med synkron SOA- 1. I tillegg ble det funnet en mer videoledende PSS når SOA- 1 var videoledende, sammenlignet med audioledende og synkron SOA- 1. Dette funnet ble kun signifikant da målet ble ekstrahert ved hjelp av en Gaussisk kurve. Da det ble brukt to S- kurver for å ekstrahere PSS, ble det ikke funnet noen signifikante endringer mellom audioledende, synkron eller videoledende SOA- 1. I motsetning til hypotesen, ble ALT mer videoledende da SOA- 1 var videoledende, sammenlignet med synkron og audioledende SOA- 1. Likevel tyder disse funnene på at det forekommer hurtig recalibrering i etterkant av presentering av asynkron AV stimuli.

Effekten av ekstraksjonsmetodene av PSS

Slik det ble formulert i hypotesen, ble det funnet en signifikant forskjell i PSS (ekstrahert fra en Gaussisk kurve) fra synkron, til videoledende og audioledende SOA- 1. Forskjellen mellom gruppene viste at når den foregående stimulusen var videoledende, ble punktet for opplevd synkronitet skiftet mot den videoledende siden ($M = 26$ ms). Den ble skiftet mot den videoledende siden i større grad enn den ble skiftet mot den audioledende siden ved presentert audioledende stimuli ($M = - 10$ ms). Det kan også bli vist ved å se på PSS verdien i gruppen det ble presentert synkron SOA- 1, som også ligger på videoledende side ($M = 8$ ms). Dette kan tyde på at på grunn av det naturlige forsinkelsen verden byr på, vil

det være mer naturlig at mennesket i en slik situasjon har lettere for å recalibrere mot den videoledende siden, fordi det er mer vant til at lys kommer før lyd (Sugita, 2003). Siden PSS ligger nærmere punktet for fysisk synkronitet ved presentert AL asynkron stimuli enn det gjør ved VL asynkron stimuli, tyder det på at større asynkroniteter er tillatt på videoledende side i opplevelsen av synkronitet og at mennesket recalibrerer i større grad etter videoledende asynkroniteter enn audioledende asynkroniteter. Funnene støttes av Roseboom (2019) sine funn om at PSS ikke vil skifte mot audioledende side ved AL SOA-1 i like stor grad som den skifter mot videoledende side ved VL SOA- 1.

Et annet interessant funn i forbindelse med analysen av PSS var at det ikke ble funnet noen signifikante funn da det ble brukt to s- kurver for å ekstrahere PSS. En mulig grunn for dette kan være at når det brukes to s- kurver for å tilpasse datasettet, vil det forekomme mye individuell varians. S- kurvene vil bedre passe til dataen kontra en Gaussisk kurve, som kun er en normalfordelingskurve (Yarrow et al., 2011). Hvor de to s- kurvene krysser, vil kunne variere veldig og avhenger av hvor spredt responsen til deltakerne er. Selv om kurvene treffer dataen mer presist enn det en Gaussisk kurve ville ha gjort, vil ekstraksjonen av PSS komme av krysningpunktet mellom de to s- kurvene, kontra midtpunktet mellom ALT og VLT på toppen av en Gaussisk kurve. I tillegg vil krysningpunktet også være avhengig av helningen på ALT- og VLT- kurvene, hvor funnene i eksperimentet indikerte en forskjell mellom dem. På grunnlag av variasjonen i PSS- målene knyttet til de to kurvetilpasningene i etterkant av SJ- oppgavene i den aktuelle studien, kan det argumenteres for hvorvidt PSS er et det beste målet på audiovisuell synkronitet.

