

THE EFFECT OF COGNITIVE FACTORS
ON CROSS-MODAL SYNCRONY
PERCEPTION

Nara Ikumi Montserrat

DOCTORAL THESIS UPF / 2016

SUPERVISOR

Dr. Salvador Soto-Faraco

Dept. de Tecnologies de la Informació i les Comunicacions



ACKNOWLEDGMENTS

En el meu pas pel departament, he tingut la oportunitat de gaudir i aprendre de molta gent. Vaig arribar sense saber en que consistia un doctorat, i menys sobre què volia dir investigar. El Phil, va iniciar-me en el món de la psicofísica, en breu, em vaig sentir atreta per un món desconegut, rodejada de llums, sons i foscor, on l'objectiu principal era que el participant fes una tasca sense saber que buscàvem. Val, ara ja sabia com havia de dissenyar-lo, com més confús fos millor! Aleshores em va arribar el torn, de plantejar el meu experiment, tenia la idea però.. com testejar-la? El Xavi em va obrir les portes a l'univers dels invents i recursos per mesurar el que necessitàvem. I això només va ser el principi, ha estat al llarg de tots aquests anys en cada un dels meus experiments. Així com la Silvia, la Cris i la Florencia, que han estat allí amb un somriure sempre que entrava atabalada per la porta.

En aquesta tesi, ha participat molta gent, des del participant més motivat per col·laborar en el coneixement de l'estudi del cervell als que s'adormien a la cadira després d'un llarg dia de feina, a tots els estudiants de màster dels qual he après coses importantíssimes. Això sí, sempre acompanyada per una colla de bons col·legues del departament que han fet pinya amb mi, en diferents moments del doctorat. Molts de vosaltres, heu realitzat i re-realitzat amb paciència els meus experiments. Crec que podria distingir perfectament algunes de les vostres corbes de simultaneïtat! Amb el Sancho i l' Alexis, he gaudit discutint experiments mentre caminaven per la muntanya, m'han ensenyat i ajudat molt en la realització dels meus estudis.

El meu món es va anar fent més ampli, amb l'Angèle, que em va fer treballar de valent, i va introduir-me a l'EEG, amb el Luis, que va ensenyar-me la realitat de la no reproducibilitat, però sobretot, s'ha

expandit durant els descansos i dinars cada dia, que m'han permès compartir inquietuts i moments important amb molts de vosaltres, molts dels quals us heu anat desplaçant cap a l'àmbit emocional. La Sara, Carol i Miguel, amb el seu humor i afecte, la Manuela, l'Alice, el Giovanni i el Manu, amb els seus ulls brillants per la ciència i per gaudir de Barcelona, la Maya, amb les seva força i empenta tan a la muntanya com al laboratori, la Marti, amb el seu amor i bon humor combinats amb tocs de dansa esbojarrada, i la Mireia, per ajudar-me incansablement, tan a nivell professional com en l'àmbit personal, m'ha fet créixer entre llibres i poesia durant el doctorat. I he comptat amb la espontaneïtat i tendresa tan natural de la Daria. Molts heu estat un pilar importantíssim durant tots aquests anys per mi, ha sigut fantàstic compartir diferents temps amb tots vosaltres, dins i fòra de la feina; fent birres, escalant, buscant bolets, ballant swing, de calçotades..

Agrair al Salva, la seva paciència i guiatge per aquest camí. He après moltíssim amb ell, i ha aconseguit crear un clima al grup, que m'ha permès créixer en molts àmbits.

I per acabar, agrair el suport de la meva família, que m'ha permès saltar moltes vegades, sense tenir por a caure. Al Bruno, perquè m'ha fet jugar i ensenyat a gaudir d'aquest procés, i al Lolo per treure'm de passeig per la ciutat.

ABSTRACT

Perception in our everyday life takes place in multisensory environments, and hence involves the processing of a multitude of signals captured by various sensory modalities. Given the different nature of the signals, understanding how all the information is combined in the brain to form an integrated percept is not straightforward. One of the paramount questions is how the considerable timing differences between sensory information processing are managed. For example, in the last years there has been a tremendous surge in interest to understand how the perceptual system elicits the subjective impression of synchrony for stimuli coming from different sensory modalities. Yet, most evidence so far concerns stimulus-related properties in simple contexts. The present dissertation addresses the influence of cognitive factors and participants' inner state (such as attention, action task demands, and ongoing brain rhythms) on synchrony perception between audio-visual events. In the first two studies of the dissertation, we have addressed the role of selective attention and action during cross-modal temporal recalibration. The results of these studies provide evidence that subjective simultaneity can be strongly modulated as a function of the focus of the observer's endogenous attention, in otherwise identical stimulation conditions. In the third study, we have recorded electroencephalographic activity while participants performed an audio-visual simultaneity judgment task for stimuli presented at different asynchronies. Our results show that the phase of low frequency neural oscillations, reflecting brain states prior to the occurrence of an audio-visual event, can predict perceptual variability in synchrony judgments. Overall, our results shed new light on how cognitive factors can modulate multisensory perception.

RESUM

La percepció del nostre entorn és multisensorial, és a dir, involucra el processament de senyals a través de diverses modalitats sensorials. Combinar aquesta informació en el cervell per tal de formar una percepció coherent i integrada és un procés complex, degut a la diferent naturalesa de les senyals. A més, això farà que el cervell hagi de resoldre diferències temporals durant el processament de la informació. En els últims anys, ha sorgit un profund interès per entendre com el sistema perceptiu genera la impressió de sincronia d'estímuls provinents de diferents modalitats sensorials. La major part dels estudis han examinat propietats de la percepció de sincronia relacionades directament amb els estímuls físics, en contextos molt simplificats. En aquesta tesi investigo la influència de factors cognitius i de l'estat intern de l'individu (com per exemple l'atenció, demandes en tasques motores, i els ritmes interns cerebrals) en la percepció de sincronia entre estímuls audiovisuals. En els primers dos estudis de la tesi, hem examinat la funció de l'atenció i les accions durant la recalibració temporal d'estímuls audiovisuals. Els resultats dels estudis mostren com la sincronia subjectiva pot ser fortament modulada en funció d'on es dirigeixi el focus atencional del participant, en condicions on l'estimulació física és idèntica. En el tercer estudi, hem enregistrat l'activitat electroencefalogràfica dels participants, mentre realitzaven una tasca de simultaneïtat. Durant aquesta tasca presentàvem diferents asincronies entre estímuls audiovisuals per tal d'estudiar la percepció de sincronia (vs. asincronia). Els resultats indiquen que la fase de les oscil·lacions neuronals, que reflecteixen estats cerebrals abans de la presència d'un estimul audiovisual, poden predir la resposta en quan a percepció de sincronia. En resum, els nostres resultats aporten coneixement sobre com alguns factors cognitius poden modular la percepció multisensorial.

PREFACE

Perception can be roughly defined as the process whereby the nervous system parses and organizes sensory information to become aware of the world surrounding us. In usual everyday life environments, perception involves the organization of information within and across sensory modalities. The process of perception requires a first stage, where physical signals from the environment are transformed into brain signals. As an illustration imagine yourself standing on a crosswalk. This multisensory experience entails seeing and hearing cars as well as people and birds passing by. At the same time, you might be smelling the scent from the coffee cup you are holding and sensing the pain of some of the muscles you stretched yesterday during the rugby match. That is, it involves the orchestration of vision, audition, touch, smell and proprioception in this case. Each sensory system has specialized receptors to detect particular external signals such as light, sound, vibrations, internal states such as muscle position and internal organs, along with distinct strategies to codify these signals into neural impulses. For instance, in the visual system, the photoreceptive cells of the retina transform light waves into neural impulses through the transduction process; whereas in the auditory system, sound waves in the cochlea are detected through mechanical vibrations and in the somatosensory system, a variety of skin receptors allows us to sense touch, vibration, pressure, temperature and pain. Further, a second stage of information processing takes place in the central nervous system. The neural impulses are transmitted across different sensory pathways and finally converge into a conscious and coherent representation of the world. At this point, we only have mentioned the so-called, bottom-up or stimulus driven processes. That is, how perception is influenced by the sensory signals coming from the external environment and captured by our sensory receptors. However,

this is only one part of the story, as perception can be also influenced by cognitive factors. Coming back to the crosswalk example, while you are standing there, a variety of internal thoughts occupy your mind: you might be worried about the meeting you have had this morning with your supervisor and, mentally trying to predict when the light is going to turn green. All these cognitive factors (or top-down processes) complicate even more the former picture of multisensory perception, but we speculate they might be relevant to help interpret how to combine the sensory signals and build a coherent representation of the environment.

Generally, it is argued that stimulus-related factors such as temporal proximity, along with spatial coincidence of events are key in order to determine how we combine (group or segregate) information across sensory modalities. However, in the audio-visual domain the exact temporal order of events at source and their subjectively perceived order might not correspond exactly, because of several sources of temporal variability. For example, in a summer storm lightning, the time when the light and the thunder roar are produced at the original source is called event time, which can be accurately measured. By the time the image of the light and the sound of the thunder reach the senses (at distances much larger than 15 m), the sound of the thunder is usually perceived a few seconds after the image of the light, making evident the delays due to differences in transmission speed of light and sound waves through air (~300 million m/s vs. ~340 m/s, respectively). In addition, alternative sources of temporal delays can take place inside the brain (brain time). These delays depend on the neural processing time of sensory information. For example, it is known that a simple sound, such as a brief tone, reaches the sensory cortex around 30-50 ms faster than a simple flash (Figure 1). Note however, that brain time can also be influenced by several factors such as the saliency of events due to contrast/luminance

changes along with the daylight conditions, expectancies generated by the prior knowledge about the causality between action and effects or by orienting of attention to specific events.

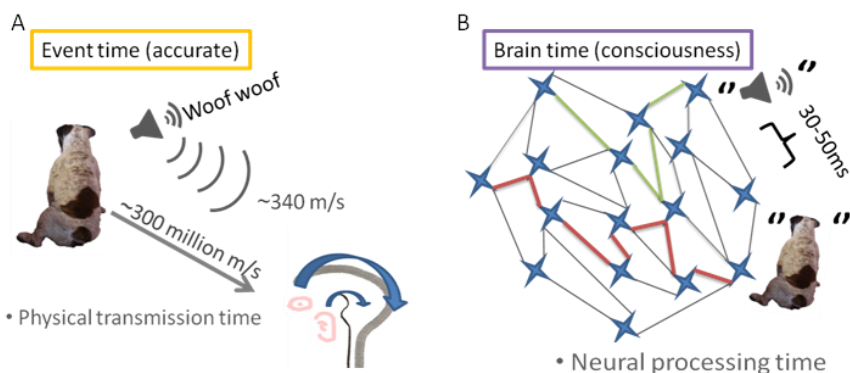


Figure 1. A, Event time is the time at which signals occur and thus, is not affected yet by the physical transmission time of the lights and sound waves. The event time can be affected by the physical transmission time when the distance between the source and the auditory and visual receptors is above 15 m, because light travels faster (~300 million m/s) than sounds (~340 m/s) through air. B, Brain time is the time at which signals are perceived. They are affected by the neural processing time of the information. For simple flashes and tones, the sound reaches faster the primary cortex (~30-50 ms) than the images.

Despite these stimulus-driven and cognitive factors influencing synchrony perception, multisensory information originating from the same event is most of the time perceived as synchronous. This strongly supports the existence of neural mechanisms to compensate temporal delays.

This thesis examines how the brain selects which cross-modal pairs would be adjusted in time, despite these sources of temporal variability. In contrast with real life environments, where multiple events can co-occur within spatial and temporal proximity, most studies to date have addressed situations in which only two relevant external events are presented. Here, I examine how these adjustment mechanisms work when additional

cognitive factors fall within the same temporal window of perception (Study 1, section 3.2 and Study 2, section 3.4). The overarching hypothesis in this dissertation is that perceptual adjustments do not only involve the passive reception of information through the senses via bottom-up parsing, but also, cognitive factors such as selective attention (Study 1, section 3.2), the current motor task performance (Study 2, section 3.4) and prior knowledge about the environment (Study 3, section 3.5). In addition, I entertain the idea that the current brain state, in general, can shape the temporal organization/structure of multisensory experience.

TABLE OF CONTENTS

ACKNOWLEDGMENTS	III
ABSTRACT	V
RESUM	VI
PREFACE	VII
OVERVIEW OF THE THESIS	1
TERMINOLOGICAL CLARIFICATIONS	3
CHAPTER 1. GENERAL INTRODUCTION	5
1.1. MULTISENSORY INTEGRATION	5
1.2. TEMPORAL COINCIDENCE: PHYSICAL SYNCHRONY VS. SYNCHRONY PERCEPTION	6
1.3. SYNCHRONY MAINTENANCE PROCESSES	11
1.4. TEMPORAL RECALIBRATION PARADIGM	12
1.4.1. Temporal recalibration of sensory-sensory pairs	12
1.4.2. Factors determining temporal recalibration of sensory- sensory pairs	16
1.5. TEMPORAL ANCHORS TO MODULATE RECALIBRATION IN MULTISENSORY CONTEXTS	18
1.5.1. The role of attention to group events in time	18
1.5.2. The role of actions to group events in time	19
1.6. NEURAL EVIDENCES UNDERLYING TEMPORAL RECALIBRATION	21
1.7. NEURAL CORRELATES OF PERCEPTUAL VARIABILITY IN SYNCHRONY PERCEPTION	21
CHAPTER 2. SCOPE OF THE THESIS	25
2.1. HYPOTHESES	25

CHAPTER 3. EXPERIMENTAL SECTION	27
3.1. OVERVIEW OF THE EXPERIMENTS	27
3.2. STUDY 1	29
Abstract	31
3.2.1. Introduction	32
3.2.2. Methods.....	35
3.2.3. Experiment 1	39
3.2.4. Experiment 2	42
3.2.5. General Discussion.....	43
3.2.6. Conclusions	46
Acknowledgments.....	48
3.2.7. References	49
3.2.8. Figures.....	54
3.2.9. Supporting Information	57
3.3. FOLLOW UP OF STUDY 1.....	69
3.4. STUDY 2	75
Abstract	77
3.4.1. Introduction	78
3.4.2. Material and methods	83
3.4.3. Procedure.....	85
3.4.4. Data analyses.....	88
3.4.5. Results	90
3.4.6. Discussion	94
Funding	99

Acknowledgements	99
3.4.7. References	100
3.4.8. Figures	105
3.4.9. Supplementary material.....	111
3.5. STUDY 3	115
Acknowledgments.....	116
Abstract	117
Significance Statement	117
3.5.1. Introduction	118
3.5.2. Materials and methods	120
3.5.3. Data analysis	123
3.5.4. Results	129
3.5.5. Discussion	135
3.5.6. References	139
3.5.7. Figures	145
CHAPTER 4. GENERAL DISCUSSION	149
4.1. SUMMARY OF THE RESULTS (TABLE).....	152
4.2. IMPLICATION OF THE PRESENT FINDINGS	157
4.2.1. Temporal recalibration	157
4.2.2. Temporal organization of perception: environmental rhythms vs. brain rhythms	162
4.3. INDIRECT IMPLICATIONS OF THE RESULTS	165
4.3.1. Multisensory perception.....	165
4.3.2. The underlying mechanisms of temporal recalibration ..	167

4.3.3.	The impact of cognitive factors.....	168
4.3.4.	Conscious perception and the processing in time through neural oscillations.....	169
4.4.	CONCLUSIONS AND FUTURE LINES OF RESEARCH..	171
BIBLIOGRAPHY		175

OVERVIEW OF THE THESIS

This dissertation is organized into three main sections. The first chapter (General Introduction, CHAPTER 1), briefly reviews general aspects of multisensory perception which are relevant to the thesis. Along these lines, I introduce the motivation underlying the study of experience-based adjustment mechanisms to maintain synchrony perception in multisensory contexts (section 1.3). In particular, the focus of this section is on a temporal recalibration paradigm on audio-visual events, which has been proved to shift synchrony perception as a function of the previous experience with asynchronous events. I describe previous related literature (section 1.4) with the purpose of reflecting the plasticity of this perceptual process, emphasizing the role played by cognitive over external factors, in determining temporal organization of events (event time vs. brain time). Finally, I review a set of studies investigating the neural correlates of the temporal organization in perception and the predictive power of current brain states to resolve perceptual variability in perception (section 1.7). The second chapter (Scope of the thesis, CHAPTER 2) introduces the scope and the specific hypotheses of the thesis. The third chapter (Experimental section, CHAPTER 3) includes three experimental studies, in the form of scientific papers (published or submitted). Study 1 (*Selective attention modulates the direction of audio-visual temporal recalibration*, section 3.2) investigates whether participants' goals can modulate perceptual adjustment processes of physically identical audio-visual events. Specifically, whether systematic relations between sensory events and selective attention can change the direction of the adjustment processes. Study 2 (*Exploring the role of action in audio-visual temporal recalibration*, section 3.4) examines whether systematic relations between motor actions and sensory events could serve as temporal anchors to group or segregate audio-visual events in time. And Study 3 (*Oscillatory*

phase in neural excitability predicts perceptual variability in cross-modal synchrony judgments, section 3.5) explores the role of ongoing oscillations predicting trial by trial variability in synchrony perception of audio-visual events. Finally, the fourth chapter (General Discussion, CHAPTER 4) summarizes the results of the three studies in a table (section 4.1) and discusses the findings in the context of previous literature on temporal recalibration (section 4.2.1) and temporal organization of perception (section 4.2.2). Then, I discuss some more speculative (or indirect) implications of our results (section 4.3) and conclude with future possible lines of research and some general remarks of the present findings (section 4.4).

TERMINOLOGICAL CLARIFICATIONS

To help clarify some of the terms/concepts that will be used throughout this thesis, a short description of some of them is provided below.

Physical synchrony: objective temporal coincidence of events according to the moment of occurrence at the signals' source.

Synchrony perception: report about the subjective perception of temporal coincidence of events (simultaneous vs. non simultaneous). Critically, two events can be perceived as being simultaneous, but not necessarily combined to be perceived as being part of the same object.

Multisensory perception: combination of signals coming from different sensory modalities that when are perceived as being part of the same object can form an integrated percept.

Integration: combination of signals coming from the same or different sensory modalities that are fused in a unique and new percept. An example of multisensory integration is the McGurk effect (McGurk and Macdonald, 1976), in which an auditory syllable (e.g. /ba/) presented with a different visual syllable (e.g. /ga/) produces the illusory perception of hearing a different syllable (e.g. /da/). Another example, is the double flash illusion, in which two flashes are reported when a flash is accompanied by two brief sounds, presented close in time (Shams et al., 2002).

Grouping or binding: combination of signals that are perceived as they were coming from the same object or event but they can be independent, that is, not integrated in a new percept. I am using these two terms to refer to the events resulting from the after-effects of a temporal recalibration paradigm. Because I have considered we cannot conclude that the signals are necessarily integrated during this recalibration process. In this

paradigm, we only have measured how perception of synchrony changes, but not if the multisensory signals have been integrated.

CHAPTER 1. GENERAL INTRODUCTION

This chapter provides a general overview of multisensory integration and the role that temporal coincidence might play in it. Temporal coincidence, or synchrony, is an elusive concept, which requires some discussion. In particular, it is important to disentangle physical synchrony from perceptual synchrony to assess its contribution to multisensory perception. Perceptual synchrony between sensory modalities can be maintained by several well known processes. I focus on temporal recalibration, which changes simultaneity perception after repeated exposure to cross-modal events appearing in a fixed (modality) order. Subsequently, I review the factors determining the re-alignment of audio-visual events, first, in contexts where only a single lag between a cross-modal pair is present during adaptation, and second, in contexts where multiple lags co-exist. Attention and action are the focus of our first and second study, respectively, which we have hypothesized, might serve as temporal anchors resolving temporal re-alignment in audio-visual contexts. Finally, I discuss the neural evidences up to date, underlying temporal recalibration and changes in perceptual synchrony on audio-visual events. The latter being the focus of our third study.

1.1. MULTISENSORY INTEGRATION

Perception of the environment involves the organization of information that arrives to the brain across different sensory modalities. Multisensory integration has been classically described as a response enhancement (or depression) relative to the responses to either of the unisensory components (see review Stein, Stanford, & Rowland, 2014). Physiologically, these responses have been extensively measured in single neurons recordings in the cat superior colliculus, characterized by receiving inputs from more than one modality. Three main factors have

been suggested to determine multisensory processes; the temporal coincidence (Meredith et al., 1987); the spatial overlap; and the principle of inverse effectiveness, which states that weaker individual inputs lead to greater multisensory response enhancement (Stein and Meredith, 1993). Behavioral, electrophysiological and neuroimaging evidences support the existence of similar processes in humans, whereby cortical and sub-cortical regions receive input from multiple senses. These processes include classic feed-forward integration pathways from unisensory brain areas to converge regions of the brain, thalamic influences and, feedback pathways from multisensory regions to sensory specific areas (e.g. Macaluso and Driver, 2005). In behavior, the integration of signals across senses has been seen to increase the accuracy and precision of signals' perception, beyond what would be expected for each sense in isolation. It is argued that signals from different senses related to the same event (congruent in time and space) might help minimizing the uncertainty of sensory estimates by providing complementary information of the same event (e.g. Ernst and Banks, 2002). Interesting for the current thesis, more and more studies highlight the relevance of multisensory exposure to ensure, not only the proper maturation of multisensory neurons during development (Wallace and Stein, 2007), but also determining the magnitude of neural responses to multisensory events (Yu et al., 2009).

1.2. TEMPORAL COINCIDENCE: PHYSICAL SYNCHRONY VS. SYNCHRONY PERCEPTION

In complex, multisensory environments, it is crucial that only the appropriate information, out of the multiple sensory inputs received, is combined across modalities to form coherent representations relying in the multisensory processes mentioned in the last section. Temporal coincidence of sensory signals, indicating a possible common source, has been considered an important factor determining the selection of which

cross-modal sources of information should be integrated, and which ones segregated (Meredith et al., 1987).