I den forstand kan et forslag være at ALT og VLT alene vil være parametere som er bedre egnet til de mer presist tilpassede s- kurvene for audiovisuelle data. Det er likevel viktig å huske at PSS er en kjent parameter som er anerkjent og hyppig brukt, selv om det nå argumenteres for at det er mulig at andre mål kan være bedre egnet. PSS kan være bra på den ene siden, hvor det er enkelt å ekstrahere ut fra en normalfordeling, på det punktet hvor det er mest sannsynlig at flest opplever synkronitet. Det som er ulempen på den andre siden, vil være at ved bruk av de enda bedre tilpassede s- kurvene, vil variasjonen i responser fra deltakerne kunne forårsake at PSS ikke alltid vil fremstå som et signifikant resultat, slik det ble funnet i dette eksperimentet. En av grunnene til et svakt PSS- resultat, var at helningen mellom ALT- grafen og VLT- grafen ikke var lik. En ulik helning kan gjøre at krysningen (der hvor PSS finner sted) ikke alltid blir midtpunktet mellom ALT og VLT slik som ved en Gaussisk kurve, men ytterligere forskjøvet til en side, noe som igjen kan påvirke resultatet.

ALT og VLT ekstrahert fra s- kurver som mål på rekalkibrering

Slik det ble nevnt ovenfor, vil s- kurver være mer presist for ALT og VLT, fordi ALT og VLT alltid vil være uavhengige av hverandre. Derfor kan det argumenteres for at bruken av s- kurver ved disse parameterne vil kunne være et bedre alternativ for å treffe datasettet mer presist. Da vil også ALT og VLT kunne ha de individuelle forskjellene de har, uten at en normalfordelingskurve skal plasseres best mulig til gjennomsnittet og utjevne forskjellene mellom verdiene til ALT og VLT (Yarrow et al., 2011). Funnet i dette eksperimentet var at VLT, slik hypotesen tilsa, ble mer videoledende hvis den foregående stimulusen var videoledende. Samtidig ble VLT lavere, altså mer mot det fysiske punktet for synkron, hvis den foregående stimulusen var audioledende. Med andre ord vil dette bety at hvis den tidligere presenterte stimulusen var videoledende, skal det mer til for individet å oppdage den videoledende asynkroniteten i den påfølgende. Altså terskelen for å oppdage asynkronitet øker og går mer mot den videoledende siden (se figur 1 og 5). Disse funnene tyder derfor på at individet har rekalkibrert og dermed tillater mer asynkroniteter i stimuliene som oppleves som synkrone. Derimot hvis den foregående stimulusen var audioledende, ble terskelen for å kunne oppdage en videoledende stimuli lavere og sensitiviteten økte litt.

Dette funnet er konsistent med ALT. Det ble funnet en lik tendens for terskelen for å oppdage audioledende stimuli (se figur 4). Terskelen økte og gikk mer mot audioledende, jo mer audioledende den foregående stimulusen var. Disse funnene tyder på at det ble funnet hurtig rekalkibrering fra den foregående stimulusen, til oppfattelsen av den neste, ved at mennesket tillater mer asynkroniteter selv om de oppfattes som synkrone, også for ALT. Samtidig ble terskelen for å oppdage audioledende stimuli mindre streng (mer sensitiv) når den foregående stimulusen var videoledende.

Det som er interessant med funnene til ALT og VLT er at i gruppe B, altså gruppen for synkrone stimuli, var VLT ($M = 260$ ms) strengere enn ALT ($M = -243$), hvor terskelen for å oppdage audioledende stimuli ligger mer mot punktet for fysisk synkronitet enn terskelen for å oppdage videoledende stimuli. Det tyder på at terskelen for både VLT og ALT er skjøvet mot videoledende side, selv ved presentert synkron stimulus. En mulig forklaring på dette kan være ved hjelp av Sugita (2003) sitt eksperiment som målte hvor mye forsinkelse lyden kunne ha før deltakerne oppdaget forsinkelsen. Der ble det funnet at lyd trengte 120 ms på å reise 40 meter, hvor terskelen for å oppdage asynkronitet mellom flash og beep var 106 ms (Sugita, 2003). Dette peker igjen mot at verden byr på utallige naturlige videoledende asynkroniteter som hjernen hele tiden vil rekalkibrere slik at det kan oppfattes som en helhet. På grunn av at mennesket er vant til at video kommer før lyd i mange tilfeller, kan det være en grunn til at