However, determining the temporal coincidence of external sensory events from their neural processing is a complex process due to the well known discrepancies in physical transmission and neural processing times across sensory modalities (e.g. King, 2005). For example, in the audio-visual domain, while physical synchrony occurs when a sound and a light are presented at the same time, perception of synchrony is usually maximal (point of subjective simultaneity or PSS) when the onset of a visual stimulus occurs slightly before the auditory onset (e.g. King and Palmer, 1985).

Moreover, the direction of the stimuli pair (vision leading audition vs. audition leading vision) is associated with different perceptual sensitivity (Conrey and Pisoni, 2006; van Wassenhove et al., 2007a; Stevenson and Wallace, 2013), being vision leading audition modality order perceived as simultaneous for a larger range of asynchronies than the reverse order.

Interestingly, some studies suggest these asymmetries towards vision leading audition depend on the task employed to measure synchrony perception (van Eijk et al., 2008). Psychophysical studies have been mainly performed using two perceptual tasks to measure synchrony; the temporal order judgments (TOJ) and the simultaneity judgments (SJ). In both tasks, participants are presented with pairs of events at a variable range of asynchronies. In temporal order judgment tasks participants are instructed to discriminate the temporal order of the two events (e.g., which came first), while in simultaneity judgments tasks, they discriminate whether the two events appear simultaneous vs. non-simultaneous (Figure 2).

From participants' responses, the PSS (point of subjective simultaneity) and the temporal window (sensitivity of perceptual discrimination, often inferred from the standard deviation or SD) of each individual can be estimated by fitting the data into a psychometric function (for example see figure 3, in which representative data from a SJ task is depicted).

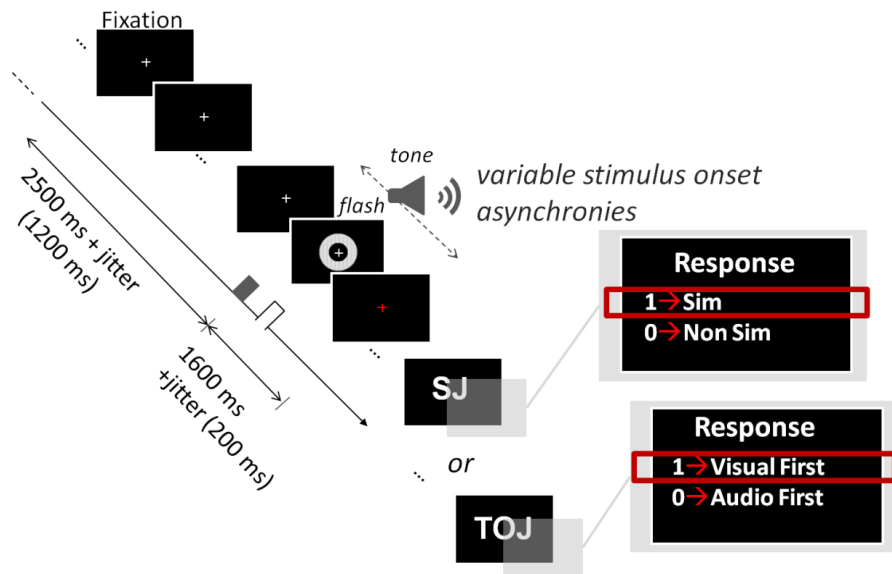


Figure 2. Simultaneity judgment (SJ) task or temporal order judgment (TOJ) task on flash-tone pairs. A white fixation cross is presented at the centre of the screen and after a variable interval, a flash-tone pair is presented at variable stimulus onset asynchronies. Subsequently, the fixation cross change to red and a response screen for the SJ task or the TOJ task is depicted, in which participants were instructed to judge either the simultaneity or the temporal order of the flash-tone pair.

Although both tasks convey the same parameters (PSS and SD), different sensitivities have been reported depending on the measure of interest (PSS or SD) and currently it is under debate whether they are actually measuring the same perceptual process (Linares and Holcombe, 2014; Miyazaki et al., 2016). Remarkably, TOJs are preferred if sensitivity in perceptual discrimination is the object of the study, while SJs are argued to be a more robust measure of PSS (van Eijk et al., 2008). We have

employed SJ tasks in our three studies as our main measure of interest has been the PSS.

These particular asymmetries between physical and perceived synchrony, whereby vision usually needs to lead audition for perception of synchrony, have been largely considered to reflect long-life environmental exposure. In particular, the experience of light reaching faster the sensor channels than sounds at distances greater than 15 meters.

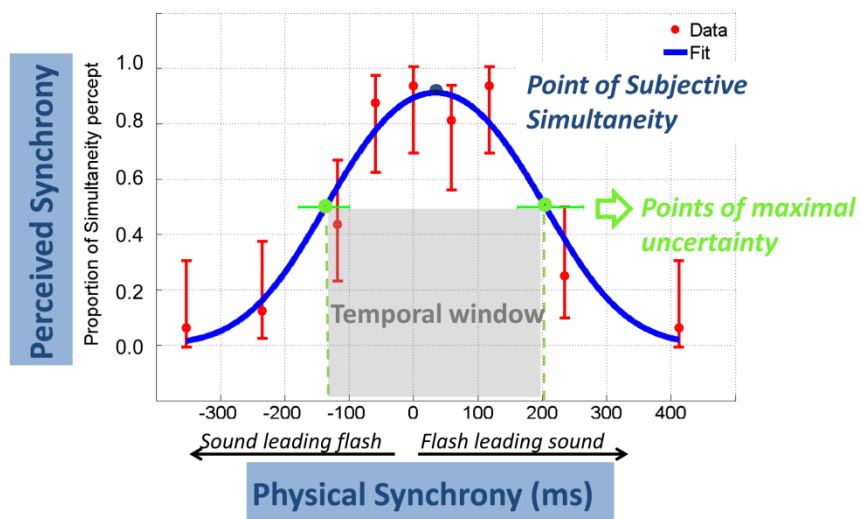


Figure 3. Representative psychometric curve obtained from a SJ task, in which perceived synchrony of flash-tone pairs is plotted as a function of the actual physical synchrony of the stimuli pair. The red dots represent the raw data with the respective confidence intervals. The blue line is the fit of the data, obtained using a maximum likelihood estimation method. The blue dot represents the estimated point of subjective simultaneity (PSS), where simultaneity is maximally perceived. Green circles represent the estimated points of maximal uncertainty, the boundaries of synchrony perception with the respective confidence intervals. Finally, the grey interval represents the estimated temporal window of simultaneity, which is the range of asynchronies that are perceived as being simultaneous.

Besides these naturally occurring delays due to differences in transmission speeds and neural processing times across sensory modalities, several

external (stimulus-driven) factors have been shown to influence the experience of synchrony. For example, spatial proximity between the events (Zampini et al., 2005), luminance contrast (White et al., 2008), duration (Boenke et al., 2009; Heron et al., 2013), intensity of the stimulus (Fister et al., 2016) and the rate of stimuli presentation (Fujisaki and Nishida, 2010).

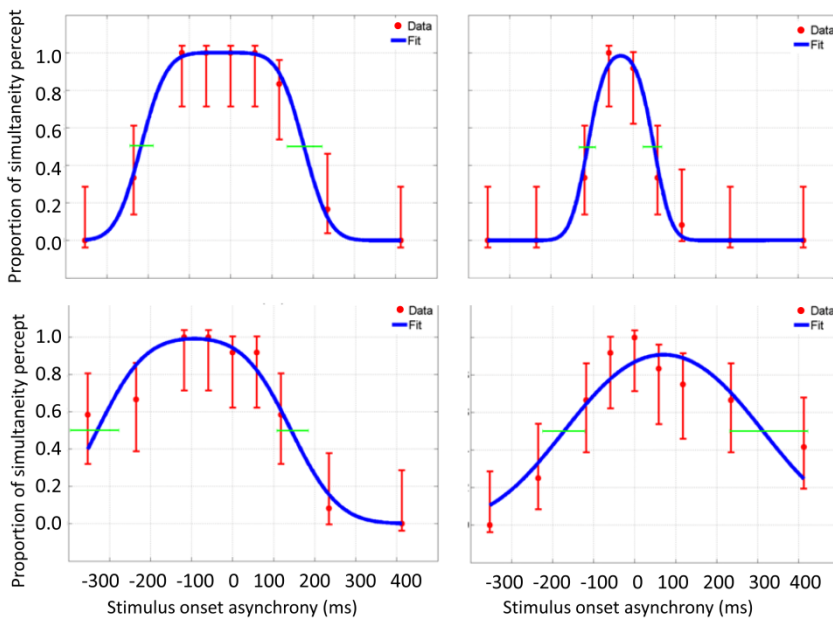


Figure 4. : Illustrative data from four different participants obtained in the Pretest of Study 2 of this Thesis. Psychometric curves obtained from a simultaneity judgment (SJ) task. We plotted the proportion of simultaneity perception as a function of the physical stimulus onset asynchronies (SOA) between stimuli. Negative values denote sound is leading the flash and positive values denote the reverse modality order. They represent the individual variability in the estimated temporal window (SD) and point of subjective simultaneity (PSS) among participants. The red dots represent the raw data with the respective confidence intervals. The blue line is the fit of the data, obtained using a maximum likelihood estimation method and green lines are the confidence intervals of the synchrony boundaries.

Likewise, several cognitive factors (non-stimulus driven) have been demonstrated to influence synchrony perception, such as prior modality

order experience (Fujisaki et al., 2004; Vroomen et al., 2004; Van der Burg et al., 2013), training feedback (Powers et al., 2009), task demands (Stevenson et al., 2014), semantic congruency (van Wassenhove et al., 2007b; Vatakis and Spence, 2007), attention (Spence and Parise, 2010), predictability (van Eijk et al., 2008) and inference of causal relations (Kohlrausch et al., 2013).

Temporal uncertainty of sensory information coming from these external and cognitive variables influencing perceived asynchronies, leads to individual variability in the temporal window and the PSS measures (see Figure 4). Thus, comparing these parameters between participants needs to be done carefully. Moreover, to deal with temporal uncertainty it is critical to use various processes to maintain synchronous perception across sensory modalities. These maintenance processes are described in the next section (1.3).

1.3. SYNCHRONY MAINTENANCE PROCESSES

Notwithstanding the different sources of delay between sensory modalities, information coming from the same event is generally subjectively perceived as simultaneous (except for very distant events leading to large asynchronies). There are several mechanisms (not necessarily exclusive of each other) involved in adjusting these temporal discrepancies among senses. One proposal refers to the existence of a temporal window, within which simultaneity of simple cross-modal events is perceived despite temporal disparities of 25-50 ms (Zampini et al., 2003). Another possible temporal adjustment mechanism is that one sensory modality attracts the other in time during perception, such as is the case of the temporal ventriloquism, where visual events are perceptually shifted towards sounds in the time dimension (Morein-Zamir et al., 2003; Vroomen and de Gelder, 2004). Another plausible mechanism

for keeping temporal coherence across modalities is the process of temporal recalibration. Numerous studies have shown that when a cross-modal asynchrony is constant throughout a period of time, the sensory system is able to recalibrate for the discrepancy, shifting perception of simultaneity according to information extracted from previous experience (Fujisaki et al., 2004; Vroomen et al., 2004). Hence, these evidences suggest that the brain is to some extent flexible in order to re-align the information received through the senses in time, to form coherent representations of the environment. In the next section (1.4) I further describe in detail the temporal recalibration paradigm.

1.4. TEMPORAL RECALIBRATION PARADIGM

1.4.1. Temporal recalibration of sensory-sensory pairs

The temporal recalibration paradigm measures changes on simultaneity perception after adaptation to a cross-modal asynchrony. For example, in the audio-visual case, the perception of simultaneity of flash-tone pairs is studied as a function of previous exposure to a fixed modality order, which can be flash leading the tone asynchrony, or vice-versa (Fujisaki et al., 2004; Vroomen et al., 2004). During this procedure, participants are exposed first to one asynchrony direction, for example, a flash leading a tone by 200 ms during 3 minutes (adaptation phase). Subsequently, there is a test phase, in which participants are presented with test trials, consisting on flash-tone asynchrony pairs spanning from several hundredths milliseconds to physical synchrony. Here, participants are instructed to judge either the temporal order of events or the simultaneity perception of the flash-tone pairs (Figure 5).

The PSS after adaptation to the flash leading a tone shifts towards vision-leading asynchronies compared to when the opposite asynchrony direction is presented during adaptation (Figure 6). The amount of change in the

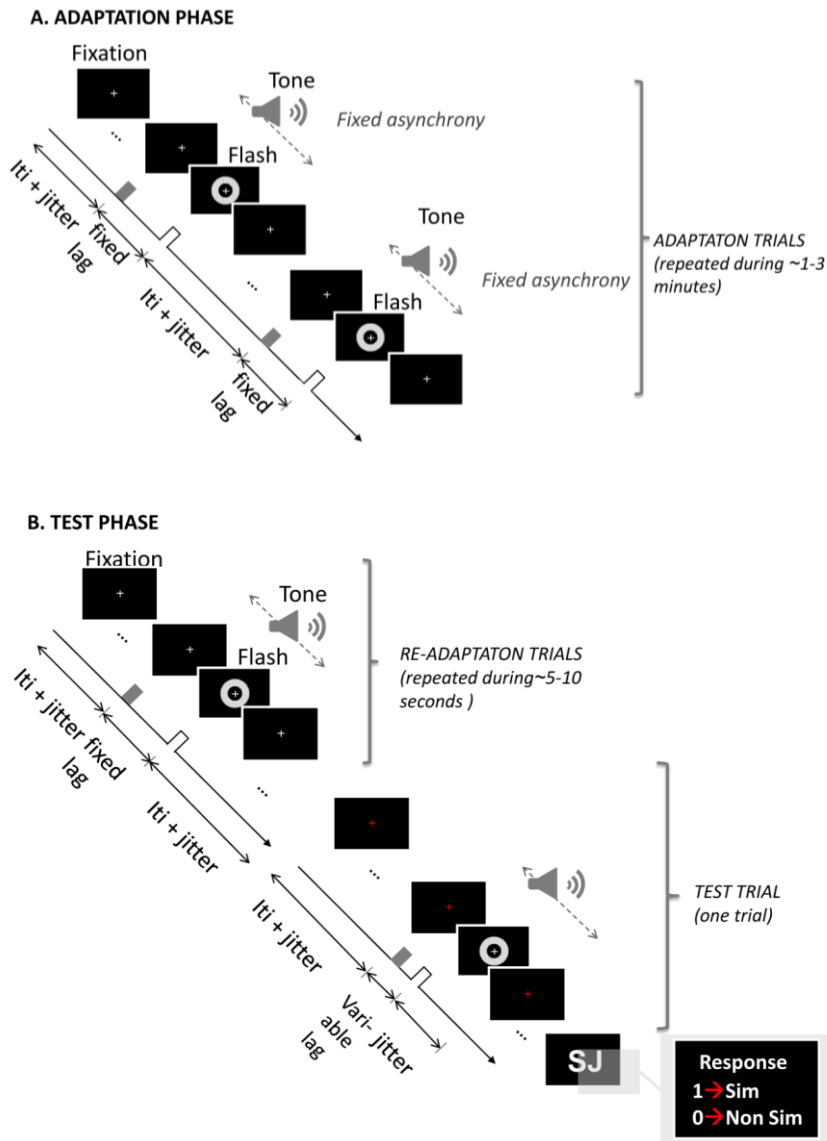


Figure 5. Timeline of a temporal recalibration paradigm. A, Adaptation phase. One trial consists on the presentation of a white fixation cross at the centre of the screen. After a variable interval, a flash-tone pair is presented always at the same fixed asynchrony (flash leading or tone leading, in separate adaptation sessions). The duration of the adaptation phase usually lasts between 1 to 3 min. B, Test phase. It consists on 5-10 s of re-adaptation trials (same trials as in the adaptation phase), followed by a test trial, which is warned by a change in color of the fixation cross (to red color). In the test trial, a flash-tone pair is presented at a variable asynchrony and after a short interval a response screen is presented, where participants were instructed to judge the simultaneity (or temporal order) of the stimuli pair by pressing two different response keys.

PSS values after adaptation is called the recalibration after-effect, whereby a re-alignment in the temporal perception of audio-visual events has taken place. That is, participants after adaptation would perceive more simultaneous the asynchrony direction of events they have been previously adapted to.

The strong contingency generated between the repeated exposure to sensory events of a given order in a recalibration procedure might subsequently induce a re-organization in the temporal structure between the mental representations of the events (in line with Miyazaki et al., 2006).

Recently, this recalibration effect has been reported to occur even after the presentation of a single exposure trial, though its interpretation is under debate (Van der Burg et al., 2013, 2015). This rapid recalibration effect could explain why the recalibration procedure needs the inclusion of re-adaptation trials during the test phase (top-up). The after-effects disappear when new temporal relations between the stimuli are presented (Machulla et al., 2012).

The mechanism underlying the re-alignment process remains elusive. On the one hand, it has been proposed that temporal recalibration consists of a common supra-modal process, independent of the implicated modalities, because recalibration occurs similarly for any modality pair combination (Hanson et al., 2008); audio-tactile, audio-visual and visuo-tactile pairs. This proposal is compatible with the existence of a centralized clock resolving the temporal re-alignments.

On the other hand, other authors suggest the re-alignment is caused by shifts in the perceptual latency of one of the modalities involved in the process. Hence, one modality might be adjusted relative to the other implying the recalibration process would differ between sensory

modalities (Harrar and Harris, 2008; Navarra et al., 2009). Failures in temporal recalibration when tactile events are involved are in agreement with the perceptual latency framework (e.g. Harrar and Harris, 2008). It has been argued that touch is less flexible than audition or vision. The reason behind is that touch is not influenced by external sources of asynchrony, and relations between physical onset and brain responses are less variable (Miyazaki et al., 2006).

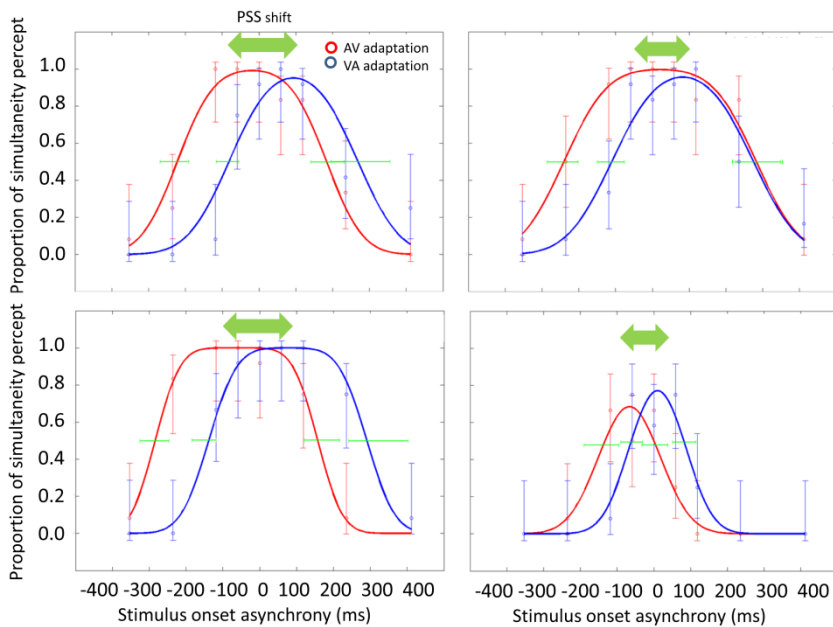


Figure 6: Illustrative results from four individual participants obtained in Study 2 of this Thesis. PSS shifts (green arrows) after adaptation to a temporal recalibration procedure. We plotted the proportion of simultaneity perception as a function of the physical stimulus onset asynchronies between stimuli. Negative values denote sound is leading the flash and positive values denote the reverse modality order. The red dots represent the raw data with the respective confidence intervals. Red dots and lines represent the psychometric curve obtained after adaptation to audition leading vision (AV). Blue dots and lines represent the psychometric curve after adaptation to vision leading audition (VA). Green intervals are the confidence intervals of the synchrony boundaries. Blue line (visual leading) is shifted to visual leading SOAs compared to red line (audition leading), showing the temporal recalibration after-effect.

Contrary, re-alignments involving audition and vision are commonly reported. For example, audition and vision re-alignments can be transferred to other non-adapted sensory modality combinations (Harrar and Harris, 2008; Di Luca et al., 2009) and can change its perceived time after adaptation (Harrar and Harris, 2008; Di Luca et al., 2009; Navarra et al., 2009). Which is the modality that will be re-aligned and which one will be used as a reference remains a controversial issue (Harrar and Harris, 2008; Navarra et al., 2009).

A current view on this controversy suggests that the more reliable modality in the recent context (e.g. according to the current parameters of the experimental setup) will be the one taken as a reference to synchronize the rest (Di Luca et al., 2009).

Other theoretical approaches propose the temporal recalibration process might be achieved through a modulation of a prior likelihood distribution, generated by contingencies presented in the environment (Yamamoto et al., 2012), by adaptation to delay sensitive neurons (Roach et al., 2011) and/or by shifts in synchrony perception at a decisional level, in particular, in the synchrony criteria (Yarrow et al., 2011a).

1.4.2. Factors determining temporal recalibration of sensory-sensory pairs

Fujisaki et al., (2004) were the first authors to examine the influence of temporal factors determining the re-alignment of audio-visual events in contexts where a single cross-modal pair is present during adaptation. They focused on the recalibration magnitude, which corresponds to the difference between PSS's after adaptation to opposite modality orders (for example; a flash leading a tone vs. a tone leading a flash). Larger recalibration magnitudes indicate stronger re-alignments between sensory modalities after adaptation.

The first factor evaluated was the size of the lag presented during adaptation. Their results suggest that at larger lags (> 400 ms) recalibration is less effective than at shorter ones (around 100 ms). However, recalibration after-effects at lags up to 700 ms have been reported when vision precedes sounds, but not for the reverse order (Navarra et al., 2013). It is argued that the canonical experience under natural conditions (of light speed being faster than sound speed) and the timings between related events (usually related events are present close in time) can account for this lag size factor.

A second fundamental temporal factor investigated was the rate of presentation of event pairs during adaptation, which determine the temporal structure of the event presentation. It has been suggested that at higher rates of presentation, recalibration might fail in some studies because a clear temporal structure is needed to re-align information in time (Roseboom et al., 2013). Indeed, discrimination of cross-modal temporal coincidence has been reported to break down at presentation rates above 2.4 to 3 Hz (Fujisaki and Nishida, 2010; Kösem and van Wassenhove, 2012).

Other factors investigated, beyond time per se, were the generalization of the re-alignment to other stimuli pairs within the same sensory modality, such as vision (Fujisaki et al., 2004) and audition (Navarra et al., 2005, 2012) as well as the transferability of the effect between different sensory modality combinations (Harrar and Harris, 2008; Di Luca et al., 2009). Furthermore, the variability in recalibration magnitudes between participants suggest that larger temporal windows of simultaneity can encompass greater recalibration (Van der Burg et al., 2013).

The studies reviewed so far tested recalibration in non-competing contexts where only one cross-modal event with a single lag was present during

adaptation. However, in real life situations multiple events are present; with a variety of lags co-existing between them. There are several factors known to help resolve recalibration when two events are present during adaptation. For example, two opposite lags presented during adaptation can be recalibrated when they are systematically presented at different spatial locations (Roseboom and Arnold, 2011; Yarrow et al., 2011b; Heron et al., 2012; Yuan et al., 2012) or characterized by different identities (Roseboom and Arnold, 2011; Yuan et al., 2012, 2015; Roseboom et al., 2013. Although see Heron et al., 2012) constrained by temporal features or by a combination of attributes of the stimuli (Yuan et al., 2015). Furthermore, it has been suggested that instructions given to the participants and therefore, participants' goals (probably mediated by selective attention) might be critical to build up new temporal structures (Roseboom et al., 2013).

Notwithstanding, the role of attention and action in temporal recalibration under competing event contexts are described below, in a separate section (1.5), as they are crucial for the scope of the present thesis.

1.5. TEMPORAL ANCHORS TO MODULATE RECALIBRATION IN MULTISENSORY CONTEXTS

1.5.1. The role of attention to group events in time

In contexts where multiple sensory stimuli compete for processing, we have proposed (see hypothesis in CHAPTER 2) that selective attention might be an effective filter to select stimuli of interest given the limited capacity of humans to determine temporal relations of cross-modal information (Fujisaki and Nishida, 2010; Kösem and van Wassenhove, 2012).

For example, in visual search tasks (at temporal rates of stimulus presentation above 1.4 Hz), detection of a visual target that changes in synchrony with a sound is gradually impaired as the number of surrounding distracters increases (Fujisaki et al., 2006; Kösem and van Wassenhove, 2012). However, the interference related to the number of distracters could be completely overcome in this and other tasks if selective attention is used to pre-cue the location or specific features of the target (Fujisaki et al., 2006; Fujisaki and Nishida, 2007). Thus, attention could be effective to mediate cross-modal synchrony discrimination in cluttered environments.

In addition, the temporal resolution in judging cross-modal temporal order is much lower than within sensory modalities (Fujisaki and Nishida, 2005, 2010). This limited temporal resolution for cross-modal temporal judgment is consistent with the existence of a central bottleneck process, probably attention-dependent, mediating synchrony perception across senses.

In Study 1 (*Selective attention modulates the direction of audio-visual temporal recalibration*, section 3.2) we investigated the role of attention on audio-visual temporal recalibration in contexts where more than two audio-visual events were presented close in time.

1.5.2. The role of actions to group events in time

In many laboratory experiments, (multisensory) perception is often assumed to be a passive process, where participants are asked to respond only after receiving some sensory information through a computer screen, LEDs, loudspeakers or other means. Contrary, in real life, perception usually takes place while we are actively interacting with the surrounding environment.

In Study 2 (*Exploring the role of actions on audio-visual temporal recalibration*, section 3.4) we hypothesized that these direct interactions throughout our life experience might allow the system to generate detailed predictions between our actions and their sensory consequences which may be important for temporal perception.

For example, when playing an instrument such as the piano, there is tactile information arising from the fingers pushing the piano keys, visual information about the hands, the keys of the piano, and even the notes of the score you are playing, the sound of the notes, and proprioceptive information about the position and movement of the arms and fingers. Some of these signals despite relating to the same event are processed at slightly different times.

Contingencies between sensory and motor signals learnt through life experience might be very useful not only to structure our own movements, but also to re-organize the temporal structure of external events via predictions.

Given the degree of uncertainty in assessing the exact temporal order between signals of different sensory modalities, already discussed in section 1.2, in Study 2 (section 3.4) we speculated that the brain might rely on reliable (self-generated) temporal anchors to re-align sensory signals. In particular, we examined the implication of motor interactions on perception of simultaneity. We hypothesized that systematic relations between actions and audio-visual events could shift perceptual simultaneity and thus, organize subsequent temporal structure of events.

1.6. NEURAL EVIDENCES UNDERLYING TEMPORAL RECALIBRATION

To our knowledge, only one study has already assessed the neural correlates of perceptual simultaneity in a temporal recalibration paradigm, on audio-visual pairs. The results of the study demonstrate a correlation between changes in the perceptual organization of audio-visual events in time and changes in the phase of 1 Hz entrained oscillations (Kösemet al., 2014). Specifically Kösem et al., (2014) compared the phase of neural oscillations over visual and auditory sensory areas at the beginning and at the end of the adaptation phase. Their findings suggest that the phase of entrained neural oscillations at 1 Hz shifted in the predicted direction of the PSS shift after adaptation, but only in the sensory area related to audition.

1.7. NEURAL CORRELATES OF PERCEPTUAL VARIABILITY IN SYNCHRONY PERCEPTION

The studies on audio-visual synchrony perception so far (Kösemet al., 2014; Kambe et al., 2015) have examined neural activity after stimulus presentation (post-stimulus), which are the evoked responses time locked to a specific event. In these studies, in order to obtain the evoked responses, the EEG (electro-encephalography)/MEG (magneto-encephalography) signal is first averaged over trials and then, it is subjected to a time-frequency analysis. The time-frequency analysis examines the strength with which various frequencies are present in the signal, at particular time intervals.

In Study 3 (*Oscillatory phase predicts perceptual variability in cross-modal synchrony judgments*, section 3.5) we presented an alternative approach that focuses on pre-stimulus activity, which is the neural oscillatory activity prior to, instead of evoked by, stimulus presentation.

In this type of analysis, the signals of interest are locked to the stimulus, but immediately preceding the onset of the stimulus. This analysis is supposed to capture endogenous variability not related to the stimulus. First, to study pre-stimulus activity, researchers apply the time-frequency analysis to each trial in the interval prior to stimulus presentation. Then, the data is averaged across trials. This analysis allows researchers to examine the brain state prior to stimulus presentation.

We argued (see hypothesis in CHAPTER 2) that ongoing brain activity that reflects cognitive factors (non-stimulus driven) could later influence different perceptual outcomes in synchrony judgments of the upcoming stimulus. Furthermore, we speculated that ongoing neural oscillations could be a good means to encode perceptual variability in synchrony perception.

Neural oscillations have been argued to reflect fluctuations in cortical excitability (Buzsáki and Draguhn, 2004; Schroeder and Lakatos, 2009; Dugue et al., 2011). An oscillation can be described by its frequency, amplitude and instantaneous phase. The amplitude of a neural oscillation is captured by the power, which depends on the total number of neurons actively firing in synchrony at a particular frequency. The instantaneous phase of an oscillation reflects the state (excitability) of a particular set of neurons (being recorded) at a particular moment in time. Both, the phase and power of an oscillation are measured at a particular frequency.

Two different neural processes have been suggested to help encoding the temporal order of sensory events. The first process is based on the entrainment of neural oscillations to external rhythms generated in the environment. A second kind of process relies on internal rhythms, such as ongoing neural oscillations. However, whether the temporal organization is actually driven by ongoing neural oscillations at specific frequencies or

rather it is determined by entrainment to external rhythms is still an open question (Köseme et al., 2014; Herbst and Landau, 2016).

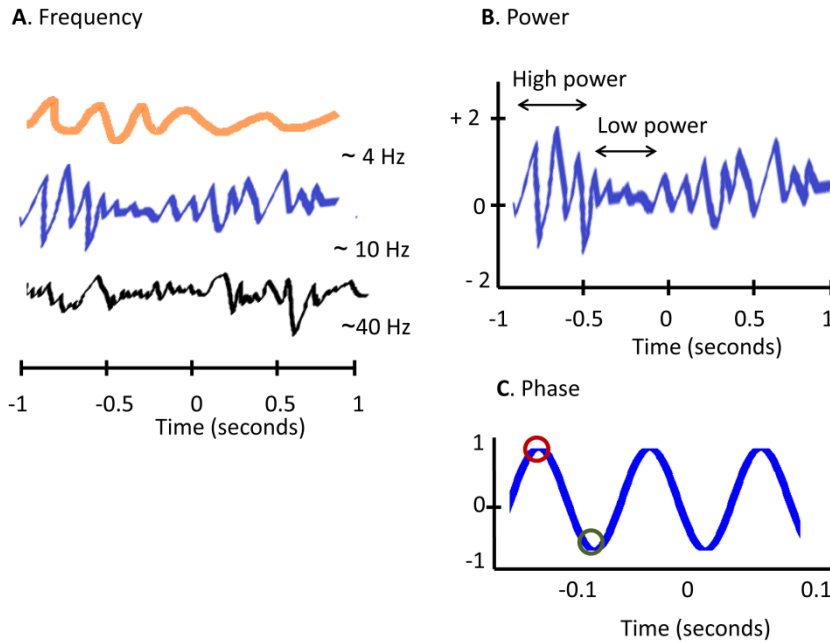


Figure 7. Illustration of neural oscillations. A. Illustration of oscillations at different frequency bands. B. Illustration of the power of an oscillation. Amplitude variation. C. Illustration of oscillatory cycles. Red and green circles signal opposite phases. Ideally a peak would correspond to a low excitability state and a trough to a high excitability state.

In Study 3 (*Oscillatory phase predicts perceptual variability in cross-modal synchrony judgments*, section 3.5) we used a non-rhythmic stimuli presentation to investigate whether ongoing neural oscillations could predict trial by trial variability in synchrony judgments of identical audio-visual events, without entrainment to external stimuli.

CHAPTER 2. SCOPE OF THE THESIS

The aim of this PhD dissertation is to examine how the brain undergoes the selection of which cross-modal information, received through different sensory modalities, is temporally adjusted in order to form coherent and unitary representations of multisensory events in the environment. In general, we have hypothesized that temporal re-alignment across audio-visual events not only results from the physical organization of the inputs in the environment, but also from endogenous factors. Hence, sections 3.2 and 3.4 present studies that illustrate how temporal processing can be resolved by attention and voluntary motor actions, acting as temporal anchors. Furthermore, in a third study (section 3.5), we examine whether pre-stimulus oscillatory activity (measured through EEG recording) can predict perception of simultaneity of audio-visual events. Specifically, we measure the phase and power of slow neural oscillations occurring prior to audio-visual events whose simultaneity is ambiguous (perceived half of the times as simultaneous and the other half as non-simultaneous). Alongside with the two behavioral studies (section 3.2 and 3.4), this third study (section 3.5) addresses the question of whether cognitive factors (not driven by external stimuli) can determine temporal organization of audio-visual events.

2.1. HYPOTHESES

- (1) We hypothesized that, in adaptation scenarios where there are multiple competing audio-visual asynchronies, selective attention to particular stimuli can determine the direction of temporal recalibration and therefore, determine the PSS shift in a temporal recalibration paradigm. We have tested this hypothesis in Study 1 (section 3.2).

- (2) We hypothesized that, in scenarios where actions occur in the context of sensory events, the nature of actions (putatively promoting grouping or segregation of audio-visual events) can modulate the magnitude of recalibration in a temporal recalibration paradigm. Specifically, we predicted that voluntary actions promoting grouping of audio-visual events will lead to stronger recalibration than actions promoting segregation of audio-visual events. We have tested this hypothesis in Study 2 (section 3.4).

- (3) We hypothesized that pre-stimulus brain activity can predict perceptual outcome in audio-visual synchrony judgments. We assumed that cognitive factors known to influence synchrony perception might reflect on pre-stimulus brain activity, specifically, on the phase and amplitude of low frequency ongoing neural oscillations. We tested this hypothesis in Study 3 (section 3.5).

CHAPTER 3. EXPERIMENTAL SECTION

3.1. OVERVIEW OF THE EXPERIMENTS

In Study 1 (3.2) we employed a variation of the classical temporal recalibration paradigm. Temporal recalibration is a mechanism that minimizes temporal discrepancies between sensory information after a repeated exposure to a constant asynchrony (Fujisaki et al., 2004). In particular, we examined whether deploying attention towards one of two currently competing audio-visually asynchronous events shifted perception of simultaneity (PSS) in the direction of the attended stimuli. To test our hypothesis, in this study participants were presented with two potential asynchronies with respect to a sound (flash-tone-flash). During the adaptation phase, we induced participants' attention directed either toward the flash preceding or toward the flash trailing the tone in separate sessions. Following adaptation, participants were tested in a simultaneity task for audio-visual events presented at 9 different asynchronies.

In Study 2 (section 3.4) we employed another variation of the temporal recalibration paradigm. In this case we presented participants with simple flash-tone asynchronous pairs during the adaptation phase. Critically, we decided to instruct participants to perform one of three alternative motor synchronization tasks toward the audio-visual events during adaptation. Independent of the direction of the asynchrony (audition leading or vision leading), each participant performed only one of the three synchronization tasks. Following adaptation (including synchronization), participants were tested for simultaneity perception, just as in Study 1. We examined whether motor task demands (promoting grouping or segregation of audio-visual events) can influence the magnitude of the temporal recalibration in the direction expected by the predicted nature of the motor

task demands. That is, we expected a larger recalibration magnitude when participants performed a motor task promoting grouping of audio-visual events compared to the groups in which the task involves segregation.

In Study 3 (section 3.5) we recorded EEG activity while participants performed a simultaneity judgment (SJ) task on audio-visual events presented at different asynchronies. Before the EEG recordings, we assessed for each participant, which audio-visual asynchronies (in either order) were perceived half of the times as simultaneous and the other half as non-simultaneous. In these particular asynchronies we then investigated whether the power and phase of slow neural oscillations in EEG prior to stimulus presentation, can predict trial by trial variability in synchrony perception of otherwise identical audio-visual stimuli.

3.2. STUDY 1

Study 1 is published in Plos One (2014).

Ikumi N, Soto-Faraco S. [Selective Attention Modulates the Direction of Audio-Visual Temporal Recalibration](#). PLoS One. 2014 Jul 8;9(7):e99311. DOI: 10.1371/journal.pone.0099311

3.3. FOLLOW UP OF STUDY 1

Considering the results obtained in the first study, we decided to further examine the interpretation of the effect of selective attention on the selection of which cross-modal event is adjusted in time. In Study 1, the preferred interpretation was that cross-modal adjustments in time can be modulated by selective attention only when stimulus-driven processes are not sufficient to drive recalibration. In Study 1 we concluded that selective attention only shifted the direction of the recalibration after-effect when participants attended to the auditory leading visual asynchrony.

In a follow up experiment, instead of presenting two competing visual stimuli and an auditory event, we exposed participants to a continuous stream composed of four different stimuli (2 auditory and 2 visual events). The stimuli were presented isochronously, alternating the sensory modalities. In order to reduce the stimulus capture of the extreme stimuli we added a fade in /out at the beginning and at the end of the adaptation stream. Here, we hypothesized that reducing the structure on the order of events determined by the physical organization will allow selective attention to modulate the temporal adjustments in both asynchrony directions (auditory leading and visual leading asynchronies). Participants were exposed to a continuous stream of audio-visual events in which the inter-stimulus interval between auditory and visual modalities was fixed at 200 ms (Figure 8).

We tested and analyzed 10 participants. We measured the magnitude and direction of recalibration by instructing participant to focus attention to one of the four possible audio-visual combinations in two separate sessions. In both adaptation sessions, the stimuli presentation was exactly the same, only the instructions given to the participants differed during adaptation. Participants were instructed to attend to a specific auditory and

visual stimuli and to ignore the unattended ones during adaptation (segregated by pitch and color).

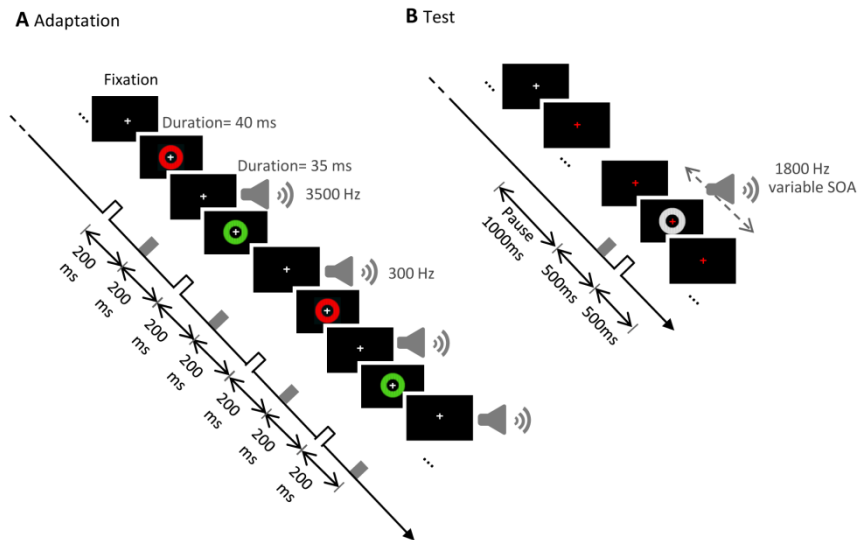


Figure 8. A) Schematic representation of a fragment of the adaptation phase in the follow up of study 1. The sequence was composed by two differently colored flashes (red and green), intermixed with two different tones (high and low pitch). The stimuli were displayed at a fixed interval (200 ms) during the adaptation phase. B) Example of a test trial in the follow up of study 1. The fixation cross changed from white to red and after a 1000 ms pause, the test trial (duration=1000 ms) started. After 500 ms one grey flash was presented in the centre of the screen and an 1800 Hz tone was displayed at one of nine asynchronies ranging from -353 to +353 ms with respect to the flash. Participants responded whether the grey flash and the 1800 Hz tone were simultaneous or not. The fixation turned back to white and a new re-adaptation sequence started.

The auditory events consisted in high and low pitch tones (~3500 Hz and ~300 Hz) equalized in loudness for each subject at the beginning of the session. The visual events consisted in green and red colored rings, equalized in luminance at the beginning of the session too.

During the adaptation participants were instructed to detect auditory and visual oddballs. Only one type of oddball (either visual or auditory) was displayed in a given set of flash-tone-flash-tone stimuli. The oddballs consisted either in a change in the tilt of the visual ring or in a change in the intensity of the tone. The test trials consisted on a SJ task; on flash-tone pairs (see Figure 8 for a more detailed description). In this case, the visual events were grey colored rings and the auditory events were ~1800 Hz tones. The test stimuli were also equalized for luminance and loudness at the beginning of the session. We did not find recalibration effects after adaptation to flash leading or flash lagging tone pairs ($Z=-0.663$, $p=0.508$) neither differences in the oddball detection ($Z=-0.204$, $p=0.838$). See results in the tables below.

Participants	PSS (ms)			
	Pretest	Flash lagging	Flash leading	PSS shift
1	29.44	50.81	73.55	22.74
2	-28.99	4.23	27.95	23.72
3	-0.10	-12.32	-10.64	1.68
4	18.48	-8.38	11.80	20.18
5	-50.66	-49.18	-16.86	32.32
6	-5.38	58.98	26.48	-32.50
7	91.78	48.04	16.88	-31.16
8	27.36	33.74	22.07	-11.67
9	-3.44	21.62	1.98	-19.64
10	-7.92	4.33	51.80	47.46
Average	7.06	15.19	20.50	5.31
Sd	38.58	33.90	27.25	

Participants	Proportion of odd detection	
	Flash lagging	Flash leading
1	0.71	0.78
2	0.78	0.72
3	0.44	0.70
4	0.69	0.10
5	0.75	0.50
6	0.45	0.38
7	0.42	0.33
8	0.75	0.83
9	0.38	0.42
10	0.70	0.86
Average	0.61	0.56
Sd	0.16	0.25

We further explored whether deploying attention towards both modality orders (visual leading audition or audition leading visual pairs) could have been different during the adaptation conditions. In order to deploy attention to the visual and auditory modalities simultaneously, participants were instructed to detect audio-visual oddballs. Both, a change in the tilt of the visual ring and in the intensity of the tone occurred consecutively. We measured detection rate, reaction times and sensitivity rates for both adaptation conditions. We hypothesized that attending to auditory leading visual asynchronies might have been easier; the auditory modality in the temporal domain is argued to be more accurate than the visual one. Thus, if attention was deployed in the two adaptation conditions in a different manner; we would have expected to find a higher detection and sensitivity rates when participants attended to the auditory leading visual asynchrony. We tested and analyzed 16 participants. We compared whether there were differences in detection rates and sensitivity depending on the attended modality order throughout the course of adaptation. However, we did not find any significant differences in the detection ($Z=-0.312$, $p=0.755$),

sensitivity rates ($Z=-0.310$, $p=0.756$) and the reaction times ($Z=-0.647$, $p=0.518$) between the adaptation conditions. See results in the table below.