den videoledende terskelen er høyere eller strengere enn den audioledende terskelen. Funnene for tersklene for å oppdage audioledende og videoledende stimuli tyder dermed på at vi er mer sensitive for audioledende asynkroniteter, fordi rekalkibrering foregår hyppig på den videoledende siden. Det gjør at hvis den foregående stimulusen var audioledende, vil ALT øke og bli litt strengere, men ikke i like stor grad som den videoledende terskelen (VLT) øker ved tidligere presentert videoledende stimulus.

Et annet poeng å ta i betraktning, sett at ALT og VLT er de prefererte målene når det kommer til AV synkronitet, kan være å se på vinduet for opplevd synkronitet og hvordan det flytter seg i henhold til SOA- 1, isteden for PSS når det brukes s- kurver. Synkronitetsvinduet vil være mellom ALT og VLT på grafene (se figur 1). Alm and Behne (2013) argumenterer for at synkronitetsvinduet er asymmetrisk, hvor det vil være enklere å oppdage asynkron audioledende stimuli ovenfor å oppfatte asynkront visuelt ledende stimuli. Samtidig har forskning funnet at mennesker er mer sensitive til audioledende stimuli enn videoledende stimuli og at asymmetrien i synkronitetsvinduet eksisterer (Cecere et al., 2016). Disse resultatene støttes av funnene i den aktuelle studien hvor terskelen er lavere for å oppdage AL asynkrone stimuli enn VL, noe som indikerer at synkronitetsvinduet ligger mer på videoledende siden enn på den audioledende siden. Det å se på et slikt vindu, kan være en like god indikator på oppfattet samtidighet som å se på PSS, fordi det viser både terskelen for hvor mye audioledende og videoledende et stimuli kan være for å detektere en asynkronitet og hvor vinduet for opplevd synkronitet ligger. Dermed viser et slikt vindu når mennesket integrerer multisensoriske stimuli til en helhet, ikke bare hvor flest individ vil svare at de opplever to AV stimuli som synkrone. Knyttet til SOA- 1 vil et alternativ til å se på hvordan PSS skifter i henhold til den foregående stimulusen derfor være å se på hvordan synkronitetsvinduet skifter plass eller størrelse.

Den audioledende terskelen hadde en større endring enn det som ble forventet. Det ble antatt at ALT ikke skulle endres på tvers av gruppe A, B og C, altså på tvers av de ulike SOA- 1 inndelingene. Tvert imot ble det funnet en større endring i ALT mellom gruppen for audioledende SOA- 1 og videolende SOA- 1 enn det ble funnet for VLT mellom de to gruppene (A og C). Dette funnet viser at s- kurven for ALT har en større helning enn kurven for VLT. Det vil si at terskelen for å oppleve audioledende stimuli ble mer påvirket av SOA- 1 enn det terskelen for å oppleve videoledende stimuli ble. Dette kan tyde på at hypotesen er basert på forskning som er gjort på eldre mennesker, hvor det ble funnet at middelaldrende voksne har en mer konservativ og stabil ALT (Alm & Behne, 2013). I tillegg ble det antatt at AL stimuli var lettere å predikere og dermed lettere å lære, som også fikk ALT til å optre