Participa nts	Proportion of detection		d'		RT (s)	
	flash lagging	flash leading	flash lagging	flash leading	flash lagging	flash leading
1	0.58	0.47	2.52	1.43	0.54	0.77
2	0.65	0.68	1.40	1.15	0.65	0.68
3	0.92	0.37	2.94	1.94	0.45	0.52
4	0.52	0.61	1.22	1.92	0.69	0.62
5	0.74	0.78	1.88	0.51	0.51	0.56
6	0.80	0.95	1.73	3.12	0.41	0.47
7	0.93	0.50	3.71	1.93	0.48	0.73
8	0.43	0.61	1.41	1.76	0.61	0.46
9	0.59	0.76	1.23	2.59	0.64	0.47
10	0.60	0.79	1.27	2.01	0.55	0.53
11	0.91	0.83	2.68	2.57	0.42	0.54
12	0.74	0.62	1.52	0.98	0.41	0.46
13	0.75	0.86	1.93	2.33	0.44	0.48
14	0.74	0.79	1.35	1.85	0.47	0.44
15	0.74	0.68	2.01	1.45	0.53	0.49
16	0.52	0.52	1.69	1.49	0.65	0.55
Average	0.70	0.68	1.91	1.81	0.53	0.55
Sd	0.15	0.16	0.72	0.66	0.10	0.10

Hence, we did not find differences in the way attention was deployed to visual leading auditory pairs and auditory leading visual pairs during the course of adaptation.

The null findings of this follow up experiment are inconsistent with the idea that selective attention can solely determine the PSS shifts in temporal recalibration studies. In these particular contexts, where physical

organization does not constrain the temporal order of events, attention seems to be equally efficiently deployed to auditory leading and visual leading pairs. Therefore, stimulus-driven processes or different selective attention deployed to the stimuli in the previous temporal recalibration paradigm (Study 1, section 3.2) might have been necessary to modulate the direction of the recalibration after-effect.

Overall, the results of Study 1 and the follow up study suggest that although selective attention can modulate the effect of recalibration, attention might not be sufficient to drive recalibration and parse audio-visual events in time.

3.4. STUDY 2

Study 2 has been submitted for publication to *Frontiers in Integrative Neuroscience* (2016).

Ikumi N, Soto-Faraco S. [Grouping and Segregation of Sensory Events by Actions in Temporal Audio-Visual Recalibration](#). *Front Integr Neurosci*. 2016 Jan 19;10:44. DOI: 10.3389/fnint.2016.00044

3.5. STUDY 3

Study 3 has been submitted for publication to Journal of Neuroscience (2016).

Oscillatory phase in neural excitability predicts perceptual variability in cross-modal synchrony judgments

Phase predicts cross-modal synchrony judgments

Nara Ikumi¹, Mireia Torralba¹, Manuela Ruzzoli¹ & Salvador Soto-Faraco^{1,2}

1. Multisensory Research Group, Center for Brain and Cognition, Universitat Pompeu Fabra, 08018 Barcelona, Spain.; 2. Institució Catalana de Recerca i Estudis Avançats (ICREA), Barcelona, Spain.

Corresponding autor: Salvador Soto-Faraco. Dept. de Tecnologies de la Informació i les Comunicacions, Universitat Pompeu Fabra. C\ Tànger, 122-140, 08018 Barcelona, Spain. s.sotofaraco@icrea.cat

Acknowledgments

This research was supported by the Ministerio de Economía y Competitividad (PSI2013-42626-P), AGAUR Generalitat de Catalunya (2014SGR856), and the European Research Council (StG-2010 263145). M.R. was supported by Juan de la Cierva postdoctoral fellowship (JCI-2012-12335) from the Ministerio de Economía y Competitividad. The authors declare no competing financial interests. We specially thank Daria Kvasova and Dr. Luis Moris Fernandez for data collection and Xavier Mayoral for methodological and technical support.

Abstract

Synchronized perception of cross-modal events in multisensory contexts (e.g. a glass crashing on the floor) involves subjective impression that stimuli in various senses occur at the same time. However, light and sound travel at different speeds through air and, in addition, neural processing of visual and auditory information takes varying amounts of time. The brain must therefore organize cross-modal signals (e.g. the sight of the glass breakage and its crashing sound) that produce the perception of a synchronous event. It has been suggested that internal rhythms, such as ongoing neural oscillations, might help encoding temporal order of sensory events. Here, we addressed the relationship between the amplitude and phase of ongoing brain oscillations in humans with perception of simultaneity across sensory modalities. Participants performed an audio-visual simultaneity judgment task for stimuli presented at different asynchronies during electroencephalographic (EEG) recording. We found that the phase of pre-stimulus oscillations at 13 ± 2 Hz, 200 ms before event presentation, was predictive of simultaneity perception in the absence of concomitant changes in power amplitude. Further, we show that that prior trial history correlates with both behavior and oscillatory power at the same frequency range (14 ± 2 Hz). Overall, these findings suggest that both, phase and power of ongoing brain activity orchestrate neural fluctuations determining synchrony perception in cross-modal events.

Significance Statement

Our study suggests that the phase of low frequency neural oscillations prior to the appearance of sensory events reliably predicts variations in the perceptual outcome of synchrony judgments across sensory modalities (audio-visual). Above and beyond the effects of cueing or entrainment of

brain states, the present finding highlights the importance of the fine temporal structure of spontaneous, ongoing neural oscillations for perceptual organization of cross-modal information in time. In addition, we examine the neural underpinnings of rapid temporal recalibration effects arising from trial-to-trial history, which reflect in power modulations in the upcoming pre-stimulus interval.

3.5.1. Introduction

Despite the varying physical and physiological sources of asynchrony for different sensory signals reaching our brain (e.g. sound, light), cross-modal stimuli arising from the same event are often subjectively perceived as synchronous (e.g., King, 2005). Hence, the brain must deal with these cross-modal time differences arising from physical transmission and neural processing in order to perceive multisensory events as synchronous. Further temporal variability resulting from factors such as attention (Spence and Parise, 2010), the task instructions (Stevenson and Wallace, 2013) and the immediate history of cross-modal temporal order presentation (Van der Burg et al., 2013), illustrate the existence of flexible, adaptive neural processes underlying synchrony perception. We argue that these factors must reflect on brain states prior to stimulus presentation. The aim of the current study is to address the pre-stimulus neural correlates which may predict synchrony perception of audio-visual events.

Specific patterns of neural oscillations reflect fluctuations in cortical excitability (Buzsáki and Draguhn, 2004; Schroeder and Lakatos, 2009; Dugue et al., 2011), which have been related to variability in trial by trial behavioral performance (van Dijk et al., 2008; Vanrullen et al., 2011). For example, the power (Thut et al., 2006; Mathewson et al., 2009) and the phase (Busch et al., 2009; Mathewson et al., 2009) of ongoing brain

activity in the alpha band (7-14 Hz) in the occipito-parietal cortex have been linked to the successful perception of near-threshold visual stimuli. Several studies highlight the importance of the intrinsic temporal structure of neural oscillations as a means for temporal organization in perception (Varela et al., 1981; VanRullen and Koch, 2003; Herzog et al., 2016). For example, Varela, Toro, John, & Schwartz (1981) found that two briefly presented flashes were perceived as simultaneous or sequential as a function of the phase of alpha fluctuations (~100 ms) (see also VanRullen and Koch, 2003; Milton and Pleydell-Pearce, 2016). Furthermore, the phase/cycle-dependency in perception has been associated to temporal prediction in visual discrimination tasks (Samaha and Postle, 2015) and to the temporal window in which cross-modal illusions are perceived (Cecere et al., 2015).

Given the theoretical and empirical emphasis on the role of neural oscillations in temporal organization during perception, we aimed at assessing whether ongoing oscillatory phase and/or power can predict trial by trial variability in the perception of simultaneity across modalities. We used a non-rhythmic stimulus presentation to avoid phase entrainment by external rhythms (Kösem et al., 2014) and temporal expectations (Lakatos et al., 2008; Cravo et al., 2013), which can drive to temporal organization of perception. This allowed us to examine the role that spontaneous neural oscillations play on temporal organization. We measured EEG activity while participants performed a simultaneity judgment (SJ) task on flash-sound pairs (Figure 1), focusing on asynchronies where perceptual uncertainty was maximal for either modality order (audition leading vision or vice-versa). We retrospectively examined pre-stimulus EEG as a function of the perceptual outcome (simultaneous vs. non-simultaneous report) for otherwise identical stimuli pairs. We found a strong phase-behavior dependency at 13 ± 2 Hz (200 ms previous to stimulus) for pairs

of stimuli in which audition leads vision. Additionally, we investigated how the pre-stimulus interval can be affected by the influence of the immediate previous trial. Indeed, in behavior it has been shown that perception of synchrony depends on the modality order of the previous trial, attributed to rapid temporal recalibration (Van der Burg, Alais & Cass, 2013). In line with this, our data showed that when the preceding presented trial was audition leading vision, participants more often judged as simultaneous pairs of stimuli of the same modality order (and vice-versa). Remarkably, we found that the direction of asynchrony in the previous trial affected oscillatory power within 14 ± 2 Hz just prior to the presentation of the subsequent trial. We therefore reasoned that the ongoing pre-stimulus activity at the current trial might take track of the effect of the previous presented trial.

3.5.2. Materials and methods

3.5.2.1 Participants

Twenty-seven volunteer participants took part in the experiment (14 female; age range: 18-30 years) in exchange for 10€/h. They had normal or corrected to normal vision and hearing, and were naive about the purpose of the experiment. All participants provided written informed consent and all procedures had been previously approved by the local ethical board (CEIC Parc de Mar).

3.5.2.2 Stimuli and apparatus

The experimental paradigm was custom programmed using Psychtoolbox toolbox (Brainard, 1997) on Matlab. The participant sat in a dimly lit quiet room at a distance of 60 cm from a CRT monitor ('PHILLIPS109B', 85 Hz, 800x600 pixels), where visual stimuli were presented. Auditory stimuli were presented through headphones (Sennheiser HD 435). Accurate timing and synchronization (± 1 ms precision) of auditory and visual events was measured using Blackbox

Toolkit (Cambridge Research Systems). The visual stimulus was a white ring (out and inner diameter, 4.2° and 2.1°, respectively) that flashed for 35 ms at the centre of a black square placeholder (9.7°) centered on the light grey background display (58.4 cd/m²). The auditory stimulus was a tone (1.8 kHz, 60 [A]dB SPL) of 35 ms duration, with 2.5 ms ramps in/out.

3.5.2.3 EEG acquisition

EEG was recorded using 60 active electrodes (actiCAP, Brain Products GmbH, Munich, Germany) placed after the 10-20 international system, with the tip of the nose as online reference and AFz as ground. The signal was recorded via BrainVision Recorder (Brain Products GmbH, Munich, Germany) at a sampling rate of 500 Hz. Electrode impedances were kept below 10 kOhm. Horizontal (Heog) and vertical (Veog) electro-oculograms were recorded by two external electrodes, and used for off-line trial exclusion due to eye blink contamination.

3.5.2.4 Procedure: Simultaneity Judgment (SJ) task

Pre-test: Once the EEG cap was on, and prior to EEG recording, we ran a SJ pretest to estimate the audio-visual asynchronies of maximal perceptual uncertainty (which stimulus asynchronies were perceived half of the times as simultaneous and the other half as non-simultaneous), for each participant and modality order (audition leading vision and vision leading audition, AV and VA, respectively). During the SJ pretest, participants judged the simultaneity of flash-tone pairs presented randomly at one out of nine possible stimulus onset asynchronies (SOAs), in a constant stimulus procedure: -353 ms; -/+235 ms; -/+118 ms; -/+59 ms; 0 ms; +412 ms, 16 trials each. The sign denotes the modality order of presentation (negative=AV; positive=VA). Each stimuli pair was presented after a variable interval (randomly chosen between 1047 and 2100 ms) after the onset of a cue, during which participants maintained fixation on a cross at

the centre of the placeholder. Following the flash-beep presentation, a response screen was displayed after a variable delay (between 688 and 1100 ms), prompting participants to report the perceived simultaneity. In order to avoid contamination in the EEG signal due to motor preparation, two keys (“1” and “0”) were randomly and equiprobably assigned to simultaneous and non-simultaneous responses at each trial. Accuracy was stressed over speed.

EEG recording: Participants were tested in a simultaneity judgment task adopting as SOAs the individual A50V (point of maximal uncertainty for auditory leading pairs) and V50A (point of maximal uncertainty for visual leading pairs) SOAs, based on the pretest session. We included trials at the individual PSS (point of subjective simultaneity) plus at two larger flash-tone asynchronies (Audio leading Filler=-353 ms; and, Visual leading Filler=+412 ms, respectively) in order to maintain equiprobable the proportion of (a)synchrony judgments and to avoid possible recalibration after-effects (Fujisaki et al., 2004). We included a jitter (Gaussian distribution, mean=0, SD=80 ms) in the PSS and filler asynchronies (but not in the A50V or V50A) to prevent memorization of the five temporal intervals. Each trial started with the presentation of a white fixation cross at the centre of the screen for a variable interval (2500-3700 ms), followed by the stimuli pair presentation. After a jittered interval (1600-1800 ms), the fixation cross color changed to red and a response screen was displayed, just as in the pretest (Figure 1). Each participant ran between 7 and 8 blocks during which EEG signal was recorded, with short resting breaks in between. In each block there were 16 measurements at each of the PSS, A50V and V50A asynchronies, plus eight Audio and eight Visual Fillers. The number of blocks was adjusted online based on the stability of participants’ simultaneity reports at the A50V and V50A SOAs. In order to minimize muscular artifacts,

participants were asked to maintain fixation during the presentation of the stimuli as much as possible. A short training run consisting of four repetitions of each SOA was run in order to familiarize participants with the SJ task, before the pretest and EEG recording sessions. No feedback about the performance was given throughout the experiment.

3.5.3. Data analysis

3.5.3.1 Behavioral analysis

Pre-test SJ data were individually analyzed prior to the EEG recording session. Individual data from the pre-test SJ task were fitted with two cumulative Gaussian functions (Yarrow et al., 2011. Matlab code downloaded from <http://www.hexicon.co.uk/Kielan/#methods>). Based on this fit, we estimated the points of maximal uncertainty (50% decision boundaries, A50V and V50A) as well as the standard deviation (SD) of the difference between the fitted cumulative Gaussians using the maximum likelihood estimation method. The point of subjective simultaneity (PSS) was taken as the mean value between the previously estimated A50V and V50A points. The EEG recording was conducted only if; (1) the bootstrap procedure for the fitting (Bca method from Efron, 1987) of pretest data succeeded to estimate the 95% confidence intervals (CI) of the model parameters (A50V and V50A), and; (2) the proportion of simultaneity responses for the Audio and Visual fillers (e.g., large asynchronies) was below 0.30. During the EEG recording session, we calculated the proportion of simultaneity responses for each of the tested audio-visual asynchronies. The proportion of simultaneity responses in the filler trials (large asynchronies) was taken as a further control measure of task performance. Six participants who failed to produce at least a 0.5 proportion of asynchrony responses for fillers were removed from the EEG analysis. A total of 21 participants were included in the final EEG data analysis.

3.5.3.2 Pre-stimulus EEG analysis

The data were analyzed in Matlab (MathWorks, Natick, MA), using Fieldtrip toolbox (Oostenveld, Fries, Maris, & Schoffelen, 2011.), CircStat toolbox (Berens, 2009) and custom-made code. An automatic artifact rejection procedure was used to exclude raw epochs in which the signal exceeded 250 μV in a window from 2500 ms pre-stimulus to 500 ms after stimulus onset for the EEG power spectrum analysis, and from 1500 ms pre-stimulus to 1500 ms post-stimulus with respect to the onset of the first stimulus, for the phase distribution analysis. The remaining epochs were visually inspected for eye movement artifacts. Only participants with at least 35 trials per condition after pre-processing were included in the subsequent analysis. This resulted in 10 and 11 participants analyzed for the power spectrum for the A50V and V50A trials, respectively. For phase analysis, the number of participants analyzed in A50V and V50A trials was 10. The signal from corrupted electrodes was discarded and their activity was reconstructed by the interpolation of the activity of the neighboring channels. To prevent sample size related biases in our measurements (e.g. phase concentration measures present a positive bias at low sample size), the number of trials between conditions of interest was equated for each participant by random picking on the most populated condition to equate the less populated condition. The most populated condition measure thus, resulted from the average of 250 random picking iterations.

The S-transform (Stockwell et al., 1996), a generalization of the Short Time Fourier Transform (STFT) based on a moving and scalable Gaussian window with a frequency-dependent resolution, was used for extracting the time evolution of oscillatory power and phase in the trials of interest (A50V and V50A, separately) in the range 1- 25 Hz (steps of 1 Hz), which

are believed to relate to trial by trial variability in perceptual judgments such as sensory detection or cross-modal time.

We compared the power spectra and phase distributions of pre-stimulus EEG for trials later classified as simultaneous vs. non-simultaneous. We did this separately for A50V and V50A conditions. In the power analysis, we compared the time evolution (2000 ms preceding the onset of the first stimulus) of the whole scalp average power spectra for each frequency between simultaneous and non-simultaneous percepts. We used the power ratio (measured in decibels, dB): $\text{Power ratio} = 10 \cdot \log_{10}(\text{Power}_S / \text{Power}_{NS})$, where Power_S is the power in the trials perceived as simultaneous and Power_{NS} the power in trials perceived as non-simultaneous.

For the phase analysis, we calculated the pre-stimulus phase distribution for trials classified as simultaneous and non-simultaneous using the bifurcation index (BI) procedure, after Busch et al., (2009). The BI is used to assess if the phase distributions of two groups of trials (here, trials judged as simultaneous and trials judged as non-simultaneous) are concentrated around different angles. The BI compares the degree of intertrial coherence (ITC) in each type of trial (ITC_S and ITC_{NS} , for trials perceived as simultaneous and non-simultaneous, respectively) to the ITC of all groups of trials collapsed ($\text{ITC}_{\text{collapsed}}$): $\text{BI} = (\text{ITC}_S - \text{ITC}_{\text{collapsed}}) \cdot (\text{ITC}_{NS} - \text{ITC}_{\text{collapsed}})$. As we expected that the effects in phase could be closer to the first stimulus onset time, the BI was calculated in 500 ms preceding the onset of the first stimulus of the pair. The BI for each electrode, time point and frequency was calculated, and the final BI values presented here correspond to average across all the electrodes in the scalp.

3.5.3.3 Analysis as a function of the previous trial

We performed a subsequent analysis motivated by evidence about rapid recalibration effects obtained in behavioral SJ paradigms (e.g. Van der Burg et al., 2013). According to these findings, judgments on the current SJ trial may depend on the temporal order of the stimulus in the previous trial. To further investigate the neural underpinnings of this fast recalibration effect, we calculated the proportion of simultaneity perception as a function of the modality order of the trials preceding A50V and V50A trials. In order to maximize the number of usable stimuli, all trials (for A50V and V50A trials, separately) following A50V and Audio Filler trials' were collapsed into the Previous AV condition and, all trials (for A50V and V50A trials, separately) following V50A and Visual Filler trials' were collapsed into the Previous VA condition. When a trial was preceded by a PSS asynchrony or it was first in a block, it was removed from the analysis. As it was done for the previous analysis, a random picking procedure was performed (1000 repetitions) in order to equate the number of trials between conditions (A50V and V50A). We measured oscillatory power across the frequencies as a function of previous trial modality order (Previous AV or Previous VA) for all the SOAs (A50V, V50A, PSS, Audio Filler and Visual Filler). Please note that in this case, the analysis of pre-stimulus EEG compares the pre-stimulus power in current trials preceded by AV trials, with the pre-stimulus power in current trials preceded by a VA trial. Breaking down by current type of trial is not meaningful, since we looked at pre-stimulus activity, therefore the actually presented trial cannot retrospectively influence ongoing oscillation. Additionally, further break down between different response categories was not possible in this analysis, due to the low number of trials per response category. EEG signals were band-pass filtered between 0.5 and 50 Hz (Butterworth filter, order 8) and re-referenced offline to the right and left mastoid channels. An automatic

artifact rejection procedure excluded band-passed filtered epochs in which the signal exceeded 100 μV in the segments of interest, from 2500 ms pre-stimulus to 500 ms post-stimulus with respect to the onset of the first stimulus, (except for one participant where the threshold at 250 μV was maintained after visual inspection of trials, as this participant showed high alpha activity), and the remaining data were screened manually for eye movement or blink artifacts. Four out of the 21 participants included based on the behavioral criteria (see above) were further removed from the analysis because of bad quality of recorded signal. The analysis was done following the rationale of van Dijk et al., (2008). The average power was calculated in the 1000 ms interval preceding the first stimulus onset. In this analysis, we calculated a STFT over the whole 1000 ms interval previous to stimulus onset (time course was not considered) because from the behavioral results we expected carry over effects coming from previous trial to be sustained in time through to the current trial. We selected the frequency bands of interest according to expected effects regarding perceptual predictions or the maintenance of a cognitive or perceptual state (Iversen et al., 2009; Engel and Fries, 2010). The frequency-bands of interest ranged therefore from 8 to 25 Hz. A multi-taper spectral estimation method was applied (3 orthogonal Slepian tappers). Each tapered data epoch was subsequently Fourier transformed and the power-spectral densities averaged. Power values were z-score normalized and the power interval was averaged with a ± 2 Hz frequency smoothing.

3.5.3.4 Statistical analysis

Differences in pre-stimulus simultaneous versus non simultaneous and in the contrast Previous AV versus Previous VA power were assessed using paired t-test (two-tailed, significance level $p < 0.05$), with correction for multiple comparisons following a cluster based randomization procedure

(1×10^4 randomizations). For each significant cluster, the measure of the cluster consisted on the sum of the t-values of the time-frequency points. We considered that a cluster was significant when the probability of obtaining a measure as big as the observed one by chance was below 5% ($p=0.05$).

Changes in phase concentration were statistically assessed following the bifurcation Index, after Busch et al. (2009). The significance of the BI was measured relative to a null distribution built from our dataset by pseudo-BI (pBI) calculation. To correct for multiple comparisons we applied the false discovery rate (fdr), which has a good control for type I errors, but not for type II. We also added a control test for outliers, in which the statistical test was repeated n (n being the number of participants) times, removing a different participant each time. Only points surviving in each and every repetition of the test ($p < 0.05$) were considered significant. Further phase statistics were performed only in the most significant electrode-frequency point of both A50V and V50A trials. The Rayleigh test (Fisher, 1996), a measure to examine uniformity in directional data, was used to check whether the phase in the selected pre-stimulus electrode/time point for simultaneous and non-simultaneous percepts were significantly concentrated or not. The phase difference index (Bonfond and Jensen, 2012) was calculated to test separated phase distribution (simultaneous or non-simultaneous perception) versus collapsed (simultaneous and non-simultaneous reports). Specifically, we compared ITC for simultaneous (N trials) and ITC for non-simultaneous percepts (N trials) vs. a null distribution built by mixing equal number of trials for simultaneous and non-simultaneous percepts (N trials).

For the analysis as a function of the previous trial, z-scored power differences in frequency-electrode pairs were assessed by means of a two-tailed paired t-test. We used a montecarlo cluster procedure as

programmed in Fieldtrip (minimum number of neighbours=2, N=10000 permutations) for multiple comparison correction.

3.5.4. Results

3.5.4.1 Behavioral pre-test

At group level (N=27), the average parameter values obtained from the pretest psychophysical data fits were PSS=6 ms, A50V=-184 ms, and V50A=196 ms. The average SD, reflecting the width of the window of synchrony perception, was 80 ms. Confidence intervals (CI) obtained from the fitting of two cumulative Gaussian functions used to model the data provide a measure of how precisely each A50V and V50A point was estimated. A paired sample t-test (two tailed) was conducted to assess possible differences in the CI for the A50V and V50A estimations. The CI (\pm the standard deviation) was wider for the V50A (82 ± 27 ms) estimate compared to the A50V (73 ± 16 ms) estimate ($t(26)=2.134$, $p=0.042$), suggesting less reliability for the estimation of vision leading trials compared to audition leading trials. This is in line with previous literature showing that synchrony perception has greater tolerance and flexibility for vision leading than audition leading asynchronies (van Eijk et al., 2008; Yarrow et al., 2011; Stevenson and Wallace, 2013; Cecere et al., 2016).

During the EEG session, the average proportion of simultaneity responses (\pm the standard deviation) for the 21 selected participants resulted in PSS= 0.8 ± 0.19 and, critically, for the A50V and V50A asynchronies simultaneous perception was close to the intended proportion of half synchronous and half asynchronous responses (0.46 ± 0.20 for A50V and 0.55 ± 0.23 for V50A). The average proportion of simultaneous responses for Audio Filler and Visual Filler trials were 0.07 ± 0.11 and 0.11 ± 0.13 , respectively, ensuring that at sufficiently large asynchronies, participants would answer asynchronous most of the time.

3.5.4.2 Pre-stimulus EEG

The EEG analysis focused on the (pre-stimulus) interval prior to the presentation of the A50V and V50A audio-visual pairs. The mean number of usable trials per participant for the A50V and V50A trials was 41 ± 9 and 46 ± 6 , respectively. We applied exactly the same analysis to the A50V and V50A trials. The analysis of time-resolved power ratios comparing trials perceived as simultaneous vs. trials perceived as non-simultaneous did not show any significant differences, neither on A50V nor V50A trials, after multiple comparison correction (time-frequency cluster).

Interestingly, regarding phase-dependency effects, we found a significant phase bifurcation index (BI; $p < 0.05$, fdr corrected) of adjacent time frequency points that survived outliers control ($p < 0.05$, fdr in all averages) in the A50V trials in the range 220 to 180 ms prior to stimulus presentation, at 13 ± 2 Hz. The most significant time-frequency point was found at 13 Hz and 200 ms prior to stimulus presentation ($p = 0.00006$). Later, we observed that the electrode with higher BI at this time-frequency point was TP9³ (BI=0.008) (Figure 2A). In the V50A trials, unlike it happened in the A50V analysis, no significant BI point survived the correction based on outliers control in the comparison between simultaneous vs. non-simultaneous trials.

A positive BI indicates that the phase (in the particular frequency, time and location) concentrates in different angles for trials that will be later perceived as simultaneous, compared to trials that will be later perceived

³ Subsequently, in a post-hoc analysis, we further investigated if the phase-dependent simultaneity perception in the selected time-frequency point extended to surrounding electrodes. Indeed, a cluster of neighboring electrodes (F7, FT7, FT9, T7, TP7, TP9) displayed a significant behavioral modulation ($p < 0.05$). Although at cluster level the set of electrodes did not survive multiple comparison correction (cluster correction $N = 10000$ randomizations; the probability of false positive being $p = 0.07$) the result confirms a trend for spatial concentration of the phase-dependent effect.

as non-simultaneous. We therefore decided to follow up on the phase distribution of the A50V trials perceived as simultaneous vs. non-simultaneous. As it would be expected, phase distribution was uniform when collapsing the phases of simultaneously and non-simultaneously perceived trials (only 1 participant displayed a p-value below 0.05). The phase difference index showed that 6 out of 10 participants displayed at least a trend of having one of the distributions more concentrated than the collapsed one. The ITC was comparable between conditions (2 tail t-test ITC simultaneous vs. ITC non-simultaneous, $p=0.28$), which is a pre-condition required for performing phase difference index measures. The mean phase angles for simultaneous and non-simultaneous perception across participants were examined. Mean phase angle for simultaneous trials, at the frequency/time/electrode of interest, was -110 degrees ($^{\circ}$) and for non-simultaneous was 28° (Figure 2C). Thus, the mean of the difference between simultaneous and non-simultaneous mean phases was $-165^{\circ}\pm 56^{\circ}$ (hence, close to opposite phases, at 180° across participants). The data from each participant was pooled into 2 phase bins and the proportion of simultaneity perception was calculated (Figure 2D, shows averaged data divided into 6 phase bins). The bins were centered on the mean phase for simultaneity perception and its opposite phase (180°). The proportion of simultaneity perception was significantly larger for the bin centered at the mean phase for simultaneity ($t(9)=3.25$, $p=0.005$, one-tailed) compared to its opposite phase bin (mean phase $+180^{\circ}$). Finally, phases from all participants were collapsed in a single dataset (as in Biau et al., 2014) in order to check whether there was a general favorable phase for simultaneous perception. The Rayleigh test indicated that only the mean phase for simultaneous percept ($p=0.03$ vs. non-simultaneous $p=0.10$) was not uniformly distributed. As expected the phase was uniformly distributed in the collapsed condition ($p=0.58$). The phase difference index showed that the distribution of phases for simultaneous

(519 trials) and non-simultaneous (606 trials) percepts were significantly different from the distribution of the collapsed dataset (p-value for simultaneous was 0.01; and for non-simultaneous 0.046). Critically, this phase-dependent effect was not accompanied by differences in power spectrum between simultaneous and non-simultaneous percepts at any frequency considered for the analysis⁴. This particular finding indicates that the phase-dependent effect observed here was not based on oscillatory amplitude differences; hence, the reliability of our phase measurements was the same in both simultaneous and non-simultaneous perceptual outcomes.

For completeness, exactly the same analysis was applied to the V50A trials in order to detect some trend toward phase-behavior correspondence. The most significant BI value, previous to outlier control ($p=0.00003$, fdr corrected) was obtained at 11 ± 2 Hz around 160 ms before first stimulus presentation. The stronger BI at this time-frequency point was found at the CP3 electrode (also at posterior left hemisphere, but more central). However, this point did not survive the correction based on outliers and it did not show any phase-behavior correspondence. Specifically, mean phase difference between simultaneous (mean phase= 103°) and non-simultaneous (mean phase= -9°) responses was $-118^\circ\pm 66^\circ$ and the proportion of synchrony perception was not greater for the bin centered at the mean phase for simultaneity ($t(9)=-0.363$, $p=0.72$) compared to its opposite phase bin. In addition, when both datasets (trials perceived as simultaneous and trials perceived as non-simultaneous) were collapsed, neither the Rayleigh test nor the phase difference index showed any

⁴ We also ran an optimized power analysis at the frequency range of interest (11-16 Hz). We calculated the contrast between simultaneous and non-simultaneous trials in a wider time window (0.5 s pre-stimulus), using a STFT with Slepian tappers and a frequency smoothing of 2 Hz. No significant clusters were found for A50V and V50A trials.

significant phase concentration for simultaneous (526 trials) and non-simultaneous percepts (567 trials).

Notably, we found a strong negative value of the BI at 4 ± 2 Hz right before stimulus presentation (from -246 ms to 0 ms for A50V trials, and from -28 ms to 0 ms for V50A trials). According to the interpretation of the BI, this would imply that one of the two perceptual outcomes (simultaneous or non-simultaneous) was strongly concentrated whereas the other would have a uniform distribution of phase angles. Further checking using the short time Fourier transform (STFT) method suggested that the activity at 4 ± 2 Hz might be related to contamination from post-stimulus (stimulus evoked) responses, and was therefore not considered any further.

3.5.4.3 Analysis as a function of the previous trial

Behavioral

To assess whether the synchrony perception for A50V and V50A current trials differed as a function of the modality order on the previous trial, we ran a two way repeated measure ANOVA with current trial (A50V and V50A) and previous trial (Previous AV and Previous VA) as within participant factors. Main effect of current trials was significant ($F(1,20)=4.7777$, $p=0.0409$), whilst the main effect of previous trial was not significant ($F(1,20)=1.5324$, $p=0.2301$). Interestingly, we found an interaction between current trial and previous trial ($F(1,20)=33.8591$, $p=0.000010801$). In order to follow up this interaction, we compared (paired-sample t-test, two tailed) the proportion of synchrony responses for Previous AV vs. Previous VA trials, within the A50V and V50A current trials separately (Figure 3). The results showed that in the A50V trials, simultaneity perception was higher ($t(20)=2.302$, $p=0.032$) when the preceding trial was AV (0.49 ± 0.20) than when it was VA (0.43 ± 0.22). In V50A trials, the effect was also significant ($t(20)=-5.474$, $p=0.00002$),

because simultaneity responses were higher when the preceding trial had the same modality order ($VA=0.61\pm0.23$), compared to when it had the opposite order ($AV=0.50\pm0.25$). Therefore, this behavioral result goes well in line with previous reports of prior trial influence on current cross-modal simultaneity judgments: When the preceding trial was the same modality order pair respect to the current trial, participants tended to perceive the pair as synchronous (Van der Burg et al., 2013).

Pre-stimulus EEG

To try and understand the possible neural underpinning of this trial to trial carry over effect observed at behavioral level, we performed a power analysis for the 1000 ms window preceding stimulus presentation. The analysis was conducted on a mean number of 147 ± 30 (Previous AV) and 148 ± 33 (Previous VA) trials. We selected the significant frequencies of interest using a paired t-test between Previous VA vs. Previous AV trials. The results indicated that the most robust pre-stimulus power difference between Previous VA and Previous AV trials was at 14 ± 2 Hz, surviving the multiple comparison correction (cluster) (Figure 4A). Specifically, a significant positive cluster ($p=0.018$) was found at 14 ± 2 Hz for a large number of neighboring channels mostly located in the right hemisphere (Fz, FC1, FC2, FC4, C2, C4, C6, T8, CP4, CP6, Pz, P2, P4, PO4, POz, PO3, O1, Oz) (Figure 4B). Hence, the power in the specified frequency range, in this cluster of electrodes, was higher when audition led vision in the previous trial, compared to when vision was leading. Finally, we classified trials in three bins as a function of z-power (normalized power values) at 14 ± 2 Hz in this cluster of electrodes (Figure 4C), and calculated the percentage of trials belonging to Previous AV and Previous VA conditions in each bin. The percentage of Previous AV trials significantly differ to the percentage of Previous VA trials across the bins in the expected direction ($F(2,17)=6.31$, $p=0.0047$): Namely, the bin with highest

14 Hz power resulted the one with higher proportion of Previous AV compared to the other two bins with less 14Hz power ($t(17)=3.2553$, $p=0.0047$, and $t(17)=-2.6124$, $p=0.0182$), which were not different between them ($t(17)=0.8398$, $p=0.4126$). Overall these results suggest a power modulation at 14 ± 2 Hz during the pre-stimulus period of a given trial, as a function of the modality order of the previously presented trial.

3.5.5. Discussion

In this study, we investigated trial to trial variability of the perception of cross-modal simultaneity by means of pre-stimulus EEG. We found that brain state before the stimulus presentation can determine perceptual variability on otherwise identical audio-visual pairs. In our study, each asynchrony direction (audition leading and vision leading) was analyzed separately, given that differences in the underlying neural processes could be expected between modality orders (Thorne and Debener, 2014; Cecere et al., 2016). One important aspect of our results is that the phase of ongoing activity at $13 (\pm 2)$ Hz, 200 ms prior to stimulus onset, is relevant for audio-visual synchrony perception, with preferential phases for synchronous and asynchronous perception being nearly opposite (close to 180°). This finding supports the idea that ongoing oscillations might provide windows for temporal organization via fluctuations of excitability, structuring the perception of audio-visual events (in line with Schroeder and Lakatos, 2009; Cravo et al., 2013; Kösem et al., 2014). This framework is compatible with previous electrophysiological studies showing that the phase/cycle of ongoing brain activity could determine subsequent temporal discrimination of pairs of visual events (Varela et al., 1981; Cecere et al., 2015; Samaha and Postle, 2015; Milton and Pleydell-Pearce, 2016). While most of these studies reported activity around 10 Hz as responsible of perceptual organization, a 13 Hz (~ 80 ms cycle) frequency was obtained in this study. A remarkable difference in our

study, compared to the other, is that we focused on audio-visual events, instead of only visual stimuli. A cycle of about 80 ms (13 Hz) matches the standard deviation value at group level (80 ms) we obtained from the analysis of behavioral data, thus reflecting the temporal window of simultaneity in our audio-visual paradigm. In a recent study using a cross-modal simultaneity judgment paradigm similar to ours but focusing on post-stimulus EEG activity, synchrony perception of identical audio-visual events correlated with stronger post-stimulus phase reset at 18 Hz (Kambe et al., 2015). Specifically, when the EEG activity from the sensory area related to the lagging modality of an audio-visual pair was phase reset, perception of synchrony was greater. We tentatively argue that the strength of this stimulus-evoked phase reset might well depend on brain states prior to stimulus presentation such as the ones reported here. In particular, the pre-stimulus phase of neural oscillations at 13 ± 2 Hz in our study could have had an impact in determining the strength of post-stimulus phase reset and hence, the perceptual outcome in synchrony perception. Please note, however, that Kambe et al. (2015) found phase resetting for both modality orders whereas in our study the phase at 13 ± 2 Hz explains perceptual variability only for audition leading vision pairs. We discuss this issue next.

We entertain two possible interpretations regarding the lack of evidence for phase-synchrony relationship in vision leading audition pairs. First, both at a behavioral and at a physiological level, we systematically found that audition leading vision events were associated to more consistent (less variable) effects. Indeed, in behavior, the confidence interval obtained for the A50V (audition leading) estimate was significantly narrower, compared to the V50A (vision leading) estimate. Given that in our study the size of the confidence intervals reflects the variance of the estimated synchrony boundaries (A50V and V50A), it is tempting to

consider that greater behavioral variance in vision leading asynchronies might have blurred the detection of possible phase effects in the EEG. Other psychophysical studies reported a greater variability from trial to trial synchrony criteria for vision leading pairs (Yarrow et al., 2011) and a lower sensitivity in asynchrony detection when vision leads audition (Conrey and Pisoni, 2006; van Wassenhove et al., 2007; Stevenson and Wallace, 2013). As a further alternative, or perhaps complementary explanation, one should also consider the possibility that synchrony perception might rely on partially different neural processes relating to modality specific sensory cortices, depending on what is the leading and lagging sensory modality (Thorne and Debener, 2014; Cecere et al., 2016). Indeed, greater improvement after perceptual training for vision leading compared to audition leading events (Powers et al., 2009; Cecere et al., 2016) have been interpreted according to separate processes governing synchrony perception as a function of the leading modality (Cecere et al., 2016). Similar asymmetries in synchrony processing across modality orders have also been found in electrophysiological correlates (Kösem et al., 2014; Kaganovich and Schumaker, 2016). Here, we cannot discern between these two types of interpretation, but definitely other factors influencing pre-stimulus activity, such as previous experience in multisensory environments might have been missed in this picture.

Another important result from our study is that we considered the influence of the previously presented trial on the current trial. Past behavioral evidence suggests possible carry-over effects from trial to trial in the domain of temporal recalibration (Van der Burg et al., 2013), and went further in the investigation of the oscillatory expression of this effect. Whilst our behavioral results go in the same direction as prior evidence (increased simultaneity judgments when the preceding trial had the same asynchrony direction as the current one), we reasoned that the

neural oscillatory state before the current trial should contain the expression of the influence of the previous trials. We expected effects in the spectral dynamic (i.e. power), instead of the fine temporal structure of the oscillatory mechanisms (i.e. phase). Indeed, if any physiological expressions of the carry-over effect could account for the differential perceptual judgment at behavioral level, it should be evident in the modulation of fluctuations of cortical excitability sustainable in time. Analysis of the pre-stimulus EEG activity as a function of the modality order of the previous trial showed that pre-stimulus power at 14 ± 2 Hz increased over the right frontal-centro-parietal-occipital channels when in the previous trial audition led vision, compared to when it was vision to lead. To the best of our knowledge, this is the first time that oscillatory power differences (at 14 ± 2 Hz) as a function of previous trials explains the behavioral shifts on audio-visual synchrony perception (Fujisaki et al., 2004; Van der Burg et al., 2013). Unfortunately, our study was not initially designed for running analysis on the pre-stimulus oscillations as a function of the previously presented trial; therefore, we could not further split the effects based on perceptual reports. However, we think this way of considering pre-stimulus on-going oscillations as the neurophysiological context influenced by the previous history (as frequently and consistently reported at behavioral level in many cognitive domains, such as expectancy (Foxe et al., 1998), sequential conflict effects (Botvinick et al., 2001)), can be promising in for understanding the nature of ongoing activity, and the link between neurophysiology and human behavior.

To conclude, this study reveals a clear correlation between the phase of the ongoing oscillations and the subjective perception of cross-modal synchrony. Moreover, our findings strongly suggest that the influence of immediate prior experience (regarding modality order) is expressed in the

amplitude of neural activity during the inter-trial period. Hence, the interplay between phase and power of ongoing oscillations might help construct temporal templates to update what we experience as cross-modal simultaneity. Here, we have explored the influence of prior context on ongoing activity, but we believe other factors known to modulate synchrony judgments such as attention and the task at hand, might exert an influence on this ongoing activity. Further exploration of these factors might help further understanding the neural correlates behind the remarkable plasticity of synchrony perception in multisensory scenarios.

3.5.6. References

- Berens P (2009) CircStat: A Matlab Toolbox for Circular Statistics. *J Stat Softw* 31 Available at: http://www.kyb.mpg.de/fileadmin/user_upload/files/publications/attachments/MPIK-TR-184_%5b0%5d.pdf [Accessed May 28, 2016].
- Biau E, Torralba M, Fuentemilla L, de Diego Balaguer R, Soto-Faraco S (2015) Speaker's hand gestures modulate speech perception through phase resetting of ongoing neural oscillations. *Cortex* 68:76–85.
- Bonnefond M, Jensen O (2012) Alpha Oscillations Serve to Protect Working Memory Maintenance against Anticipated Distracters. *Curr Biol* 22:1969–1974 Available at: <http://linkinghub.elsevier.com/retrieve/pii/S0960982212009943>.
- Botvinick MM, Braver TS, Barch DM, Carter CS, Cohen JD (2001) Conflict monitoring and cognitive control. *Psychol Rev* 108:624–652 Available at: <http://www.ncbi.nlm.nih.gov/pubmed/11488380> [Accessed July 29, 2016].
- Brainard D (1997) The psychophysics toolbox. *Spat Vis* 10:433–436 Available at: <http://bbs.bioguider.com/images/upfile/2006-4/200641014348.pdf> [Accessed May 28, 2016].
- Busch NA, Dubois J, VanRullen R (2009) The phase of ongoing EEG oscillations predicts visual perception. *J Neurosci* 29:7869–7876.
- Buzsáki G, Draguhn A (2004) Neuronal oscillations in cortical networks. *Science* 304:1926–1929.

- Cecere R, Gross J, Thut G (2016) Behavioural evidence for separate mechanisms of audiovisual temporal binding as a function of leading sensory modality. *Eur J Neurosci* Available at: <http://www.ncbi.nlm.nih.gov/pubmed/27003546> [Accessed April 4, 2016].
- Cecere R, Rees G, Romei V (2015) Individual differences in alpha frequency drive crossmodal illusory perception. *Curr Biol* 25:231–235 Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=4300399&tool=pmcentrez&rendertype=abstract> [Accessed April 12, 2016].
- Conrey B, Pisoni DB (2006) Auditory-visual speech perception and synchrony detection for speech and nonspeech signals. *J Acoust Soc Am* 119:4065–4073 Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3314884&tool=pmcentrez&rendertype=abstract> [Accessed May 12, 2016].
- Cravo AM, Rohenkohl G, Wyart V, Nobre AC (2013) Temporal expectation enhances contrast sensitivity by phase entrainment of low-frequency oscillations in visual cortex. *J Neurosci* 33:4002–4010 Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3638366&tool=pmcentrez&rendertype=abstract> [Accessed August 12, 2013].
- Dugue L, Marque P, VanRullen R (2011) The Phase of Ongoing Oscillations Mediates the Causal Relation between Brain Excitation and Visual Perception. *J Neurosci* 31:11889–11893 Available at: <http://www.jneurosci.org/cgi/doi/10.1523/JNEUROSCI.1161-11.2011>.
- Efron B (1987) Better Bootstrap Confidence Intervals. *J Am Stat Assoc* 82:171–185 Available at: <http://www.tandfonline.com/doi/abs/10.1080/01621459.1987.10478410> [Accessed June 2, 2016].
- Engel AK, Fries P (2010) Beta-band oscillations—signalling the status quo? *Curr Opin Neurobiol* 20:156–165 Available at: <http://linkinghub.elsevier.com/retrieve/pii/S0959438810000395>.
- Fisher NI (1996) *Statistical Analysis of Circular Data*. Cambridge University Press. Available at: <http://www.cambridge.org/us/academic/subjects/physics/general-and-classical-physics/statistical-analysis-circular-data> [Accessed May 29, 2016].