mer stabilt enn VLT. Utvalget i dette eksperimentet var NTNU- studenter med en gjennomsnittsalder på 23 år, mens studien det ble tatt utgangspunkt i undersøkte middelaldrende voksne. I studien ble det argumentert for at ALT ble mer stabil på grunn av økt erfaring med audiovisuell binding med alderen (Alm & Behne, 2013). En økt ALT-stabilitet med alderen, kan derfor stemme med at det ble funnet en desto mer ustabil ALT som varierte mer enn VLT i det aktuelle eksperimentet, på grunn av det unge utvalget, som ikke enda har hatt nok erfaring med VL asynkroniteter til at toleransen for AL har blitt mindre. Hadde det samme eksperimentet blitt gjort på et eldre utvalg, hadde kanskje hypotesen hatt et bedre grunnlag for å bli beholdt. Det kan også se ut til at hypotesen ble basert på Roseboom (2019) sine funn om at PSS ikke ville endres ved AL stimuli, noe som ikke vil bety det samme som at ALT ikke ville endres.

Konsistente funn og videre forskning

Tross funnene for PSS cross, ble funnet for PSS avg funnet å være konsistent med andre beep flash studier, hvor et beep som kommer etter et flash gjør at PSS skifter mot samme retning, noe som får neste beep som kommer etter flash til å oppfattes som synkront (Yarrow et al., 2011). På lik måte, støtter Van Der Burg et al. (2013) sin antakelse om PSS sitt skifte mot den ledende sansen våre funn. Med tanke på diskusjonen vedrørende om PSS var et godt mål hvis det ble brukt s- kurver isteden for Gaussiske kurver, ble synkronitetsvinduet mellom ALT og VLT skiftet på lik måte som antakelsene om PSS. Funnene for både ALT og VLT indikerte et skiftet mot den ledende sansen, som forklarer hvordan vinduet for opplevd synkronitet skifter i takt med recalibreringen knyttet til de foregående stimuliene.

Det som kan bli betraktet i videre forskning, er hvordan stimuliene som kommer enda tidligere har betydning for recalibreringen av asynkronitetene. I denne studien ble det kun satt søkelys på betydningen stimulusen like før hadde, uten å måle i hvilken grad de foregående stimuliene kunne ha påvirket den opplevde synkroniteten. I tillegg kunne videre forskning ha inkludert en gruppe med eldre mennesker i eksperimentet, for å bedre måle hvordan stabiliteten til ALT utvikler seg ved å se på forskjellen i endringen i ALT mellom eldre og yngre mennesker. Da har det vært mulig og testet hvordan langtidserfaring av VL stimuli kan ligge i grunn før recalibreringen og hvordan tersklene deretter flytter seg. Den aktuelle studien hadde basert hypotesen for ALT på et feil grunnlag, noe som førte til andre funn enn forventet. Likevel ga alle funnene indikatorer på hvordan recalibrering foregår og hvordan tersklene for å oppdage asynkroniteter endrer seg ved eksponering av asynkrone audiovisuelle stimuli.

Konklusjon

Studien viser at den tidligere presenterte stimulusen har påvirkning på oppfattelsen av den neste, ved at hjernen foretar en hurtig recalibrering i etterkant av presentering av asynkron AV stimuli. Dette stemmer overens med funn fra tidligere forskning som viser at mennesket recalibrerer i etterkant av eksponering av asynkroniteter, men at det også foreligger en forskjell mellom den audioledende og videoledende asynkronene stimulusen på grunn av erfaring med videoledende AV stimuli. I samsvar med tidligere forskning ble det funnet at både ALT og VLT blir strengere ved eksponering av den tilhørende ledende sansen. Det skal altså for eksempel mer VL asynkronitet til for å oppdage den, i etterkant av presentasjon av VL asynkron stimulus. Samtidig tyder funnene på at terskelen for å oppdage audioledende stimuli blir mer sensitiv når videoledende asynkronene stimuli blir presentert og den videoledende terskelen blir mer sensitiv ved presentasjon for audioledende stimuli. Videre tyder funnene for PSS på at to tilpassede s- kurver vil passe bedre til ALT og VLT slik at de behandles uavhengig av hverandre og ikke tvinges sammen i en Gaussisk kurve, men at dette ikke nødvendigvis gjør PSS til det beste målet, på grunn av at helningen til ALT- kurven og VLT- kurven er ulik og dermed krysser på varierende punkter. Derfor ble det foreslått at vinduet for opplevd synkronitet kan være en mer passende indikator på synkronitet enn PSS. Studien viser også at det er konsekvente funn på tvers av beep flash og tale stimuli. Resultatene gir implikasjoner på at hurtig recalibrering forekommer, men på grunn av en skjevhet i synkronitetsvinduet (mer mot videoledende side) vil det også tyde på at langtidserfaringer med videoledende stimuli vil ligge til grunn, selv om det forekommer en hurtig recalibrering i henhold til stimulusen som ble presentert i forkant.

Litteraturliste

- 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>
- Arnold, D. H., Tear, M., Schindel, R., & Roseboom, W. (2010). Audio-Visual Speech Cue Combination. *PLoS ONE*, 5(4), e10217. <https://doi.org/10.1371/journal.pone.0010217>
- Boersma, P., & Weenink, D. (2009). *PRAAT: Doing phonetics by computer (Version 5.1) (Computer program)*. In www.praat.org
- 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/10.1111/ejn.13242>
- Fujisaki, W., Shimojo, S., Kashino, M., & Nishida, S. Y. (2004). Recalibration of audiovisual simultaneity. *Nature Neuroscience*, 7(7), 773-778. <https://doi.org/10.1038/nn1268>
- Harrar, V., & Harris, L. R. (2008). The effect of exposure to asynchronous audio, visual, and tactile stimulus combinations on the perception of simultaneity. *Experimental Brain Research*, 186(4), 517-524. <https://doi.org/10.1007/s00221-007-1253-0>
- 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. Cognit. Brain Res.*, 25, 499.
- Oldfield, R. C. (1970). The assessment and analysis of handedness: the Edinburgh inventory. *Neuropsychologia*, 9, 97- 113.
- Pöppel, E., Schill, K., & von Steinbüchel, N. (1990). Sensory integration within temporally neutral systems states: A hypothesis. *Naturwiss.*, 77, 89.
- Redaksjonen for norsk APA- stil. (2021). Norsk APA- manual: En nasjonal standard for norskspråklig APA- stil basert på APA 7th (Versjon 1.7). Unit. In.
- 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>
- Sugita, Y. S., Yôiti. (2003). Audiovisual perception Implicit estimation of sound-arrival time. *421 (6926)*, 911-911.
- Van Der Burg, E., Alais, D., & Cass, J. (2013). Rapid Recalibration to Audiovisual Asynchrony. *Journal of Neuroscience*, 33(37), 14633-14637. <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(1), 53-59. <https://doi.org/10.1007/s00221-014-4085-8>
- Vatakis, A., Navarra, J., Soto-Faraco, S., & Spence, C. (2008). Audiovisual temporal adaptation of speech: temporal order versus simultaneity judgments. *Experimental Brain Research*, 185(3), 521-529. <https://doi.org/10.1007/s00221-007-1168-9>
- Vatakis, A., & Spence, C. (2006). Audiovisual synchrony perception for music, speech, and object actions. *Brain Research*, 134- 142.
- Vilares, I., & Kording, K. (2011). Bayesian models: the structure of the world, uncertainty, behavior, and the brain. *Annals of the New York Academy of Sciences*, 1224(1), 22-39. <https://doi.org/10.1111/j.1749-6632.2011.05965.x>
- Vroomen, J., & Keetels, M. (2010). Perception of intersensory synchrony: a tutorial review. *Atten Percept Psychophys*, 72(4), 871-884. <https://doi.org/10.3758/APP.72.4.871>
- Watt, W. S. (2003). How Visual Acuity Is Measured. <https://lowvision.preventblindness.org/2003/10/06/how-visual-acuity-is-measured/>
- 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. <https://doi.org/10.1016/j.concog.2011.07.003>