- Foxe JJ, Simpson G V, Ahlfors SP (1998) Parieto-occipital approximately 10 Hz activity reflects anticipatory state of visual attention mechanisms. *Neuroreport* 9:3929–3933.
- Fujisaki W, Shimojo S, Kashino M, Nishida S (2004) Recalibration of audiovisual simultaneity. *Nat Neurosci* 7:773–778 Available at: <http://www.ncbi.nlm.nih.gov/pubmed/15195098> [Accessed March 10, 2012].
- Herbst SK, Landau AN (2016) Rhythms for cognition: the case of temporal processing. *Curr Opin Behav Sci* 8:85–93 Available at: <http://linkinghub.elsevier.com/retrieve/pii/S2352154616300213>.
- Herzog MH et al. (2016) Time Slices: What Is the Duration of a Percept? *PLOS Biol* 14:e1002433 Available at: <http://dx.plos.org/10.1371/journal.pbio.1002433> [Accessed July 2, 2016].
- Iversen JR, Repp BH, Patel AD (2009) Top-down control of rhythm perception modulates early auditory responses. *Ann N Y Acad Sci* 1169:58–73 Available at: <http://www.ncbi.nlm.nih.gov/pubmed/19673755> [Accessed March 14, 2016].
- Kaganovich N, Schumaker J (2016) Electrophysiological correlates of individual differences in perception of audiovisual temporal asynchrony. *Neuropsychologia* 86:119–130 Available at: <http://www.sciencedirect.com/science/article/pii/S0028393216301233> [Accessed May 8, 2016].
- Kambe J, Kakimoto Y, Araki O (2015) Phase reset affects auditory-visual simultaneity judgment. *Cogn Neurodyn* 9:487–493 Available at: <http://www.ncbi.nlm.nih.gov/pubmed/26379799> [Accessed April 28, 2016].
- King AJ (2005) Multisensory integration: strategies for synchronization. *Curr Biol* 15:R339–R341 Available at: <http://www.ncbi.nlm.nih.gov/pubmed/15886092> [Accessed August 14, 2012].
- Kösem A, Gramfort A, van Wassenhove V (2014) Encoding of event timing in the phase of neural oscillations. *Neuroimage* 92:274–284 Available at: <http://www.sciencedirect.com/science/article/pii/S1053811914001013> [Accessed November 20, 2014].

- Lakatos P, Karmos G, Mehta AD, Ulbert I, Schroeder CE (2008) Entrainment of neuronal oscillations as a mechanism of attentional selection. *Science* 320:110–113 Available at: <http://www.ncbi.nlm.nih.gov/pubmed/18388295> [Accessed June 8, 2016].
- Mathewson KE, Gratton G, Fabiani M, Beck DM, Ro T (2009) To see or not to see: prestimulus alpha phase predicts visual awareness. *J Neurosci* 29:2725–2732.
- Milton A, Pleydell-Pearce CW (2016) The phase of pre-stimulus alpha oscillations influences the visual perception of stimulus timing. *Neuroimage* 133:53–61 Available at: <http://www.sciencedirect.com/science/article/pii/S1053811916001749> [Accessed March 5, 2016].
- Oostenveld R, Fries P, Maris E, Schoffelen J-M (2011) FieldTrip: Open source software for advanced analysis of MEG, EEG, and invasive electrophysiological data. *Comput Intell Neurosci* 2011:156869 Available at: <http://dl.acm.org/citation.cfm?id=1945410.1945411> [Accessed December 10, 2015].
- Powers AR, Hillock AR, Wallace MT (2009) Perceptual training narrows the temporal window of multisensory binding. *J Neurosci* 29:12265–12274 Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2771316&tool=pmcentrez&rendertype=abstract> [Accessed July 16, 2012].
- Rohenkohl G, Cravo AM, Wyart V, Nobre AC (2012) Temporal expectation improves the quality of sensory information. *J Neurosci* 32:8424–8428 Available at: <http://www.ncbi.nlm.nih.gov/pubmed/22699922> [Accessed August 7, 2013].
- Samaha J, Postle BR (2015) The Speed of Alpha-Band Oscillations Predicts the Temporal Resolution of Visual Perception. *Curr Biol* 25:2985–2990 Available at: <http://www.ncbi.nlm.nih.gov/pubmed/26526370> [Accessed November 1, 2015].
- Schroeder CE, Lakatos P (2009) Low-frequency neuronal oscillations as instruments of sensory selection. *Trends Neurosci* 32:9–18.
- Spence C, Parise C (2010) Prior-entry: a review. *Conscious Cogn* 19:364–379 Available at:

<http://www.sciencedirect.com/science/article/pii/S1053810009001871> [Accessed April 4, 2016].

- Stevenson RA, Wallace MT (2013) Multisensory temporal integration: task and stimulus dependencies. *Exp brain Res* 227:249–261 Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3711231&tool=pmcentrez&rendertype=abstract> [Accessed April 2, 2016].
- Stockwell RG, Mansinha L, Lowe RP (1996) Localization of the complex spectrum: the S transform. *IEEE Trans Signal Process* 44:998–1001 Available at: <http://ieeexplore.ieee.org/articleDetails.jsp?arnumber=492555> [Accessed February 15, 2016].
- Thorne JD, Debener S (2014) Look now and hear what 's coming : On the functional role of cross-modal phase reset. *Hear Res* 307:144–152 Available at: <http://dx.doi.org/10.1016/j.heares.2013.07.002>.
- Thut G, Nietzel A, Brandt SA, Pascual-Leone A (2006) Alpha-band electroencephalographic activity over occipital cortex indexes visuospatial attention bias and predicts visual target detection. *J Neurosci* 26:9494–9502 Available at: <http://www.ncbi.nlm.nih.gov/pubmed/16971533> [Accessed July 25, 2016].
- Van der Burg E, Alais D, Cass J (2013) Rapid Recalibration to Audiovisual Asynchrony. *J Neurosci* 33:14633–14637 Available at: <http://www.jneurosci.org/content/33/37/14633.abstract> [Accessed September 11, 2013].
- van Dijk H, Schoffelen J-M, Oostenveld R, Jensen O (2008) Prestimulus oscillatory activity in the alpha band predicts visual discrimination ability. *J Neurosci* 28:1816–1823 Available at: <http://www.jneurosci.org/content/28/8/1816.long> [Accessed March 29, 2016].
- van Eijk RLJ, Kohlrausch A, Juola JF, van de Par S (2008) Audiovisual synchrony and temporal order judgments: effects of experimental method and stimulus type. *Percept Psychophys* 70:955–968 Available at: <http://www.ncbi.nlm.nih.gov/pubmed/18717383> [Accessed July 2, 2013].
- van Wassenhove V, Grant KW, Poeppel D (2007) Temporal window of integration in auditory-visual speech perception. *Neuropsychologia*

45:598–607 Available at:
<http://www.ncbi.nlm.nih.gov/pubmed/16530232> [Accessed May 23, 2013].

Vanrullen R, Busch NA, Drewes J, Dubois J (2011) Ongoing EEG Phase as a Trial-by-Trial Predictor of Perceptual and Attentional Variability. *Front Psychol* 2:60 Available at:
<http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3110813&tool=pmcentrez&rendertype=abstract> [Accessed May 5, 2016].

VanRullen R, Koch C (2003) Is perception discrete or continuous? *Trends Cogn Sci* 7:207–213.

Varela F, Toro A, John E, Schwartz E (1981) Perceptual framing and cortical alpha rhythm. :675–686.

Yarrow K, Jahn N, Durant S, Arnold DH (2011) Shifts of criteria or neural timing? The assumptions underlying timing perception studies. *Conscious Cogn* 20:1518–1531 Available at:
<http://www.ncbi.nlm.nih.gov/pubmed/21807537> [Accessed April 23, 2012].

3.5.7. Figures

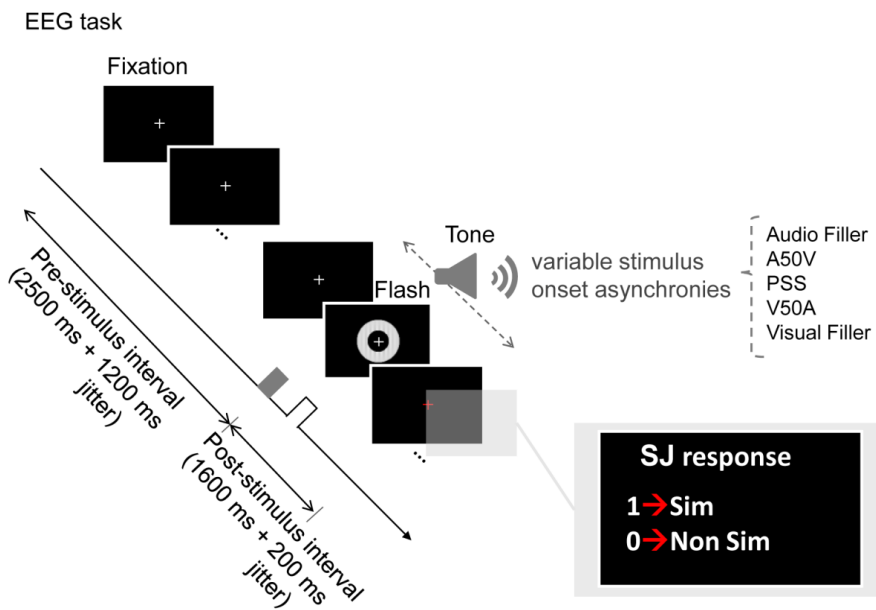


Figure 1. Simultaneity Judgment task. Schematic representation of the timeline of the task during the EEG recording. Trials started with the presentation of a white fixation cross at the centre of the screen followed by a jittered interval after which a flash-tone asynchrony was presented. The flash-tone asynchrony could be an Audio filler (tone leading a flash by 353 ms), a Visual filler (flash leading a tone by 412 ms), the individually estimated point of maximal simultaneity (PSS) or one of the estimated points of maximal uncertainty (A50V and V50A trials corresponding to Audio leading and Visual leading asynchronies, respectively). Subsequently, the fixation cross turned red and a response screen was displayed after a variable interval. Participants were instructed to report the perception of simultaneity of the flash-tone pair, by pressing one of the two keys (“1” or “0”), signaling simultaneity and non-simultaneity reports.

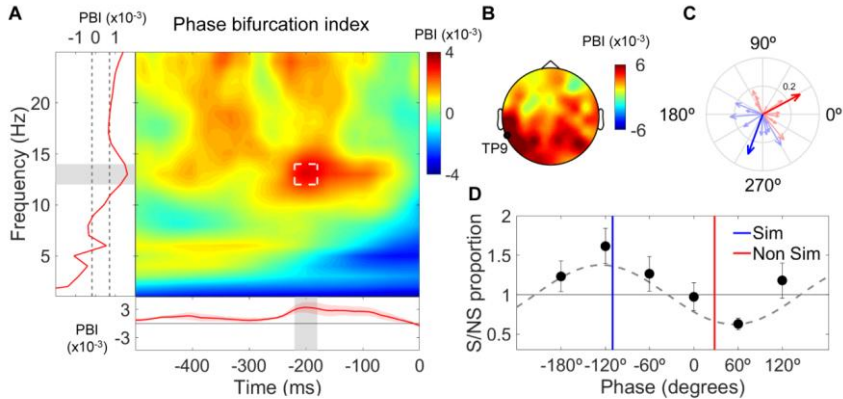


Figure 2. Effects of oscillatory phase for the A50V trials. A, Phase bifurcation index (BI) averaged across all channels and participants. Positive values indicate that phases for simultaneous and non-simultaneous percept are concentrated and different. Negative values indicate that for only one of the two possible simultaneity percepts the phase is concentrated. Left inset depicts BI averaged across time points in the pre-stimulus interval and the two dashed lines represent the 95% confidence interval). The main difference was found in the 13 ± 2 Hz frequency range. Bottom inset depicts BI at 13Hz, with a maximal BI around 200 ms preceding the first stimulus onset. B, The topography showed the distribution of the 13 ± 2 Hz, 200 ms preceding the first stimulus onset. Activity was higher at the TP9 electrode. C, Mean phase angle for simultaneous (blue arrow) and non-simultaneous (red arrow) perception and the corresponding individual phase angles (shadowed arrows). Length indicates the degree of concentration and the direction of the average phase. D, Mean phase-behavior correspondence. Relationship between phase at 13 ± 2 Hz, 200 ms pre-stimulus and standardized performance on audio-visual synchrony perception. 6 bins were chosen, centered at the mean phase for simultaneous perception (blue vertical line) and its opposite phase (red vertical line).

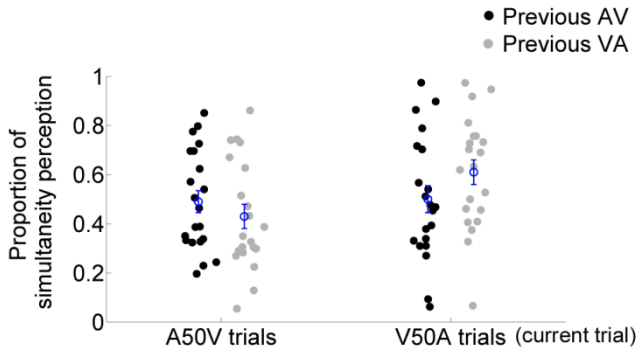


Figure 3. Previous trial individual behavioral data. Proportion of synchrony perception for the maximal uncertain asynchronies (A50V: audition leading; V50A: vision leading) modulated as a function of the previous trial modality order (Previous AV: audition leading vision; Previous VA: vision leading audition). As predicted, synchrony perception was higher for A50V trials when in the preceding trial an auditory event was leading; and for the V50A trials when a visual event was leading in the trial before. Blue circles depict the mean and the corresponding SEM for each condition.

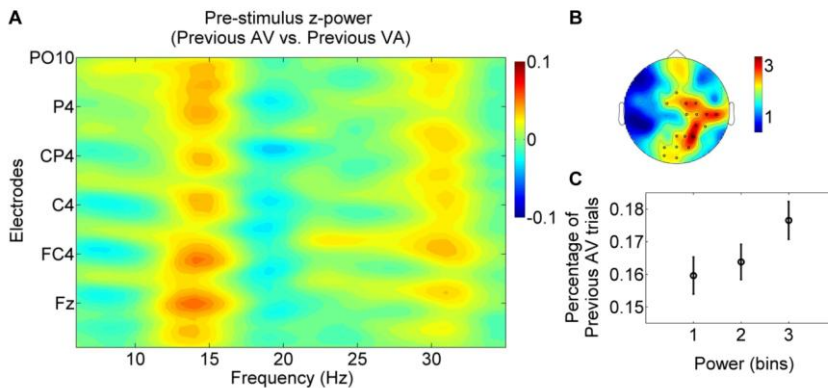


Figure 4. Effects of oscillatory power on the previous trial EEG analysis. A, Difference in spectral z-power (normalized power values) averaged across a time window of interest (1 second preceding first stimulus onset) and participants. We calculated the z-power difference between Previous AV (audition was leading in the preceding trial) and Previous VA (vision was leading in the preceding trial) condition. The main difference was found in the 14 ± 2 Hz frequency range. B, The topography showed the distribution of the 14 ± 2 Hz at 1 second interval preceding first stimulus onset. Activity was higher at a large cluster of electrodes (indicated by black dots). C, Previous AV and Previous VA trials were sorted according to the pre-stimulus 14 ± 2 Hz power and binned into 3 bins. The percentage of Previous AV trials was calculated for each bin. As expected, the percentage of Previous AV trials was higher as the power at 14 ± 2 Hz increased. Error bars depict the confidence intervals.

CHAPTER 4. GENERAL DISCUSSION

The three studies presented in this dissertation aim at understanding how humans organize the temporal structure of external sensory events during perception, in particular between vision and audition. *Is perception of synchrony mainly driven by external factors or rather it can be determined by internal factors, such as the distribution of the observer's attention, action plans or brain states via rhythms in neural activity?* To answer this question, we have conducted two purely behavioral studies and an EEG / behavioral one. The behavioral studies consisted of a temporal recalibration paradigm, differing in the instructions given to the participants during the adaptation phase. In the EEG study we examined the pre-stimulus neural correlates of synchrony (vs. asynchrony) perception on a SJ task.

In the first study (section 3.2) we investigated the role of attention in a temporal recalibration process, whereby synchrony perception shifts after adaptation. Based on previous studies, attention can help resolve temporal coincidence of audio-visual events in temporally cluttered contexts. We created a context with various competing events, which rendered various possible audio-visual asynchronies close in time. This context was accomplished by presenting triplets consisting of flash-tone-flash sequences during adaptation. Participants were instructed to direct attention to one of the flashes (flash leading a tone or flash lagging a tone, in separate sessions) to measure PSS shifts as a function of the attended asynchrony direction, despite the sequence of stimulus events was unaltered between conditions.

In the second study (section 3.4) we investigated the role of action on a temporal recalibration process, on audio-visual pairs. Previous studies have demonstrated that systematic relations between actions and sensory

events can successfully re-organize the temporal structure of events. We examined whether action structure during adaptation can modulate the magnitude of recalibration between audio-visual pairs. To this aim, we asked participants to synchronize actions with one or two stimuli of the sensory pair during adaptation, putatively promoting grouping or segregation of the stimuli pair. We hypothesized that actions promoting grouping of audio-visual events will lead to larger recalibration magnitudes compared to actions promoting segregation.

In the third study (section 3.5) we examined whether neural oscillations prior to stimuli presentation can predict perceptual variability in SJ tasks. We recorded EEG activity and analyzed the phase and power of neural oscillations as a function of the perceptual outcome (simultaneous vs. non-simultaneous percept) for both modality orders (vision leading and audition leading events). Further behavioral and EEG analysis explored the influence of the modality order of the previous trial on the next trial perception of synchrony.

This section summarizes the results of these three studies in a table (section 4.1) and discusses the findings in the context of the current literature on temporal recalibration (section 4.2.1) and temporal organization of perception (section 4.2.2). Section 4.2.1 introduces the findings of the first study by discussing the impact of attention on temporal recalibration studies (section 4.2.1.1). At last, I summarize the findings of the second study and discuss the impact of actions, attention and the predictive processes on temporal recalibration (section 4.2.1.2). Section 4.2.2 summarizes and discusses the findings of the third study based on the context of the current views of how time is encoded in the brain. Section 4.3 examines the indirect implications of our results in the frame of multisensory perception (section 4.3.1). There, I discuss the importance of the temporal coincidence of cross-modal events in

multisensory perception and the implications derived from our hypothesis. The second part of section 4.3 contextualizes our results within the current mechanisms proposed to account for temporal recalibration (section 4.3.2). The third part of the section 4.3 provides a general overview of the impact of cognitive factors on temporal organization of perception (section 4.3.3). The last part of section 4.3 introduces potential frameworks for the role of neural oscillations in the temporal structuring of conscious perception (section 4.3.4). Finally, in section 4.4 I summarize the main conclusions of this thesis and suggest future lines of research.

4.1. SUMMARY OF THE RESULTS (TABLE)

STUDY 1								
Type	Paradigm	Participants	Scope of the study	Conditions	PSS shift	SD	Oddball detection	Conclusions
Behavioral	Temporal recalibration	Experiment 1=18	Determine whether selective attention can determine which particular audio-visual events would be recalibrated in time in a temporal recalibration paradigm, when competing asynchronies are present.	Pretest	-Selective attention to particular stimuli during the adaptation can induce a PSS shift (attend leading vs. attend lagging flash). -Recalibration asymmetry towards vision leading asynchronies when comparing both attend conditions with a pretest baseline.	None of the comparisons resulted significant	None of the comparisons resulted significant.	-PSS is shifted in the direction of selective attention, when audition was leading vision. -We speculate that the stronger saliency of the flash leading was capturing the temporal adjustment. -An alternative explanation is that selective attention is more efficiently deployed when audition was
		Experiment		Find the origin of this recalibration				

		2=22	asymmetry towards vision leading asynchronies.	Attend leading flash	flash condition was equivalent to the PSS for the attend-alternate baseline, whereas the PSS in the attend-lagging flash shifted in the direction of attention.	resulted significant	resulted significant.	leading vision.
				Attend lagging flash				
				Attend alternate flash				

STUDY 2								
Type	Paradigm	Participants	Scope of the study	Conditions	Recalibration magnitude	SD	Action synchronization (sync) task	Conclusions
Behavioral	Temporal recalibration	Participants per condition=16	Determine whether actions that induce audio-visual grouping or segregation can determine the strength of the temporal recalibration adjustment.	Grouping	-Recalibration for all the conditions.	None of the terms resulted significant.	-Lower sync performance in the segregation group compared to the grouping and control group. -No differences between grouping and control group.	-Recalibration magnitude difference between the grouping and segregation conditions cannot be attested to the task difficulty. -Motor actions or the allocation of attention to the planning/execution of concurrent actions can modulate temporal recalibration of audio-visual events.
				Segregation	-Larger magnitude of recalibration in the grouping than in the segregation condition.			
				Control	-Recalibration magnitude for the control was similar to the segregation, and much smaller than in the grouping condition.			

STUDY 3

Type	Paradigm	Participants	Scope of the study	Conditions	Phase	Power	Time	Electrode	Conclusions
EEG	Simultaneity Judgment	27	Correlation between phase or power of low frequency ongoing neural oscillations and synchrony (vs. asynchrony) perception	A50V	-Phase of pre-stimulus oscillations at 13 Hz, was predictive of simultaneity perception, with preferential phases for synchronous and asynchronous perception being nearly opposite (close to 180°).	None of the terms resulted significant	200 ms before stimulus presentation	The highest BI at this time-frequency point was found at TP9.	- Phase of ongoing activity at 13 Hz, 200 ms prior to stimulus onset can predict synchrony perception, for audition leading vision pairs. - Prior trial history correlates with both behavior and oscillatory power at the same frequency
				V50A	None of the terms resulted significant	None of the terms resulted significant	None of the terms resulted significant	None of the terms resulted significant	

				Previous AV vs. Previous VA		Power at 14 Hz correlate with the modality order of the preceding trial	During 1 second interval before stimulus presentation	A cluster was found at Fz, FC1, FC2, FC4, C2, C4, C6, T8, CP4, CP6, Pz, P2, P4, PO4, POz, PO3, O1, Oz.	range (14 Hz).
--	--	--	--	-----------------------------	--	-------------------------------------------------------------------------	-------------------------------------------------------	--------------------------------------------------------------------------------------------------------	----------------

4.2. IMPLICATION OF THE PRESENT FINDINGS

4.2.1. Temporal recalibration

In the studies presented in section 3.2 and section 3.4, we addressed the hypotheses that general temporal anchors, such as attention and action, might contribute to resolve the re-alignment between sensory modalities in a temporal recalibration paradigm with multiple possible cross-modal pairings. The results obtained highlighted attention during adaptation as an important factor modulating the direction of the PSS (Study 1) or the magnitude of recalibration (Study 2).

In the next section I discuss the particular role of attention on the recalibration procedure found in Study 1 and 2 and discuss whether attention actually helps the resolution of audio-visual temporal recalibration processes or might only contribute by modulating the perception of time.

4.2.1.1 The impact of attention

As pointed out in the introduction (section 1.2), synchrony perception is not a simple reflection of the physical processing times between the neural signals, but an interpretation of the external events. This can be manifested by the changes in synchrony perception of identical cross-modal pairs after adaptation to two opposite modality orders. In the first study (section 3.2) we showed that the process of selecting information from the environment and adjusting it in time, not only results from the physical organization in the environment, but also from the focus of attention.

Attention can modulate perception of simultaneity between cross-modal pairs by the well known prior entry effect (Spence et al., 2001). Prior entry is characterized by an earlier perception of stimuli in the attended

modality compared to stimuli from unattended modalities. Fujisaki et al., (2004) were the first who examined the effect of prior entry on a temporal recalibration paradigm, on flash-tone pairs. During the test trials, the authors instructed participants to first, report the time of appearance of one of the two sensory modalities (audition), and second, judge the simultaneity of the flash-tone pairs. Their results showed that when participants attended to one sensory modality, this particular modality was perceived faster. Therefore, to report simultaneity of audio-visual events, the attended modality had to be presented slightly after than when it was not attended.

Unlike prior entry effect studies, in section 3.2 we instructed participants to attend to either the first or the second flash (in separate sessions) during adaptation. Thus, in both attention conditions participants attended to the same sensory modality (vision) and only instructions to judge the simultaneity of flash-tone pairs were given during the test trials. Hence, prior entry might have speeded up the perceived time in both attended conditions (flash leading and flash lagging) and could not explain the PSS differences found in section 3.2. We tested whether selective attention under conditions of multiple possible pairings in time (sequence of flash-tone-flash triplets) can shift the PSS in the direction of the attended asynchrony direction (flash leading or flash lagging). The results of the first experiment showed that despite the physical stimulation was identical in both attention conditions, the PSS shifted in the direction of the attended flash. Interestingly, the sign of the PSS's after perceptual adaptation was always positive in both attention conditions, that is, vision has to be presented always preceding audition in order to perceive the events as simultaneous, compared to the Pretest measurement (SJ test previous to any adaptation).

To further explore this recalibration outcome, we ran a second experiment to control for possible stimulus-driven factors present in the adaptation phase which might be independent of the attentional effect. Participants were instructed to alternatively attend to the flash leading and lagging during adaptation in order to cancel away the effect of attention. The resulting PSS was similar to the PSS when participants attended to the flash leading, suggesting the first flash might have been capturing the recalibration adjustment in absence of any specific effect of endogenous attention. Therefore, only when participants attended to the second flash (flash lagging respect to the tone), the positive PSS shift was reduced towards a more audition leading vision asynchrony direction.

This asymmetry in recalibration was discussed in section 3.2, along with two possible interpretations. The preferred proposal, mentioned in the previous paragraph, was the possible stronger impact of the first flash on temporal recalibration due to having more salient properties than the second flash. An alternative interpretation was that the effectiveness of attention deployment over the first and second flash could have been different. This would have explained why the effect of attention only appeared in the flash lagging condition.

Overall, we concluded that the temporal structure of events determined by the physical stimulation (stronger saliency of the first flash) drove the temporal recalibration. However, when attention was directed away from the recalibrated stimulus (towards the second flash) the PSS was more similar to the PSS obtained in the Pretest baseline, decreasing the temporal adjustment. This interpretation is in line with Heron et al., (2010) findings. Heron et al., (2010) considered situations in which attention was directed away from the stimuli (to the fixation cross or the inter-trial intervals). Even in these cases recalibration was observed,

indicating a strong stimulus-driven effect that could be in turn, modulated by the focus of attention.

In a follow up experiment (section 3.3) we wanted to determine the contribution of attention when physical organization did not enforce a structure on the order of events (equal salience). In view of the null findings we speculated that the role of attention is limited, being not sufficient to resolve the temporal structure of events. This hypothesis is in agreement with the limited capacity to determine temporal relations of cross-modal information (Fujisaki and Nishida, 2010; Kösem and van Wassenhove, 2012).

An interesting future research line would be to explore at which state endogenous attention can influence the temporal organization of perception; whether it concerns each sensory modality separately, concurrent with the perceptual latency framework, or rather influences a more centralized supra-modal organization of time.

4.2.1.2 The impact of actions, attention and the predictive system

Thanks to the feedback obtained by the active interaction with the environment it is possible to correct the interpretation about the timings of sensory events with respect to behavior. Hence, sensory-motor perception has an important advantage with respect to sensory perception; the possibility for rapid correction of the time of the signals. This feedback from actions, for example, allows a tennis player to accurately predict the time at which a ball has to be hit, by invoking internal forward models helping to anticipate the time of events (Driver and Frith, 2000). According to Schubotz (2007), the motor system is critical for perception, as it helps generating temporal predictions.

In Study 2 (section 3.4) we speculated that the predictive process of the motor system could be similar in terms of the dynamic organization of the

temporal recalibration process, as both might involve internally generated states of outcome predictions. Specifically, we hypothesized that actions could serve as a temporal anchor, helping resolve the most appropriate alignment between sensory modalities. Although this is not a new idea (Parsons et al., 2013), we used a novel approach to test it. We employed a temporal recalibration paradigm and during the adaptation phase we instructed participants to synchronize voluntary actions with (some of the) events containing audio-visual asynchronies. The motor synchronization task was intended to promote systematic relations between actions and sensory events. Synchronization of actions to sensory events is characterized by an anticipation of the actions, as if the sensory events were taken as a consequence of the action (Yarrow and Obhi, n.d.). In particular, we employed actions putatively promoting grouping or segregation of the audio-visual pairs.

Our results showed that the magnitude of recalibration shifted according to the action type (involving grouping vs. segregation of audio-visual signals) performed during adaptation. However, a further control for action execution suggested that planning or execution of actions at the time of stimuli presentation could have modulated the magnitude of recalibration between audio-visual pairs. After consideration of several alternatives (see discussion in section 3.4.6), we took our results as to indicate that attention may have acted in this occasion as a mediator for the impact of the motor system demands. To understand the role of attention in Study 2 (section 3.4), it is necessary to examine the interplay between attention and the motor predictive system.

Morillon et al., (2014) suggested that rhythmic actions engage the focus of attention, enhancing the temporal selection process of sensory information (audition). In particular, they found that motor activity facilitates perception of relevant information by filtering the relevant from the

distracting information. Furthermore, they reported that the impact of the motor actions on perception depended on the alignment between motor and attention rhythms.

These results are in line with the interpretation of the findings in Study 2 (section 3.4), where we proposed that actions could generate attention episodes helping to segregate or group events in time. Moreover, this explanation may be able to encompass the alternative proposed in the discussion of Study 2. There, we suggested that in the conditions with more actions (in the segregation and control conditions), attention would have been deployed away of the sensory events in order to execute the actions. Deploying attention away would have had an impact decreasing the magnitude of temporal recalibration.

4.2.2. Temporal organization of perception: environmental rhythms vs. brain rhythms

Considerable controversy remains concerning how the temporal organization of perception is encoded, and represented in the brain. One proposal is that time perception is driven by neural entrainment to external rhythms, helping the calibration of the brain time to the temporal structure of the environment. A second proposal is that the intrinsic structure of brain rhythms (neural oscillations) can encompass the representation of time (Herbst and Landau, 2016; van Wassenhove, 2016).

The phase entrainment of oscillations at 1 Hz during rhythmic presentation of events has been shown to correlate with changes in perceptual organization of audio-visual events in time (Kösemet al., 2014). However, in Kösemet al.,'s (2014) study, the frequency of the external rhythmic stimulation (1 Hz), coincided with the 1 Hz neural activity, which has been found to correlate with the synchrony perception. For this reason, as pointed out by the authors of the study, the specific role

of the 1 Hz frequency on the temporal organization of perception remains unclear: whether this neural oscillation is a specific frequency relevant to encode time in the brain or rather, it results by the entrainment to the 1 Hz rhythmic event presentation. This fact hinders the disentanglement between the two mechanisms mentioned above, that are proposed to encode temporal organization (entrainment vs. endogenous rhythms).

An empirical question resulting from this study is whether the phase of the 1 Hz oscillation would still correlate with the changes in temporal order perception when exposing participants to a different rhythmic presentation of events or as we did in Study 3 (section 3.5), to non-rhythmic event presentation.

Specifically, in Study 3 (section 3.5), we examined whether neural oscillations prior to the stimulus, could predict synchrony perception on audio-visual information in the absence of rhythmic information. We used a SJ task and instead of exposing participants to rhythmic presentation of the stimuli, the intervals between our stimuli presentation were jittered, to avoid any external source of temporal structure. We determined in each participant the audio-visual asynchronies for which perceptual outcome of synchrony perception was maximally uncertain, so that they were perceived half of the times as simultaneous and the other half as non-simultaneous, for both modality orders (audition leading vision and vision leading audition). The phase and power of slow neural oscillations prior to stimulus presentation were examined for these particular asynchronies. The method employed was to record EEG activity while participants ran a SJ task on these and other asynchronies.

Our results showed that the phase at 13 ± 2 Hz, 200 ms before event presentation could predict synchrony perception of audio-visual events, though this resulted reliable only when audition was leading vision.

Surprisingly, in the opposite modality order the effect was not observed. Yet, similar asymmetries on audio-visual modalities have been reported by other EEG studies (Kösemet et al., 2014; Kaganovich and Schumaker, 2016). We proposed that vision might be more variable tracking the time of events compared to audition (measure reflected by the larger length of the confidence intervals needed to estimate the vision leading boundary), and this could have blurred possible phase effects. The other possibility we discussed was that the temporal perception might be differently organized as a function of the modality leading the audio-visual pair. This account is supported by behavioral studies showing asymmetric improvements after perceptual training as a function of the leading sensory modality (Cecere et al., 2016). Further investigations need to be done to examine the nature of this interesting asymmetry between vision leading and audition leading events.

In this same Study, we also explored audio-visual synchrony perception as a function of the previous trial modality order. As predicted by the rapid temporal recalibration effect (Van der Burg et al., 2013), we found that perception of simultaneity increased in a given trial when the preceding trial contained the same modality order as the current trial. This rapid temporal recalibration effect was found for both modality orders. In a further analysis within Study 3 (section 3.5), we investigated the neural correlates behind this rapid temporal recalibration. We examined the amplitude of neural oscillations before stimulus presentation as a function of the modality order presented in the preceding trial. We found that the power at 14 ± 2 Hz during the 1 second interval before stimulus presentation correlates with the previous trial modality order. Thus, we found correlates of the previous history trial on the ongoing activity of the subsequent trial.

Critically, the findings that ongoing oscillations can carry over the prior history of the external context makes us wonder about the nature of the ongoing activity, often so-called “spontaneous activity” or “noise”. Along with other authors (Fuxe et al., 1998) we speculated that ongoing activity might be conveying information related to expectancy or other cognitive factors, known to modulate perception. Additional studies are required to further correlate the amplitude of 14 ± 2 Hz oscillations, the previous modality order and the perceptual outcome in cross-modal synchrony judgments.

Overall, our findings support previous proposals suggesting a specific role of the phase of slow ongoing neural oscillations on temporal order perception (Varela et al., 1981; Vanrullen et al., 2011; Milton and Pleydell-Pearce, 2016). In addition, the results indicating that ongoing activity might convey information from prior history trials, and thus, from the environment, illustrate an interesting example of how ongoing activity could be used to predict information that might help humans adapt and interact with the world. I speculate that both mechanisms (entrainment and endogenous rhythms) might be at play during the resolution of the temporal organization of cross-modal events in perception, as they might contain complementary information about time.

4.3. INDIRECT IMPLICATIONS OF THE RESULTS

4.3.1. Multisensory perception

Classically, it has been argued that humans might recover the exact temporal order of events in multisensory perception (Kopinska and Harris, 2004). However, some authors such as Holcombe (2013) suggest that simple grouping and segregation processes might be more important than the exact temporal order between sensory signals to organize perception.

This approach is consistent with our first and second study, where we concluded that attention and action could serve as temporal anchors to resolve the alignment between cross-modal events. This basic idea has important theoretical implications on the study of time perception.

First, the idea of general temporal anchors organizing temporal order, demands further exploration of the tasks used to measure temporal organization of perception. While TOJ calls for the segregation of signals in order to accomplish the task, SJ tasks induce the grouping of signals. In all our experiments we have employed SJ tasks and thus, one straightforward question is whether TOJ and SJ are measuring the same perceptual process. The results of several studies indicate that resolving “non-simultaneous or successiveness” is a necessary, but not sufficient condition for the retrieval of the temporal order of two events (for example see Miyazaki et al., 2016). In agreement with this observation, nowadays, more and more studies emphasize that there might be distinct cognitive and neural processes behind TOJ and SJ tasks (van Eijk et al., 2008; Linares and Holcombe, 2014; Miyazaki et al., 2016), although classically the results obtained from these two tasks have been indistinctly compared among researchers.

Second, we believe that the different sources of temporal uncertainty discussed in the introduction (section 1.2) favor the implication of flexible processes in order to determine the temporal organization of events. Thus, temporal order in multisensory contexts cannot only depend on the exact event time but also on experience-based knowledge about the environment and participants’ goals amongst others. Indeed, a good example of this experience-dependent process is given by the malleability of the window of simultaneity (Powers et al., 2009; Stevenson and Wallace, 2013). The width of the simultaneity window can be narrowed (hence gaining in precision) after training (Powers et al., 2009) and,

change (narrow or wide) depending on stimulus properties and task demands amongst others (Stevenson and Wallace, 2013).

Overall, these observations, far from attesting a direct correspondence between the objective (physical) temporal order of signals and perceptual organization of time, acknowledge the paramount role of experience and goal-based processes. That is, grouping and segregation processes acquired through the previous experience or shaped by the current goals might have helped dealing with the temporal uncertainty present in multisensory environments.

A directly related question concerns to how the idea of temporal anchors organizing the time of events relates to the current mechanisms proposed to underlie the temporal recalibration process. This question is addressed in the next section (section 4.3.2) and related to the findings of Study 1 (section 4.2.1.1) and Study 2 (section 4.2.1.2).

4.3.2. The underlying mechanisms of temporal recalibration

Heron et al., (2010) investigated the role of attention in temporal recalibration, providing in my opinion, compelling evidence of how the grouping and segregation processes can act as temporal anchors. Heron et al., (2010) showed that attending to the temporal relationship between the signals during adaptation in a temporal recalibration paradigm increased the magnitude of recalibration, compared to when attention was directed toward the two stimuli separately, or away from the pair altogether.

Attending to the temporal relationship might have involved grouping the signals as being part of a single event in time. Attention thus, might have been acting as a temporal anchor between signals.

Despite it is unclear how the temporal anchor account is related to the proposed mechanisms underlying temporal recalibration (section 1.4.1);

the results obtained in our work shed some light on the cognitive factors that can modulate the re-alignment between signals. On the one hand, the temporal anchor account seems more consistent with the framework of a supra-modal temporal organization process (Hanson et al., 2008) because it can predict the re-alignment of both sensory modalities, not accounted for the perceptual latency framework (Harrar and Harris, 2008; Di Luca et al., 2009; Navarra et al., 2009).

On the other hand, the experience of event time based on temporal anchors rather than the exact order of events might involve also the co-existence of multiple representations of time in the brain. These multiple representations appear to be incompatible with the existence of a single centralized-comparator or clock, at the essence of the supra-modal framework. However, I speculate that the singularity of the centralized clock might not be critical to access the supra-modal framework. Indeed, it is unclear how other authors' models (which are supposed to favor the supra-modal framework), such as the ones based on changes at a decisional level (Yarrow et al., 2011a) and the ones based on multimodal neurons (Roach et al., 2011) tuned to specific delays can encompass the existence of multiple re-alignments during adaptation (Roseboom and Arnold, 2011; Heron et al., 2012; Yuan et al., 2012, 2015; Roseboom et al., 2013). Hence, we argue for a more flexible definition at this particular point of the supra-modal framework, allowing the existence of multiple factors regulating one or multiple comparators.

4.3.3. The impact of cognitive factors

The results of this dissertation highlight the influence of cognitive factors on multisensory perception. By studying different perceptual outcomes on identical audio-visual pairs, we were able to show that a variety of cognitive factors, such as the focus of attention, action planning and even

the intrinsic rhythms of the brain can shape synchrony perception. In this sense, perception would not only convey information from the stimuli themselves, but, rather would involve a wide variety of signals conveying information related to the current state of the individual.

Another complicating factor, being mentioned in several of our studies was task difficulty. In the first and second study, we were concerned about task difficulty and tried to control for similar levels of task difficulty. This was important because different levels of task difficulty could have promoted different levels of mental effort or cognitive load between the conditions of interest. To this aim, the percentage of oddball (deviant) detection and percentage of action synchronization measures were employed in the first and second study, respectively. The extent to which these measures of task difficulty can capture differences in cognitive load is a matter of debate and falls beyond the scope of the present thesis. Such differences could have influenced the efficiency in deploying attention in our adaptation conditions. Further studies are thus needed to assess how differences in cognitive load may pose a limitation to the interpretation of our results. Investigation with more sensitive measures of cognitive and/or perceptual load is needed to confirm that it is the focus of attention the factor modulating the direction and magnitude of the temporal adjustment.

4.3.4. Conscious perception and the processing in time through neural oscillations

In a contemporary dance performance, the audience can perfectly follow the dancers' trajectories as a continuous stream of movements, at each single moment in time. However, this does not directly imply sensory information should be processed in a continuous manner by the brain. In fact, there are frameworks suggesting that neural oscillations can help structuring information in time in a visual scene by prioritizing processing

by inhibition of neuronal firing (Jensen et al., 2014) or by confining events to discrete moments such as in a film (Varela et al., 1981; VanRullen and Koch, 2003), despite conscious perception appears to be continuous.

For example, Jensen et al., (2014) addressed how information can be correctly combined when several visual events are presented simultaneously. That is, they studied which neural processes could be used to transform a spatial representation into a temporal code. They proposed that the phase of neural oscillations (specifically, inhibitory pulses of alpha oscillation, 10 Hz) might determine this sequential processing of information in vision. In particular, alpha activity might determine the order of neuronal firing according to the level of excitability of each individual representation. Interestingly, the authors of the study propose that the level of excitability could be determined by stimulus-driven and top down factors, such as spatial attention.

Other examples of frameworks come from experimental studies showing a correlation between the intrinsic temporal structure of neural oscillations and the temporal organization of perception (Varela et al., 1981; VanRullen and Koch, 2003; Herzog et al., 2016). For example, Varela et al., (1981) showed that two discrete flashes presented in rapid succession were perceived as simultaneous or sequential depending on the phase of the alpha oscillation (~10 Hz) at the moment of stimuli presentation. Thus, the phase might determine these discrete snapshots, such that all the stimuli that fall inside a certain snapshot would be perceived as occurring at a single moment in time.

In multisensory contexts, the cycle of neural oscillations (~100 ms) has been suggested to determine the size of the temporal window in which some cross-modal illusions are perceived (Cecere et al., 2015).

Congruent with these evidences, in Study 3 (section 3.5) we found that the size of the temporal window was similar to the cycle of the 13 Hz frequency, which phase predicted simultaneity perception (vs. non-simultaneity) in audio-visual events. Further investigation is needed to confirm if this relation is also present at the individual level. However, I speculate this relation might not be straightforward in our study because the phase was only predicting the perceptual outcome in synchrony judgments in one of the asynchrony directions; when audition was leading vision. If this is the case, the results would highlight the importance to consider separate neural processes to determine synchrony perception depending on the leading modality (e.g. Cecere et al., 2016).

4.4. CONCLUSIONS AND FUTURE LINES OF RESEARCH

In this dissertation I have reported a series of experiments illustrating the involvement of cognitive factors/participants' inner state during the temporal organization of cross-modal events. In particular, our results highlight the role of attention on temporal recalibration, modulating both, the direction and the magnitude of the temporal re-alignment between sensory modalities, under otherwise physically identical circumstances. In addition, we found that the phase of slow ongoing neural oscillations prior to the sensory events can play a role in determining synchrony perception across sensory modalities. Yet, the evidence here reported is not sufficient to reveal the nature of the relationship between external inputs and participants' inner state resolving synchrony across senses. The main conclusions of this thesis, directly related to the hypotheses listed in section 2.1 are the following:

- (1) Selective endogenous attention can modulate the adjustment of the audio-visual temporal recalibration when competing

asynchronies are presented during adaptation. This was conclusively shown only for audition leading vision asynchronies. Future lines of research are needed to evaluate whether this asymmetry can be attributed to the stronger saliency of vision leading asynchronies in our paradigm or to a more efficient deployment of attention when audition was leading vision than in the opposite modality order.

- (2) The preparation/execution of actions can modulate the adjustment of the audio-visual temporal recalibration, possibly via attention. Further experiments are required to confirm it is the nature of actions, or else this is another example of an attention mediated effect. Either actions, or actions via attention, would have been used as a temporal anchor to group or segregate audio-visual events in time.
- (3) The phase of neural oscillations at around 13 Hz, before the stimulus presentation, can predict trial by trial variability on the judgments about audio-visual synchrony. We found preferential phases for synchronous and asynchronous perception being nearly opposite (close to 180°). The amplitude at around the same frequencies (14 Hz) during the inter-trial period seemed to be influenced by immediate prior experience (regarding modality order), which in turn modulates the perception of synchrony. New EEG experiments are necessary to relate the power at 14 Hz and the prior modality order context with perceived simultaneity.

In conclusion, the process of multisensory perception is strongly shaped by our inner states, determined by goals, focus of attention, actions and previous experience. A larger range of flexibility, possibly integrating all

sorts of sensory and motor operations might be needed to re-define multisensory perception as a multidimensional learning process.

BIBLIOGRAPHY

- Boenke LT, Deliano M, Ohl FW (2009) Stimulus duration influences perceived simultaneity in audiovisual temporal-order judgment. *Exp brain Res* 198:233–244 Available at: <http://www.ncbi.nlm.nih.gov/pubmed/19590862> [Accessed April 2, 2016].
- Buzsáki G, Draguhn A (2004) Neuronal oscillations in cortical networks. *Science* 304:1926–1929.
- Cecere R, Gross J, Thut G (2016) Behavioural evidence for separate mechanisms of audiovisual temporal binding as a function of leading sensory modality. *Eur J Neurosci* Available at: <http://www.ncbi.nlm.nih.gov/pubmed/27003546> [Accessed April 4, 2016].
- Cecere R, Rees G, Romei V (2015) Individual differences in alpha frequency drive crossmodal illusory perception. *Curr Biol* 25:231–235 Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=4300399&tool=pmcentrez&rendertype=abstract> [Accessed April 12, 2016].
- Conrey B, Pisoni DB (2006) Auditory-visual speech perception and synchrony detection for speech and nonspeech signals. *J Acoust Soc Am* 119:4065–4073 Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3314884&tool=pmcentrez&rendertype=abstract> [Accessed May 12, 2016].
- Di Luca M, Machulla T-K, Ernst MO (2009) Recalibration of multisensory simultaneity: cross-modal transfer coincides with a change in perceptual latency. *J Vis* 9:7.1–16 Available at: <http://www.ncbi.nlm.nih.gov/pubmed/20053098> [Accessed June 9, 2013].
- Driver J, Frith C (2000) Shifting baselines in attention research. *Nat Rev Neurosci* 1:147–148 Available at: <http://www.ncbi.nlm.nih.gov/pubmed/11252778> [Accessed August 18, 2016].
- Dugue L, Marque P, VanRullen R (2011) The Phase of Ongoing Oscillations Mediates the Causal Relation between Brain Excitation and Visual Perception. *J Neurosci* 31:11889–11893 Available at: <http://www.jneurosci.org/cgi/doi/10.1523/JNEUROSCI.1161-11.2011>.

- Ernst MO, Banks MS (2002) Humans integrate visual and haptic information in a statistically optimal fashion. *Nature* 415:429–433.
- Fister JK, Stevenson RA, Nidiffer AR, Barnett ZP, Wallace MT (2016) Stimulus intensity modulates multisensory temporal processing. *Neuropsychologia*:1–9 Available at: <http://linkinghub.elsevier.com/retrieve/pii/S0028393216300513>.
- Foxe JJ, Simpson G V, Ahlfors SP (1998) Parieto-occipital approximately 10 Hz activity reflects anticipatory state of visual attention mechanisms. *Neuroreport* 9:3929–3933.
- Fujisaki W, Koene A, Arnold D, Johnston A, Nishida S (2006) Visual search for a target changing in synchrony with an auditory signal. *Proc Biol Sci* 273:865–874 Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1560215&tool=pmcentrez&rendertype=abstract> [Accessed March 1, 2012].
- Fujisaki W, Nishida S (2005) Temporal frequency characteristics of synchrony-asynchrony discrimination of audio-visual signals. *Exp Brain Res* 166:455–464 Available at: <http://www.ncbi.nlm.nih.gov/pubmed/16032402> [Accessed June 3, 2012].
- Fujisaki W, Nishida S (2007) Feature-based processing of audio-visual synchrony perception revealed by random pulse trains. *Vision Res* 47:1075–1093 Available at: <http://linkinghub.elsevier.com/retrieve/pii/S0042698907000144> [Accessed May 5, 2012].
- Fujisaki W, Nishida S (2010) A common perceptual temporal limit of binding synchronous inputs across different sensory attributes and modalities. *Proc Biol Sci* 277:2281–2290 Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2894908&tool=pmcentrez&rendertype=abstract> [Accessed March 14, 2012].
- Fujisaki W, Shimojo S, Kashino M, Nishida S (2004) Recalibration of audiovisual simultaneity. *Nat Neurosci* 7:773–778 Available at: <http://www.ncbi.nlm.nih.gov/pubmed/15195098> [Accessed March 10, 2012].
- Hanson JVM, Heron J, Whitaker D (2008) Recalibration of perceived time across sensory modalities. *Exp Brain Res* 185:347–352 Available at: <http://www.springerlink.com/content/t416551035021867/> [Accessed December 28, 2011].

- Harrar V, Harris LR (2008) The effect of exposure to asynchronous audio, visual, and tactile stimulus combinations on the perception of simultaneity. *Exp Brain Res* 186:517–524 Available at: <http://www.ncbi.nlm.nih.gov/pubmed/18183377> [Accessed March 19, 2012].
- Herbst SK, Landau AN (2016) Rhythms for cognition: the case of temporal processing. *Curr Opin Behav Sci* 8:85–93 Available at: <http://linkinghub.elsevier.com/retrieve/pii/S2352154616300213>.
- Heron J, Hotchkiss J, Aaen-Stockdale C, Roach NW, Whitaker D (2013) A neural hierarchy for illusions of time: duration adaptation precedes multisensory integration. *J Vis* 13 Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3852255&tool=pmcentrez&rendertype=abstract> [Accessed April 4, 2016].
- Heron J, Roach NW, Hanson JVM, McGraw P V, Whitaker D (2012) Audiovisual time perception is spatially specific. *Exp Brain Res* 218:477–485 Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3324684&tool=pmcentrez&rendertype=abstract> [Accessed July 26, 2012].
- Heron J, Roach NW, Whitaker D, Hanson JVM (2010) Attention regulates the plasticity of multisensory timing. *Eur J Neurosci* 31:1755–1762 Available at: <http://www.ncbi.nlm.nih.gov/pubmed/20584179> [Accessed August 25, 2011].
- Herzog MH et al. (2016) Time Slices: What Is the Duration of a Percept? *PLOS Biol* 14:e1002433 Available at: <http://dx.plos.org/10.1371/journal.pbio.1002433> [Accessed July 2, 2016].
- Holcombe A (2013) The Temporal organization of perception. :1–20.
- Jensen O, Gips B, Bergmann TO, Bonnefond M (2014) Temporal coding organized by coupled alpha and gamma oscillations prioritize visual processing. *Trends Neurosci* 37:357–369 Available at: <http://dx.doi.org/10.1016/j.tins.2014.04.001>.
- Kaganovich N, Schumaker J (2016) Electrophysiological correlates of individual differences in perception of audiovisual temporal asynchrony. *Neuropsychologia* 86:119–130 Available at: <http://www.sciencedirect.com/science/article/pii/S0028393216301233> [Accessed May 8, 2016].

- Kambe J, Kakimoto Y, Araki O (2015) Phase reset affects auditory-visual simultaneity judgment. *Cogn Neurodyn* 9:487–493 Available at: <http://www.ncbi.nlm.nih.gov/pubmed/26379799> [Accessed April 28, 2016].
- King AJ (2005) Multisensory integration: strategies for synchronization. *Curr Biol* 15:R339–R341 Available at: <http://www.ncbi.nlm.nih.gov/pubmed/15886092> [Accessed August 14, 2012].
- King AJ, Palmer AR (1985) Integration of visual and auditory information in bimodal neurones in the guinea-pig superior colliculus. *Exp Brain Res* 60:492–500 Available at: <http://www.ncbi.nlm.nih.gov/pubmed/4076371> [Accessed July 2, 2013].
- Kohlrausch A, van Eijk R, Juola JF, Brandt I, van de Par S (2013) Apparent causality affects perceived simultaneity. *Atten Percept Psychophys* 75:1366–1373 Available at: <http://www.ncbi.nlm.nih.gov/pubmed/23943500> [Accessed June 25, 2016].
- Kopinska A, Harris LR (2004) Simultaneity constancy. *Perception* 33:1049–1060 Available at: <http://www.ncbi.nlm.nih.gov/pubmed/15560507> [Accessed June 16, 2016].
- Kösem A, Gramfort A, van Wassenhove V (2014) Encoding of event timing in the phase of neural oscillations. *Neuroimage* 92:274–284 Available at: <http://www.sciencedirect.com/science/article/pii/S1053811914001013> [Accessed November 20, 2014].
- Kösem A, van Wassenhove V (2012) Temporal structure in audiovisual sensory selection. *PLoS One* 7:e40936 Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3400621&tool=pmcentrez&rendertype=abstract> [Accessed July 2, 2013].
- Linares D, Holcombe AO (2014) Differences in perceptual latency estimated from judgments of temporal order, simultaneity and duration are inconsistent. *Iperception* 5:559–571.
- Macaluso E, Driver J (2005) Multisensory spatial interactions: A window onto functional integration in the human brain. *Trends Neurosci* 28:264–271.

- Machulla T-K, Di Luca M, Froehlich E, Ernst MO (2012) Multisensory simultaneity recalibration: storage of the aftereffect in the absence of counterevidence. *Exp Brain Res* 217:89–97 Available at: <http://www.ncbi.nlm.nih.gov/pubmed/22207361> [Accessed April 5, 2012].
- MCGURK H, MACDONALD J (1976) Hearing lips and seeing voices. *Nature* 264:746–748 Available at: <http://www.nature.com/doi/10.1038/264746a0> [Accessed August 15, 2016].
- Meredith MA, Nemitz JW, Stein BE (1987) Determinants of multisensory integration in superior colliculus neurons. I. Temporal factors. *J Neurosci* 7:3215–3229 Available at: <http://www.ncbi.nlm.nih.gov/pubmed/3668625> [Accessed July 2, 2013].
- Milton A, Pleydell-Pearce CW (2016) The phase of pre-stimulus alpha oscillations influences the visual perception of stimulus timing. *Neuroimage* 133:53–61 Available at: <http://www.sciencedirect.com/science/article/pii/S1053811916001749> [Accessed March 5, 2016].
- Miyazaki M, Kadota H, Matsuzaki KS, Takeuchi S, Sekiguchi H, Aoyama T, Kochiyama T (2016) Dissociating the neural correlates of tactile temporal order and simultaneity judgements. *Sci Rep* 6:23323 Available at: <http://www.nature.com/articles/srep23323>.
- Miyazaki M, Yamamoto S, Uchida S, Kitazawa S (2006) Bayesian calibration of simultaneity in tactile temporal order judgment. *Nat Neurosci* 9:875–877 Available at: <http://www.ncbi.nlm.nih.gov/pubmed/16732276> [Accessed March 16, 2012].
- Morein-Zamir S, Soto-Faraco S, Kingstone A (2003) Auditory capture of vision: examining temporal ventriloquism. *Brain Res Cogn Brain Res* 17:154–163 Available at: <http://www.ncbi.nlm.nih.gov/pubmed/12763201>.
- Morillon B, Schroeder CE, Wyart V (2014) Motor contributions to the temporal precision of auditory attention. *Nat Commun* 5:5255 Available at: <http://www.nature.com/doi/10.1038/ncomms6255> [Accessed June 8, 2016].
- Navarra J, Fernández-Prieto I, Garcia-Morera J (2013) Realigning thunder

- and lightning: temporal adaptation to spatiotemporally distant events. *PLoS One* 8:e84278 Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3877269&tool=pmcentrez&rendertype=abstract> [Accessed February 1, 2014].
- Navarra J, García-Morera J, Spence C (2012) Temporal adaptation to audiovisual asynchrony generalizes across different sound frequencies. *Front Psychol* 3:152 Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3351678&tool=pmcentrez&rendertype=abstract> [Accessed July 2, 2013].
- Navarra J, Hartcher-O'Brien J, Piazza E, Spence C (2009) Adaptation to audiovisual asynchrony modulates the speeded detection of sound. *Proc Natl Acad Sci U S A* 106:9169–9173 Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2695059&tool=pmcentrez&rendertype=abstract> [Accessed June 9, 2013].
- Navarra J, Vatakis A, Zampini M, Soto-Faraco S, Humphreys W, Spence C (2005) Exposure to asynchronous audiovisual speech extends the temporal window for audiovisual integration. *Brain Res Cogn Brain Res* 25:499–507 Available at: <http://www.ncbi.nlm.nih.gov/pubmed/16137867> [Accessed May 27, 2013].
- Parsons BD, Novich SD, Eagleman DM (2013) Motor-sensory recalibration modulates perceived simultaneity of cross-modal events at different distances. *Front Psychol* 4:46 Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3582016&tool=pmcentrez&rendertype=abstract> [Accessed August 17, 2013].
- Powers AR, Hillock AR, Wallace MT (2009) Perceptual training narrows the temporal window of multisensory binding. *J Neurosci* 29:12265–12274 Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2771316&tool=pmcentrez&rendertype=abstract> [Accessed July 16, 2012].
- Roach NW, Heron J, Whitaker D, McGraw P V (2011) Asynchrony adaptation reveals neural population code for audio-visual timing. *Proc Biol Sci* 278:1314–1322 Available at: <http://www.ncbi.nlm.nih.gov/pubmed/20961905> [Accessed June 26, 2016].
- Roseboom W, Arnold DH (2011) Twice upon a time: multiple concurrent temporal recalibrations of audiovisual speech. *Psychol Sci* 22:872–877 Available at: <http://www.ncbi.nlm.nih.gov/pubmed/21690312>

[Accessed April 25, 2012].

- Roseboom W, Kawabe T, Nishida S (2013) Audio-Visual Temporal Recalibration Can be Constrained by Content Cues Regardless of Spatial Overlap. *Front Psychol* 4:189 Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3633943&tool=pmcentrez&rendertype=abstract> [Accessed September 17, 2013].
- Schroeder CE, Lakatos P (2009) Low-frequency neuronal oscillations as instruments of sensory selection. *Trends Neurosci* 32:9–18.
- Shams L, Kamitani Y, Shimojo S (2002) Visual illusion induced by sound. *Cogn Brain Res* 14:147–152 Available at: <http://www.sciencedirect.com/science/article/pii/S0926641002000691> [Accessed May 10, 2016].
- Spence C, Parise C (2010) Prior-entry: a review. *Conscious Cogn* 19:364–379 Available at: <http://www.sciencedirect.com/science/article/pii/S1053810009001871> [Accessed April 4, 2016].
- Spence C, Shore DI, Klein RM (2001) Multisensory prior entry. *J Exp Psychol Gen* 130:799–832 Available at: <http://www.ncbi.nlm.nih.gov/pubmed/11757881> [Accessed July 18, 2013].
- Stein BE, Meredith MA (1993) *The merging of the senses*. Cognitive neuroscience., Cambridge,.
- Stein BE, Stanford TR, Rowland BA (2014) Development of multisensory integration from the perspective of the individual neuron. *Nat Rev Neurosci* 15:520–535 Available at: <http://dx.doi.org/10.1038/nrn3742> [Accessed October 3, 2015].
- Stevenson RA, Wallace MT (2013) Multisensory temporal integration: task and stimulus dependencies. *Exp brain Res* 227:249–261 Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3711231&tool=pmcentrez&rendertype=abstract> [Accessed April 2, 2016].
- Stevenson RA, Wallace MT, Altieri N (2014) The interaction between stimulus factors and cognitive factors during multisensory integration of audiovisual speech. *5*:1–3.
- Van der Burg E, Alais D, Cass J (2013) Rapid Recalibration to

- Audiovisual Asynchrony. *J Neurosci* 33:14633–14637 Available at: <http://www.jneurosci.org/content/33/37/14633.abstract> [Accessed September 11, 2013].
- Van der Burg E, Alais D, Cass J (2015) Audiovisual temporal recalibration occurs independently at two different time scales. *Sci Rep* 5:14526 Available at: <http://www.nature.com/articles/srep14526>.
- van Eijk RLJ, Kohlrausch A, Juola JF, van de Par S (2008) Audiovisual synchrony and temporal order judgments: effects of experimental method and stimulus type. *Percept Psychophys* 70:955–968 Available at: <http://www.ncbi.nlm.nih.gov/pubmed/18717383> [Accessed July 2, 2013].
- van Wassenhove V (2016) Temporal cognition and neural oscillations. *Curr Opin Behav Sci* 8:1–7 Available at: <http://dx.doi.org/10.1016/j.cobeha.2016.02.012>.
- van Wassenhove V, Grant KW, Poeppel D (2007a) Temporal window of integration in auditory-visual speech perception. *Neuropsychologia* 45:598–607 Available at: <http://www.ncbi.nlm.nih.gov/pubmed/16530232> [Accessed May 23, 2013].
- van Wassenhove V, Grant KW, Poeppel D (2007b) Temporal window of integration in auditory-visual speech perception. *Neuropsychologia* 45:598–607 Available at: <http://www.sciencedirect.com/science/article/pii/S002839320600011X> [Accessed May 18, 2016].
- Vanrullen R, Busch NA, Drewes J, Dubois J (2011) Ongoing EEG Phase as a Trial-by-Trial Predictor of Perceptual and Attentional Variability. *Front Psychol* 2:60 Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3110813&tool=pmcentrez&rendertype=abstract> [Accessed May 5, 2016].
- VanRullen R, Koch C (2003) Is perception discrete or continuous? *Trends Cogn Sci* 7:207–213.
- Varela F, Toro A, John E, Schwartz E (1981) Perceptual framing and cortical alpha rhythm. :675–686.
- Vatakis A, Spence C (2007) Crossmodal binding: Evaluating the “unity assumption” using audiovisual speech stimuli. *Percept Psychophys* 69:744–756 Available at:

- <http://www.springerlink.com/index/10.3758/BF03193776> [Accessed July 30, 2016].
- Vroomen J, de Gelder B (2004) Temporal ventriloquism: sound modulates the flash-lag effect. *J Exp Psychol Hum Percept Perform* 30:513–518 Available at: <http://www.ncbi.nlm.nih.gov/pubmed/15161383> [Accessed March 27, 2012].
- Vroomen J, Keetels M, de Gelder B, Bertelson P (2004) Recalibration of temporal order perception by exposure to audio-visual asynchrony. *Brain Res Cogn Brain Res* 22:32–35 Available at: <http://www.ncbi.nlm.nih.gov/pubmed/15561498> [Accessed June 13, 2013].
- Wallace MT, Stein BE (2007) Early experience determines how the senses will interact. *J Neurophysiol* 97:921–926 Available at: <http://jn.physiology.org/content/97/1/921> [Accessed April 1, 2016].
- White AL, Linares D, Holcombe AO (2008) Visuomotor timing compensates for changes in perceptual latency. *Curr Biol* 18:R951–R953 Available at: <http://www.cell.com/article/S0960982208010993/fulltext> [Accessed April 2, 2016].
- Yamamoto S, Miyazaki M, Iwano T, Kitazawa S (2012) Bayesian calibration of simultaneity in audiovisual temporal order judgments. *PLoS One* 7:e40379 Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3392227&tool=pmcentrez&rendertype=abstract> [Accessed July 17, 2012].
- Yarrow K, Jahn N, Durant S, Arnold DH (2011a) Shifts of criteria or neural timing? The assumptions underlying timing perception studies. *Conscious Cogn* 20:1518–1531 Available at: <http://www.ncbi.nlm.nih.gov/pubmed/21807537> [Accessed April 23, 2012].
- Yarrow K, Obhi SS (n.d.) Temporal perception in the context of action. Available at: file:///C:/Users/user/Downloads/Temporal_perception_in_the_context_of_action_refs.pdf [Accessed March 26, 2015].
- Yarrow K, Roseboom W, Arnold DH (2011b) Spatial grouping resolves ambiguity to drive temporal recalibration. *J Exp Psychol Hum Percept Perform* 37:1657–1661 Available at: <http://www.ncbi.nlm.nih.gov/pubmed/21688937> [Accessed July 2, 2013].

- Yu L, Stein BE, Rowland BA (2009) Adult plasticity in multisensory neurons: short-term experience-dependent changes in the superior colliculus. *J Neurosci* 29:15910–15922 Available at: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2824179&tool=pmcentrez&rendertype=abstract> [Accessed April 1, 2016].
- Yuan X, Bi C, Huang X (2015) Multiple concurrent temporal recalibrations driven by audiovisual stimuli with apparent physical differences. *Atten Percept Psychophys* Available at: <http://www.ncbi.nlm.nih.gov/pubmed/25772103> [Accessed March 25, 2015].
- Yuan X, Li B, Bi C, Yin H, Huang X (2012) Audiovisual temporal recalibration: space-based versus context-based. *Perception* 41:1218–1233 Available at: <http://www.ncbi.nlm.nih.gov/pubmed/23469702> [Accessed June 26, 2016].
- Zampini M, Guest S, Shore DI, Spence C (2005) Audio-visual simultaneity judgments. *Percept Psychophys* 67:531–544 Available at: <http://www.ncbi.nlm.nih.gov/pubmed/16119399> [Accessed July 2, 2013].
- Zampini M, Shore DI, Spence C (2003) Audiovisual temporal order judgments. *Exp Brain Res* 152:198–210 Available at: <http://www.ncbi.nlm.nih.gov/pubmed/12879178> [Accessed July 2, 2013].